EUROPEAN CHEMICALS AGENCY

GUIDANCE

Guidance on Information Requirements and Chemical Safety Assessment

Chapter R.7a: Endpoint specific guidance

Version 4.1

October 2015

Legal Notice

This document aims to assist users in complying with their obligations under the REACH Regulation. However, users are reminded that the text of the REACH Regulation is the only authentic legal reference and that the information in this document does not constitute legal advice. Usage of the information remains under the sole responsibility of the user. The European Chemicals Agency does not accept any liability with regard to the use that may be made of the information contained in this document.

Guidance on Information Requirements and Chemical Safety Assessment
Chapter R.7a: Endpoint specific guidance

Reference: ECHA-15-G-04.1-EN
ISBN: 978-92-9247-411-9
Publication date: October 2015
Language: EN

If you have questions or comments in relation to this document please send them (indicating the document reference, issue date, chapter and/or page of the document to which your comment refers) using the Guidance feedback form. The feedback form can be accessed *via* the ECHA Guidance website or directly *via* the following link:
https://comments.echa.europa.eu/comments_cms/FeedbackGuidance.aspx

European Chemicals Agency

Mailing address: P.O. Box 400, FI-00121 Helsinki, Finland

Visiting address: Annankatu 18, Helsinki, Finland

Preface

This document describes the information requirements under REACH with regard to substance properties, exposure, uses and risk management measures, and the chemical safety assessment. It is part of a series of guidance documents that are aimed to help all stakeholders with their preparation for fulfilling their obligations under the REACH Regulation. These documents cover detailed guidance for a range of essential REACH processes as well as for some specific scientific and/or technical methods that industry or authorities need to make use of under REACH.

The guidance documents were drafted and discussed within the REACH Implementation Projects (RIPs) led by the European Commission services, involving stakeholders from Member States, industry and non-governmental organisations. After acceptance by the Member States Competent Authorities the guidance documents had been handed over to ECHA for publication and further maintenance. Any updates of the guidance are drafted by ECHA and are then subject to a consultation procedure, involving stakeholders from Member States, industry and non-governmental organisations. For details of the consultation procedure, please see:

http://echa.europa.eu/documents/10162/13608/mb_63_2013_revision_consultation_procedure_guidance_en.pdf

The guidance documents can be obtained *via* the website of the European Chemicals Agency:

http://echa.europa.eu/guidance-documents/guidance-on-reach

This document relates to the REACH Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006[1].

[1] Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC (OJ L 396 of 30 December 2006, p. 1; corrected by OJ L 136, 29.5.2007, p. 3).

Version	Changes	Date
Version 1.0	First edition	May 2008
Version 2.0	Full revision of the Introduction and Section R.7.1 “Physicochemical properties” within Chapter R.7a: “Endpoint specific guidance” addressing structure and content. The Introduction and Section R.7.1 have been revised by updating, correcting or deleting mistakes and inconsistencies related to actual interpretation and application of generic aspects of the REACH Regulation (EC No 1907/2006) and the overall process for determining physicochemical information requirements in order to fulfil the registration requirements for a substance under the REACH Regulation. The content has been reworked with the aim to help registrants to establish a link between the REACH Regulation and the CLP Regulation (EC No 1272/2008) and guide them on how to comply with both of these Regulations when preparing a chemical safety assessment. As some physicochemical properties – notably explosive, flammable and oxidising properties – are intimately linked to physical hazards and there is thus a link between the physical hazards classification and the respective information requirements on explosive, flammable and oxidising properties it was decided to inclorporate the content of the former IR&CSA Guidance Chapter R.9: “Physico-chemical hazards” into relevant sub-sections of Section R.7.1 “Physicochemical properties” of the present document. The original Chapter R.9: “Physico-chemical hazards” of the IR&CSA Guidance will therefore be obsoleted when the present document is published. For the purposes of structuring the updated Guidance document according to CLP but nevertheless allowing the assignment to the respective information requirements of Annexes VII to XI to REACH, an updated and completely revised structure of Section R.7.1 has been implemented. Furthermore, to give the registrants further guidance when applying the general rules for adaptation of the standard testing regime set out in Annexes VII to X of the REACH Regulation a specific sub-section covering further guidance on this topic has been included in the revised text for every endpoint. Similarly an additional sub-section giving advice on how to provide Endpoint specific information in the registration dossier/IUCLID has been included in each relevant section. Information already covered by technical manuals, content falling under the scope of other guidance document or other internationally recognised recommendations has been removed and link to it has instead been provided. The update includes the following: • revision of section Introduction, by eliminating and amending out of date information. • revision of section R.7.1 Physicochemical properties, by reorganising the text in order to reflect the	November 2012

	Guidance structure update. The order of subsections has been modified and several sub-sections added if deemed necessary or deleted where information was identified as redundant. • Addition of a Table showing correlations between the Information requirements as specified in Annexes VII to IX to REACH and corresponding test methods according to the Test Method Regulation and CLP. • Complete revision of content and structure of sections R.7.1.2 – R.7.1.18. • Addition of new sections R.7.1.19 and R.7.1.20 in order that a link with new Appendices addressing recommendations for nanomaterials applicable to physicochemical properties could be established. • Addition of a new section R.7.1.21 in order to remind registrants which further information for classification and labelling in hazard classes of the substance in accordance with Article 10 (a) (iv) of REACH must be included in a REACH registration dossier. • Deletion of Appendices R.7.1-1 “Comments on thermodynamic consistency of physico-chemical properties”, R.7.1-2 “pH correction of partition coefficients for ionisable substances” and R.7.1-3 “Temperature correction” and an update of Appendix R.7.1-1 [before R.7.1-4] “Henry’s law and evaporation rate”.	
Version 2.1	Corrigendum covering the following: • Addition of a new footnote 8 on page 26 with a reference to a comprehensive review paper with the title: “QSPR prediction of physico-chemical properties for REACH” in sub-chapter R.7.1.1.3 Evaluation of available information on physicochemical properties.	August 2013
Version 2.2	Corrigendum correcting the page numbers within the reference in footnote 8 on page 26.	August 2013
Version 2.3	Corrigendum covering the following: • new formatting for the entirety of the R.7a guidance; • new pathfinder figure on the p.6; • addition of a title for a table R.7.1-2: ‘CLP Regulation hazard classes for which the REACH Regulation does not require the generation of information’; • a new footnote below tables R.7.1-1, R.7.1-2, R.7.1-7 and R.7.1-15 reminding the reader about changes introduced by the 4th ATP No 487/2013; • a new footnote in chapters R.7.1.10.1 and R.7.1.21.2 reminding the reader about changes introduced by the 4th ATP No 487/2013; • updated *Guidance on the Application of the CLP Criteria* references to reflect the changes of the Version 4.0 published in November 2013.	December 2013

Version 2.4	Corrigendum correcting a value for water density in chapter R.7.1.4.2 and a reference to REACH Annex in chapter R.7.1.16.6 and R.1.18.6.	February 2014
Version 3.0	Full revision addressing the content of sub-sections R.7.7.1 to R.7.7.7 related to Mutagenicity. The update includes the following: • Update of the information on non-testing methods in sub-section R.7.7.3.1, in particular with regard to the prediction models for mutagenicity and the OECD QSAR toolbox; • Update of the information on new/revised OECD test guidelines for genotoxicity testing in sub-section R.7.7.3.1, in particular with regard to the Transgenic rodent (TGR) somatic and germ cell gene mutation assays and the *in vivo* comet assay; • Amendment of sub-section R.7.7.4 on *Evaluation of available information on mutagenicity* based on the updated information on non-testing and testing methods; • Amendment of sub-section R.7.7.6 on *Integrated Testing Strategy (ITS) for mutagenicity* to take into account the new/revised OECD test guidelines for genotoxicity testing, in particular with regard to the recommended follow-up *in vivo* genotoxicity tests; • Clarification of the similarities and differences between this Guidance and other authoritative Guidance documents with regard to the recommended testing strategy for genotoxicity testing; • Clarification of the Registrant's obligation to submit a testing proposal to ECHA for any test mentioned in REACH Annex IX or X independendtly from the registered tonnage; • Clarification of the use of genotoxicity test results for Classification and Labelling; • Update of Figure R.7.7-1 on the recommended mutagenicity testing strategy in line with the amended Guidance text; • Update of table R.7.7-5 with addition of a missing title, insertion of a new row presenting a new example case, amendment of outdated information in line with the amended Guidance text; • Update of hyperlinks to ECVAM and ECVAM DB-ALM webpages in different sections across Chapter R.7a.	August 2014
Version 4.0	Minor revision and correction of the *Introduction* to Chapter R.7a to better reflect the structure of the updated sections of Chapter R.7a, in particular for the Human Health endpoints.	July 2015

	Update of two sections in Chapter R.7a: 1. Section R.7.2 *Skin corrosion/irritation, Serious eye damage/eye irritation, and Respiratory tract corrosion/irritation* Full revision addressing the content of Section R.7.2. The update includes the following: • Modification of Section R.7.2 structure and subdivision by endpoint: Skin corrosion/irritation (Sections R.7.2.2 to R.7.2.6), Serious eye damage/eye irritation (Sections R.7.2.7 to R.7.2.11) and Respiratory tract corrosion/irritation (Sections R.7.2.12 to R.7.2.14). • Update of the information on new/revised EU test methods and OECD test guidelines for skin corrosion/irritation and serious eye damage/eye irritation; • Update of the information on respiratory tract corrosion/irritation assessment; • Replacement of the terms "eye corrosion" by "serious eye damage" and "respiratory irritation" by "respiratory tract corrosion/irritation"; • Update of the information on non-testing methods, in particular in Appendices R.7.2-2 *QSARs and expert systems for skin corrosion and irritation* and R.7.2-3 *QSARs and expert systems for serious eye damage and eye irritation*; • Update of the recommended testing and assessment strategy for skin corrosion/irritation and serious eye damage/eye irritation in Sections R.7.2.6 and R.7.2.11 respectively; • Replacement of the terms "Integrated Testing Strategy (ITS)" by "testing and assessment strategy" to account for the non-testing part of the evaluation strategy; • Update of the information on Classification and Labelling to reflect changes arising from the 2nd and 4th Adaptations to Technical and Scientific Progress of the CLP Regulation, and to align the text with the revised Sections 3.2 *Skin corrosion/irritation* and 3.3 *Serious Eye damage/Eye irritation* of the *Guidance on the Application of the CLP Criteria* (version 4.0, November 2013). 2. Section R.7.6 *Reproductive toxicity* Full revision addressing the content of Section R.7.6. The update includes the following: • The new test method, the Extended One-Generation Reproductive Toxicity Study (EOGRTS), has been added to the Guidance text, including four new Appendices to support the text: o A checklist for information that contributes to EOGRTS design, o EOGRTS Study Design,	

<table>
<tr><td></td><td><li style="list-style-type:circle">o Premating exposure duration in EOGRTS;
o Evaluation of Triggers.
Update of the text on prenatal developmental toxicity (PNDT) (second species) following a decision of the Board of Appeal;
The entire section has been re-organised (within the overall structure of R7a) to present a more logical order for improved understanding and clarification;
Update of the section R.7.6.7 on Integrated Testing Strategy in line with the re-organised section and a new supporting Appendix for testing approaches and adaptations for Stage 3.</td><td></td></tr>
<tr><td>Version 4.1</td><td>Corrigendum covering the following:
Appendix R.7.6-2, point 4) Inclusion of Cohort 3, third bullet point on “(respiratory) sensitisation”: addition of text for clarification to avoid misinterpreation or misunderstanding. The point was discussed and agreed at the Partner Expert Group but the text accidentally omitted in drafting.</td><td>October 2015</td></tr>
</table>

Convention for citing the REACH and the CLP Regulations

Where the REACH and the CLP Regulations are cited literally, this is indicated by text in italics between quotes.

Table of Terms and Abbreviations

See Chapter R.20

Pathfinder

The figure below indicates the location of part R.7(a) within the Guidance Document

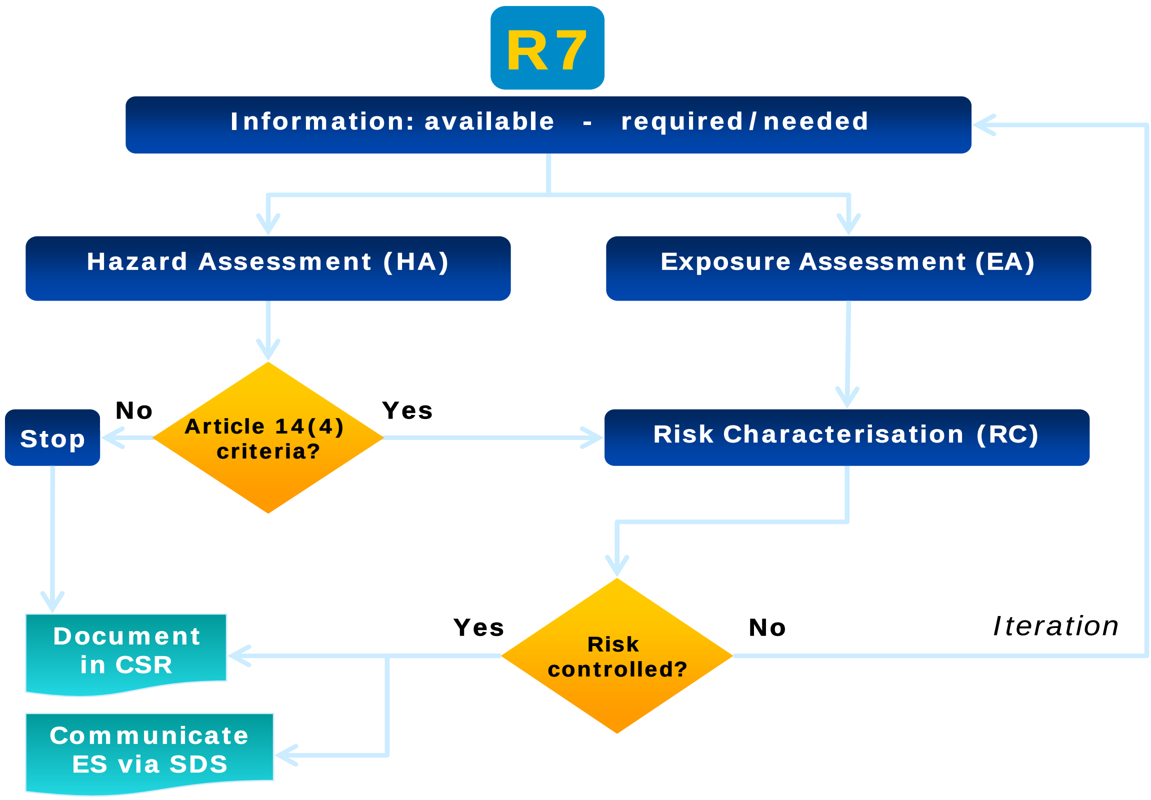


Table of Contents

R.7 ENDPOINT SPECIFIC GUIDANCE ... 18

Introduction ... 18

R.7.1 Physicochemical properties ... 26

R.7.1.1 Introduction on physicochemical properties ... 26
R.7.1.1.1 Information requirements on physicochemical properties ... 26
R.7.1.1.2 Available information on physicochemical properties ... 31
R.7.1.1.3 Evaluation of available information on physicochemical properties ... 32
R.7.1.1.4 Overall consistency of the physicochemical information ... 35
R.7.1.1.5 References for introduction of Physicochemical properties ... 35
R.7.1.2 Melting point/freezing point ... 36
R.7.1.2.1 Type of property ... 36
R.7.1.2.2 Definition ... 36
R.7.1.2.3 Test method(s) ... 36
R.7.1.2.4 Adaptation of the standard testing regime ... 36
R.7.1.2.5 Impurities; uncertainties ... 37
R.7.1.2.6 Endpoint specific information in the registration dossier/ IUCLID ... 37
R.7.1.3 Boiling point ... 38
R.7.1.3.1 Type of property ... 38
R.7.1.3.2 Definition ... 38
R.7.1.3.3 Test method(s) ... 39
R.7.1.3.4 Adaptation of the standard testing regime ... 39
R.7.1.3.5 Impurities; uncertainties ... 41
R.7.1.3.6 Endpoint specific information in the registration dossier / in IUCLID ... 41
R.7.1.4 Relative density ... 42
R.7.1.4.1 Type of property ... 42
R.7.1.4.2 Definition ... 42
R.7.1.4.3 Test method(s) ... 42
R.7.1.4.4 Adaptation of the standard testing regime ... 43
R.7.1.4.5 Impurities; uncertainties ... 44
R.7.1.4.6 Endpoint specific information in the registration dossier / in IUCLID ... 44
R.7.1.5 Vapour pressure ... 45
R.7.1.5.1 Type of property ... 45
R.7.1.5.2 Definition ... 45
R.7.1.5.3 Test method(s) ... 45
R.7.1.5.4 Adaptation of the standard testing regime ... 45
R.7.1.5.5 Impurities; uncertainties ... 48
R.7.1.5.6 Endpoint specific information in the registration dossier / in IUCLID ... 48
R.7.1.5.7 References on vapour pressure ... 48
R.7.1.6 Surface tension ... 50
R.7.1.6.1 Type of property ... 50
R.7.1.6.2 Definition ... 50
R.7.1.6.3 Test method(s) ... 50
R.7.1.6.4 Adaptation of the standard testing regime ... 50
R.7.1.6.5 Impurities; uncertainties ... 51
R.7.1.6.6 Endpoint specific information in the registration dossier / in IUCLID ... 51
R.7.1.7 Water solubility ... 52
R.7.1.7.1 Type of property ... 52
R.7.1.7.2 Definition ... 53
R.7.1.7.3 Test method(s) ... 53
R.7.1.7.4 Adaptation of the standard testing regime ... 55
R.7.1.7.5 Impurities; uncertainties ... 58
R.7.1.7.6 Endpoint specific information in the registration dossier / in IUCLID ... 58
R.7.1.7.7 References on water solubility ... 59
R.7.1.8 Partition coefficient n-octanol/water ... 61
R.7.1.8.1 Type of property ... 61
R.7.1.8.2 Definition ... 61
R.7.1.8.3 Test method(s) ... 61
R.7.1.8.4 Adaptation of the standard testing regime ... 63

R.7.1.8.5 Impurities; uncertainties........ 64
R.7.1.8.6 Endpoint specific information in the registration dossier / in IUCLID........ 66
R.7.1.9 Flash point........ 68
R.7.1.9.1 Type of property........ 68
R.7.1.9.2 Definition........ 68
R.7.1.9.3 Test method(s)........ 68
R.7.1.9.4 Adaptation of the standard testing regime........ 68
R.7.1.9.5 Impurities; uncertainties........ 70
R.7.1.9.6 Endpoint specific information in the registration dossier / in IUCLID........ 70
R.7.1.10 Flammability........ 71
R.7.1.10.1 Flammable gases........ 72
R.7.1.10.2 Flammable liquids........ 75
R.7.1.10.3 Flammable solids........ 79
R.7.1.10.4 Self-reactive substances and mixtures........ 82
R.7.1.10.5 Pyrophoric liquids........ 85
R.7.1.10.6 Pyrophoric solids........ 88
R.7.1.10.7 Self-heating substances and mixtures........ 90
R.7.1.10.8 Substances which in contact with water emit flammable gases........ 93
R.7.1.10.9 Organic peroxides........ 96
R.7.1.11 Explosive properties........ 100
R.7.1.11.1 Explosives........ 103
R.7.1.11.2 Self-reactive substances and mixtures........ 109
R.7.1.11.3 Organic peroxides........ 109
R.7.1.12 Self-ignition temperature........ 110
R.7.1.12.1 Auto-ignition........ 111
R.7.1.12.2 Self-heating substances........ 113
R.7.1.13 Oxidising properties........ 114
R.7.1.13.1 Oxidising gases........ 115
R.7.1.13.2 Oxidising liquids........ 117
R.7.1.13.3 Oxidising solids........ 121
R.7.1.14 Granulometry........ 125
R.7.1.14.1 Type of property........ 125
R.7.1.14.2 Definitions........ 126
R.7.1.14.3 Test methods........ 127
R.7.1.14.4 Adaptation of the standard testing regime........ 133
R.7.1.14.5 Impurities; uncertainties........ 134
R.7.1.14.6 Endpoint specific information in the registration dossier / in IUCLID........ 135
R.7.1.14.7 Further information / references........ 136
R.7.1.15 Adsorption/Desorption........ 137
R.7.1.15.1 Type of property........ 137
R.7.1.15.2 Definition........ 137
R.7.1.15.3 Test method(s)........ 138
R.7.1.15.4 Adaptation of the standard testing regime........ 139
R.7.1.15.5 Impurities;uncertainties........ 141
R.7.1.15.6 Endpoint specific information in the registration dossier/ in IUCLID........ 141
R.7.1.15.7 Further information/references........ 144
R.7.1.16 Stability in organic solvents and identity of relevant degradation products........ 146
R.7.1.16.1 Type of property........ 146
R.7.1.16.2 Definition........ 146
R.7.1.16.3 Test method(s)........ 147
R.7.1.16.4 Adaptation of the standard testing regime........ 147
R.7.1.16.5 Impurities; uncertainties........ 148
R.7.1.16.6 Endpoint specific information in the registration dossier / in IUCLID........ 148
R.7.1.16.7 Further information / references........ 149
R.7.1.17 Dissociation constant........ 149
R.7.1.17.1 Type of property........ 149
R.7.1.17.2 Definition........ 150
R.7.1.17.3 Test method(s)........ 151
R.7.1.17.4 Adaptation of the standard testing regime........ 151
R.7.1.17.5 Impurities; uncertainties........ 154
R.7.1.17.6 Endpoint specific information in the registration dossier / in IUCLID........ 155
R.7.1.17.7 References on dissociation constant........ 156
R.7.1.18 Viscosity........ 156

R.7.1.18.1 Type of property 156
R.7.1.18.2 Definition 156
R.7.1.18.3 Test method(s) 157
R.7.1.18.4 Adaptation of the standart testing regime 157
R.7.1.18.5 Impurities; uncertainties 158
R.7.1.18.6 Endpoint specific information in the registration dossier / in IUCLID 159
R.7.1.19 Shape 159
R.7.1.20 Surface area 160
R.7.1.21 Further information to be submitted for classification and labelling in hazard classes of the substance in accordance with article 10 (a) (iv) REACH 160
R.7.1.21.1 Flammable aerosols 161
R.7.1.21.2 Gases under pressure 161
R.7.1.21.3 Corrosive to metals 161

R.7.2 Skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation 169

R.7.2.1 Introduction 169
R.7.2.1.1 Definitions of skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation 169
R.7.2.1.2 Objective of the guidance on skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation 172
R.7.2.2 Information requirements on skin corrosion/irritation 173
R.7.2.2.1 Information requirements for quantities of ≥1 tpa (Annex VII to the REACH Regulation) 173
R.7.2.2.2 Information requirements for quantities of ≥10 tpa (Annex VIII to the REACH Regulation) 174
R.7.2.3 Information sources on skin corrosion/irritation 175
R.7.2.3.1 Non-human data on irritation/corrosion 175
R.7.2.3.2 Human data on skin corrosion/irritation 182
R.7.2.4 Evaluation of information on skin corrosion/irritation 182
R.7.2.4.1 Non-human data on skin corrosion/irritation 182
R.7.2.4.2 Human data on skin corrosion/irritation 188
R.7.2.4.3 Exposure considerations for skin corrosion/irritation 188
R.7.2.4.4 Remaining uncertainty on skin corrosion/irritation 189
R.7.2.5 Conclusions on skin corrosion/irritation 189
R.7.2.5.1 Concluding on suitability for Classification and Labelling 189
R.7.2.5.2 Concluding on suitability for Chemical Safety Assessment 189
R.7.2.5.3 Information not adequate 190
R.7.2.6 Testing and assessment strategy for skin corrosion/irritation 190
R.7.2.6.1 Objective / General principles 190
R.7.2.6.2 Testing and assesment strategy for skin corrosion/irritation 192
R.7.2.7 Information requirements for serious eye damage/eye irritation 201
R.7.2.7.1 Information requirements for quantities of ≥1 tpa (Annex VII to the REACH Regulation) 201
R.7.2.7.2 Information requirements for quantities of ≥10 tpa (Annex VIII to the REACH Regulation) 202
R.7.2.8 Information sources on serious eye damage/eye irritation 203
R.7.2.8.1 Non-human data on serious eye damage/eye irritation 203
R.7.2.8.2 Human data on serious eye damage/eye irritation 210
R.7.2.9 Evaluation of information on serious eye damage/eye irritation 211
R.7.2.9.1 Non-human data for serious eye damage/eye irritation 211
R.7.2.9.2 Human data on serious eye damage/eye irritation 217
R.7.2.9.3 Exposure considerations for serious eye damage/eye irritation 217
R.7.2.9.4 Remaining uncertainty on serious eye damage/eye irritation 217
R.7.2.10 Conclusions on serious eye damage/eye irritation 218
R.7.2.10.1 Concluding on suitability for Classification and Labelling 218
R.7.2.10.2 Concluding on suitability for Chemical Safety Assessment 218
R.7.2.10.3 Information not adequate 218
R.7.2.11 Testing and assessment strategy for serious eye damage/eye irritation 219
R.7.2.11.1 Objective / General principles 219
R.7.2.11.2 Testing and assessment strategy for serious eye damage/eye irritation 221
R.7.2.12 Information sources on respiratory tract corrosion/irritation 228

R.7.2.12.1 Animal data 228
R.7.2.12.2 Human data 228
R.7.2.13 Evaluation of information on respiratory tract corrosion/irritation 229
R.7.2.13.1 Animal data 229
R.7.2.13.2 Human data 230
R.7.2.14 Conclusions on respiratory tract corrosion/irritation 231
R.7.2.14.1 Concluding on suitability for Classification and Labelling 231
R.7.2.14.2 Concluding on suitability for Chemical Safety Assessment 231
R.7.2.15 References on skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation 232

R.7.3 Skin and respiratory sensitisation 259

R.7.3.1 Introduction 259
R.7.3.1.1 Definition of skin and respiratory sensitisation 259
R.7.3.1.2 Objective of the guidance on skin and respiratory sensitisation 259
R.7.3.1.3 Mechanisms of immunologically-mediated hypersensitivity 260
R.7.3.2 Information requirements for skin and respiratory sensitisation 261
R.7.3.3 Information for skin sensitisation and its sources 262
R.7.3.3.1 Non-human data for skin sensitisation 262
R.7.3.3.2 Human data on skin sensitisation 267
R.7.3.4 Evaluation of available information on skin sensitisation 267
R.7.3.4.1 Non-human data on skin sensitisation 268
R.7.3.4.2 Human data on skin sensitisation 272
R.7.3.5 Information and its sources on respiratory sensitisation 273
R.7.3.5.1 Non-human data on respiratory sensitisation 273
R.7.3.5.2 Human data on respiratory sensitisation 275
R.7.3.6 Evaluation of available information for respiratory sensitisation 276
R.7.3.6.1 Non-human data for respiratory sensitisation 276
R.7.3.6.2 Human data for respiratory sensitisation 277
R.7.3.7 Conclusions on skin and respiratory sensitisation 278
R.7.3.7.1 Remaining uncertainty on sensitisation 278
R.7.3.7.2 Concluding on suitability for Classification and Labelling 278
R.7.3.7.3 Additional considerations 280
R.7.3.7.4 Information not adequate 281
R.7.3.8 Integrated testing strategy (ITS) for sensitisation 281
R.7.3.8.1 Objective / General principles 281
R.7.3.8.2 Preliminary considerations 281
R.7.3.8.3 Testing strategies for sensitisation 281
R.7.3.9 References on skin and respiratory sensitisation 283

R.7.4 Acute toxicity 292

R.7.4.1 Introduction 292
R.7.4.1.1 Definition of acute toxicity 292
R.7.4.1.2 Objective of the guidance on acute toxicity 292
R.7.4.2 Information requirements for acute toxicity 293
R.7.4.3 Information and its sources on acute toxicity 293
R.7.4.3.1 Non-human data on acute toxicity 293
R.7.4.3.2 Human data on acute toxicity 298
R.7.4.3.3 Exposure considerations for acute toxicity 298
R.7.4.4 Evaluation of available information on acute toxicity 298
R.7.4.4.1 Non-human data on acute toxicity 298
R.7.4.4.2 Human data on acute toxicity 301
R.7.4.4.3 Exposure considerations on acute toxicity 302
R.7.4.4.4 Remaining uncertainty on acute toxicity 302
R.7.4.5 Conclusions on acute toxicity 302
R.7.4.5.1 Concluding on suitability for Classification and Labelling 302
R.7.4.5.2 Concluding on suitability for Chemical Safety Assessment 303
R.7.4.5.3 Information not adequate 304
R.7.4.6 Integrated Testing Strategy (ITS) for acute toxicity 305
R.7.4.6.1 Objective / General principles 305
R.7.4.6.2 Preliminary considerations 305
R.7.4.6.3 Testing strategy for acute toxicity (see Figure R.7.4–1) 306
R.7.4.7 References on acute toxicity 311

R.7.5 Repeated dose toxicity ... 313
R.7.5.1 Introduction ... 313
R.7.5.1.1 Definition of repeated dose toxicity ... 313
R.7.5.1.2 Objective of the guidance on repeated dose toxicity ... 314
R.7.5.2 Information requirements for repeated dose toxicity ... 314
R.7.5.3 Information and its sources on repeated dose toxicity ... 315
R.7.5.3.1 Non-human data on repeated dose toxicity ... 315
R.7.5.3.2 Human data on repeated dose toxicity ... 320
R.7.5.3.3 Exposure considerations on repeated dose toxicity ... 321
R.7.5.4 Evaluation of available information on repeated dose toxicity ... 321
R.7.5.4.1 Non-human data on repeated dose toxicity ... 321
R.7.5.4.2 Human data on repeated dose toxicity ... 331
R.7.5.4.3 Exposure considerations for repeated dose toxicity ... 331
R.7.5.4.4 Remaining uncertainty on repeated dose toxicity ... 333
R.7.5.5 Conclusions on repeated dose toxicity ... 335
R.7.5.5.1 Concluding on suitability for Classification and Labelling ... 336
R.7.5.5.2 Concluding on suitability for Chemical Safety Assessment ... 336
R.7.5.5.3 Information not adequate ... 337
R.7.5.6 Integrated Testing Strategy (ITS) for repeated dose toxicity ... 337
R.7.5.6.1 Objective / General principles ... 337
R.7.5.6.2 Preliminary considerations ... 337
R.7.5.6.3 Testing strategy for repeated dose toxicity ... 338
R.7.5.7 References on repeated dose toxicity ... 351

R.7.6 Reproductive toxicity ... 353
R.7.6.1 Introduction ... 353
R.7.6.2 Information requirements and testing approaches for reproductive toxicity ... 354
R.7.6.2.1 REACH information requirements ... 355
R.7.6.2.2 Key objectives and information produced by the test methods referred to in REACH 360
R.7.6.2.2.1 Reproduction/Developmental Toxicity Screening Test ... 360
R.7.6.2.2.2 Prenatal developmental toxicity study ... 360
R.7.6.2.2.3 Extended one-generation reproductive toxicity study ... 360
R.7.6.2.3 Adaptation and testing approaches ... 363
R.7.6.2.3.1 Overview ... 363
R.7.6.2.3.2 Procedure for adaptations and testing approaches ... 365
R.7.6.3 Information sources on reproductive toxicity ... 378
R.7.6.3.1 Information on reproductive toxicity from non-animal approaches ... 378
R.7.6.3.2 Information on reproductive toxicity in humans ... 379
R.7.6.3.3 Information on reproductive toxicity from *in vivo* animal studies ... 379
R.7.6.4 Evaluation of available information for reproductive toxicity ... 381
R.7.6.4.1 Non-animal data ... 381
R.7.6.4.1.1 Physico-chemical properties ... 382
R.7.6.4.1.2 (Q)SAR ... 382
R.7.6.4.1.3 *In vitro* data and Adverse Outcome Pathways (AOPs) ... 382
R.7.6.4.2 Animal data ... 384
R.7.6.4.2.1 Reproduction/developmental toxicity screening test ... 384
R.7.6.4.2.2 Prenatal developmental toxicity study ... 386
R.7.6.4.2.3 Extended one-generation reproductive toxicity study ... 387
R.7.6.4.2.4 Two-generation reproductive toxicity study ... 391
R.7.6.4.2.5 One-generation reproductive toxicity study ... 392
R.7.6.4.2.6 Developmental neurotoxicity studies ... 393
R.7.6.4.2.7 Developmental immunotoxicity studies ... 394
R.7.6.4.2.8 Repeated-dose toxicity studies ... 395
R.7.6.4.2.9 *In vivo* assays for endocrine disruption mode of action ... 396
R.7.6.4.3 Human data on reproductive toxicity ... 397
R.7.6.4.4 Derivation of DNELs and DMELs ... 398
R.7.6.5 Classification and labelling ... 398
R.7.6.6 Conclusions on reproductive toxicity ... 399
R.7.6.7 Integrated Testing Strategy (ITS) for reproductive toxicity ... 399
R.7.6.8 References on reproductive toxicity ... 400

R.7.7 Mutagenicity and carcinogenicity ... 436

R.7.7.1 Mutagenicity 436
R.7.7.1.1 Definition of mutagenicity 436
R.7.7.1.2 Objective of the guidance on mutagenicity 436
R.7.7.2 Information requirements on mutagenicity...437
R.7.7.3 Information and its sources on mutagenicity 439
R.7.7.3.1 Non-human data on mutagenicity 439
R.7.7.3.2 Human data on mutagenicity... 445
R.7.7.4 Evaluation of available information on mutagenicity...445
R.7.7.4.1 Non-human data on mutagenicity 446
R.7.7.4.2 Human data on mutagenicity... 448
R.7.7.4.3 Remaining uncertainty on mutagenicity 448
R.7.7.5 Conclusions on mutagenicity...448
R.7.7.5.1 Concluding on Classification and Labelling... 448
R.7.7.5.2 Concluding on suitability for Chemical Safety Assessment 448
R.7.7.5.3 Information not adequate 450
R.7.7.6 Integrated Testing Strategy (ITS) for mutagenicity..450
R.7.7.6.1 Objective / General principles... 450
R.7.7.6.2 Preliminary considerations 451
R.7.7.6.3 Testing strategy for mutagenicity.. 453
R.7.7.7 References on mutagenicity 467
R.7.7.8 Carcinogenicity 470
R.7.7.8.1 Definition of carcinogenicity 470
R.7.7.8.2 Objective of the guidance on carcinogenicity 471
R.7.7.9 Information requirements on carcinogenicity 471
R.7.7.10 Information and its sources on carcinogenicity 472
R.7.7.10.1 Non-human data on carcinogenicity 472
R.7.7.10.2 Human data on carcinogenicity.. 477
R.7.7.10.3 Exposure considerations for carcinogenicity.. 478
R.7.7.11 Evaluation of available information on carcinogenicity..................................479
R.7.7.11.1 Non-human data on carcinogenicity 479
R.7.7.11.2 Human data on carcinogenicity.. 483
R.7.7.11.3 Exposure considerations for carcinogenicity.. 484
R.7.7.11.4 Remaining uncertainty on carcinogenicity 484
R.7.7.12 Conclusions on carcinogenicity...485
R.7.7.12.1 Concluding on suitability for Classification and Labelling 485
R.7.7.12.2 Concluding on suitability for Chemical Safety Assessment 485
R.7.7.12.3 Information not adequate 486
R.7.7.13 Integrated Testing Strategy (ITS) for carcinogenicity....................................487
R.7.7.13.1 Objective / General principles... 487
R.7.7.13.2 Preliminary considerations 488
R.7.7.13.3 Testing strategy for carcinogenicity... 490
R.7.7.14 References on carcinogenicity 493

Tables

Table R.7.1–1 Information requirements as specified in Annexes VII to IX to REACH and corresponding tests methods according to the Test Method Regulation and CLP 28
Table R.7.1–2 CLP Regulation hazard classes for which the REACH Regulation does not require the generation of information 31
Table R.7.1–3 Test methods for determining density........ 42
Table R.7.1–4 Group contribution approach and vapour pressure 47
Table R.7.1–5 Test methods for the determination of water solubility........ 55
Table R.7.1–6 Methods for determination of partition coefficient n-octanol/water 62
Table R.7.1–7 Assignment of CLP hazard classes to the information requirement ‘Flammability’ according to REACH, Annex VII and correlation between the Test Method Regulation and the test method according to CLP and supporting link with the *Guidance on the Application of the CLP criteria*. 71
Table R.7.1–8 Assignment of CLP hazard classes to the information requirement ‘Explosive properties’ according to REACH, Annex VII and correlation between the Test method Regulation and the test method according to CLP and supporting link with the *Guidance on the Application of the CLP criteria*. 102
Table R.7.1–9 Assignment of CLP hazard classes to the information requirement ‘Self ignition temperature’ according to REACH, Annex VII and the Test Method Regulation. 110
Table R.7.1–10 Assignment of CLP hazard classes to the information requirement ‘Oxidising properties’ according to REACH, Annex VII and correlation between the Test method Regulation and the test method according to CLP and supporting link with the *Guidance on the Application of the CLP criteria*. 114
Table R.7.1–11 Methods to determine particle size distribution of a material........ 129
Table R.7.1–12 Methods to generate/sample airborne dispersed or nebulised particles....... 131
Table R.7.1–13 Methods that measure inhalable fractions only or that give no detailed distributions........ 133
Table R.7.1–14 Methods for the measurement of adsorption........ 138
Table R.7.1–15 Information to be submitted for general registration purposes according to Article 10 (a) (iv) REACH, CLP hazards classes and corresponding tests methods according to the Test Method Regulation and CLP 160
Table R.7.1–16 Experimental approaches for the determination of HLC 164
Table R.7.1–17 Conditions that have influence on HLC values 165
Table R.7.2–1 Overview of available (Q)SARs for skin corrosion/irritation. See Appendix R.7.2–2 for more information on these models. 178
Table R.7.2–2 Accepted *in vitro* test methods for skin corrosion/irritation 180
Table R.7.2–3 Overview of available (Q)SARs for serious eye damage/eye irritation. See Appendix R.7.2–3 for more information on these models. 205
Table R.7.2–4 Accepted *in vitro* test methods for serious eye damage/eye irritation........ 207
Table R.7.2–5 Categories of irritant substances and their typical mode of action in eye irritation. 240
Table R.7.2–6 Available literature-based models for skin corrosion/irritation. 245
Table R.7.2–7 Available literature-based models for serious eye damage/eye irritation....... 254
Table R.7.3–1 Examples of structural alerts for respiratory sensitisation........ 274
Table R.7.5–1 Overview of other *in vivo* test guideline studies giving information on repeated dose toxicity 319
Table R.7.5–2 Overview of *in vivo* repeated dose toxicity test guideline studies........ 327
Table R.7.5–3 Methods for investigation of neurotoxicity 350
Table R.7.6–1 Summary of information requirements for reproductive toxicity in REACH (Annexes VII to X). 359
Table R.7.6–2 Overview of *in vivo* EU test methods and OECD test guidelines for reproductive toxicity referred to in REACH 362

Table R.7.7–1 REACH information requirements for mutagenicity 438
Table R.7.7–2 *In vitro* test methods.. 442
Table R.7.7–3 *In vivo* test methods, somatic cells 443
Table R.7.7–4 *In vivo* test methods, germ cells 444
Table R.7.7–5 Examples of different testing data sets and follow-up procedures to conclude on genotoxicity/mutagenicity according to the mutagenicity testing strategy. 461
Table R.7.7–6 Standard information requirements for carcinogenicity and the specific rules for adaptation of these requirements.. 472

Figures

Figure R.7.1–1 Comparison of the DSD and the CLP classification 78
Figure R.7.1–2 Example 2,2'-Azodi (isobutyronitrile)... 85
Figure R.7.1–3 Example: Di-tert-butyl peroxide... 99
Figure R.7.1–4 Results from application of the class 1 acceptance procedure 107
Figure R.7.1–5 Results from the application of the class 1 assignment procedure 108
Figure R.7.1–6 Integrated testing strategy for granulometry....................................... 128
Figure R.7.1–7 Integrated testing strategy for dissociation constant 152
Figure R.7.2–1 Overview of the testing and assessment strategy for skin corrosion/irritation 192
Figure R.7.2–2 Testing and assessment strategy for evaluating the skin corrosion/irritation potential of substances (footnotes a to h are detailed below the figure). 194
Figure R.7.2–3 Schematic presentation of Top-down and Bottom-up approaches for Skin Corrosion/irritation.. 199
Figure R.7.2–4 Overview of the testing and assessment strategy for serious eye damage/eye irritation 220
Figure R.7.2–5 Testing and assessment strategy for evaluating the serious eye damage/eye irritation potential of substances (footnotes a to f are detailed below the figure). 222
Figure R.7.3–1 Integrated testing strategy for skin sensitisation................................... 282
Figure R.7.3–2 Integrated evaluating strategy for respiratory sensitisation data* 283
Figure R.7.4–1 ITS for acute toxicity endpoint.. 309
Figure R.7.4–2 ITS for acute inhalation toxicity endpoint (see also draft OECD GD 39)....... 310
Figure R.7.5–1 Integrated Testing Strategy for repeated dose toxicity............................ 339
Figure R.7.7–1 Flow chart of the mutagenicity testing strategy 460
Figure R.7.7–2 Integrated Testing Strategy for carcinogenicity...................................... 492

Appendices

Appendix R.7.1–1 Henry's law constant and evaporation rate.. 163
Appendix R.7.2–1 Mechanisms of local toxicities: skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation 238
Appendix R.7.2–2 (Q)SARs and expert systems for skin corrosion and irritation 244
Appendix R.7.2–3 (Q)SARs and expert systems for serious eye damage and eye irritation.. 253
Appendix R.7.5–1 Testing strategy for specific system/organ toxicity............................ 346
Appendix R.7.6–1 A check list for information that contributes to EOGRTS design............. 406
Appendix R.7.6–2 EOGRTS Study Design 408
Appendix R.7.6–3 Premating exposure duration in the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) 417
Appendix R.7.6–4 Procedure for testing approaches and adaptation; Stage 3 - Stages 3.1.1 – 3.1.8.. 424
Appendix R.7.6–5 Evaluation of triggers 429

R.7 Endpoint specific guidance

Introduction

The previous sections of the *Guidance on information requirements and chemical safety assessment* (IR&CSA) provide advice on the interpretation and application of generic aspects of the Regulation describing the overall process that should be followed in finding, assembling and evaluating all the relevant information that is required for the registration of a chemical under Regulation (EC) No 1907/2006 (the REACH Regulation). The chapters also describe factors that may have an influence on the information requirements and give advice on how the information collected from different sources could be integrated and used in an approach to allow a conclusion on whether or not the available information is sufficient for regulatory purposes, i.e. hazard assessment and risk assessment.

Under Regulation (EC) No 1272/2008 (CLP Regulation or CLP), this approach is called a *Weight-of-Evidence* (WoE) determination. According to CLP, an evaluation by applying WoE determination (i.e. all available information relevant for the evaluation of the specific hazard is considered together) using expert judgement, must always be carried out where the criteria cannot be applied directly (Article 9(3), CLP). This WoE determination should not be confused with the use of Weight of Evidence according to Annex XI, 1.2 of REACH, an adaptation rule for standard information requirements where sufficient weight of evidence may allow the conclusion/assumption that a substance has or has not a particular dangerous property.

The guidance given thus far is applicable across the field and comprises the general rules that should be followed.

Structure of Chapter R.7a

In this chapter, specific guidance on meeting the information requirements set out in Annexes VI to XI to the REACH Regulation is provided. The information requirements relate both to those physicochemical properties that are relevant for exposure and fate considerations as well as to physical hazards, human health hazards and environmental hazards. The guidance for each specified property or hazard has been developed as a specific “sub-chapter” (referred to as a “Section”) in this guidance, addressing the aspects of collection, generation and evaluation of information to help registrants provide adequate and relevant information for registration under REACH.

All data sources, including non-testing data, have to be taken into account when doing the chemical safety assessment. Most of the reports follow a logical common format that complements the generic guidance and the general decision making frameworks detailed in the first paragraph above.

R.7.1 Physicochemical properties

This first “sub-chapter”, underwent a guidance revision process between 2011 and 2012 and therefore follows a revised structure. The Section R.7.1 covers both classification and non-classification related properties, where the sections covering the physicochemical properties each have six or seven “sub-sections” (also referred to as “sections”), depending on the need for information on references and the sections covering the physical hazards have seven “sub-sections” (also referred to as “sections”).

In the physicochemical properties sections:

- the first section details the type of property;
- the second section provides the definition of the property;
- the third section lists the preferred test method(s);

- the fourth section deals with adaptation of the standard testing regime, namely adaptation options that can be explored under each specific physicochemical property;
- the fifth section deals with impurities and uncertainties, and
- the last section outlines what kind of property-specific information should be given in the registration dossier (note that sometimes an additional section is added where relevant references are provided).

By contrast the physical hazard sections:

- start with the definition section;
- followed by a second section on classification criteria and relevant information;
- the third section explores various adapation options, namely how the standard testing regime can be adapted;
- the fourth section outlines the impurities and uncertainties;
- the fifth section aims to help in concluding on the Directive 67/548/EEC (Dangerous Substances Directive - DSD) classification, repealed by Regulation (EC) No 1272/2008 (CLP Regulation or CLP);
- the sixth section outlines the physical hazards-specific information to be included in the registration dossier and in IUCLID, and
- the seventh section gives relevant further information and used references.

R.7.2 Human health properties or hazards

Chapters tackling human health properties or hazards in R.7a remain generally unchanged, using a similar structure. However as each section is updated the information may be re-organised to be presented in a clearer and more constructive order. In these chapters there are seven main sections to the guidance on each property or hazard:

- the introduction section (R.7.X.1 Introduction) provides an introduction in which the property or hazard is described, further defined and an explanation given as to its importance in the context of human health, or environmental fate and effect of a given substance;

- the second section (R.7.X.2 Information requirements …) details the specific information requirements for the endpoint of interest; these will depend on the tonnage band of the substance, its usage pattern and other considerations including data on other endpoints and on related substances. Endpoint[2] specific guidance can be thought of as logical steps that should be taken to assemble the information that is detailed under the second section; thus

- the third section (R.7.X.3 Information sources on…) provides an inventory of all the types of data that could potentially provide useful information on the endpoint of interest and, most importantly the sources of that information;

---

[2] REACH uses the term “endpoint” both to denote a physicochemical property (example: Annex VII to REACH, Column 1 standard information required: 7.3 Boiling point, and 7.4 Relative density) and to denote hazardous properties (example: Annex VII to REACH, Column 1 standard information required: 7.11 Explosive properties and 7.13 Oxidising properties), which are subject to classification according to the applicable EU legislation. In the following, the wording of Part 7(a) of this guidance document will differentiate between these different types of properties where this appears appropriate, in order to facilitate the identification of properties which serve the regulatory purpose of classification.

- in the fourth section (R.7.X.4 Evaluation of available information for…) guidance is given on how to evaluate the information that might be available for a given substance; this advice focuses on providing the criteria to aid in the judgement and ranking of the available data for their adequacy and completeness. This section also provides an indication of the remaining uncertainty inherent in the different types of data for the given endpoint;

- the fifth section (R.7.X.5 Conclusions on…) describes how conclusions may be drawn for a given substance on the suitability of the available information for regulatory purposes. Chemical safety assessment within REACH is fundamentally dependent on an adequate conclusion on classification and PBT/vPvB assessment since exposure assessment and risk characterisation are triggered by classification and fulfilment of PBT/vPvB criteria. Therefore data need to be adequate for both classification & labelling and for chemical safety assessment if the latter is required;

- the sixth section (R.7.X.6 Integrated Testing Strategy (ITS) for…) comprises an Integrated Testing Strategy (ITS), also referred to in some sections as Testing and assessment strategy, for the given endpoint(s), providing guidance on how to define and generate relevant information on substances in order to meet the requirements of REACH. It is noteworthy that all experiments using vertebrate animals shall be designed to avoid distress and unnecessary pain and suffering to experimental animals, in accordance with Article 7(4) of Directive 86/609/EEC.

  The proposed testing strategies are guidance for data generation in a stepwise approach. The strategies build on the concept that if the available information is not sufficient to meet the regulatory needs, further gathering of information at a succeeding step in the testing strategies is needed. On the other hand, if the available information is adequate and the standard information requirements are met, no further gathering of information is necessary. Standard information requirements will not need to be fulfilled by standard tests, where the available information is judged to be sufficient to adapt the standard information requirements in accordance with REACH Annex XI or an applicable Column 2 provision of REACH Annexes VII to X;

- the seventh and final section (R.7.X.7 References on…) lists all used references on the given endpoints.

Additional considerations

The following additional considerations apply generally to the endpoint specific guidance given in this chapter:

Information requirements in the light of the applicable classification regime

The main regulatory purpose of the information requirements set out in Annexes VI to X to the REACH Regulation is to assess hazards and risks related to substances and to develop and recommend appropriate risk management measures, as highlighted in Recital 19 of the REACH Regulation. According to Recital 26: '*in order to undertake chemical safety assessments of substances effectively, manufacturers and importers of substances should obtain information on these substances, if necessary by performing new tests*'. The chemical safety assessment (CSA) should be performed in accordance with the provisions set out in Annex I to the REACH Regulation. According to Section 0.6 of Annex I, the first three steps of the CSA require the carrying out of a human health hazard assessment, a human health hazard assessment of physicochemical properties and an environmental hazard assessment, including determining the classification of substances. When the REACH Regulation was adopted, the DSD was the applicable classification regime (see, more in particular, the transitional provisions set out in Article 61 of Regulation (EC) No 1272/2008). Accordingly, many REACH information requirements are inspired by the categories of danger under DSD such as points 7.10, 7.11

and 7.13 in Column 1 of REACH Annex VII (*i.e.* flammability, explosive properties and oxidising properties, respectively).

On 20 January 2009 Regulation (EC) No 1272/2008 (CLP Regulation or CLP) entered into force. The CLP Regulation has amended certain parts of the REACH Regulation (see Article 58 of CLP for amendments applicable from 1 December 2010 and Article 59 of CLP for amendments applicable from 1 June 2015). Nevertheless, the terminology used in REACH currently still comprises terms which were used under the DSD (for substances) and still apply (for mixtures until 1 June 2015) under Directive 1999/45/EC (Dangerous Preparations Directive – DPD). With respect to the updated physicochemical part of this guidance and the section dealing with the exploration of adaptation possibilities of the standard testing regime, the term 'dangerous' can be interpreted in a broader context (particularly, in certain contexts within this document, to include 'hazardous' as defined under CLP) as it does not refer strictly to the DSD.

According to the requirements of Article 10(a)(iv) of the REACH Regulation, the technical dossier required for registration purposes includes the classification and labelling of the substance as specified in Section 4 of Annex VI to REACH, resulting from the application of Titles I and II of the CLP Regulation. From 1 December 2010 until 1 June 2015 substances must be classified in accordance with both DSD and CLP and they must be labelled and packaged in accordance with CLP (Article 61(3) of CLP). Similarly, until 1 June 2015 Safety Data Sheets (SDS) must include information on classifications according to both CLP and DSD for substances and component substances in mixtures until 1 June 2015 (see updates to REACH *via* Commission Regulation (EU) No 453/2010 and the ECHA *Guidance on the compilation of Safety Data Sheets*.

Use of data derived from EU or other international standardised test methods

For the purposes of determining whether any of the physical hazards referred to in Part 2 of Annex I of CLP apply to a substance (or a mixture), the manufacturer, importer or downstream user must perform the tests required by the above mentioned Part 2, unless there is adequate and reliable information available (see Article 8(3) of CLP). Further in this guidance for each relevant physical hazard a reference to the corresponding test according to UN Recommendations on the Transport and Dangerous Goods, Manual of Test and Criteria (UN-MTC), starting with a UN test method name will be provided.

According to Article 8(5) of CLP, where new tests for **physical hazards** are carried out for classification and labelling purposes, they must be performed in compliance with a relevant recognised quality system (e.g. GLP) or by laboratories complying with a relevant recognised standard (e.g. with EN ISO/IEC 17025), at the latest from January 2014.

For the purpose of determining whether a substance or mixture fulfils the criteria for classification in any of the **human health and/or environmental hazard classes** (and differentiations within a hazard class, if applicable), there is no similar testing requirement. If there is already adequate and reliable information available (see Article 8(2) of CLP), this must be used. Provided that the manufacturer, importer or downstream user has exhausted all other means of generating information, new tests may however be performed (Article 8(1), CLP).

Where new tests for **human health or environmental hazards** are carried out for classification purposes, they must be performed in compliance with a relevant recognised quality system (e.g. GLP) or by laboratories complying with a relevant recognised standard (e.g. with EN ISO/IEC 17025), at the latest from January 2014. (Article 8(5), CLP). Further requirements for tests performed for the purpose of CLP are given in Article 8 of CLP.

Further, according to Article 13(3) of REACH, tests for generating information on intrinsic properties of substances must be conducted in accordance with the test methods laid down in

Commission Regulation (EC) 440/2008 (Test Methods Regulation)[3] or in accordance with other international test methods recognised by the Commission or the Agency as being appropriate, such as European Standards (EN) (http://www.cen.eu/Pages/default.aspx) or the OECD guidelines (http://www.oecd.org/chemicalsafety/testing/oecdguidelinesforthetestingofchemicals.htm). Regulation (EC) 440/2008 lays down the test methods to be applied for the purposes of REACH. Thus, in the following sections on specific endpoints, references given for each test method will include the OECD Test Guideline (TG) number and, where available, the test method number, as defined in the Test Methods Regulation.

According to Recital 37 of the REACH Regulation, if tests are performed, they should comply with the relevant requirements for protection of laboratory animals, as set out in Council Directive 86/609/EEC[4]. Article 13(4) of REACH states that ecotoxicological and toxicological tests and analyses must be carried out in compliance with the principles of good laboratory practice (GLP) provided for in Directive 2004/10/EC[5] or other international standards recognised as being equivalent by the Commission or the Agency and with the provisions of Council Directive 86/609/EEC, if applicable.

Interdependence of endpoints in hazard assessment

Although guidance is provided for each specific endpoint separately, it should be remembered that different endpoints are related to each other. Information collected within one endpoint may influence hazard/risk assessment of other endpoints, for example, information on rapid primary degradation of a parent substance may result in including the degradation products in the overall assessment of the toxicity of a substance. Regarding the physicochemical properties of a substance, for example boiling point and flash point are properties used for the classification of flammable liquids, and therefore these properties are important for physical hazard assessment. Similarly, information on toxicity/specific mode of action in one endpoint may indicate possible adverse effects for organisms considered for assessment of other endpoints, for example, endocrine disrupting mode of action in mammals may indicate the same mode of action in fish. Another example may be when data on toxic effects measured in one group of organisms may be directly used in more than one endpoint, for example, data from a repeated dose toxicity study may also be used in assessment of risk for secondary poisoning of mammals exposed *via* the food chains.

Adequacy of methods for generating additional information

Before (proposing) additional animal testing, use of all other options should be considered. It is important to emphasise that testing on vertebrate animals must only be conducted or proposed as a last resort, when all other data sources have been exhausted (see Recital 47 of the REACH Regulation, Article 25 of REACH and Step 4 of REACH Annex VI). Therefore, it is

[3] Council Regulation (EC) No 440/2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) [OJ L 142, 31.5.2008, p. 1].

[4] Council Directive of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes (86/609/EEC).

[5] Directive 2004/10/EC of the European Parliament and of the Council of 11 February 2004 on the harmonisation of laws, regulations and administrative provisions relating to the application of the principles of good laboratory practice and the verification of their applications for tests on chemical substances.

important to first consider all issues that may impact upon this decision whether and how to perform the testing, such as:

- applicable information requirements pursuant to REACH;
- adaptation possibilities of REACH Annex XI and Column 2 of REACH Annexes VII to X, e.g.:
  - classifications that may allow for adaptations,
  - available data on a category, a group or on individual substances for which the physicochemical and toxicological properties are likely to be similar,
  - assumption/conclusion on presence or absence of a particular dangerous property of a substance in a weight of evidence approach based on animal or human data,several independent sources,
  - absence or no significant exposure based on exposure scenarios;
- substance properties;
- available *in vitro* and *in vivo* data;
- available toxicokinetic and toxicodynamic information;
- any trigger/alert that may require testing going beyond the applicable minimum information requirements.

All these issues should be considered, not only to design fit-for-purpose *in vivo* tests, but also for justifying why an *in vivo* study is not needed under certain circumstances. Animal tests must comply with the provisions laid down in Council Directive 86/609/EEC[6].

Degradation products and metabolites

In the context of evaluating substances for their effects, it is important to note that, once released into the environment or taken up by animals, a substance may be transformed through degradation or metabolism. These processes and their outcome may need to be taken into account in the overall assessment.

Degradation products may be formed as a result of transformation processes in the environment, either biotic or abiotic. For distinguishing the substance undergoing degradation from the degradation products, the former is often referred to as the parent substance.

Degradation products may be formed as a result of abiotic environmental processes such as hydrolysis, direct or indirect photolysis or oxidation. They may also be formed as a result of aerobic or anaerobic biodegradation, i.e. due to microbial activity. Degradation products require further investigation if the Chemical Safety Assessment indicates the need, i.e. if stable degradation products are formed in the environment within a relevant time frame, as deduced from the test system, or if they fulfil the PBT/vPvB criteria. Likewise it may be considered to assess whether degradation products fulfil the environmental hazard classification criteria (see Section R.7.9 in Chapter R.7b of the *Guidance on IR&CSA*).

Metabolites refer to transformation products, which are formed due to biodegradation (and then the term metabolite is synonymous with the term biodegradation product) or formed as a result of biotransformation (metabolism) within exposed organisms after uptake of the parent substance. Metabolic pathways and hence the identity of metabolites may or may not be fully known. The latter is frequently the case. Moreover for the same substances metabolic

---

[6] Council Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes [OJ L 358, 18.12.1986, p. 1].

pathways may or may not differ between various organisms belonging to different phyla and/or trophic levels. However, the toxicity of metabolites formed within the duration of laboratory tests will be reflected by their parent substance, with the exception of delayed effects which are only evident after the observation time of the tests. Knowledge of metabolic pathways and metabolites may increase planning and focussing of toxicity testing and understanding of toxicological findings (see Section R.7.12 in Chapter R.7c of the Guidance on IR&CSA). Therefore, in some cases it may be possible to use grouping approaches for structurally closely-related substances, which undergo similar metabolic transformation (see Section R.6.2 in Chapter R.6 *Guidance on QSARs and grouping of substances* of the Guidance on IR&CSA).

When biotransformation processes include oxidation, metabolites are often less hydrophobic than the parent substance. This is a very general rule of thumb and may not always apply; however, when it does, often this has implications for the hazard profile of the metabolites. For example more polar metabolites created after oxidation processes have normally a lower adsorption potential, and thus the relevance of the metabolites for the soil and sediment compartments is normally lower than that of the parent substance. Such less hydrophobic metabolites also tend to be excreted more rapidly from organisms than the parent substance. Hence both their bioaccumulative potential and narcotic toxicity tend to be lower.

Similarities in metabolic pathways of structurally-related substances may serve as an indication for waiving for further investigation, depending on the case and nature of the metabolites.

It should be noted that metals, and in particular metal substances, do not degrade in the environment in the same way as organic substances. They transform usually through dissolution to the dissolved form.

Selection of the appropriate route of administration for toxicity testing

Having established the need for additional toxicity testing to meet the requirements of REACH for a given substance, for certain endpoints, notably acute or repeated dose toxicity but also reproductive toxicity, chronic toxicity and carcinogenicity, a decision must be made on which route(s) of exposure is(are) most appropriate. The overall objective of such testing is to determine the potential hazard of the test substance to human beings. Humans may normally be exposed to substances by one or more of three routes: inhalation, dermal and oral. In general, the final decision on which route of exposure is to be considered in a particular test should be taken in the light of the requirements for the particular endpoint concerned, the recommendation given in the respective test methods, all available information including physicochemical properties of the substance, human exposure, structure-activity relationships (SAR) or the data from available toxicity tests on the substance itself.

If no adequate experimental effect data using the relevant route of administration is available, route-to-route extrapolation might be an alternative method for evaluating the hazard. However this approach should only be used for systemic effects, and not for local effects such as irritation of the lungs following inhalation of a substance. Route-to-route extrapolation is recommended only under conditions where route-specific effects are not expected. Therefore, route-to-route extrapolation should be considered on a case-by-case basis taking into account the additional uncertainties. It is to be noted that route-to-route extrapolation is associated with a high degree of uncertainty and should be conducted with caution relying on expert judgement. In a subsequent risk assessment the uncertainties introduced through route-to-route extrapolation should be taken into account, for example by adjusting the assessment factor in the determination of the DNEL (see Section R.8.4.3, Chapter R.8 *Characterisation of dose [concentration]-response for human health* of the Guidance on IR&CSA). Further guidance on this strategic approach to toxicity testing is given in Chapter R.8 of the Guidance on IR&CSA.

Assessment of the environmental impact of a substance

With regard to the evaluation of the environmental impact of a substance, the interaction of that substance with the environment is an important consideration. The fate and behaviour of a substance are largely governed by its inherent physicochemical properties. The knowledge of the physicochemical properties of the substance, together with results from multimedia fate and transport models (e.g. Mackay level 3 models), enables the identification of the environmental compartment(s) of primary concern. Such information will also determine the prioritisation of higher tiered tests. More extensive guidance and considerations on this aspect are given in Chapter R.16 *Environmental Exposure Estimation*.

R.7.1 Physicochemical properties

Advice to registrants with regard to nanomaterials characterisation can be found in *Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the *Guidance on IR&CSA*, section 2 Recommendation for physicochemical properties arising from RIP-oN 2 for nanomaterials.

R.7.1.1 Introduction on physicochemical properties

According to Article 12 of the REACH Regulation, for registration purposes all physicochemical information that is relevant and available to the registrant must be included in the technical dossier, i.e. information such as:

- data on intrinsic properties of the substance (e. g. melting point/freezing point, boiling point, vapour pressure, density);
- data necessary to assess the physical hazards of a substance (e. g. flammability), with the view to determine its classification and labelling according to CLP (and until 1 June 2015, according to DPD, see Article 61 of CLP);
- supplementary data for hazard assessment and health and environmental classification (e. g. viscosity, n-octanol/water partition coefficient).

Some physicochemical properties - notably explosive, flammable and oxidising properties - are intimately linked to physical hazards. The most straight-forward way of assessing these properties is through the classification of the substance for the corresponding physical hazards. There is thus a link between the physical hazards classification and the information requirements on explosive, flammable and oxidising properties. This is further elaborated below (see Table R.7.1–1) and in the various chapters addressing these endpoints. For substances manufactured or imported in quantities of 100 tonnes or more per annum, some additional physicochemical data are required; in accordance with Annex IX to REACH (see also Table R.7.1–1).

Further details are given in the sections dedicated to specific endpoints.

R.7.1.1.1 Information requirements on physicochemical properties

Commission Regulation (EU) No 252/2011[7] has amended Annex I to REACH in order to adapt the chemical safety assessment provisions to the criteria for classification laid down in the CLP Regulation. The relevant amendments have been applied since 5 May 2011; however, for registrations submitted prior to this date, the chemical safety report shall be updated in accordance with Regulation No 252/2011 by 30 November 2012 at the latest.

The information needed under Article 12, REACH on one hand and according to section 4 of Annex VI to REACH on the other (namely hazard classification according to Title I and II CLP) is often complementary but in some cases may be different. The reason is that the classification criteria and/or test methods under DSD and CLP regimes are different. This is

[7] Commission Regulation (EU) No 252/2011 of 15 March 2011 amending Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards Annex I.

also expressed by the fact that CLP classifications are distributed over a different grid of hazard classes and categories compared to the DSD regime, e.g. substances and mixtures classified as explosive under DSD may be classified as explosives or self-reactives or organic peroxides under CLP, or they may even be classified as flammable solids, oxidizing solids or not at all. A translation table from DSD to CLP classification is provided in Annex VII, CLP and an indication of potential classification outcomes under CLP compared to DSD classifications is provided by Table 1.7.2.1(a) in the *Guidance on the Application of the CLP criteria*.

The CLP classification regime is not explicitly considered in Annex VII to REACH and therefore has to be understood as part of the information requirements under REACH. In particular, certain headlines set out in column 1 of Annex VII to REACH, namely 'explosive properties', 'flammability' and 'oxidizing properties', must be interpreted as covering the CLP hazard classes that are referred to in Article 58(11) of CLP.

The physical hazard classes according to CLP are structured differently from the corresponding classifications according to DSD. Despite this, most of the CLP physical hazard classes can unambigously be assigned to specific heading of the information requirements according to Annexes VI to IX to REACH. However, for some CLP physical hazard classes - notably the hazard class 'self-reactive substances and mixtures' and the hazard class 'organic peroxides' – the assignment to a specific heading is not straight-forward, since they may have both explosive and/or flammable properties. Therefore, some of the hazard classes are listed twice in Table R.7.1–1 below. It should be noted that this assignment is provided only as example and is done for the purposes of structuring this guidance document according to CLP but nevertheless also allowing the assignment to the respective information requirements according to Annexes VII to IX to REACH.

According to Article 1(6) CLP, CLP Regulation does not apply to the transport of dangerous goods by air, sea, road, rail or inland waterways (save where the specific rules for labelling of packaging applies under Article 33 of CLP). The transport of dangerous goods is, covered by the UN Model Regulations for Transport of Dangerous Goods (UN-RTDG) and related legal instruments (ADR, RID, ADN, IMDG Code and ICAO TI); the criteria listed in these instruments and in CLP Regulation for classification purposes are intended to be the same. Thus, a substance (or mixture) classified in a hazard class which is common to both CLP and the transport legislation will normally be classified the same according to both systems. Therefore the transport classification of a substance could be a source of information for the classification and labelling of substance (or a mixture) under CLP for physical hazards. However it should be kept in mind that the transport classifications do not cover all hazard categories which are relevant for CLP and it may be based on other considerations than just the test data and criteria (e.g classifications which are based on experience rather than testing or which apply only in connections with certain special provisions). As a result, the transport classifications may be different for the classification according to CLP. Similarly, the absence of a transport classification does not necessarily mean the substance (or mixture) should not be classified under CLP. Consequently in the case of a substance which has been tested for the purposes of the UN-RTDG and for which the same procedure was followed as required by the CLP Regulation, the same information could be used to comply with the REACH Regulation on a case-by-case basis. The limitations to the approach described above are described in detail in the *Guidance on the Application of the CLP criteria*, Section 1.7.2.1.

For the preparation of the registration dossier, registrants are required to submit all the information listed in Article 10 of REACH. Article 14(1) in conjunction with Annex I and Article 10(a)(vii) of the REACH Regulation, require the provision of a Robust Study Summary (RSS) for information derived from the application of Annexes VI to XI to REACH. In order to facilitate the evaluation conducted by the European Chemicals Agency and the Member States, as well as to save registrant's resources in case of a tonnage update, it **is recommended** that registrants also use the RSS for covering physicochemical endpoints under section 4 of the IUCLID file. This guidance includes under each physicochemical property chapter a list of detailed information to be given for each respective endpoint. Note that no further guidance is provided on the general aspects related to information common for all endpoints in IUCLID. For these aspects, further guidance can be found in *Practical guide 3: How to report robust study summaries* available on the ECHA Website (at: http://echa.europa.eu/practical-guides) and in the *IUCLID 5 End User Manual* available on the IUCLID Website at: http://iuclid.eu/index.php?fuseaction=home.documentation#usermanual.

Those endpoints, such as explosive properties and oxidising properties, which are intimately linked to classification, should be assessed according to CLP. For these endpoints, the test methods of CLP should preferably be used, in order to avoid double testing. For endpoints not linked to classification the preferred test methods are those found in the Test Method Regulation. For some endpoints (for example flammability), more than one test procedure is referred to in the standard test method reported in the Test Method Regulation. The one chosen should be suitable for the substance in question and be operating within its validity range.

Note that in the table below (Table R.7.1–1) in order to distinguish the physicochemical properties that are directly linked to physical hazard classifications from those that are not, the former have been shaded in gray and that in addition the preferred test methods for the different endpoints have been put in bold text.

Table R.7.1–1 Information requirements as specified in Annexes VII to IX to REACH and corresponding tests methods according to the Test Method Regulation and CLP

Information requirement according to Art. 10 (a) (vi) of the REACH Regulation (EC) No. 1907/2006 (the no. in brackets is the respective no. in the table in Annexes VII to IX to REACH)	**Corresponding test method according to The Test Method Regulation No. 440/2008**	**Chapter in revised R.7(a) guidance**	**CLP Regulation (EC) No. 1272/2008** (the no. in brackets is the respective chapter no. in Annex I to CLP)	**Corresponding test method according to CLP Regulation**
Melting/ Freezing point (7.2)	**A.1 Melting/Freezing temperature**	R.7.1.2	n.a.	n.a.
Boiling point (7.3)	**A.2 Boiling temperature**	R.7.1.3	n.a.	n.a.
Relative density (7.4)	**A.3 Relative density**	R.7.1.4	n.a.	n.a.
Vapour pressure (7.5)	**A.4 Vapour pressure**	R.7.1.5	n.a.	n.a.
Surface tension (7.6)	**A.5 Surface tension**	R.7.1.6	n.a.	n.a.
Water solubility (7.7)	**A.6 Water solubility**	R.7.1.7	n.a.	for metals - Transformation/Dissolution Protocol (Annex 10 to UN GHS)

Partition coefficient n-octanol/water (7.8)	**A.8 Partition coefficient**	R.7.1.8	n.a.	n.a.
Flash point (7.9)	A.9 Flash-point	R.7.1.9	n.a.	**CLP Annex I chapter 2.6.4.4**
Flammability (7.10)	A.11 Flammability (gases)	R.7.1.10.1	Flammable gases[8] (2.2)*	**ISO 10156** **EN 1839**
	for liquids: see Flash point	R.7.1.10.2	Flammable liquids (2.6)*	**see CLP, Annex I, Chapter 2.6.4.4, Table 2.6.3**
	A.10 Flammability (solids)	R.7.1.10.3	Flammable solids (2.7)*	**UN Test N.1**
	n.a.	R.7.1.10.4	Self-reactive substances and mixtures (2.8)*	**UN Test series A to H**
	A.13 Pyrophoric properties of solids and liquids	R.7.1.10.5	Pyrophoric liquids (2.9)*	**UN Test N.3**
		R.7.1.10.6	Pyrophoric solids (2.10)*	**UN Test N.2**
	n.a.	R.7.1.10.7	Self-heating substances and mixtures (2.11)*	**UN Test N.4**
	A.12 Flammability (Contact with water)	R.7.1.10.8	Substances and mixtures which in contact with water emit flammable gases (2.12)*	**UN Test N.5**
	n.a.	R.7.1.10.9	Organic peroxides (2.15)*	**UN Test series A to H**
Explosive properties (7.11)	A.14 Explosive properties	R.7.1.11.1	Explosives (2.1)*	**UN Test series 1 to 3 (further test series 4 to 6 are necessary for classification)**
	n.a.	R.7.1.11.2 see	Self-reactive substances and	**A.14 (existing data only)**

8 The Commission Regulation (EU) No 487/2013 of 8 May 2013 amending, for the purposes of its adaptation to technical and scientific progress, Regulation (EC) No 1272/2008 of the European Parliament and of the Council on classification, labelling and packaging of substances and mixtures (hereinafter referred to as '4th Adaptation to Technical Progress (ATP) to the CLP Regulation') amends the criteria in the CLP Annex I, Section 2.2 Flammable gases by including subclassifications for chemically unstable gases. The 4th ATP to the CLP Regulation will apply in respect of substances from 1 December 2014 and in respect of mixtures from 1 June 2015.

		R.7.1.10.4	mixtures (2.8)*	
	n.a.	R.7.1.11.3 See R.7.1.10.9	Organic peroxides (2.15)*	**A.14 (existing data only)**
Self ignition temperature (7.12)	A.15 Auto-ignition temperature (liquids and gases)	R.7.1.12.1	For gases and liquids*	n.a.
	A.16 Relative self-ignition temperature for solids	R.7.1.12.2, R.7.1.10.7	For solids* Note: the UN Test N.4 is preferable to generate the information for this endpoint. Refer to R.7.1.10.7.	n.a.
Oxidising properties (7.13)	n.a.	R.7.1.13.1	Oxidising gases (2.4) *	**ISO 10156**
	A.21 Oxidising properties (liquids)	R.7.1.13.2	Oxidising liquids (2.13) *	**UN Test O.2**
	A.17 Oxidising properties (solids)	R.7.1.13.3	Oxidising solids (2.14) *	**UN Test O.1**
Granulometry (7.14)	n.a.	R.7.1.14	n.a.	n.a.
Adsorption/Desorption (7.15)	n.a.	R.7.1.15	n.a.	n.a.
Stability in organic solvent and degradation products (7.16)	n.a.	R.7.1.16	n.a.	n.a.
Dissociation constant (7.17)	n.a.	R.7.1.17	n.a.	n.a.
Viscosity (7.18)	n.a.	R.7.1.18	n.a.	n.a.

* Note that regardless of whether the hazard class or category is listed in Article 14 (4) (a) of REACH, the chemical safety assessment (when required) must be performed in accordance with Article 14 (3) of REACH. Furthermore, according to Article 10 (a) (iv) of REACH the technical dossier of a registration for a substance under the REACH Regulation must include information on classification and labelling of the substance as specified in section 4 of Annex VI to the REACH Regulation.

In addition the CLP Regulation has the following hazard classes (Table R.7.1–2) for which the REACH Regulation does not require the generation of information (Article 10(a)(vi) and (vii) REACH):

Table R.7.1–2 CLP Regulation hazard classes for which the REACH Regulation does not require the generation of information

CLP Regulation (EC) No. 1272/2008 (the no. in brackets is the respective chapter no. in Annex I to CLP)	Corresponding test method according to the Test Method Regulation	Chapter in revised R.7(a) guidance	Information requirement according to Art. 10(a)(vi) of the REACH Regulation	Corresponding test method according to CLP Regulation
Flammable aerosols[9] (2.3)	n.a.	R.7.1.21.1	n.a.	Test methods according to 75/324/EC amended by 2008/47/EC (harmonised with **UN-MTC Section 31)**
Gases under pressure (2.5)	n.a.	R.7.1.21.2	n.a.	n.a.
Corrosive to metals (2.16)	n.a.	R.7.1.21.3	n.a.	**UN Test C.1 (UN-MTC Section 37.4)**

In order to comply with the REACH information requirements, registrants have to take due account of specific rules for adaptation according to column 2 of the tables in Annexes VII to XI to REACH, including the provisions given within the individual test methods of the Test Method Regulation, which have to be interpreted and applied in relation to the appropriate CLP hazard class. Further adaptations according to Annex XI to REACH must then be read together with the adaptation possibilities provided for by Article 8(2) of CLP and the CLP criteria themselves, namely those in Part 2 of Annex I to CLP.

Physicochemical data are mostly numeric and should be provided in SI units. Normally a numeric value or range is required. Where relevant, additional information should be provided on test conditions, such as temperature and/or pressure and/or concentration level or range etc., and estimated uncertainty in the numerical value. Furthermore details of any observations made during testing should be reported, e.g. decomposition during melting or boiling, emulsion formation during partitioning.

R.7.1.1.2 Available information on physicochemical properties

There are many published sources of data for basic substance characterisation and of supplementary information for hazard assessment. The relevant references are listed under the respective endpoint.

---

[9] The 4th ATP to the CLP Regulation amends the criteria in the CLP Annex I, Section 2.3 Flammable aerosols by changing the scope and title to Section 2.3 Aerosols.

R.7.1.1.3 Evaluation of available information on physicochemical properties

Advice to registrants with regard to nanomaterials characterisation can be found in *Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the *Guidance on IR&CSA*, section 2.1.3 Evaluation of available information.

Experimental data

Further, according to Article 13(3) of the REACH Regulation, tests to generate information on intrinsic properties of substances must be conducted in accordance with the test methods laid down in a Commission Regulation or in accordance with other international test methods recognised by the Commission or the Agency as appropriate, such as european standards (http://www.cen.eu/Pages/default.aspx) or OECD guidelines (http://www.oecd.org/chemicalsafety/testing/oecdguidelinesforthetestingofchemicals.htm). Data obtained from the tests in accordance with section 1.1.1 of Annex XI to REACH can be considered to be equivalent to data generated by the corresponding test methods referred in Article 13 (3) of REACH. Data for the purpose of physical hazard classification can be obtained using the test methods specified in the Articles 5 (1) and 8 (3) CLP. The test methods for the physicochemical properties are described in Regulation (EC) No 440/2008, whereas preferred tests for the purposes of physical hazard classification are referred to in Part 2 of Annex I to CLP, *via* references to the UN-MTC and to applicable standards. In Table R.7.1–1, the preferred test method for each endpoint is highlighted in bold. The test methods referred to in the CLP Regulation are also used for the transport of dangerous goods. Therefore, available information on physicochemical properties and physical hazards may also originate from tests that were carried out for the purposes of classification for transport. Such test data may be used for the information requirements according to the REACH Regulation. It should, however, be kept in mind that the classification for transport does not cover all hazard categories which are relevant for CLP and it may be based on other considerations than just the test data and criteria (e.g. classifications which are based on experience rather than testing or which apply only in connection with certain special provisions). As a result transport classifications may be different from the classification according to CLP. Such limitations are described in detail in the *Guidance on the Application of the CLP criteria*, Section 1.7.2.1.

Where relevant recognised standards for testing are applicable, the use of the most recent updates is advised; they are accessible *via* numerous websites, for example:

1. EN standards;
2. ISO standards;
3. IEC standards.

The national editions of the EN or ISO standards are available *via* the national standardization organizations accessible *via* the CEN Website.

Measured values which are evaluated in reviews and assigned recommended values are given precedence over calculated values. The major criteria that characterise the analysis of the available information are:

- **Experimental data**. When assessing physicochemical properties, priority is given to first hand experimental results (primary references) provided that the methods are suitable for the substance under investigation and that they operate within their validity range. Proper documentation on the methods and the inherent uncertainty of the measurements should also be provided.
- **Non-testing information**. If the information described in point (a) is not available, QSPRs, read-across or secondary data sources (e.g. handbook data) can be used in

accordance with the limitations described in the individual endpoint chapters (7.2 to 7.19 in this guidance) instead, and within the constraints of Annex XI to REACH.

Measurement uncertainty

Test data have an uncertainty of measurement. Some test methods include information about their uncertainty, which then may be referred to for test data generated using these test methods. Where the uncertainty of measurement is not specified by the test method, it is recommended to determine uncertainty by generally accepted processes of measurement uncertainty estimation (e.g. according to ISO/IEC Guide 98-3:2008).

Quality assurance for the determination of physicochemical properties

Test data on physicochemical properties should be of sufficient quality i.e. they must be reliable. Normally this can only be achieved by testing that is carried out in compliance with a relevant recognised quality system (e.g. GLP) or by laboratories complying with a relevant recognised standard (e.g. EN ISO/IEC 17025). Under Article 8 (5) of CLP, where new tests for physical hazards are required for the purposes of CLP they have to be carried out in compliance with a relevant recognised quality system or by laboratories complying with a relevant recognised standard at the latest from 1 January, 2014.

Non-experimental data

Quantitative Structure Property Relationships (QSPR) models exist for some of the physicochemical endpoints[10]. Where applicable, the details of any specific QSPR models are given under each endpoint.

The majority of QSPR models have been built using training sets of substances. The model will have been optimised to calculate values for the training substances that most closely match measured ones. Therefore, the use of QSPR estimation techniques requires expert judgement. The calculated values need to be checked to ensure that they are reasonable and that the model used is appropriate.

A valid model will give values that are in reasonably close agreement with the measured ones for your chosen analogue substances (i.e. the substance with a data gap should have similar substances in the training set of the model). The models may not predict very well the properties of substances which are too dissimilar to the reference set for the model. Thus, the model can be used to provide a predicted value for your substance without the need for testing. Another check is that the values are realistic. This can be done by cross-referencing the calculated value to measured values for similar substances and related endpoints. If a QSAR method is used as a stand alone method to determine a value to meet the endpoint data requirements, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Assessing the quality of QSPR models

The European Commission and the OECD member countries adopted five principles for the validation of (Q)SAR/(Q)SPR models in 2004 (OECD[a], 2004). According to these principles, a valid (Q)SPR model should have 1) a defined endpoint whose experimental conditions are clearly specified; 2) an unambiguous algorithm; 3) a defined domain of applicability that defines for what kind of substances predictions can be made; 4) appropriate measures of goodness-of-fit, robustness and predictivity; and 5) a mechanistic interpretation if possible.

---

[10] A comprehensive review paper with a title: “QSPR prediction of physico-chemical properties for REACH” was published in the SAR and QSAR in Environment Research in 2013 (Dearden, J.C., Rotureau, P., Fayet G. (2013). QSPR prediction of physico-chemical properties for REACH, SAR and QSAR in Environmental Research, Vol. 24, No.4, 279-318).

These principles are outlined on the ECB website and more extensively covered in the *Guidance on IR&CSA* Chapter R.6: 'QSAR and grouping of substances', Section R.6.1.3. Moreover, a practical overview of these principles is given in the report from the expert group on (Q)SARs (OECD[b] 2004).

Assessing the quality of read-across predictions

This paragraph reports the basic concepts of a read-across approach. Thorough information on this topic can be found in the guidance on the grouping of substances (see *Guidance on IR&CSA* Chapter R.6: 'QSAR and grouping of chemicals', Section R.6.2).

A read-across/analogue approach assesses the relevance of a given property on one or more chemical structures and then makes some assessment (qualitative or quantitative) on the relevance of this information for another substance (see Annex XI, REACH). Since a read-across may involve two substances[11] it is of paramount importance to detail the reasoning behind the inference on the substance whose property is unknown. An analogue must:

- contain the same major structural features and the same functional groups as the substance under investigation;
- have a physicochemical profile comparable to that of the substance under examination as far as the known physicochemical properties are concerned;
- have comparable values for the relevant molecular descriptors (i.e. excess molar refractivity and hydrogen bond donor and acceptor abilities for water solubility predictions) generally used for the quantification of the property of interest;
- have approximately the same molecular weight.

The interpretative analysis of a read-across is usually the result of an expert judgement evaluation and detailed documentation should therefore always be provided to support the conclusions. It is important to point out that, in practice, read-across for physicochemical properties is not generally recommended, since reliable data should normally be available or easily obtainable. This is particularly true for physical hazard related physicochemical properties for which reliable test data must be available according to Article 8 (2) of CLP. Therefore, if read-across is used as a stand alone method to generate a value to meet the endpoint data requirements, the criteria given in section 1.5 of Annex XI to REACH must be met.

Use of secondary and historical data sources for physicochemical properties

The reliability of data must be demonstrated by providing information on the identity and purity of the test substance, the methodology used to make the measurement, and whether or not this was performed in compliance with a relevant recognised quality system (e.g. GLP) (Annex VI, REACH).

Numerical physicochemical data is particularly prone to data recycling (transfer from one database to another, often with loss of the original source and contextual information). Data from secondary and historical sources must be adequate and is especially important where the

[11] A read-across can also involve more than two substance: one-to-one (one analogue used to make an estimation for a single substance) b) many-to-one (two or more analogues used to make an estimation for a single substance c) one-to-many (one analogue used to make estimations for two or more substances) d) many-to-many (two or more analogues used to make estimations for two or more substances).

endpoint is relevant for classification, PBT/vPvB assessment, is the basis of waiving arguments for other endpoints, or has a large influence on the outcome of the risk assessment. The criteria in section 1.1.1 of Annex XI or section 1.2 of Annex XI to REACH must be met.

R.7.1.1.4 Overall consistency of the physicochemical information

The physicochemical data for a given substance cannot contain incompatible values for two or more properties (i.e. high boiling point and high vapour pressure at normal temperature). This consistency check should be always done and it can turn out to be particularly useful when measured values are significantly at odds with predictions from QSPR models. Indeed, in this case a wider assessment of the known physicochemical properties should be performed in order to determine the possible cause of the inconsistencies.

Concluding on classification and labelling and chemical safety assessment

Data on physicochemical properties not only determine the presence or absence of a physical hazard but also have also an impact on the sections of the chemical safety assessment concerning the environment and human health. The assessment determines the risk posed to humans and the environment from all stages of the substance's lifecycle. This includes its manufacture, transfer, use and disposal. Firstly, the physicochemical data set provides the input parameters for the purpose of the human and environmental exposure estimation. For example, the vapour pressure and particle size information are required to estimate the likely exposure of humans, both in the workplace and in consumer use as well as to estimate the likelihood of forming flammable/explosive vapour/dust-air mixtures. The volatility (vapour pressure) or the size and nature of particles are indicators of the potential for inhalation exposure. Particle size is also important for determining the likely dermal exposure and the presence of a dust explosion hazard. Viscosity is a key parameter in determining aspiration hazards. The physical state of a substance at the process temperature is important for determining possible hazards. Further, physico-chemical data are essential for the correct planning of (eco)toxicological studies and for the optimisation of the test conditions.

R.7.1.1.5 References for introduction of Physicochemical properties

Recommendations on the transport of dangerous goods, Manual of Test and Criteria, United Nations. http://www.unece.org/trans/danger/publi/manual/manual_e.html

Guidance on the Application of the CLP Criteria, Version 4.0 - 2013, ECHA. http://echa.europa.eu/web/guest/guidance-documents/guidance-on-clp

OECD[a] (2004) Principles for the Validation of (Q)SARs http://www.oecd.org/chemicalsafety/risk-assessment/validationofqsarmodels.htm

OECD[b] (2004) series on testing and assessment Number 49 The report from the expert group on (quantitative) structure activity relationships [(Q)SARs] on the principles for the validation of (Q)SARs. 2nd Meeting of the ad hoc Expert Group on QSARs http://www.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2004)24&doclanguage=en

R.7.1.2 Melting point/freezing point

R.7.1.2.1 Type of property

The melting point contributes to the indentification of a substance and to the designation of its physical state (liquid or solid[12]) of a substance. A number of physical hazard classes are distinguished based on the physical state. Therefore the melting point of a substance and the consequent designation as liquid or solid has also consequences for the assignment of the correct hazard class. Furthermore, the melting/freezing point together with vapour pressure serves as an indicator for the physical state (liquid or solid) of a substance under specific conditions (e.g environmental conditions, manufacturing process conditions). As a result, with regard to environmental relevance the melting point can give an indication of the distribution of the substance within and between the environmental media (water, soil and air).

R.7.1.2.2 Definition

The melting temperature is defined as the temperature at which the phase transition from the solid to the liquid state occurs at atmospheric pressure and this temperature ideally corresponds to the freezing temperature. As the phase transition of many substances takes place over a temperature range, it is often described as the melting range. For some substances, the determination of the freezing or solidification point is more appropriate. Where, due to the particular properties of the substance, none of the above parameters can be conveniently measured, a pour point may be appropriate.

R.7.1.2.3 Test method(s)

Method A.1 of Regulation (EC) 440/2008 or OECD Test Guideline 102 should be generally used for testing. Any procedure given in A.1 may be used within the scope and applicability specifications. However, it is advisable to use the Differential Scanning Calorimetry (DSC) or Differential Thermo-Analysis (DTA) method since they give additional information about the thermal stability of the substance like decomposition onset and energy. If decomposition occurs during the melting point study, determination of the boiling point need not be carried out. In this case, if DSC has been used, conducting the experiment under inert gas should be considered.

R.7.1.2.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for melting/freezing point:

'The study does not need to be conducted below a lower limit of - 20°C.'

Therefore, Annex VII to REACH does not require determination of the melting point below a lower limit of -20°C. The lower limit should be confirmed through testing, except where a (Q)SAR indicates a melting point of -50°C or lower.

---

[12] Definitions of physical states can be found in Section 1.0. of Annex I to the CLP Regulation.

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria of Annex XI, section 1.1.1, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in Annex XI, section 1.3.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria of Annex XI, section 1.1.1 or where several sources of similar reliability with deviating results exist, a weight-of-evidence approach may be used. The criteria of Annex XI, section 1.2 must then be met[13].

(Q)SAR

For the determination of the melting point, (Q)SAR approaches are discouraged, because the accuracy is not sufficient (± 25°C or more) for the purposes of classification / risk assessment.

Grouping of substances and read-across approach

For the determination of the melting point read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

Some substances will decompose or sublime before the melting point is reached.

Further adaptation possibilities

Not foreseen.

R.7.1.2.5 Impurities; uncertainties

Impurities can have a significant influence on the melting point, as they will generally lower the melting point noticeably. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

R.7.1.2.6 Endpoint specific information in the registration dossier / IUCLID

Materials and methods

- type of method or reference to the standard or the test method applied.

Results and discussion

- melting point value (°C) as measured;

---

[13] National Institute of Standards and Technology (NIST) have a useful statistical approach which has been used for the evaluation of literature melting point data (ref.: http://webbook.nist.gov/chemistry/site-cal.html#AVG).

- rate of temperature increase if available;
- decomposition or sublimation temperature (if applicable);
- measurement uncertainty if available;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on melting point/freezing point can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.2	VII 7.2	Melting point/freezing point	E.4.3	3.2

R.7.1.3 Boiling point

R.7.1.3.1 Type of property

The boiling point is a property:

- which contributes to the characterisation of a substance and to the designation of its physical state (gas or liquid);
- which is the basis for the assignment of the correct hazard class because a number of physical hazard classes are distinguished based on the physical state;
- which is needed for the classification of flammable liquids into categories;
- which gives an indication of the distribution of the substance within and between the environmental compartments (air, soil and water);
- which have correlations with vapour pressure and therefore gives indications whether a substance may be available for inhalation as a vapour or may form flammable/explosive vapour-air mixtures, too;
- which is important for physical hazard assessment.

R.7.1.3.2 Definition

The normal boiling point is the temperature at which the vapour pressure of a liquid equals 101.3 kPa.

Note: If the vapour pressure equals 101.3 kPa or more at a given temperature this means the substance is completely gaseous at that temperature. If this is the case at temperatures ≤20°C the substance is a gas also according to the CLP Regulation.

R.7.1.3.3 Test method(s)

Method A.2 of Regulation (EC) 440/2008 or OECD Test Guideline 103 should be used for testing. Any determination method may be used within the scope and applicability specifications. DSC allows the determination of the melting and boiling point in a single test. Likewise, for some substances a single test can be used to determine both ‘boiling point’ and ‘vapour pressure’, as when the dynamic method is applied.

For high-boiling liquids or liquids which may decompose, auto-oxidize etc. before the boiling point at 101.3 kPa or more is reached, it is recommended to determine the boiling point either under inert gas or at reduced pressures, in order to derive the boiling point at reduced pressures from the vapour pressure curve.

If explosive substances, pyrophoric substances or self-reactive substances are to be characterized, determination of the boiling point is in general not practicable. For pyrophoric substances testing under inert gas or reduced pressures should be considered.

Where standards are applicable, the use of the most recent updates is advised; they are accessible *via* numerous websites, see above in Section R.7.1.1.3.

R.7.1.3.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for boiling point:

‘The study does not need to be conducted:

- *for gases; or*
- *for solids which either melt above 300 °C or decompose before boiling. In such cases the boiling point under reduced pressure may be estimated or measured; or*
- *for substances which decompose before boiling (e.g. auto-oxidation, rearrangement, degradation, decomposition, etc.).’*

Therefore the Annex VII to REACH does not require determination of the boiling point if:

- the substance is a gas;

However, for some gases the boiling point may be relevant. In the CLP Regulation, the boiling point is the main criterion to distinguish gases from liquids (see Annex I, section 1.0: Gas means a substance which (i) at 50°C has a vapour pressure greater than 300 kPa (absolute); or (ii) is completely gaseous at 20°C at a standard pressure of 101.3 kPa). Therefore it is important to report the boiling point in borderline cases where the transition from liquid to gas occurs close to 20°C.

- the melting point of the substance is above 300°C or when any chemical change occurs during the melting point study;
- the substance decomposes before boiling at ambient pressure.

In such cases the boiling point under reduced pressure (down to 0.2 kPa) should be determined if possible without decomposition.

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria of Annex XI, section 1.1.1, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in Annex XI, section 1.3.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria of Annex XI, section 1.1.1 or where several sources of similar reliability with deviating results exist, a weight-of-evidence approach may be used. The criteria of Annex XI, section 1.2 must then be met[14].

(Q)SAR

For the determination of the boiling point, (Q)SAR approaches are discouraged for the purpose of classification / risk assessment, except when the mean absolute error of the method is lower than 2 K.

Grouping of substances and read-across approach

For the determination of the boiling point read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

Testing is not possible if:

- the substance is an explosive;
- the substance is self-reactive;
- any chemical change occurs during the melting point study;
- the liquid decomposes before the boiling point is reached even at reduced pressures below 0.2 kPa.

In such cases the decomposition temperature in relation to the (reduced) pressure should be reported, in order to allow determination of whether it is the substance itself or its decomposition products that should be considered under environmental conditions for the purpose of risk assessment. The details of the determination method should also be reported.

Further adaptation possibilities

Data generated with the same tests and classification principles as specified in the CLP Regulation on boiling point generated in conjunction with transport classification can be

[14] The NIST have a useful statistical approach which has been used for the evaluation of literature boiling point data (ref.: http://webbook.nist.gov/chemistry/site-cal.html#AVG).

deemed to satisfy the REACH requirements on a case-by-case basis. As stated in Annex IX of the REACH Regulation, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

R.7.1.3.5 Impurities; uncertainties

Impurities can have a significant influence on the boiling point. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

R.7.1.3.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

- type of method or reference to the standard or the test method applied.

Results and discussion

- boiling point value (°C) as measured;
- pressure value and unit;
- rate of temperature increase if available;
- decomposition (if applicable);
- measurement uncertainty if available;
- boiling point value in ºC (corrected to standard pressure, except where the boiling point has been determined at specified reduced pressures) (as above, but in a separate block of fields);
- if testing is waived, the reasons for waiving must be documented in the dossier.

Note: In cases where the boiling point is determined at reduced pressure a determination at ambient pressure is obviously not possible. A boiling point at standard pressure could then only be derived by extrapolation of the vapour pressure curve in cases where a vapour pressure curve is known. Even in such cases this corrected/extrapolated boiling point could only be nominal one and would be potentially misleading because it is not possible to determine it at ambient pressure.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on boiling point can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.3	VII 7.3	Boiling point	E.4.4	3.3

R.7.1.4 Relative density

R.7.1.4.1 Type of property

For gaseous materials, relative density is of value in determining the tendency to settle or to disperse when discharged at high concentrations into the atmosphere. The relative density of gaseous substances can be calculated from molecular weight using the Ideal Gas Law.

For insoluble liquids and solids, (absolute) density will be a determining factor in the settling of the substance.

R.7.1.4.2 Definition

Density (ρ) of a substance is the quotient of the mass m and its volume V:

$$\rho = m/V \qquad \text{SI units } (kg/m^3)$$

The relative density is related to a standard, the density of which is set to 1. It has no dimension. For gases air is used as standard so that gases with a relative density of less than 1 are lighter than air (and and those with a value above 1 heavier).

The relative density, D_4^{20}, of solids or liquids is the ratio between the mass of a volume of substance to be examined, determined at 20 ºC, and the mass of the same volume of water, determined at 4 ºC (at which temperature, water has its maximum density, i.e. 999.975 $kg/m^{3)}$.

R.7.1.4.3 Test method(s)

Test methods for determining (absolute) density are applicable to solids and liquids. Table R.7.1–3 lists the respective test methods.

Table R.7.1–3 Test methods for determining density

Method	Applicability	Maximum Dynamic Viscosity (Liquids only) / Pa.S
Hydrometer	Liquids	5
Hydrostatic balance	Solids and Liquids	5
Immersion ball	Liquids	20
Pycnometer	Solids and Liquids	500
Air comparison pycnometer	Solids	-
Oscillating densitimeter	Liquids	5

EU Test guideline A.3 for relative density Regulation (EC) No 440/2008 includes a list of standards with technical information about the different methods and actual measuring of different types of substances.

R.7.1.4.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for relative density:

'The study does not need to be conducted if:

- *the substance is only stable in solution in a particular solvent and the solution density is similar to that of the solvent. In such cases, an indication of whether the solution density is higher or lower than the solvent density is sufficient; or*
- *the substance is gaseous at room temperature. In this case, an estimation based on calculation can be made from its molecular weight and the Ideal Gas Laws.'*

For liquids, it is useful to have some indication of the dynamic viscosity as this can affect the choice of method. The physical state of test substances should always be homogeneous, this is particularly relevant for highly viscous substances where internal bubbles can be formed; in these cases, the test substance should be allowed to rest until all internal bubbles have disappeared.

The summary should include the numerical value for density and temperature at which it was measured, test material identity, purity of the sample used, physical state, method and guideline used and reference substance (if any).

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met[15].

(Q)SAR

(Q)SAR is generally not applicable for determination of relative density.

For this endpoint there are often experimental measurements and therefore QSPR models for this property have not received special attention in the environmental literature. Several software programs can be used to calculate the density of a given substance but the documentation and validation of the methods is limited.

---

[15] The NIST have a useful statistical approach which has been used for the evaluation of literature data (ref.: http://webbook.nist.gov/chemistry/site-cal.html#AVG).

Grouping of substances and read-across approach

For the determination of the relative density read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies. Waiving relative density testing on the basis of not being technically possible is not applicable.

Further adaptation possibilities

Not foreseen.

R.7.1.4.5 Impurities; uncertainties

Impurities can have a significant influence on the density. This influence depends on the amount and density of the impurity; thus, the higher the amount of impurity and the higher the difference between the densities of the main component and the impurity, the higher the influence. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

Density is temperature dependant. Whenever possible, determinations should be performed at 20°C.

R.7.1.4.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

- type of method or reference to the standard or the test method applied.

Results and discussion

- temperature (°C);
- relative (for gases)/ absolute (for liquids and solids) density value (dimensionless);
- measurement uncertainty if available;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on relative density can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.4	VII 7.4	Relative density	E.4.5	3.4

R.7.1.5 Vapour pressure

R.7.1.5.1 Type of property

Vapour pressure is a property:

- for substance characterisation;
- which serves as a key parameter for assessing some toxicological and environmental hazards;
- which gives indications whether a substance may be available for inhalation as a vapour or may form flammable/explosive vapour-air mixtures;
- which allows determination of the volatility of a substance from an aqueous medium or soil, in terms of the Henry's Law constant (Appendix R.7.1–1) and partition coefficient air/soil, respectively;
- which allows determination of the right container/vessel to ensure safety during storage, transport and use;
- which is importiant for physical hazard assessment.

R.7.1.5.2 Definition

The vapour pressure of a substance is defined as the saturation pressure above a solid or a liquid substance at constant temperature. At the thermodynamic equilibrium, the vapour pressure of a pure substance is a function of temperature only.

R.7.1.5.3 Test method(s)

Method A.4 of Regulation (EC) 440/2008 or OECD Test Guideline 104 (Vapour pressure) should be used for testing. It is useful to have preliminary information on the structure, the melting point and the boiling point of the substance to perform this test.

There is no single measurement procedure applicable to the entire range of vapour pressure values. Therefore, several methods are recommended to be used for the measurement of vapour pressure from < 10^{-10} to 10^{5} Pa. For the selection of the test method the scope and applicability specifications have to be taken into account. The results should be checked for consistency with other physical data like boiling point, flash point etc.

It is recommended to determine the vapour pressure at least for two temperatures, for volatile substances (boiling point up to 150°C) preferably at 20°C and at 50°C.

Where standards are applicable, the use of the most recent updates is advised, please check Section R.7.1.1.3 for further information.

R.7.1.5.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for vapour pressure:

'The study does not need to be conducted if the melting point is above 300°C.

If the melting point is between 200°C and 300°C, a limit value based on measurement or a recognised calculation method is sufficient.'

Vapour pressure testing is also not required for substances with a standard boiling point of < 30 ºC, as these substances will have vapour pressures above the limit of measurement (i.e. 10^5 Pa).

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

For the determination of the vapour pressure, (Q)SAR approaches may be used if determination by experiment is not possible.

The vapour pressure depends on the temperature. This dependence was modelled by Grain (Grain, 1982), based on thermodynamic principles. The estimation methods differ for vapour pressure that can be applied for compounds that are liquid or gaseous at the temperature of interest, and for solid and liquid compounds. The former can be estimated by the Antoine equation, while the latter could be predicted by the Watson correlation, which accounts also for the heat of vaporisation. Another method, described by Mackay *et al*. (1982), is applicable only for hydrocarbons and halogenated hydrocarbons. Further, the Grain model was modified to be applicable for all solids, liquids, and gases. These methods are still in practical use today.

The OECD guideline 104 reports that the Watson correlation is applicable over the pressure range from 10^5 Pa to 10^{-5} Pa. It should in any case be pointed out that estimated values for vapour pressure can be subjected to great uncertainty if the computed pressure is lower than 1 Pa, especially when the boiling point has not been experimentally determined (OECD monograph 67). The uncertainty is even greater if the estimated value is used together with water solubility in order to estimate the Henry's Law constant.

The environment monograph 67 of the OECD describes all of the above mentioned methods and the OECD guideline 104 supports the use of the Watson correlation for the calculation of vapour pressure, but does not specifically reject other calculation methods.

The handbook for estimating the physico-chemical properties of organic compounds (Reinhard and Drefahl, 1999) reports another method based on thermodynamic properties and elaborated by Mishra and Yalkowsky that discussed the application of the method of Mackay (Mackay *et al*., 1982).

The equation by Mishra and Yalkowsky gave significantly better estimates than the method of Mackay on the same data set (Mishra and Yalkowsky, 1991).

Another methodology that proved to be effective in estimating vapour pressure relies on group contribution approaches. Several models using this strategy have been proposed (Reinhard and Drefahl, 1999; see Table R.7.1–4).

Table R.7.1–4 Group contribution approach and vapour pressure

Compounds	Authors	Methodology	Statistics
Alkyl aromatic compounds	Amidon and Anik	Group contribution approach	Standard error … 1.1 kJ on the estimation for the free energy of vaporisation
Mono-, di-, tri- and tetra substituted	Hoshino *et al.*	Group contribution approach	Average error 3.7% Max. Error 30.9%
Perfluorinated saturated hydrocarbons	Kelly *et al.*	Group contribution approach	Arithmetic mean deviation <0.5%

Numerous other models are available for the estimation of vapour pressure, and Schwarzenbach *et al.* (1993), Delle Site (1996), Sage and Sage (2000) and Dearden (2003) have reviewed many of these. The descriptors used in vapour pressure QSPRs include physico-chemical, structural and topological descriptors, and group contributions. Katritzky *et al.* (1998) used 4 CODESSA descriptors to model the vapour pressure (in atmospheres at 25°C) of 411 diverse organic chemicals, with r^2 = 0.949 standard error = 0.331 log unit. A number of studies (Andreev *et al.* 1994, Kühne *et al.* 1997, Yaffe & Cohen 2001) allow of the estimation of vapour pressures over a range of temperatures.

Grouping of substances and read-across approach

For the determination of vapour pressure read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

Vapour pressure testing is not required for substances with a standard boiling point of < 30ºC, as these substances will have a vapour pressure value above the limit of measurement (i.e. 10^5 Pa).

For substances which decompose during measurement or which are unstable or explosive, determination of the vapour pressure may not be technically possible. This also applies to self-reactive substances and organic peroxides.

Pyrophoric substances may be difficult to handle experimentally. If fully inert conditions cannot be maintained during sample preparation and measurement, use of an appropriate calculation method is recommended.

A calculation method should also be applied in the case of some corrosive substances which would destroy essential metallic parts of the measurement apparatus.

Further adaptation possibilities

Not foreseen.

R.7.1.5.5 Impurities; uncertainties

Impurities can have a large influence on vapour pressure. The influence depends on the amount of the impurity and the vapour pressure of that impurity. Small amounts of volatile impurities may increase the vapour pressure by several orders of magnitude. This has to be kept in mind when performing the measurements and for the interpretation of results. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

Where there are volatile impurities in the sample which could affect the result, the substance may be purified. Test method A.4 states that it may also be appropriate to quote the vapour pressure for the technical material. However, in consideration of the large effect that impurities may have (see above), doing so is strongly discouraged.

R.7.1.5.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

- type of method or description of the apparatus or reference to the standard or the test method applied.

Results and discussion

- if testing is waived, the reasons for waiving must be documented in the dossier;
- measured value of the vapour pressure for at least two temperatures;
- estimate of the vapour pressure at 20 or 25ºC (if not measured at these temperatures);
- if a transition (change of state, decomposition) is observed, the following should be noted:
    - nature of change;
    - temperature at which change occurs.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on vapour pressure can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.6	VII 7.5	Vapour pressure	E.4.7	3.6

R.7.1.5.7 References on vapour pressure

OECD Guidelines for the Testing of Chemicals / Section 1: Physical-Chemical properties, Test No. 104: Vapour Pressure, OECD Code: 979910401E1, July 2006.

Andreev N.N, Kuznetsov S.E, Storozhenko S.Y. (1994) Prediction of vapour pressure and boiling points of aliphatic compounds. Mendeleev Commun. 173-174.

Grain C.F., (1982) Handbook of chemical property estimation methods. New York, Mc Graw-Hill

Delle Site A. (1996) The vapour pressure of environmentally significant organic chemicals: a review of methods and data at ambient temperature. J. Phys. Chem. Ref. Data 26:157-93.

Dearden JC. (2003) Quantitative structure-property relationships for prediction of boiling point, vapour pressure, and melting point. Environ Toxicol Chem 22(8):1696-709.

Katritzky AR, Y. W, Sild S, Tamm T, Karelson M. (1998) QSPR studies on vapour pressure, aqueous solubility, and the prediction of water-air partition coefficients. J. Chem. Inf. Comput. Sci. 38:720-5.

Kühne R, Ebert RU, Schüürmann G. (1997) Estimation of vapour pressures for hydrocarbons and halogenated hydrocarbons from chemical structure by a neural network. Chemosphere 34:671-86.

Mackay D, Bobra A, Chan W, Shiu WY. (1982) Vapour pressure correlation for Low-Volatility Environmental Chemicals. Environ. Sci. Technol. 16:645-9.

Mishra DS, Yalkowsky SH. Estimation of vapour pressure of some organic compounds. Ind. Eng. Chem. Res. 1991;30:1609-12.

OECD Guidelines for Testing of Chemicals, Method 104 "Vapour Pressure Curve"

Reinhard M, Drefahl (1999).A. Handbook for Estimating Physico-Chemical Properties of Organic Compounds. New York: Wiley.

Sage M.L, Sage G.W.(2000) Handbook of Property Estimation Methods for Chemicals. Boca Raton, FL: Lewis.

Schwartzenbach, R.P., Gswend, P.M., Imboden, D.M. (1993). Environmental Organic Chemistry. John Wiley and Sons.

Yaffe D, Cohen, Y (2001) Neural network based temperature-dependent quantitative structure property relationships (QSPRs) for predicting vapour pressure of hydrocarbons. J. Chem. Inf. Comput. Sci. 41:463-477.

R.7.1.6 Surface tension

R.7.1.6.1 Type of property

Surface tension measurements of aqueous solutions are significant since decreasing the surface tension of water may impact on the properties of the solution and other physicochemical measurements.

R.7.1.6.2 Definition

- Surface tension:

'The free surface enthalpy per unit of surface area is referred to as surface tension' (Council Regulation (EC) No 440/2008).

The surface tension is given as: N/m (SI unit) or mN/m (SI sub-unit). 1 N/m = 10^3 dyne/cm or 1mN/m = 1 dyne/cm in the obsolete cgs system.

The surface tension of an aqueous solution of a substance can be used to determine whether the substance is surface active.

- Surface active substance (surfactant):

''Surfactant' means any organic substance and/or preparation [mixture] *used in detergents, which has surface-active properties and which consists of one or more hydrophilic and one or more hydrophobic groups of such a nature and size that it is capable of reducing the surface tension of water, and of forming spreading or adsorption monolayers at the water-air interface, and of forming emulsions and/or microemulsions and/or micelles, and of adsorption at water-solid interfaces'* (see Article 2(6) of Council Regulation (EC) No 648/2004).

R.7.1.6.3 Test method(s)

Testing should be done in accordance with one of the methods specified under section A.5 of Regulation (EC) No 440/2008. These methods are applicable to most chemical substances.

It is useful to have preliminary information on the water solubility, the structure, the hydrolysis properties and the critical concentration for micelles formation of the substance before performing the test.

Surface tension measurements require a test material that is stable against hydrolysis during the test period and soluble in water at concentrations of > 1 mg/l. Measurements should be performed on a solution at either 90 % of the solubility limit or 1 g/l (where viscosity permits), whichever is smaller.

R.7.1.6.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for surface tension:

'The study need only be conducted if:

- *based on structure, surface activity is expected or can be predicted; or*

- *surface activity is a desired property of the material.*

If the water solubility is below 1mg/l at 20°C the test does not need to be conducted.'

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

At the time of writing, no reliable (Q)SAR methods exist for sufficiently accurate predictions of surface tension.

Grouping of substances and read-across approach

For the determination of the surface tension read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies. Testing may not be possible for reactive substances which react with water or air (hydrolyse, are pyrophoric, evolve gas, etc).

Further adaptation possibilities

Not foreseen.

R.7.1.6.5 Impurities; uncertainties

For the measurement of surface tension the ring or plate tensiometer methods are preferred. The error on the measurement is in the order of 0.1–0.3 mN/m. Use of the standard protocols and GLP procedures are recommended. Surface active impurities in substances may in some cases lead to false-positive surface tension measurements.

R.7.1.6.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied;

- test material identity: apart from general issues, if surface tension of active impurities affects results, it should be noted.

Results and discussion

- surface tension value and unit (preferably mN/m or N/m but other units are also acceptable);
- concentration of the solution*[16];
- age of solution*;
- type of water or solution used*;
- results from repeated measurements with varied equilibrium time (of the solution);
- several measurement results should be provided to assess the possible time-dependency of the measurement. Equilibration times may vary from minutes to hours. Measurements should be sufficient to prove that a constant surface tension was reached;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on surface tension can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.10	VII 7.6	Surface tension	E.4.11	3.9

R.7.1.7 Water solubility

Advice to registrants with regard to nanomaterials characterisation of water solubility can be found in *Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the Guidance on IR&CSA, section 2.2.1 Water solubility.

R.7.1.7.1 Type of property

Water solubility is a significant parameter for a number of reasons:

- the mobility of a test substance is largely determined by its solubility in water. In general, highly soluble substances are more likely to be distributed by the hydrological cycle;

- water soluble substances gain access to humans and other living organisms;

- knowledge of the water solubility is a prerequisite for setting up test conditions for a range of fate (e.g. biodegradation, bioaccumulation) and effects studies;

[16] *As indicated in test A.5. Surface tension is described in Council Regulation (EC) No 440/2008.

- it is also used to derive other environmental parameters, such as K_{ow}, K_{oc} and Henry's Law Constant (Appendix R.7.1–1). It is also used as input for some QSAR models;

- water solubility is used as a regulatory trigger for waiving certain physicochemical and ecotoxicological endpoints.

R.7.1.7.2 Definition

'The solubility of a substance in water is specified by the saturation mass concentration of the substance in water at a given temperature. The solubility in water is specified in units of mass per volume of solution. The SI unit is kg/m³ (grams per litre may also be used)' (see Regulation (EC) No 440/2008, A.6, section 1.2).

Mixtures of organic compounds, e.g. petroleum substances, behave differently from their single constituent compounds when brought into contact with water. Petroleum substances are typically hydrophobic and exhibit low solubility in water. However, reflecting the range of structures, constituent hydrocarbons will exhibit a wide range of water solubility. Therefore, water solubility measurements for these substances are loading rate dependent due to their complex composition. This water solubility behaviour impacts on both the conduct and interpretation of aquatic toxicity tests for these complex substances. The complex composition, and generally low water solubility, impact also on the choice and conduct of biodegradation studies.

Consequently, the above definition for solubility of a single substance in water is not applicable to substances which are multi-component, such as multi-constituent or UVCB substances, i.e. complex substances. The usually accepted meaning of 'solubility' in such cases is 'the composition of the aqueous solution formed at equilibrium under a defined set of conditions'. Temperature and the amount of substance added per unit volume of water (i.e. the 'loading') are the main factors to consider. It may not always be possible to establish that equilibrium of all components has been achieved; in these cases, time and type of agitation of the test vessels must also be described.

Similar testing issues also apply to inorganic compounds. Water solubility among compounds of the same metal may differ by several orders of magnitude. Differences in the solubility of metal compounds are related to the metal species and the characteristics of the aqueous medium. Highly soluble inorganic metal compounds can be assessed through the normal procedures. For sparingly soluble metal compounds, a solubility product can be calculated thermodynamically (e.g. by using the Facility for Analysis of Chemical Thermodynamics ('F*A*C*T', FACT-Win version 3.05). Although metals are generally insoluble, metals in the elemental state may react with water or a dilute aqueous electrolyte to form soluble or sparingly soluble cationic or anionic products. During this process the metal will oxidise, or transform, from the neutral or zero oxidation state to a higher oxidation state. The OECD Test Guidance on transformation/dissolution of metals and sparingly soluble metal compounds (OECD, 2001) can be used to determine the rate and extent to which metals and sparingly soluble metal compounds can produce soluble bioavailable ionic and other metal-bearing species in aqueous media under a set of standard laboratory conditions representative of those generally occurring in the environment. The outcomes of the transformation/dissolution tests are to be used for aquatic environmental hazard classification purposes.

R.7.1.7.3 Test method(s)

No single method is available to cover the whole range of solubility values in water, from relatively soluble to very low soluble substances. General test guidelines (OECD Method 105; EU Method A.6, Regulation (EC) No 440/2008) include two test methods which cover the whole range of solubility values but are not applicable to volatile substances. Water solubility determinations are normally run at 20 ºC in distilled water according to standard test guidelines (OECD Method 105; EU Method A.6). Solubility data determined using these standard physico-chemical guidelines may differ if the test material is solubilised in either aqueous solutions containing salts or at different test temperatures (or both) (e.g. ecotoxicological test media).

The methods should be applied to essentially pure substances that are stable in water. Details of suitable methods are shown in Table R.7.1–5.

A number of standardised methods are available for the determination of single substances and complex mixtures of liquids and solids. For metals and sparingly soluble inorganic metal compounds a specific water solubility approach was designed to measure transformation to the dissolved fraction under standard conditions. The test methods are not applicable to volatile substances. Care should be taken to ensure that the test substances examined are as pure as possible and their solubility levels are determined analytically using a specific analytical method wherever possible. Precautions should be taken to minimise degradation of the test substance, in particular if long periods of equilibration are required (e.g. 'slow stir' methods).

Measurement of water solubility does not usually impose excessive demands on chemical techniques. However, measurement of the solubility of sparingly soluble compounds requires extreme care to generate saturated solutions of the material without the introduction of dispersed material; invariably specific methods of analysis are able to determine the low levels (sub ppb-ppm) in solution. Reported water solubility data for such compounds can often contain appreciable errors.

Table R.7.1–5 Test methods for the determination of water solubility

Method details	Applications and requirements	Repeatability and sensitivity
Column elution method Based on elution of the test substance with water from a micro-column which is charged with an inert carrier material such as glass beads, silica gel or sand and an excess of test substance. The water solubility is determined when the mass concentration of the eluate is constant. The mass concentration of the test substance is determined analytically	Applicable to essentially pure substances only Used for low solubilities (< 10^{-2} g/l) Organic substances, but not mobile oils or liquids	< 30% ; down to 1 µg/l
Flask method The test substance is dissolved in water at a temperature somewhat above the test temperature. When saturation is achieved the mixture is cooled and kept at the test temperature, stirring as long as necessary to reach equilibrium The mass concentration of the test substance is determined analytically	Applicable to essentially pure substances and also complex substances. Use of fast stirring techniques (300-400 rpm) appropriate for higher solubility (> 10^{-2} g/l) test substances. Use of slow-stirring techniques (<100 rpm) appropriate for low solubility (< 10^{-2} g/l) test substances (Letinski et al, 2002) Requires equilibration study to determine the time taken to equilibrate the test substance and water	< 15%; down to 1 µg/l
OECD series on Testing and Assessment Number 29 - Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous media.	Applicable to all metals and sparingly soluble inorganic metal compounds	/

R.7.1.7.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for water solubility:

'The study does not need to be conducted if:

- *the substance is hydrolytically unstable at pH 4, 7 and 9 (half-life less than 12 hours); or*
- *the substance is readily oxidisable in water.*

If the substance appears 'insoluble' in water, a limit test up to the detection limit of the analytical method shall be performed.'

For ionising substances, the pH-dependence of the water solubility should be known. At least the pH of the test water needs to be identified. In the context of marine risk assessment, when the pK_a is close to 8 it may be necessary to obtain realistic measurements using seawater.

For volatile compounds, it can be useful to have information on the vapour pressure.

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

Most physical properties, such as molecular weight, melting point, boiling point, density and water solubility can be obtained from commonly used environmental Handbooks, such as Verschueren's Handbook of Environmental Data on Organic Chemicals (1983), Howard's Handbook of Environmental Fate and Exposure Data, Vol. I and II (1990), Lide's CRC Handbook of Physics and Chemistry, Lange's Handbook of Chemistry, the Merck Index, the Aldrich Catalog, Kirk-Othmer Encyclopaedia of Chemical Technology and other handbook compilations such as Riddick *et al*. (1986).

Alternatively, searching on various environmental databases, such as HSDB (http://www.toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB), will provide summaries of chemical and physical properties of substances.

It is not unusual to find in the literature a wide range of solubilities for the same product. The oldest literature generally yields the highest solubility values: this is due to the fact that products were originally not as pure as they are nowadays and also non-specific methods were used which would not differentiate between the dissolved product and any impurities. Reported water solubility data for such compounds can often contain appreciable errors. Therefore, the reliability of data must be demonstrated by providing information on the identity and purity of the test substance, the methodology used to make the measurement, and whether or not this was performed to GLP standards.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Secondary data sources can be used in a WoE approach and they can collectively support the choice of a specific value for the water solubility. These secondary sources have to be based on a critical evaluation of peer-reviewed data and a consequent selection of a reliable and representative value for the water solubility. The use of Klimisch codes, can be extended to these secondary sources and a reliability code of (2) valid with restrictions should be assigned when using an authoritative secondary source.

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

For an organic solute to dissolve in water, firstly, the solute molecules must be separated from one another. Secondly, the solvent molecules must become partially separated from one another to create a cavity large enough to accommodate the solute. Once the solute occupies the cavity, there will be new attractive forces between solute and solvent. Finally, the water molecules in the solvation shell will form extra H-bonds to neighbouring water molecules. Thus, the water solubility depends not only on the affinity of a solute for water, but also on its affinity for its own structure. Molecules that are strongly bound to each other require considerable energy to separate them. This also means that such compounds have high melting points (for solids). Generally, solids with a high-melting temperature have poor solubility in any solvent.

Removal of a molecule from its crystal lattice means an increase in entropy, and this can be difficult to model accurately. For this reason, as well as the fact that the experimental error on solubility measurements can be quite high (generally reckoned to be about 0.5 log unit), the prediction of aqueous solubility is not as accurate as is the prediction of octanol/water partitioning. Nevertheless, many papers (Dearden 2006) and a book (Yalkowsky & Banerjee 1992) have been published on the prediction of aqueous solubility, as well as a number of reviews (Lyman 1990, ECETOC 1998, Reinhard & Drefahl 1999, Mackay 2000, Schwarzenbach *et al.* 2003, Dearden 2006). There are also a number of software programs available for that purpose (ECETOC 2003, Dearden 2006). Livingstone (2003) has discussed the reliability of aqueous solubility predictions from both QSPRs and commercial software.

It should be noted that there are various ways that water solubilities can be reported: in pure water, at a specified pH, at a specified ionic strength, as the undissociated species (intrinsic solubility), or in the presence of other solvents or solutes. Solubilities are also reported in different units, for example g/100 ml, mole/litre, mole fraction. The use of mole/litre is recommended, as this provides a good basis for comparison.

For solids, work has to be done to remove molecules from their crystal lattice, and the simplest way to account for this is to use what Yalkowsky and co-workers have termed the general solubility equation (GSE), which incorporates a melting point term to account for the behaviour of solids (Sanghvi *et al* 2003):

$$\log S_{aq} = 0.5 - \log K_{ow} - 0.01(MP - 25)$$

where MP is the melting point (^oC). The melting point term is taken as zero for compounds melting at or below 25^oC. Calculated log K_{ow} and MP values can be used in the GSE, although measured values are preferred. Aqueous solubilities of 1026 non-electrolytes, with a log S_{aq} range of – 13 to + 1 (S in mole L^{-1}), calculated with the GSE had a standard error of 0.38 log unit.

Good predictions for a large diverse data set have been obtained by the use of linear solvation energy descriptors (Abraham & Le 1999). These included two terms for polarity/polarisability, the sums of hydrogen bond donor and acceptor abilities of the solute molecule, and an expression of molecular volume

According to the Abraham and Le equation, the main factors controlling aqueous solubility seem to be hydrogen bond acceptor ability and molecular size, both of which are important elements in the molecular mechanisms of solubility.

Solubility can vary considerably with temperature, and it is important that solubility data are reported at a given temperature.

Grouping of substances and read-across approach

For the determination of the water solubility read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

For this endpoint, testing should almost always be possible and water solubility should usually be determined experimentally. Nonetheless, testing by the flask method might be precluded when the high viscosity of the saturated test solutions prevent from normal stirring. If it is technically not possible to conduct the study as a consequence of the properties of the substance (e.g. substances flammable in contact with water or substances readily oxidisable in water), testing may be omitted according to general rules for adaptation of the standard testing regime described in REACH Annex XI, Section 2.

Further adaptation possibilities

Not foreseen. However, for complex substances the information obtained from such testing is not relevant or of practical use, and therefore conducting the test may be waived where the data is irrelevant for subsequent assessments.

R.7.1.7.5 Impurities; uncertainties

The water solubility of the test substance can be considerably affected by the presence of impurities.

For a complex substance, the measured solubility is dependent on the amount of test substance added. In practical terms, solubility data are generated using at least two loading rates (e.g. 100 mg/l and 1000 mg/l). Accuracy in determining water solubility decreases as the water solubility of a test substance is reduced (e.g. as shown for reference substance data in OECD Method 105). When dealing with test substances with water solubilities of the order of < 10 µg/l, precautions need to be taken to avoid the introduction of dispersed material into the final extract.

Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

R.7.1.7.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied;
- results from preliminary test (if any);
- chemical identity and impurities (preliminary purification step, if any);
- water temperature during saturation process;
- analytical method employed;
- any evidence of chemical instability;
- all information relevant for the interpretation of the results.

If Column Elution method:

- concentrations, flow rates and pH for each sample;
- mean and standard deviation of five samples at least;
- average for each of two successive runs at least;
- nature and loading of support material;
- solvent used.

If Flask method:

- pH of each sample;
- individual analytical determinations and the average;
- average of the values for different flasks.

Results and discussion & Applicant's summary and conclusion

- water solubility in (mg/l) at temperature (°C);
- pH value and concentration of test substance;
- description of solubility (if relevant);
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on water solubility can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.8	VII 7.7	Water solubility	E.4.9	3.8

R.7.1.7.7 References on water solubility

Abraham M.H. and Le J. (1999) The correlation and prediction of the solubility of compounds in water using an amended solvation energy relationship. *J. Pharm. Sci.* 88, 868-880.

Dearden J.C. (2006) *In silico* prediction of aqueous solubility. *Expert Opinion on Drug Discovery* 1, 31-52.

EC Method A6. "Water Solubility", Dir 92/69/EEC, Official Journal of the European Communities, O.J. L383 A)

EC Method A7. "Hydrolysis ", Dir 92/69/EEC, Official Journal of the European Communities, O.J. L383 A)

ECETOC Technical Report No. 74: *QSARs in the Assessment of the Environmental Fate and Effects of Chemicals*. ECETOC, Brussels, 1998.

ECETOC Technical Report No. 89: (Q)SARs: *Evaluation of the Commercially Available Software for Human Health and Environmental Endpoints with Respect to Chemical Management Applications.* ECETOC, Brussels, 2003.

Livingstone D.J. (2003) Theoretical property predictions. *Current Topics in Med. Chem.* 3, 1171-1192.

Letinski, D.J., Connolly, M.J., Peterson, D.R. and Parkerton, T.F. (2002) "Slow-stir water solubility measurements of selected alcohols and diesters", *Chemosphere*, 48, 257 – 265).

Mackay D. Solubility in water. In Boethling R.S. and Mackay D. (Eds.), *Handbook of Property Estimation Methods for Chemicals: Environmental and Health Sciences.* Lewis, Boca Raton, FL, 2000, pp. 125-139.

OECD Environmental Health and Safety Publications, 2000. Number 23. Guidance document on aquatic toxicity testing of difficult substances and mixtures.

OECD Environment, Health and Safety Publications, 2001. Series on Testing and Assessment, No. 29, Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media

OECD Guidelines for Testing of Chemicals – Method 105 "Water Solubility"

OECD Guidelines for Testing of Chemicals – Method 111 "Hydrolysis as a Function of pH"

Reinhard M. and Drefahl A. Handbook for Estimating Physico-chemical Properties of Organic Compounds. Wiley, New York, 1999.

Sanghvi T., Jain N., Yang G. and Yalkowsky S.H. (2003) Estimation of aqueous solubility by the general solubility equation (GSE) the easy way. *QSAR Comb. Sci.* 22, 258-262.

Schwarzenbach R.P., Gschwend P.M. and Imboden D.M. (2003) *Environmental Organic Chemistry*, 2nd edition, Wiley, Hoboken, NJ.

Yalkowsky S.H. and Banerjee S. (1992). Aqueous Solubility: Methods of Estimation for Organic Compounds. Marcel Dekker, New York.

R.7.1.8 Partition coefficient n-octanol/water

Advice to registrants with regard to nanomaterials characterisation of water solubility can be found in *Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the Guidance on IR&CSA, section 2.2.2 Partition coefficient n-octanol/water.

R.7.1.8.1 Type of property

The n-octanol/water partition coefficient (K_{ow}) is one of the key physicochemical parameters, and it is used in numerous estimation models and algorithms for environmental partitioning, sorption, bioavailability, bioconcentration, bioaccumulation and also human toxicity and ecotoxicity. As such K_{ow} is a critical parameter for chemical safety assessment, classification and labelling, and PBT assessment/screening (where required).

The generation of a K_{ow} value is required at all tonnage bands (i.e. > 1 t/y; information requirements according to REACH Annexes VII-X).

R.7.1.8.2 Definition

The n-octanol/water partition coefficient (K_{ow}) is defined as the ratio of the equilibrium concentrations of a dissolved substance in a two-phase system consisting of the largely immiscible solvents n-octanol and water. The property is moderately temperature-dependent and typically measured at 25°C. For further information on definition and units please see the Test Methods Regulation ((EC) No 440/2008), test method A.8, section 1.2.

R.7.1.8.3 Test method(s)

EU test method A.8 of the Test Methods Regulation ((EC) No 440/2008) describes two test procedures; a direct measurement *via* the Shake Flask method (OECD Test Guideline 107) and a correlation approach using the HPLC method (OECD Test Guideline 117). The Shake Flask method falls within the logKow range -2 to 4 and the HPLC method within the range 0 to 6. The applicability of the methods differ depending on the substance type and the amount of impurities in the test substance. Neither of the methods is applicable to surface active materials, for which an estimated value based on individual solubilities, or a calculated value along with calculation details should be provided. As with any endpoint and predictive method, the documentation and training set of the predictive method should be examined carefully to decide whether it is applicable to special categories of substances, such as zwitterionic or surface active substances.

Regardless of the method used, highly accurate measurements of log K_{ow} > ~5 are complicated by the fact that small amounts of octanol are entrained in the aqueous phase, leading to a potential underestimation of the measured log K_{ow} values. All of the direct methods for measuring log K_{ow} require quantifying the test material in either octanol or water and preferably in both matrices.

In addition, the OECD test guideline 123, Slow-stirring method, can be used to generate data for this endpoint.

Table R.7.1–6 Methods for determination of partition coefficient n-octanol / water

Method details	Repeatability	Applicability range
Shake Flask Method (EU A.8, OECD TG 107) The Shake Flask method is the default procedure. It is considered to give accurate results for low to medium hydrophobic substances. For substances with a high expected log K_{ow}, alternative methods are recommended. A suitable analytical method is needed to determine the concentration of the test material in the octanol and water phases. By applying mass balance considerations, it may be possible to measure the test material in only the less-soluble phase. However, this approach significantly decreases the reliability in the reported value. This technique is not suitable for surface active compounds (surfactants), or compounds that hydrolyse rapidly.	Three replicates should fall within +/- 0.3 log K_{ow}	-2 < log K_{ow} < 4
HPLC Method (EU A.8, OECD TG 117) This is a relatively quick way of estimating log K_{ow}. It is not measured directly, but from a correlation between log k (capacity factor) and log K_{ow} for a series of reference substances. It therefore depends on the quality of the log K_{ow} measurement of reference substances (often measured by the shake flask method). A series of reference compounds with similar chemical functionality to the test material should be used to generate the log k: log K_{ow} correlation. In general, the HPLC method is less sensitive to impurities than the shake flask method. The RP-HPLC is not recommended for strong acids and bases, metal complexes or surface active agents, or for measurements across very different classes of substances. The HPLC method is also very suitable for measuring the K_{ow} of mixtures of chemical homologues.	Three replicates should fall within +/- 0.1 log K_{ow}	0 < log K_{ow} < 6
Slow-Stirring Method (OECD TG 123) This is a more recent method developed as an alternative to the shake flask procedure (OECD TG 107, EU A.8). The advantage of slow stirring versus shaking is that emulsion formation will be reduced. The method requires a few days to reach equilibrium. The method may be difficult to adapt to a high throughput approach. As with the other direct methods, a suitable analytical method is needed to measure the concentration of the test material in the octanol and water phases. NB: Radiolabelled substances – which may be synthesised for use in other tests – can be very useful for accurate log K_{ow} determination.	Intralaboratory median standard deviation from 0.15 – 0.3 Log K_{ow} (Tolls et al, 2003).	Validation has shown that this method can also be used for very hydrophobic substances, up to Log K_{ow} 8.3 (OECD 2003, Tolls *et al*, 2003).
Estimation method based on individual solubilities in EU A.8 This method enables partition coefficients to be estimated based on the ratio of the solubility of the material in octanol and water. For some substances (e.g. some surfactants and pigments) it is technically not feasible (or good practice) to measure an octanol-water		

partition coefficient by OECD 107. For such substances it may be possible to obtain a ratio of the saturated water solubility (OECD 105) and saturated octanol solubility (no guideline currently available but based on the principles of OECD 105). This method however has the drawback that it does not include the interaction between the water and solvent phase (i.e. a substance with high Kow is rather 'pushed out of the water' than 'pulled into octanol'). This explains the poor correlation typically observed between octanol solubility and K_{ow} (Dearden, 1990, Sijm *et al*., 1999). The ratio was found to be somewhat more representative if one uses octanol/saturated water and water/saturated octanol. As such, a ratio estimation would be a less preferred yet acceptable alternative for the octanol/water partition coefficient (Kow), but must be treated with caution as it would not have been derived in the same manner as other K_{ow}s (OECD TG 107).		

R.7.1.8.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for n-octanol water partition coefficient:

'The study does not need to be conducted if the substance is inorganic. If the test cannot be performed (e.g. the substance decomposes, has a high surface activity, reacts violently during the performance of the test or does not dissolve in water or in octanol, or it is not possible to obtain a sufficiently pure substance), a calculated value for log P as well as details of the calculation method shall be provided.'

If experimental testing including estimation from the individual solubilities is not possible, log K_{ow} must normally be calculated by an appropriate numeric method based on the molecule's structure.

In case of rapid hydrolysis the registrant needs to provide evidence in the form of a hydrolysis endpoint study record (study summary) and should consider testing for the hydrolysis products instead, as information on the properties of (environmentally and toxicologically) relevant degradation products are needed for conducting the risk assessment of the substance to be registered.

Adaptation possibilities according to Annex XI to REACH

The reporting of the K_{ow} information cannot usually be waived (except for inorganic substances), because it is essential for CSA, classification and labelling and PBT assessments.

Use of existing data: Data on physicAL-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

Log K_{ow} is a commonly documented property in substance databases, such as IUCLID (http://iuclid.eu/index.php?fuseaction=home.iuclidHome). Additional sources are the Canadian National Committee for CODATA (CNC/CODATA) database with evaluated log K_{ow} values for over 20000 substances (http://logkow.cisti.nrc.ca/logkow/) and the QSAR Toolbox (http://www.qsartoolbox.org).

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Measured values are given precedence over calculated values. For organic substances experimentally derived high-quality K_{ow} values, or values which are evaluated in reviews and assigned *recommended values*, are preferred over other determinations of K_{ow}. Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

When no experimental data of high quality are available, or if experimental methods are known to be unreliable, valid (Q)SARs for log K_{ow} may be used e.g. in a weight-of-evidence approach. Due to the availability of large number of measured log K_{ow} values and robust QSAR models for this property, the QSARs can, in some cases, predict the partition coefficient of a molecule with higher accuracy compared to a single test. Such valid QSAR models may be used if they are restricted to substances for which their applicability is well characterised. In order to be used as a stand alone source of values to meet the data requirements of Annex VII, 7.8, the QSARs must meet the criteria set out in Annex XI, 1.3.

Grouping of substances and read-across approach

For the determination of the partition coefficient n-octanol/water read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies.

Further adaptation possibilities

Not foreseen.

R.7.1.8.5 Impurities; uncertainties

The effect of impurities in the test substance are discussed in the referenced test guidelines.

Difficult to test substances:

There are certain structural or physico-chemical properties that can make the accurate determination of Kow or its measurement difficult. Difficult to test substances include poorly soluble, volatile, surface active, ionisable substances, mixtures of substances, as well as substances subject to rapid degradation due to such processes as phototransformation, hydrolysis, oxidation, or biotic degradation.

Guidance on regulatory compliant Kow determination for ionisable substances and salts:

The Kow is typically defined as the partition coefficient of the neutral, undissociated form of a substance. However, the relative extent to which an ionisable substance is likely to be dissociated in the environment (with pH usually in the range 5-9) can have a marked effect on its physicochemical properties, especially the octanol-water partition coefficient and water solubility, which in turn affect fate and behaviour. As log Kow is routinely used to predict bioconcentration/bioaccumulation potential, this aspect is especially important in a PBT context. For substances which dissociate within an environmentally relevant pH range (pK_a 5-9), values for Kow must be derived for the neutral form, and preferably also for the dissociated form. In some cases a factor 4-5 has been recorded between the log Kow of both species. The value for the dissociated molecule determined around a pH of 7 (sometimes referred to as Dow) is considered more realistic for PBT and chemical safety assessment.

Based on practical experience the following guidance is provided:

Simple acids and bases in the normal pH range:

- The HPLC method is to be applied to acids and bases in their non-ionised forms, although the pH should be kept in the range 2 to 9 (however pH 5 to 9 is preferred).
- For the shake-flask method, the approach must be followed in which the study is conducted at a pH where the substance is not ionised, if possible, or at a pH where the extent of ionisation is minimised.
- Validated QSAR estimations may be useful for acids and bases.

Zwitterionic substances:

- For zwitterions, the shake-flask method should be used to develop a valid Kow value. Even if the ionic charge pattern of the compound in octanol is not known, the value represents a practical and useful parameter. It is not justifiable to expect a full description of all the equilibria in both water and octanol. The pH of such a study should be 7 or the iso-electric point (pH value at which the molecule has no net electrical charge), as long as that point is in the range pH 5 to 9, so as to maximise the possibility of partition into octanol. There is no need to give both pH values.
- The HPLC method must not be used. The usual estimation methods should be valid, but particular care should be exercised.
- QSAR estimations may be useful provided that they are validated.

Salts of organic compounds:

- The shake-flask method should be used, usually at pH 7, or at any pH in the range 5 to 9 which maximises the potential for partition into octanol. For salts, the nature of the analytical method compared to the chemical composition will have to be considered. The ideal is to monitor cation and anion** individually in both phases. When only one half can be analysed, then the result must be understood as partial, even if it is the best that is achievable.
- Estimation by HPLC is not valid for the whole salt.
- QSAR methods will be valuable in assessing the properties of each half of the salt. Current estimation methods cannot estimate the Kow of the ion pair.

Guidance on regulatory compliant Kow determination for surfactants:

In many cases a calculated Kow value based on the octanol and water solubilities will be the first choice for surfactants. It is also useful to compare a calculated with a measured value. For the calculation approaches, one needs to consider the pH of the system (which determines the ionisation of the surfactant – see Section R.7.1.17). None of the experimental methods is very well suited for determining the Kow of surface active substances. The shake flask method is the least suitable experimental method for surfactants. HPLC methodology may fail due to secondary interactions, and is sensitive to fluctuations of ionic strength. The slow stirring method in theory is the best, but still not demonstrated to be perfect. If using slow stirring, one needs to demonstrate a consistent result when starting with the surfactant in either phase, not just in the octanol. A working approach for surfactants might be the comparison of measured solubilities in octanol and water. However, it would then be prudent to take the critical micelle concentration in water (CMC) as a solubility limit, in order to avoid the artefact of unrealistically low K_{ow} values.

Guidance on regulatory compliant Kow determination for mixtures:

It is possible that different components of mixtures have significantly different behaviour in the physico-chemical tests and therefore also *in vivo* and in the environment. It is therefore important to ensure that the results presented for the physico-chemical tests represent each component rather than the mixture being treated as a single component. For simple mixtures where the components are known and easily identifiable, this may mean presenting individual values for K_{ow}. For complex mixtures, the HPLC method is ideal for determination of K_{ow}, and a defined range of values should be presented, with an indication of the proportion of substance within a given range (e.g. > 90% of components have log K_{ow} in the range 4-5), to allow the significance of these results to be reflected in the risk assessment. The HPLC method is also recommended for petroleum products, which are typically mixtures.

R.7.1.8.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

Shake-flask method (EU A.8 / OECD TG 107):

- equilibrium concentrations of the test substance in both phases;
- relative volumes of the two phases;
- analytical method(s).

Calculation method (EU A.8):

- identification of the method;
- working principle of the method;
- reference to the method;
- information on source chosen to justify K_{ow} values of fragments being manipulated;
- applicability of the method.

HPLC method (EU A.8 / OECD TG 117):

- column(s) used;
- mobile phase (composition, buffer, pH);
- reference substances with respective Kow values from the literature;
- concentrations measured.

Slow-stirring method (OECD TG 123):

- label purity of labelled substances and molar activity (where appropriate);
- sampling times;
- description of the test vessels and stirring conditions;
- number of replicates;
- temperature during the experiment;
- volumes of 1-octanol and water at the beginning, during and remaining after the test;

- determined concentrations of the test substance in 1-octanol and water as a function of time;
- description of the test vessels and stirring conditions (geometry of the stirring bar and of the test vessel, vortex height in mm, and when available: stirring rate) used;
- analytical methods used to determine the test substance (its repeatability and sensitivity) and the method limit of quantification;
- sampling times;
- pH of the aqueous phase and of the buffers used, when pH is adjusted for ionisable molecules;
- number of replicates;
- demonstration of mass balance;
- temperature and standard deviation or the range of temperature during the experiment;
- the regression of concentration ratio against time.

Results and discussion

- final value for log Kow;
- Kow values and their mean;
- standard deviation of individual Kow values;
- theoretical value when it has been calculated;
- temperature of the test solutions (°C);
- pH value(s) of the aqueous solution(s);
- composition and concentration of buffers;
- concentration of the stock solution;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used and reasons for it or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on partitition coefficient can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.7	VII 7.8	Partitition coefficient	E.4.8	3.7

R.7.1.9 Flash point

R.7.1.9.1 Type of property

The flash point is a property:

- for substance characterization;
- for the classification of flammable liquids;
- which is importiant for physical hazard assessment.

R.7.1.9.2 Definition

The flash point is the lowest temperature of the liquid (as measured in a prescribed manner) at a pressure corrected to 101.325 kPa, at which application of an ignition source causes the vapour of the liquid to ignite momentarily and the flame to propagate across the surface of the liquid under the specified conditions of test (see section 1.2, Test Method A.9).

R.7.1.9.3 Test method(s)

The EU test method A.9 – Flash point from the Regulation (EC) 440/2008 can be used. Suitable methods are listed in the CLP Regulation Annex I, 2.6.4.4, Table 2.6.3.

The method to be used has to be chosen taking into account the properties of the liquid (viscosity, halogenated compounds present) and the scope of the standard.

For substances with a high decomposition potential, a method using small amounts of liquid (e.g. EN ISO 3679: Determination of flash point - Rapid equilibrium closed cup method) is recommended to reduce the amount of substance under test.

For classification purposes it is recommended to use the mean of at least two test runs. If the experimentally determined flashpoint is found to be within ± 2°C of the limiting criterion for classification or assigning a category when using a non-equilibrium method, it is recommended to repeat the determination with an equilibrium method.

R.7.1.9.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for flash point:

'The study does not need to be conducted if:

- *the substance is inorganic;*
- *the substance only contains volatile organic components with flash-points above 100°C for aqueous solutions; or*
- *the estimated flash-point is above 200°C; or*
- *the flash-point can be accurately predicted by interpolation from existing characterised materials.'*

The first point has to be further specified as:

- The substance is inorganic except where there are covalent bonds;

because some inorganic liquids with covalent bonds are flammable e.g. CS_2, N_2H_2, HCN.

The third point should only be applied when a well validated estimation model was used.

The fourth point should only be applied when there are enough reliable experimental data from existing characterised materials to be able to accurately interpolate to estimate the flash point.

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) which meet the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

For the determination of the flash point, weight of evidence is not possible.

(Q)SAR

For the determination of the flash point, QSAR approaches are discouraged for the purpose of classification / risk assessment, except where the mean absolute error of the QSAR is less than 2°C.
For non-halogenated liquids calculation based on the vapour pressure curve and lower explosion limit of the substance can be used as a screening test and a flashpoint need not be determined experimentally if the calculated value is at least 5°C higher than the relevant classification criterion.

Grouping of substances and read-across approach

For the determination of the flash-point read-across is usually not possible. However interpolation may still be possible within homologous series.

Testing is technically not possible

This applies if:

- the liquid is an explosive;
- the liquid is pyrophoric or self-reactive;
- decomposition occurs during the melting point study;
- some impurities have an inpact on the ignition source in such a way as to distort/invalidate the results.

Testing should always be considered, if none of the waiving possibilities applies.

Further adaptation possibilities

The flash point does not need to be determined experimentally if conclusive and consistent literature data are available.

Data generated with the same tests and classification principles as specified in the CLP Regulation for flash point generated in conjunction with transport classification can satisfy the REACH requirements, but this needs to be checked on a case by case basis.

R.7.1.9.5 Impurities; uncertainties

Impurities can have a significant influence on the flash point. The influence depends on the amount and the vapour pressure of the impurity. Even if their concentration is below 0.5 %, especially if their boiling point is substantially lower, they may have a strong effect on the flash point. Impurities with a higher boiling point will normally have no effect on the flashpoint. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

R.7.1.9.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

- reference to the standard or the test method applied;
- open cup or closed cup (for classification purposes only the closed cup methods are allowed);
- equilibrium or non-equilibrium method.

Results and discussion

- corrected flashpoint and unit;
- data on repeatability and reproducibility as given in the method;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flash point can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.11	VII 7.9	Flash point	E.4.12	3.10

R.7.1.10 Flammability

Some of the information requirements according to REACH Annex VII were phrased in a way that they correspond to 'indications of danger' as given in Annex II of the DSD. For substances, classification and labelling according to CLP Regulation has been mandatory since 1 December 2010 (and will become mandatory for mixtures (preparations) from 1 June 2015, when the DPD will be repealed). Consequently properties associated with flammability are covered by classification of the substance according to the CLP Regulation. However, the physical hazards according to the CLP Regulation are structured completely differently from the physicochemical properties according to the DSD (and therefore also REACH, Annex VII). This means that for some of the CLP hazard classes an unambiguous assignment to one of the headlines (information requirements) in Annex VII to REACH is not possible. The assignment of hazard classes to the headline 'Flammability' as shown in the table below (Table R.7.1–7) must therefore only be understood as a means to structure this document in accordance with Annex VII to REACH. It has to be noted that self-reactive substances and organic peroxides are assigned to the headline 'Flammability' and only a cross reference is added under the headline 'Explosive properties' because these two hazard classses can have explosive and/or flammable properties.

Table R.7.1–7 Assignment of CLP hazard classes to the information requirement 'Flammability' according to REACH, Annex VII and correlation between the Test Method Regulation and the test method according to CLP and supporting link with the *Guidance on the Application of the CLP criteria*.

Information requirement according to Art. 10 (a) (vi) of the REACH Regulation (EC) No. 1907/2006 (the no. in brackets is the respective no. in the table in Annexes VII to IX to REACH)	**CLP Regulation (EC) No. 1272/2008** (the no. in brackets is the respective chapter no. in Annex I to CLP)	**Chapter in revised R.7(a) guidance**	**Corresponding test method according to The Test Method Regulation Regulation (EC) No 440/2008**	**Corresponding test method according to CLP Regulation**	**Chapter in the Guidance on the application of the CLP Criteria (ex RIP 3.6)**
Flammability (7.10)	Flammable gases[17] (2.2)*	R.7.1.10.1	A.11 Flammability (gases)	**ISO 10156** **EN 1839**	2.2
	Flammable liquids (2.6)*	R.7.1.10.2	for liquids: see Flash point	**see CLP, Annex I, Chapter 2.6.4.4, Table 2.6.3**	2.6
	Flammable solids (2.7)*	R.7.1.10.3	A.10 Flammability (solids)	**UN Test N.1**	2.7

[17] The 4th ATP to the CLP Regulation amends the criteria in the CLP Annex I, Section 2.2 Flammable gases by including subclassifications for chemically unstable gases.

	Self-reactive substances and mixtures (2.8)*	R.7.1.10.4	n.a.	**UN Test series A to H**	2.8
	Pyrophoric liquids (2.9)*	R.7.1.10.5	A.13 Pyrophoric properties of solids and liquids	**UN Test N.3**	2.9
	Pyrophoric solids (2.10)*	R.7.1.10.6		**UN Test N.2**	2.10
	Self-heating substances and mixtures (2.11)*	R.7.1.10.7	n.a.	**UN Test N.4**	2.11
	Substances and mixtures which in contact with water emit flammable gases (2.12)*	R.7.1.10.8	A.12 Flammability (Contact with water)	**UN Test N.5**	2.12
	Organic peroxides (2.15)*	R.7.1.10.9	n.a.	**UN Test series A to H**	2.15

* Note that regardless of whether the hazard class or category is listed in Article 14(4)(a) REACH the chemical safety assessment (where required) must be performed in accordance with Article 14(3) REACH. Furthermore, according to Article 10(a)(iv) of REACH the technical dossier of a registration of a substance under the REACH Regulation must include information on classification and labelling of the substance as specified in section 4 of Annex VI to the REACH Regulation.

In addition, it has to be noted that some substances have flammable properties which do not result in classification. Examples are the following:

- gases that do not have a flammable range at 20°C and standard pressure (and therefore are not classified as flammable gases) might have a flammable range at higher temperatures and/or pressure (e.g. ammonia);
- liquids that do not have a flash point (and therefore are not classified as flammable liquids) might have an explosion range (especially halogenated hydrocarbons).

Information about such properties should also be indicated in the dossier.

R.7.1.10.1 Flammable gases

Definition

'Flammable gas means a gas or gas mixture having a flammable range with air at 20°C and a standard pressure of 101.3 kPa' (Annex I to CLP, Section 2.2.1).

Classification criteria and relevant information

Flammable gases are classified into two categories depending on their flammability range (Annex I to CLP, Section 2.2.2. Table 2.2.1).

Detailed guidance on the classification criteria and the test method(s) can be found in the *Guidance on the Application of the CLP criteria*, section 2.2[18].

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for flammability:

'The study does not need to be conducted:

- *if the substance is a solid which possesses explosive or pyrophoric properties. These properties should always be considered before considering flammability;or*
- *for gases, if the concentration of the flammable gas in a mixture with inert gases is so low that, when mixed with air, the concentration is all time below the lower limit; or*
- *for substances which spontaneously ignite when in contact with air.'*

The relevant points can be paraphrased (first point is not relevant for this chapter), namely the study does not need to be conducted:

- if the concentration of the flammable gas in a mixture when mixed with air is below the lower limit;
- if the gas spontaneously ignites when in contact with air.

Gases that spontaneously ignite in contact with air are pyrophoric and are therefore flammable gases.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Many gases are classified in Annex VI to CLP either as Flam. Gas 1 or Flam. Gas 2, and additional flammable gases are listed in the UN-RTDG whose classifications correspond to Flam. Gas 1 according to CLP.

---

18 The 4th ATP to the CLP Regulation amends the criteria in the CLP Annex I, Section 2.2 Flammable gases by including subclassifications for chemically unstable gases. Consequently the *Guidance on the Application of the CLP criteria*, Part 2: Physical hazards has been restructured to take account of the 4th ATP, which applies to substances from 1 December 2014 and to mixtures from 1 June 2015. When the 4th ATP is applied a Guidance corrigendum will be made to delete the outdated sub-chapter 2.2.1 Flammable gases in the *Guidance on the Application of the CLP criteria*.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

For gases that are not classified in Annex VI to the CLP Regulation nor in the UN-RTDG, there is ample scientific literature giving the flammability range for most gases (e.g. IEC 60079-20-1 *Data for flammable gases and vapours, relating to the use of electrical apparatus* – (under revision).

- (Q)SAR

At present (Q)SAR is generally not applicable for determination of explosion (/flammability) limits of gases.

- Grouping of substances and read-across approach

For the determination of the flammable gases read-across is usually not possible. However interpolation may still be possible within homologous series.

- Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies.

Further adaptation possibilities

Further adaptation is possible for gases that are known to be non-flammable. Examples are nitrogen, the noble gases (helium, neon, argon, krypton, xenon), carbon dioxide and sulphur hexafluoride. As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

Impurities; uncertainties

Tests should be performed with the lowest concentration of impurities in the gas encountered in the normal manufacturing process and the moisture content should be less than or equal to 0.01 mol %. Utmost care should be taken in the selection of the key study(ies) and/or use of weight-of-evidence approaches that the data selected is representative of the substance being registered by the respective companies.

How to conclude on the DSD classification

All gases with a flammability range in air are classified 'Extremely flammable F+ ; R12' according to DSD, unless classified differently according to Annex VI, Table 3.2 of the CLP Regulation. This means that all gases classified as flammable gases according to CLP (either Category 1 or 2) are classified as 'Extremely flammable F+; R12'.

Endpoint specific information in the registration dosser/in IUCLID

Material and methods:

- description of the apparatus and dimensions or reference to the standard or the test method applied;

- test temperature;
- tested concentrations.

Results and discussion & Applicant's Summary and conclusion (interpretation of results)

- indicate lower and upper explosion limits in % volume;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flammability can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.13	VII 7.10	Flammability	E.4.14	3.12

Further information / references

For the testing of flammable gases according to CLP classification requirements, refer also to the *Guidance on the Application of the CLP criteria*, section 2.2, and in Directive 2008/47/CE.

R.7.1.10.2 Flammable liquids

Definition

Flammable liquid means a liquid which may form flammable/explosive vapour-air mixtures. Within the CLP Regulation 'Flammable liquid' means a liquid having a flashpoint of not more than 60°C (see CLP Annex I, section 2.6.1).

Classification criteria and relevant information

Flammable liquids are classified in three categories according to the criteria of the CLP Regulation (see CLP Annex I, section 2.6, table 2.6.1) based on their boiling point and their flash point. Derogation is possible (see CLP Annex I, section 2.6.4.5) for Flam. Liquid Cat. 3 having a flashpoint above 35°C based on the information on sustained combustibility. Furthermore, gas oils, diesel and light heating oils having a flash point between ≥ 55°C and ≤ 75°C may be regarded as Category 3 flammable liquids according to the CLP Regulation (CLP Annex I, section 2.6, footnote to table 2.6.1).

In addition EUH018 - *'In use may form flammable/explosive vapour-air mixture'* has to be assigned to substances classified under the CLP Regulation which may form flammable/explosive vapour-air mixtures although they do not have a flash point e. g. CH_2Cl_2, $C_2H_3Cl_3$. In such cases it is possible to make the decision on whether flammable/explosive vapour-air mixture may be formed based on either the determination of explosion limits according to EN 1839 or the determination of explosion points according to EN 15794. It is sufficient to determine either the lower explosion limit or the lower explosion point.

Detailed guidance on the classification criteria and the test method(s) can be found in the *Guidance on the Application of the CLP criteria*, section 2.6.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

The entries ‘flammability’ (7.10), ‘boiling point’ (7.3) and ‘flashpoint’ (7.9) are the relevant ones. For the latter two entries, see their respective relevant sections in this document.

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for flammability:

‘The study does not need to be conducted:

- *if the substance is a solid which possesses explosive or pyrophoric properties. These properties should always be considered before considering flammability;or*
- *for gases, if the concentration of the flammable gas in a mixture with inert gases is so low that, when mixed with air, the concentration is all time* [i.e. ‘always’] *below the lower limit; or*
- *for substances which spontaneously ignite when in contact with air.’*

The relevant points can be paraphrased (first two points are not relevant for this chapter), namely the 3rd point specifies that for flammability, Annex VII to REACH does not require testing for substances which spontaneously ignite when in contact with air.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

To be used as a stand alone value to meet the data requirements of Annex VII, 7.8, QSAR models must meet the criteria set out in Annex XI, 1.3. The entries ‘boiling point’ (7.3) and ‘flashpoint’ (7.9) are also the relevant ones, therefore please check under each respective QSAR sub-section for more information.

Sustained Combustibility:

No (Q)SAR exists currently.

For further reference see also *the Guidance on the Application of the CLP criteria*, section 2.6.

- Grouping of substances and read-across approach

The entries ‘boiling point’ (7.3) and ‘flashpoint’ (7.9) are again the relevant ones. For both these entries, see their respective sections in this document.

Sustained Combustibility:

For the determination of the sustained combustibility read-across is usually not possible. However interpolation may still be possible within homologous series.

- Testing is technically not possible

Testing is not possible if:

- the liquid is an explosive;
- the liquid is pyrophoric or self-reactive.

Testing should always be considered if none of the waiving possibilities applies.

Further adaptation possibilities

Use of data on boiling point, flashpoint when determined with a closed cup method, explosion limits or lower explosion point from validated literature (see below chapter Further information/ references) is possible. Data on boiling point generated in relation to transport classification may also satisfy the Annex XI requirements. Data on flashpoint generated in relation to with transport classification may satisfy the Annex XI requirements if closed cup methods have been used. However care has to be taken in cases where there is no transport classification as ‘flammable liquid’, because certain substances can form flammable/explosive vapour-air mixtures although they do not have a flash point.

As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

Impurities; uncertainties

Boiling point:

Impurities will influence the boiling point of the main component. The influence depends on the amount and boiling point of the impurity. The higher the amount and the higher the difference between the boiling points of the main component and the impurity, the higher the influence.

Flashpoint:

Special care has to be taken when a sample contains impurities with a lower boiling point than the main component. Even if their concentration is below 0.5%, especially if their boiling point is substantially lower, they may have a strong effect on the test result. Impurities with a higher boiling point will normally have no effect on the flashpoint.

Sustained combustibility:

Impurities with lower boiling point may influence the ability to sustain combustion. However it is not yet possible to quantify the influence of impurities.

How to conclude on the DSD classification

Based on the data on boiling point and flashpoint the DSD classification according to the respective DSD criteria is possible. Simplified direct translation between CLP classification and DSD classification is not possible, see figure below (Figure R.7.1–1).

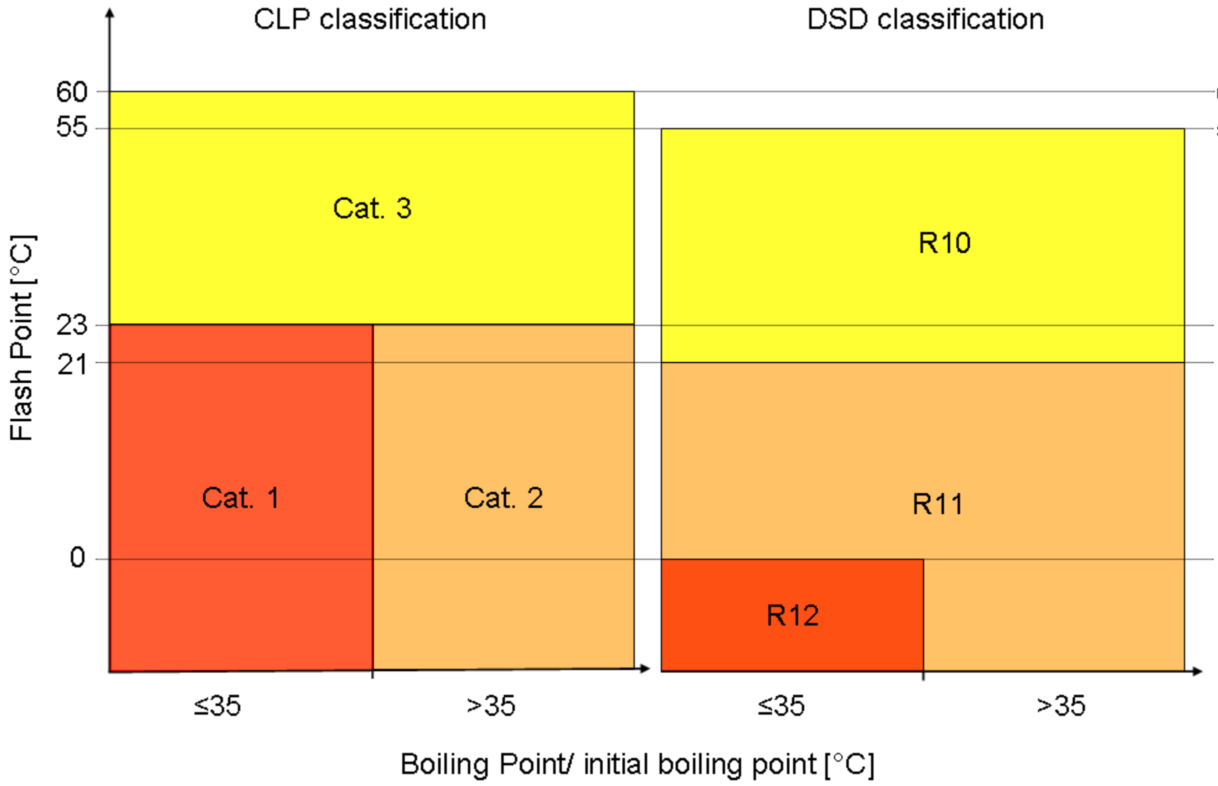


Figure R.7.1–1 Comparison of the DSD and the CLP classification

Substances exempted from classification in Cat. 3 because of their flashpoint and behaviour when tested for sustained combustibility can be exempted from being classified under DSD as R10, if they don't show additional dangerous properties relevant for classification.

Endpoint specific information in the registration dosser / in IUCLID

The physicochemical studies reporting data relevant for classification as a flammable liquid (flashpoint and boiling point) are to be reported in the relevant IUCLID endpoint records.

Material and methods

See chapter R.7.1.9 Flash point and R.7.1.3 Boiling point.

Results and discussion

- corrected flashpoint and unit;
- data on repeatability and reproducibility as given in the method;
- boiling point value (°C) as measured;
- pressure value and unit;
- rate of temperature increase;
- decomposition (if applicable);
- measurement uncertainty if available;
- boiling point value in ºC (corrected to standard pressure, except where the boiling point was determined at reduced pressures) (as above, but in a separate block of fields);
- if available explosion limits;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flammability can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.13	VII 7.10	Flammability	E.4.14	3.12

Further information / references

See also R.7.1.3 Boiling point and 0 Flash point. For testing of flammable liquids according to CLP classification requirements refer also to the *Guidance on the Application of the CLP criteria*, section 2.6.

R.7.1.10.3 Flammable solids

Definition

'A flammable solid means a solid which is readily combustible, or may cause or contribute to fire through friction. Readily combustible solids are powdered, granular, or pasty substances or mixtures which are dangerous if they can be easily ignited by brief contact with an ignition source, such as a burning match, and if the flame spreads rapidly' (see CLP Regulation, Annex I, section 2.7.1).

Classification criteria and relevant information

Solid substances and mixtures are classified as flammable in two categories according to their burning behaviour (see the CLP Regulation, Annex I, section 2.7) using UN Test N.1 as described in section 33.2.1 of the UN-MTC.

Chapter 2.7 of the *Guidance on the Application of the CLP criteria* gives detailed information on the CLP classification of flammable solids, the UN Test N.1 and the relation to the DSD and the UN-RTDG regulations.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for flammable solids:

'The study does not need to be conducted:

- *if the substance is a solid which possesses explosive or pyrophoric properties. These properties should always be considered before considering flammability;or*
- *for gases, if the concentration of the flammable gas in a mixture with inert gases is so low that, when mixed with air, the concentration is all time* [i.e. always] *below the lower limit; or*
- *for substances which spontaneously ignite when in contact with air.'*

Concerning the first indent, testing for flammability of a solid is a part of classification in CLP Regulation. Refer also to the *Guidance on the Application of the CLP criteria*, section 2.7 on classification requirements. For substances having explosive properties, testing for a classification as a flammable solid may be waived. This applies to substances and mixtures classified as explosives, organic peroxides and self-reactive substances and mixtures.

Second indent is not applicable for this endpoint.

With regards to the third indent, substances which spontaneously ignite when in contact with air are pyrophoric substances as defined by the CLP Regulation (see the *Guidance on the Application of the CLP criteria*, section 2.10). Such substances are not classified as flammable solids but as pyrophoric solids under the CLP Regulation.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

Literature data – even if available – should not be used since flammability strongly depends on particle size, surface treatment and other parameters.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

If available data from an A.10 test method indicate that a classification as a flammable solid does not apply (result: not highly flammable), no more testing is necessary. However, if the A.10 test method has come to the conclusion ‘highly flammable’, it will be necessary to also determine the influence of the wetted zone as described in the UN Test N.1.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

At present (Q)SAR is generally not applicable for flammable solids. Application of (Q)SAR is not possible.

- Grouping of substances and read-across approach

At present, grouping and read across are not applicable.

- Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies.

Further adaptation possibilities

As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

If a suitable screening test clearly shows that the substance is not flammable, further testing is not necessary (see also the *Guidance on the Application of the CLP criteria*, section 2.7.4.2). An example for a suitable screening test is the burning index as described in VDI guideline (VDI Guideline, 1990) if a burning index of 3 or less is found, the substance should not be classified as a flammable solid and no further testing is required.

Substances and mixtures classified according to the CLP Regulation as explosives, organic peroxides, self-reactive substances and mixtures as well as pyrophoric or oxidising solids should not be considered for classification as flammable solids (see the *Guidance on the Application of the CLP criteria*, section 2.7.3).

However, if a substance gives a positive result in UN Test Series 1 or 2 as described in the UN-MTC, but is exempted from classification as an explosive on the basis of UN Test Series 6, a test for classification as a flammable solid should be performed.

Impurities; uncertainties

Impurities do not tend to have a large effect on the flammability of a solid. However, if a solid which is not flammable in the pure state contains flammable organic liquids or organometallic impurities it may burn more rapidly and thus become flammable. Therefore utmost care should be taken in the selection of the key study(ies) and during use of weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

How to conclude on the DSD classification

Any substance found to be a flammable solid according to CLP Regulation has to be classified as ‘F; R11’ according to the DSD.

Endpoint specific information in the registration dosser/in IUCLID

Material and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied.

Solid flammability:

- indicate if preliminary and/or main test performed;
- moisture content;
- particle size and distribution (if available) (see R.7.1.14 Granulometry).

Results and discussion

- indicate burning time;
- pass/non pass of the wetted zone (in the case of the UN Test N.1);
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flammability can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.13	VII 7.10	Flammability	E.4.14	3.12

Further information / references

VDI guideline 2263, part 1, (1990): 'Test methods for the Determination of the Safety Characteristics of Dusts'.

For testing of flammable solids according to CLP classification requirements, refer also to the *Guidance on the Application of the CLP criteria*, section 2.7.

R.7.1.10.4 Self-reactive substances and mixtures

In the CLP Regulation self-reactive substances are a distinct hazard class. Self-reactive substances are classified into one of the seven categories of 'Types A to G' according to the classification criteria given in section 2.8.2.3 of Annex I of CLP. In the Dangerous Substances Directive (67/548/EEC) no hazard class for 'self-reactive substances' is defined. Nevertheless, self-reactive substances were also classified as dangerous according to the DSD, e.g. as flammable or as substances with explosive properties.

As mentioned below under the sub-section 'Definition', self-reactive substances are excluded from testing as explosives according to Test Series 1 to 8 in Part I of the UN-MTC (see R.7.1.11.1 Explosives). In Test Series A to H however, no tests on sensitivity to impact (solids and liquids) and friction (solids only) are included. For the risk assessment and the safe use and handling, data according to the EU test method A.14 as described in Regulation (EC) No 440/2008, if available, or UN Test 3 (a) (ii) BAM Fallhammer and Test 3 (b) (i) BAM friction apparatus (see R.7.1.11) should be part of the hazard communication in the registration dossier (REACH Annex VII, 7.11) and the safety data sheet.

Definition

The definition of a self-reactive substance is given in section 2.8.1 of Annex I to CLP Regulation:

'Self-reactive substances or mixtures are thermally unstable liquid or solid substances or mixtures liable to undergo a strongly exothermic decomposition even without participation of oxygen (air). This definition excludes substances and mixtures classified according to this Part as explosives, organic peroxides or as oxidising. A self-reactive substance or mixture is regarded as possessing explosive properties when in laboratory testing the formulation is liable to detonate, to deflagrate rapidly or to show a violent effect when heated under confinement.'

Background information and guidance on the definition is given in the *Guidance on the Application of the CLP criteria*, sections 2.8.1 and 2.8.2.

Classification criteria and relevant information

Classification principles are given in CLP Regulation Annex I, sections 2.8.2 and 2.8.4. Background information and guidance on relevant aspects regarding the classification is given in the *Guidance on the Application of the CLP criteria*, sections 2.8.4, 2.8.5 and 2.8.6.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Only self-reactive substances, as defined in the section definition, have to be tested according to the UN-MTC, Part II test series A - H.

CLP Annex I, section 2.8.2.1 provides the following specific rules for adaptation of the standard information requirement for self-reactive substances and mixtures.

'Any self-reactive substance or mixture shall be considered for classification in this class as a self-reactive substance or mixture unless:

a. *they are explosives, according to the criteria given in 2.1;*

b. *they are oxidising liquids or solids, according to the criteria given in 2.13 or 2.14, except that mixtures of oxidising substances, which contain 5 % or more of combustible organic substances shall be classified as self-reactive substances according to the procedure defined in 2.8.2.2;*

c. *they are organic peroxides, according to the criteria given in 2.15;*

d. *their heat of decomposition is less than 300 J/g; or*

e. *their self-accelerating decomposition temperature (SADT) is greater than 75 º C for a 50 kg package[19].'*

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

A number of already tested and classified substances and mixtures are listed in UN-RTDG, section 2.4.2.3.2.3. Available information may originate from the classification for transport. More details are given in the *Guidance on the Application of the CLP criteria*, sections 1.7.2.1 and 2.8.6.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

For the determination of the self-reactive substances and mixtures, weight of evidence is not possible.

- (Q)SAR

At present (Q)SAR is generally not applicable for determination of self-reactive substances. Application of (Q)SAR is not possible.

---

[19] See UN RTDG, sub-sections 28.1, 28.2, 28.3 and Table 28.3.

- Grouping of substances and read-across approach

At present grouping and read-across are not applicable.

- Testing is technically not possible

A few of substances can, for safety reasons, only be handled and tested in diluted form, see the substances and mixtures listed in UN-RTDG, section 2.4.2.3.2.3.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

Minor impurities can have an influence on thermal stability. Background information and guidance on these aspects is given in the *Guidance on the Application of the CLP criteria*, section 2.8.4.3.

How to conclude on the DSD classification

In the DSD self-reactive substances are not covered. They may be classified in other DSD classes (e.g. explosive substance, flammable solid or liquid). See also the *Guidance on the Application of the CLP criteria*, section 2.8.6.1.

What information is required in the registration dossier in IUCLID

Material and methods

- see UN-MTC, Part II, classification procedures and test series A-H.

Results and discussion

The following data on self-reactive substances should be submitted:

- type of self-reactive substance;
- decomposition energy (value and method of determination);
- SADT (Self accelerating decomposition temperature) together with the volume the SADT relates to;
- detonation properties (Yes/Partial/No);
- deflagration properties (Yes rapidly/Yes slowly/No);
- effect of heating under confinement (Violent/Medium/Low/No);
- explosive power if applicable (Not low/Low/None).

For assigning the type of self-reactive substance, the list of currently assigned self-reactive substances according to the 2.4.2.3.2.3 of the UN-RTDG can be used, in cases where the assignment was based on test(s) according to the UN-MTC. The relevant underlying test data may be collected from the respective UN documents from the UN Committee of experts on the transport of dangerous goods, from test reports produced by competent authorities or industry, or from other reliable sources.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

The following example (Figure R.7.1–2)shows how the data mentioned above could be documented in the chemical safety report (CSR):

Figure R.7.1–2 Example 2,2'-Azodi (isobutyronitrile)

UN Test Series A to H	Test method	Results + Evaluation	Remarks
Propagation of detonation	A.5	"yes"	Apparent density (kg/m^3): 366 Fragmented length (cm): 40
Propagation of deflagration #1	C.1	"yes, slowly"	68 ms
Propagation of deflagration #2	C.2	"no"	
Effect of heating under defined confinement #1	Koenen E.1	"violent"	Limiting diameter 3.0 mm Type of fragmentation: F
Effect of heating under defined confinement #2	DPVT E.2	"medium"	Limiting diameter 5.5 mm
Explosive power	F.4	"not Low"	Average net expansion (cm^3): 18
SADT	H.4	50°C	500 ml Dewar vessel
Competent Authority approval number	*Example from UN Recommendations on the Transport of Dangerous Goods, Manual of Tests and Criteria*		

Reference to other ECHA Guidance Documents

A template data set does not currently exist in IUCLID for the hazard class 'self-reactive substances'. As long as there is no specific section available in IUCLID the test results in IUCLID section 4.23 'Additional physico-chemical information' under the endpoint title 'Self-reactive substances' should be inserted. In the CSR the information should be included under flammability.

Further information / references

Background information and guidance on classification testing, additional testing and available information is given in the *Guidance on the Application of the CLP criteria*, section 2.8.

R.7.1.10.5 Pyrophoric liquids

Definition

The definition of a pyrophoric liquid is given in the section 2.9.1 of Annex I to CLP Regulation:

'Pyrophoric liquid means a liquid substance or mixture which, even in small quantities, is liable to ignite within five minutes after coming into contact with air.'

Background information and guidance on the definition is given in the *Guidance on the Application of the CLP criteria*, sections 2.9.1 and 2.9.2.

Classification criteria and relevant information

Classification principles are given in CLP Regulation Annex I, section 2.9.2.

The criterion for a pyrophoric liquid is as follows: '*The liquid ignites within 5 min when added to an inert carrier and exposed to air, or it ignites or chars a filter paper on contact with air within 5 min.'*

Background information and guidance on relevant aspects regarding the classification is given in the *Guidance on the Application of the CLP criteria*, sections 2.9.1, 2.9.2, 2.9.3 and 2.9.4.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Other flammability tests do not have to be performed as well as the determination of the self-ignition temperature, if the substance is a pyrophoric substance. However, flammability in contact with water may be relevant.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

The UN Test N.3 of the UN-MTC is also used for classification according to the regulations on the transport of dangerous goods (ADR and RID). If the liquid in question has been classified as belonging to Class 4.2, packing group I of the ADR/RID on the basis of UN Test N.3 results, it is a pyrophoric liquid according to CLP criteria. Packing group I of the ADR/RID directly corresponds to Category 1 of the CLP.

According to the DSD, the A.13 method of Regulation (EC) 440/2008 is used for the assessment of pyrophoric properties for liquids and liquids. This method is identical to the UN Test N.3.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

Application of (Q)SAR is not possible, however assessment of the chemical structure may be used to exclude pyrophoric properties of a substance. Such an assessment of chemical structure, in conjunction with experience in manufacture and handling, could also formally form part of a weight-of-evidence argument.

- Grouping of substances and read-across approach

Assessment of the chemical structure may be used to anticipate pyrophoric properties of a substance.

- Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies. Due to pyrophoric properties a number of other tests on physicochemical, toxicological and eco-toxicological endpoints cannot be conducted.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

More background information and guidance on this and other aspects is given in the *Guidance on the Application of the CLP criteria*, section 2.9.

How to conclude on the DSD classification

Because the test methods of DSD and CLP are identical for this endpoint there is no difference in classification, see also the *Guidance on the Application of the CLP criteria*, section 2.9.6.

Endpoint specific information in the registration dossier IUCLID

Material and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied.

Note that in this case the experience in handling may be sufficient.

Results and discussion

- whether ignition occurs when poured or whether the filter paper is charred;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flammability can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.13	VII 7.10	Flammability	E.4.14	3.12

Further information / references

Background information and guidance on classification testing, additional testing and available information is given in the *Guidance on the Application of the CLP criteria*, section 2.9.

R.7.1.10.6 Pyrophoric solids

Definition

The definition of a pyrophoric solid is given in CLP Regulation Annex I, section 2.10.1.

'Pyrophoric solid means a solid substance or mixture which, even in small quantities, is liable to ignite within five minutes after coming into contact with air.'

Background information and guidance on the definition is given in the *Guidance on the Application of the CLP criteria*, sections 2.10.1 and 2.10.2.

Classification criteria and relevant information

Classification principles are given in CLP Regulation Annex I, section 2.10.2.

The criterion for a pyrophoric solid is as follows: '*The solid ignites within 5 minutes of coming into contact with air.'*

Background information and guidance on relevant aspects regarding the classification is given in the *Guidance on the Application of the CLP criteria*, sections 2.10.1, 2.10.2, 2.10.3 and 2.10.4.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Other flammability tests do not have to be performed in addition to the determination of the self-ignition temperature, if the substance is a pyrophoric substance. However, flammability in contact with water may be relevant.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

The UN Test N.2 of the UN-MTC is also used for classification according to the regulations on the transport of dangerous goods (ADR and RID). If the solid in question has been classified as belonging to Class 4.2, packing group I of the ADR/RID on the basis of UN Test N.2 results, it is a pyrophoric solid according to CLP Regulation criteria. Packing group I of the ADR/RID directly corresponds to Category 1 of CLP.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

According to the DSD, the A.13 method of Regulation (EC) 440/2008 is used for the assessment of pyrophoric properties for solids and liquids. This method is identical to the N.2 test method.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

Application of (Q)SAR is not possible, however assessment of the chemical structure may be used to exclude pyrophoric properties of a substance. Such an assessment of chemical structure, in conjunction with experience in manufacture and handling, could also formally form part of a weight-of-evidence argument.

- Grouping of substances and read-across approach

Assessment of the chemical structure may be used to anticipate pyrophoric properties of a substance.

- Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies. Due to pyrophoric properties a number of other tests on physicochemical, toxicological and ecotoxicological endpoints cannot be conducted.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

Particle size may play an important role. More background information and guidance on this and other aspects is given in the *Guidance on the Application of the CLP criteria*, section 2.10.

How to conclude on the DSD classification

Because the test methods of DSD and CLP Regulation are identical for this endpoint there is no difference in classification, see also the *Guidance on the Application of the CLP criteria*, section 2.10.6.

Endpoint specific information in the registration dosser/in IUCLID

Material and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied;
- particle size and distribution (if practicable);

Note that in this case experience in handling may be sufficient.

Results and discussion

- whether ignition occurs when poured;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flammability can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.13	VII 7.10	Flammability	E.4.14	3.12

Further information / references

Background information and guidance on classification testing, additional testing and available information is given in the *Guidance on the Application of the CLP criteria*, section 2.10.

R.7.1.10.7 Self-heating substances and mixtures

Definition

For solids and liquids adsorbed onto a large surface, self-heating may occur by reaction with air with subsequent ignition. According to the section 2.11.1.1 of Annex I to CLP Regulation:

'A self-heating substance or mixture is a liquid or solid substance or mixture, other than a pyrophoric liquid or solid, which, by reaction with air and without energy supply, is liable to self-heat; this substance or mixture differs from a pyrophoric liquid or solid in that it will ignite only when in large amounts (kilograms) and after long periods of time (hours or days).'

Classification criteria and relevant information

Self-heating substances and mixtures are classified in two categories according to the criteria of the CLP Regulation (see section 2.11, table 2.11.1). In general, self-heating occurs only for solids in contact with air. The *Guidance on the Application of the CLP criteria*, section 2.11 gives detailed background information about this phenomenon.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of the REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for self-ignition temperature.

'The study does not need to be conducted:

- *if the substance is explosive or ignites spontaneously with air at room temperature; or*
- *for liquids non flammable in air, e.g. no flash point up to 200°C, or*
- *for gases having no flammable range, or*
- *for solids, if the substance has a melting point < 160°C, or if preliminary results exclude self-heating of the substance up to 400°C.'*

The first indent specifies that no data is required for substances which is explosive or ignites spontaneously with air at room temperature.

Second and third indent are not applicable for this endpoint.

With regards to fourth indent, for the purposes of REACH, no data are required for solids classified as:

- pyrophoric; or

- explosive, unstable or division 1.1 to 1.6; or
- organic peroxide; or
- self-reactive substance.

Further, no data are required for substances with a melting point below 160°C. This means also that liquids do not have to be tested for this endpoint for the purposes of this regulation. Annex VII of REACH also allows waiving *'if preliminary results exclude self-heating of the substance up to 400°C'.* This refers to Test Method Regulation 440/2008, method A.16. However, the criteria are not very clear, and therefore it is recommended to instead refer to the CLP Regulation classification criteria, if applicable, and to waive otherwise.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

Literature data – even if available – should not be used since self-heating strongly depends on particle size, surface treatment and other parameters.

The use of existing data is possible provided that the test has been carried out by a qualified institution. If available data from a test according to method A.16 indicate that a classification as a self-heating substance does not apply, no more testing is necessary. However, the interpretation of the A.16 test method data in terms of the CLP criteria requires appropriate expert knowledge.

- Weight of evidence

For the determination of the self-heating substances and mixtures, weight of evidence is not possible.

- (Q)SAR

At present (Q)SAR is generally not applicable for self-heating substances and mixtures. Application of QSAR is not possible.

- Grouping of substances and read-across approach

At present grouping and read-across are not applicable.

- Testing is technically not possible

In some cases, exothermic decomposition may occur when performing the test, and special care will be necessary with respect to performing the tests and interpreting the results; see the *Guidance on the Application of the CLP criteria*, section 2.11.4.4.3. In such cases, it may not be possible to determine these properties.

Further adaptation possibilities

According to the UN-MTC, the classification procedure for self-heating substances or mixtures need not be applied if the results of a screening test can be adequately correlated with the classification test and an appropriate safety margin is applied. Examples of screening tests are:

a. the Grewer Oven test (VDI guideline, 1990) with an onset temperature 80 K above the reference temperature for a volume of 1 litre;

b. the Bulk Powder Screening Test (Gibson *et al.*, 1985) with an onset temperature 60 K above the reference temperature for a volume of 1 litre.

As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used. The *Guidance on the Application of the CLP criteria*, section 2.11.4.2 should be consulted for details about waiving and screening criteria.

Impurities; uncertainties

Particle size may play an important role. More background information and guidance on this and other aspects is given in the *Guidance on the Application of the CLP criteria*, section 2.11.

How to conclude on the DSD classification

This hazard class is not defined in DSD, therefore translation is not possible.

Endpoint specific information in the registration dosser/in IUCLID

Material and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied;
- indicate if preliminary and/or main test performed;
- moisture content;
- particle size and distribution (if available).

Results and discussion

- indicate temperature rise obtained for the individual tests and classification result.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flammability can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.13	VII 7.10	Flammability	E.4.14	3.12

Further information / references

ECHA guidance document *the* *Guidance on the Application of the CLP criteria* gives in section 2.11 detailed information on the self-heating property, the CLP-classification, the relevant test method and the relation to the DSD and the UN-RTDG.

VDI guideline 2263, part 1 (1990): ‘Test methods for the Determination of the Safety Characteristics of Dusts’.

Gibson, N. Harper, D.J. Rogers (1985): ‘Evaluation of the fire and explosion risks in drying powders’, Plant Operations Progress, 4 (3), 181-189.

R.7.1.10.8 Substances which in contact with water emit flammable gases

Definition

The CLP Regulation, Annex I, section 2.12.1 provides the following definition:

'Substances or mixtures which, in contact with water, emit flammable gases means solid or liquid substances or mixtures which, by interaction with water, are liable to become spontaneously flammable or to give off flammable gases in dangerous quantities.'

Classification criteria and relevant information

Classification according to the CLP Regulation is required if the gas produced upon contact with water ignites spontaneously and/or if the reaction rate with which the flammable gas is produced is ≥ 1 l/kgh.

If the gas produced ignites spontaneously, this does not necessarily imply that the gas produced is pyrophoric but this generally is the case if the heat of reaction is sufficient to result in ignition of the gas.

The test method for classification of substances and mixtures which in contact with water emit flammable gases is described in the UN-MTC (UN Test N.5, see Section 33.4). This method is referred to in Annex I, Part 2 of the CLP Regulation and it is strongly recommended to use this method and not to apply test method A.12 of the Test Methods Regulation if new testing is carried out. UN Test N.5 foresees dividing into three categories depending on the violence and rate of the reaction whereas test method A.12 does not allow any further dividing of the substances. Furthermore, the results of both methods might differ slightly due to some differences in the testing procedure (for these differences see the *Guidance on the Application of the CLP criteria*, Section 2.12.6). Therefore unnecessary testing can be avoided by applying only UN Test N.5 because it leads to more detailed information (and has in any case to be applied for other purposes such as classification and transport).

Data which is based on the classification according to DSD may be available. There are, however, differences between the methods UN Test N.5 and A.12 which should be considered. They are described in detail in the *Guidance on the Application of the CLP criteria*, section 2.12.6.

Detailed guidance on the test method itself can be found in the *Guidance on the Application of the CLP criteria*, section 2.12.4.4.1.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for flammability.

'The study does not need to be conducted:

- *if the substance is a solid which possesses explosive or pyrophoric properties. These properties should always be considered before considering flammability; or*
- *for gases, if the concentration of the flammable gas in a mixture with inert gases is so low that, when mixed with air, the concentration is all time below the lower limit; or*
- *for substances which spontaneously ignite when in contact with air.'*

The first point is valid with regard to explosive substances because they are not classified as substances which in contact with water emit flammable gases. In that case testing can be waived.

The other waiving possibilities are not applicable with regard to substances which in contact with water emit flammable gases.

The first point is not correct with regard to pyrophoric substances because pyrophoric substances can be classified as substances which in contact with water emit flammable gases based on UN Test N.5 which is referred to by CLP. UN Test N.5 explicitly requires testing of pyrophoric substances under nitrogen (see UN-MTC, section 33.4.1.3.1).

The second point is not applicable because gases do not fall under the hazard class of substances which in contact with water emit flammable gases.

For the same reasons, the last point (waiving would be possible for substances which spontaneously ignite when in contact with air) is also not valid in this case.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

There are currently no QSPR models for predicting whether a substance in contact with water emits flammable gases and if so what the gas evolution rate is.

- Grouping of substances and read-across approach

At present grouping and read-across are not applicable.

- Testing is technically not possible

Testing should always be possible if none of the waiving possibilities applies. If the substance is known to be soluble in water to form a stable solution, or if it is clearly known that it does not react with water, e.g. because it is manufactured or washed with water, testing is not necessary.

Further adaptation possibilities

Classification in certain hazard classes do not foresee the assignment of further physical hazard classes or at least normally do not match with classification in this hazard class:

Substances that are classified as explosives, self-reactives or organic peroxides are not classified in this hazard class (or any other physical hazard class). For explosives this is considered through the first point of the adaptation possibilities according to REACH Annex VII, column 2 (see above).

Oxidizing substances are generally not considered for flammability and therefore are also not classified in this hazard class (there may be some exceptions, however).

As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

Impurities; uncertainties

The descriptions of the methods UN Test N.5 and A.12 are not very detailed and therefore allow for technical variations such as with regard to the apparatus used or the procedure. In particular, the testing protocol does not prescribe a specific method for measuring the gas evolution rate. An interlaboratory comparison for this test method has shown that laboratories - based on the freedom the description of the test methods gives - apply different approaches when performing this test. Furthermore, the interlaboratory comparison showed that the test results vary in a rather wide range. It therefore has to be kept in mind that this test method has a non-negligible uncertainty with regard to trueness and precision. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

Sea water may be a particular case of interest (in case of maritime transport).

How to conclude on the DSD classification

Substances which in contact with water emit flammable gases would be classified as ‘F; R15’ under DSD (the sum of categories 1 to 3 corresponds to ‘F; R15’).

Endpoint specific information in the registration dosser/in IUCLID

Material and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied;
- partice size and distribution.

Results and discussion

- indicate whether full test was performed or whether it was terminated at a particular step/stage;
- substance identity of evolved gas;
- indicate whether the gas evolved ignites spontaneously;
- rate of gas evolution (unless the test has been terminated);
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on flammability is found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.13	VII 7.10	Flammability	E.4.14	3.12

Further information / references

The ECHA document *Guidance on the Application of the CLP criteria* gives in its section 2.12 detailed information on substances and mixtures which, in contact with water, emit flammable gases, their CLP-classification, the relevant test method and the relation to the DSD and the transport of dangerous goods regulations.

Janès *et al.*, 'Towards the improvement of UN N.5 test method intended to the characterization of substances which in contact with water emit Flammable Gases', submitted in revised form to the Journal of Loss Prevention in the Process Industries.

Interlaboratory test on the method UN Test N.5 / EC A.12 'Substances which, in contact with water, emit flammable gases' 2007, Kunath, K., Lüth, P., Uhlig, S., ISBN 978-3-9814634-1-5, http://www.bam.de/de/service/publikationen/publikationen_medien/short__report_rv_un_n_5.pdf.

R.7.1.10.9 Organic peroxides

In the Dangerous Substances Directive (DSD) (67/548/EEC) organic peroxides were classified on the basis of their chemical structure either as explosive or as oxidising. In general, organic peroxides have only weak oxidising properties or do not show oxidizing properties at all. In the CLP Regulation organic peroxides are a distinct hazard class. Organic peroxides are classified in one of the seven categories of 'Types A to G' according to the classification criteria given in Section 2.15.2 of Annex I, of CLP.

As mentioned below under sub-section Definition, organic peroxides are excluded from testing as explosives according to Test Series 1 to 8 in Part I of the UN-MTC (see R.7.1.11.1 Explosives). In Test Series A to H however, no tests on sensitivity to impact (solids and liquids) and friction (solids only) are included. For the risk assessment and the safe use and handling, data according to the EU test method A.14 as described in Regulation (EC) No 440/2008, if available, or UN Test 3 (a) (ii) BAM Fallhammer and Test 3 (b) (i) BAM friction apparatus (see R.7.1.11) should be part of the hazard communication in the registration dossier (REACH Annex VII, 7.11) and in the safety data sheet.

Definition

The definition of an organic peroxide is given in CLP Annex I, section 2.15.1:

'*Organic peroxides means liquid or solid organic substances which contain the bivalent -O-O- structure and may be considered derivatives of hydrogen peroxide, where one or both of the hydrogen atoms have been replaced by organic radicals. The term organic peroxide includes organic peroxide mixtures (formulations) containing at least one organic peroxide. Organic peroxides are thermally unstable substances or mixtures, which can undergo exothermic self-accelerating decomposition. In addition, they can have one or more of the following properties:*

(i) be liable to explosive decomposition;

(ii) burn rapidly;

(iii) be sensitive to impact or friction;

(iv) react dangerously with other substances.

An organic peroxide is regarded as possessing explosive properties when in laboratory testing the mixture (formulation) is liable to detonate, to deflagrate rapidly or to show a violent effect when heated under confinement.'

Background information and guidance on the definition is given in *Guidance on the Application of the CLP criteria*, sections 2.15.1 and 2.15.2.

Classification criteria and relevant information

The Classification principles are given in CLP Annex I, sections 2.15.2 and 2.15.4. Background information and guidance on relevant aspects regarding the classification is given in *Guidance on the Application of the CLP criteria*, sections 2.15.3, 2.15.4, 2.15.5, 2.15.6 and 2.15.7.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Only organic peroxides, as defined in CLP, Annex I, section 2.15.1 definition, have to be tested according to the UN-MTC, Part II test series A - H.

CLP Annex I, section 2.15.2.1. provides the following specific rules for adaptation of the standard information requirement for organic peroxides.

'Any organic peroxide shall be considered for classification in this class, unless it contains:

(a) not more than 1.0% available oxygen from the organic peroxides when containing not more than 1.0% hydrogen peroxide; or

(b) *not more than 0.5% available oxygen from the organic peroxides when containing more than 1.0% but not more than 7.0% hydrogen peroxide.*

NOTE: The available oxygen content (%) of an organic peroxide mixture is given by the formula:

$$16 \times \sum_{i}^{n} \left(\frac{n_i \times c_i}{m_i} \right)$$

where:

n_i = number of peroxygen groups per molecule of organic peroxide i;

c_i = concentration (mass %) of organic peroxide i;

m_i = molecular mass of organic peroxide i.'

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

A number of already tested and classified substances and mixtures are listed in the UN-RTDG, 2.5.3.2.4.

Available information may especially originate from the classification for transport. In the DSD organic peroxides were classified as oxidizing substances, by definition. More details are described in the *Guidance on the Application of the CLP criteria*, sections 1.7.2.1 and 2.15.6. If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

For the determination of the organic peroxides, weight of evidence is not possible. Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

At present QSAR is generally not applicable for organic peroxides. Application of (Q)SAR is not possible.

- Grouping of substances and read-across approach

At present grouping and read across are not applicable.

- Testing is technically not possible

A number of substances can, for safety reasons, only be handled and tested in diluted form, see the substances and mixtures listed in UN TDG, 2.5.3.2.4. Testing should always be considered if none of the waiving possibilities applies.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

Minor impurities can have an influence on thermal stability. Background information and guidance on these aspects is given in *Guidance on the Application of the CLP criteria*, section 2.15.4.

How to conclude on the DSD classification

In the DSD organic peroxides are classified as oxidizing substances and a few of them as having explosive properties.

Endpoint specific information in the registration dossier / in IUCLID

Material and methods

- See UN MTC, Part II, classification procedures and test series A-H.

Results and discussion

The following data on organic peroxides should be submitted:

- if testing is waived, the reasons for waiving must be documented in the dossier;
- type of organic peroxide;
- SADT (Self accelerating decomposition temperature) together with the volume the SADT related to;
- detonation properties (Yes/Partial/No);
- deflagration properties (Yes rapidly/Yes slowly/No);
- effect of heating under confinement (Violent/Medium/Low/No);
- explosive power, if applicable (Not low/Low/None).

The following example (Figure R.7.1–3) shows how data mentioned above could be documented in the CSR:

Figure R.7.1–3 Example: Di-tert-butyl peroxide

UN Test Series A to H	Test method	Results + Evaluation	Remarks
Propagation of detonation	A.1	"No"	Fragmented length (cm): 16
Propagation of deflagration #1	C.1	"Yes, slowly "	Maximum pressure (kPa): > 2070 Time for a pressure rise from 690 to 2070 kPa (ms): 100
Propagation of deflagration #2	C.2	"No"	deflagration rate (mm/s): 0.27
Effect of heating under defined confinement #1	Koenen E.1	"No"	Limiting diameter (mm): < 1.0 Type of fragmentation (and pieces): O
Effect of heating under defined confinement #2	DPVT E.2	"Medium"	Limiting diameter (mm): 3.5
Explosive power	F.3	"Not Low"	Expansion (cm^3/10 g test sample): 28
Explosive power	F.4	"Not Low"	Average net expansion (cm^3): 12
SADT	H.4	80°C	500 ml Dewar vessel
Competent Authority approval number	*Example from UN Recommendations on the Transport of Dangerous Goods, Manual of Tests and Criteria*		

For assigning the Type of organic peroxide, the list of currently assigned organic peroxides according to section 2.5.3.2.4 of the UN RTDG can be used, in case the assignment was based on a test according to the UN MTC. The relevant underlying test data may be collected from the respective UN documents from the UN Committee of experts on the transport of dangerous goods, from test reports produced by either competent authorities or industry, or from other reliable sources (such as e.g. the dedicated database 'DATATOP').

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

A Template data set in IUCLID does not exist for the hazard class 'organic peroxides'. As long as there is no specific section in IUCLID the test results in section 4.23 'Additional physico-chemical information' should be inserted under the endpoint title 'organic peroxides'. The information on organic peroxides should not be included in IUCLID section 4.15 'Oxidising properties'. In the registration dossier the information should be included under flammability.

Further information / references

Background information and guidance on classification testing, additional testing and available information is given in *Guidance on the Application of the CLP criteria*, section 2.15.

Data from the 'DATATOP' database can be obtained from the gatekeeper of this database TNO, Department Energetic Materials, Lange Kleiweg 137, 2288GJ, Rijswijk The Netherlands.

Various national guidelines which provide guidance and outline safe standards for handling and storage of organic peroxides for the assignment of organic peroxides to storage groups are available e.g. Netherlands Directive: Publication Series on Dangerous Substances 8 (PGS 8) Storage of Organic Peroxides, UK HSE: The storage and handling of organic peroxides - Guidance Note CS21 or German guideline: BGV B4.

R.7.1.11 Explosive properties

Some of the information requirements according to the REACH Regulation, Annex VII were phrased such that they correspond to 'indications of danger' as given in Annex II of DSD. For substances, classification and labelling according to the CLP Regulation has been mandatory since 1 December 2010 (and will become mandatory for mixtures (preparations) from 1 June 2015, when the DSD and DPD will be repealed). Consequently, explosive properties are covered by classification of the substance according to the CLP Regulation. However, the physical hazards according to CLP are structured completely differently from the physico-chemical properties according to the DSD (and therefore also REACH, Annex VII). This means that for some of the CLP hazard classes an unambiguous assignment to one of the headlines (information requirements) in Annex VII to REACH is not possible. The assignment of hazard classes to the headline 'Explosive properties' as shown in Table below (

Table R.7.1–8) must therefore only be understood as a means to structure this document in accordance with Annex VII to REACH. It has to be noted that self-reactive substances and organic peroxides are primariliy assigned to the headline 'Flammability' and only a cross reference to corresponding sub-chapter under headling 'Flammability' is included in the sub-chapters on 'Explosive properties' below because these two hazard classses can have explosive and/or flammable properties.

Table R.7.1–8 Assignment of CLP hazard classes to the information requirement 'Explosive properties' according to REACH, Annex VII and correlation between the Test method Regulation and the test method according to CLP and supporting link with the *Guidance on the Application of the CLP criteria*.

Information requirement according to Art. 10 (a) (vi) of the REACH Regulation (EC) No. 1907/2006 (the no. in brackets is the respective no. in the table in Annexes VII to IX to REACH)	**CLP Regulation (EC) No. 1272/2008** (the no. in brackets is the respective chapter no. in Annex I to CLP)	**Chapter in revised R.7(a) guidance**	**Corresponding test method according to the Test Method Regulation, Regulation (EC) No. 440/2008**	**Corresponding test method according to the CLP Regulation**	**Chapter in the Guidance on the Application of the CLP Criteria (ex RIP 3.6)**
Explosive properties (7.11)	Explosives (2.1)*	R.7.1.11.1	A.14 Explosive properties	**UN Test series 1 to 3 (further test series 4 to 6 are necessary for classification)**	2.1
	Self-reactive substances and mixtures (2.8)*	R.7.1.11.2 See R.7.1.10.4	n.a.	**A.14 (existing data only)**	2.8
	Organic peroxides (2.15)*	R.7.1.11.3 See R.7.1.10.9	n.a.	**A.14 (existing data only)**	2.15

* Note that regardless of whether the hazard class or category is listed in Article 14(4)(a) REACH the chemical safety assessment must be performed in accordance with Article 14 (3) of REACH. Furthermore, according to Article 10(a)(iv) of REACH the technical dossier of a registration of a substance under the REACH Regulation must include information on classification and labelling of the substance as specified in section 4 of Annex VI to the REACH Regulation.

In addition, it has to be noted that some substances have explosive properties which do not result in classification. Examples are the following:

- substances with a positive result in UN Test Series 1 or 2 but which are exempted from the classification as explosives based on their packaging in UN Test Series 6;
- substances which are mechanically sensitive only. These are substances with a sensitiveness to impact (determined by UN Test Series 3 (a) (ii)) of 40 J or less and/or a sensitiveness to friction (determined by Test Series 3 (b) (i)) of 360 N or less for substances and mixtures which may have explosive properties based on the screening procedure according to Appendix 6, Part 3 of the UN-MTC and which are not classified as explosives, self-reactive or organic peroxide.

Such substances may be classified in other hazard classes (e.g. as flammable solids, oxidizing solids, corrosive to metals) or even not at all. Information about such explosive properties should be indicated in the dossier as well.

R.7.1.11.1 Explosives

Please note that explosive atmospheres as, for example, created by flammable liquids and by powders are not the subject of this chapter.

Definition

The following definitions are provided in CLP Annex I, section 2.1.1:

'An explosive substance or mixture is a solid or liquid substance or mixture of substances which is in itself capable by chemical reaction of producing gas at such a temperature and pressure and at such a speed as to cause damage to the surroundings. Pyrotechnic substances are included even when they do not evolve gases.

A pyrotechnic substance or mixture is a substance or mixture of substances designed to produce an effect by heat, light, sound, gas or smoke or a combination of these as the result of non-detonative self-sustaining exothermic chemical reactions.

An unstable explosive is an explosive substance or mixture which is thermally unstable and/or too sensitive for normal handling, transport and use.

An explosive article is an article containing one or more explosive substances or mixtures.

A pyrotechnic article is an article containing one or more pyrotechnic substances or mixtures.

An intentional explosive is a substance, mixture or article which is manufactured with a view to producing a practical, explosive or pyrotechnic effect.'

Organic Peroxides and Self Reactive Substances may also have explosive properties and should be screened. See chapter R.7.1.11.3 for Organic peroxides and chapter R.7.1.11.2 for Self Reactive Substances and Mixtures.

Intentional explosive

Council Directive 93/15/EEC of 5 April 1993 lays down rules for the harmonisation of the provisions relating to the placing on the market and supervision of explosives for civil uses.

Directive 2007/23/ EC on the placing on the market of pyrotechnic articles establishes rules designed to achieve the free movement of pyrotechnic articles in the internal market while, at the same time, ensuring a high level of protection of human health and public security and the protection and safety of consumers and taking into account the relevant aspects related to environmental protection. Pyrotechnic articles (CLP, Annex I, Section 2.1.1.2) are classified as explosives for CLP and as class 1 for transport (see UN-RTDG). Accoding to Article 9 and Annex II of Directive 2007/23/EC the conformity assessment procedures are carried out by notified bodies, which have to issue an EC type-examination certificate to the applicant. All data included in the EC type-examination certificate are sufficient for the information requirements under the REACH Regulation.

Classification criteria and relevant information

Substances, mixtures and articles of this class are classified as an unstable explosive on the basis of the flowchart in Annex I to CLP Regulation, Figure 2.1.2. The test methods are described in Part I of the UN-MTC.

Explosives, which are not classified as an unstable explosive, must be classified in one of the six Divisions referred to in paragraph 2.1.2.2 of Annex 2.1 to the CLP Regulation, based on the results of the tests laid down in Table 2.1.1 on Test Series 2 to 8 in Part I of the UN-MTC. If explosives are unpackaged or repacked in packaging other than the original or similar packaging, they must be retested. If a substance gives a positive result in any of the test series 1 or 2 this should be mentioned in the REACH registration dossier for the substance, even if it would not be classified as an 'Explosive' in Test Series 6.

The test methods used for deciding on provisional acceptance into the class of explosives are grouped into four series, numbered 1 to 4 (see CLP Annex I, Figure 2.1.2).

It may be important for the safety of testers that certain tests, using small amounts of material, be conducted first before proceeding to test with larger quantities. Therefore it is highly recommended to start the testing procedure with Test Series 3, because these tests involve relatively small sample sizes, which reduces the risk to personnel.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for explosive properties.

'The study does not need to be conducted if:

- *there are no chemical groups associated with explosive properties present in the molecule, or*
- *the substance contains chemical groups associated with explosive properties which include oxygen and the calculated oxygen balance is less than -200, or*
- *the organic substance or a homogenous mixture of organic substances contains chemical groups associated with explosive properties, but the exothermic decomposition energy is less than 500 J/g and the onset of exothermic decomposition is below 500°C, or*
- *for mixtures of inorganic oxidising substances (UN Division 5.1) with organic materials, the concentration of the inorganic oxidising substance is:*
    - *less than 15%, by mass, if assigned to UN Packaging Group I (high hazard) or II (medium hazard),*
    - *less than 30%, by mass, if assigned to UN Packaging Group III (low hazard).*

Note: Neither a test for propagation of detonation nor a test for sensitivity to detonative shock is required if the exothermic decomposition energy of organic materials is less than 800 J/g.'

Note on the use of the Oxygen Balance:

The oxygen balance is calculated for the chemical reaction:

$$C_xH_yO_z + [x + (y/4) - (z/2)]\ O_2 \rightarrow x\ CO_2 + (y/2)\ H_2O$$

Using the formula:

$$\text{Oxygen balance} = -1600\ [2x + (y/2) - z]/\text{molecular weight};$$

The oxygen balance was developed for compounds containing only nitrate groups and it applies only to organic substances. Extending its use to molecules with other oxygen containing groups should be done with care. As an example the presence of hydroxyl-groups will strongly affect the oxygen balance towards higher values, whereas this group does not contribute to explosive properties. In addition the presence of for instance halogens tends to decrease the flammability and explosivity but this is not accounted for.

Please also check Appendix 6, Section 3 of the UN-MTC.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met. Application of weight of evidence is possible with substances where explosive properties can clearly be excluded. Weight of evidence should be accompanied with extensive and reliable literature references.

- (Q)SAR

There is currently no QSPR/(Q)SAR software known with sufficient accuracy and reliability to assist in assessing (potential) explosive properties. DSC testing is cheap and fast and is strongly recommended to identify potential hazards connected with the substance.

- Grouping of substances and read-across approach

An assessment of chemical structure would formally form part of a column 2 waiver. For further information please refer to the *Guidance on the Application of the CLP criteria*, Part 2: Physical Hazards, Section 2.1 Explosives.

- Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies. Testing for explosives may be omitted if it is technically not possible to conduct the study as a consequence of the properties of the substance.

Further adaptation possibilities

Testing may be waived if there are no chemical groups associated with explosive properties present in the molecule. The potential generation of explosive atmospheres by flammable gases/liquids or combustible solids is not considered an explosive property and should therefore not be reported under this heading.

As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

Impurities; uncertainties

Small amounts of other compounds may enhance or suppress the chemical reaction that gives the explosive property to a substance. Therefore impurities may considerably influence the explosive properties of a substance. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

How to conclude on the DSD classification

For DSD explosives are substances and preparations which may explode under the effect of flame or which are more sensitive to shocks or friction than dinitrobenzene.

Reclassification of substances classified as explosive according to DSD:

Under the regime of the old DSD, testing of explosive properties was achieved by performing test method A.14. For classification purposes under the CLP Regulation this test is not adequate in the case of a negative result for thermal sensitivity. The test method A.14 stops with a limiting diameter of 2 mm, while UN Test E.1 proceeds to down to a 1 mm orifice. Testing according to the CLP Regulation is the same as that described in Part I of the UN-MTC. This is why the translation table of Annex VII of the CLP Regulation states that there is no direct translation possible for classification from (E, R2) and (E, R3) to CLP criteria.

Therefore, if the screening procedure of section 2.1.4.2 of the CLP Regulation identifies a substance or mixture to be a potential explosive, appropriate data are required for classification.

Moreover, if data from performing test method A.14 or the UN Test series 3 tests 3a or 3b indicate that a substance is sensitive to impact or friction such information should be provided in the REACH registration dossier.

Endpoint specific information in the registration dosser/in IUCLID

Material and methods

- reference to the standard and the test method applied;
- description of the substance that was tested.

Results and discussion

- if testing is waived, the reasons for waiving must be documented in the dossier;
- if testing is not waived then the tests done according to the UN Test Manual and the outcome (explosive or not explosive) must be documented in the dossier. The mechanical sensitivity test according to UN Test Series 3a and 3b must be done and documented if UN Test Series 1 or 2 give a positive result. If data according to test method A.14 are available, then the results can be used instead of UN Test series 3a and 3b.

An example is given below (Figure R.7.1–4) of how summarised results from the application of the class 1 procedure for the hypothetical substance 'New explosive substance' could be presented.

Figure R.7.1–4 Results from application of the class 1 acceptance procedure

1. Name of substance	New explosive substance
2. General data	2.1 Composition : technically pure 2.2 Physical form : Fine crystalline powder 2.3 Colour : Yellow
3. Box 2	Is the substance manufactured with the view to producing a practical explosive or pyrotechnic effect? 3.1 Answer : **No**
4. Box 3	4.1 Propagation of Detonation : UN-Test A.1 Result : "-", no propagation of detonation 4.2 Effect of heating under confinement: 4.2.1 Koenen test (test 1(b)) Result : "+", 4.2.2 Time/pressure test (test 1(c)(i)) Result : "-", no effect on ignition under confinement 4.5 Exit : Go to Box 4
5. Box 4	Is it an explosive substance? 5.1 Answer from Test Series 1 : **Yes** 5.2 Exit : Go to box 5
6. Box 5	6.1 Sensitivity to shock : based on the test result of UN-Test A.1 Result "-" 6.2 Effect of heating under confinement: Koenen test (test 2(b)): limiting diameter 2,5 mm Result: "+" 6.3 Exit : Go to Box 6
7. Box 6 :	Is the substance too insensitive for acceptance into Class 1? 7.1 Answer from Test Series 2 : **No** 7.2 Conclusion : Substance to be considered for Class 1 (box 8) 7.3 Exit : Go to Box 9
8. Box 9	Test Series 3 8.1 Thermal Stability: based on the DSC measurement data Result: thermally stable 8.2 Impact sensitivity : BAM fallhammer test (test 3(a)(ii)) Result : "-", not too dangerous to transport in form tested 8.3 Friction sensitivity : BAM friction test (test 3(b)(i))

	Result : "-", not too dangerous to transport in form tested 8.4 Exit : Go to box 10
9. Box 10	Is the substance thermally stable? 9.1 Answer from DSC data : **Yes** 9.2 Exit : Go to box 11
10. Box 11	Is the substance too dangerous for transport in the form in which it was tested? 10.1 Answer from Test Series 3 (a)(ii) and 3 (b)(i): **No** 10.2 Exit : Go to box 18
11. Conclusion	**PROVISIONALLY ACCEPT INTO CLASS 1** 11.1 Exit : Apply the Class 1 assignment procedure

Figure R.7.1–5 Results from the application of the class 1 assignment procedure

1. Box 19	Is the substance a candidate for Division 1.5? 1.1 Answer : **No** 1.2 Exit : Go to box 25
2. Box 25	2.1 UN-Tests 6(a) and 6(c) were not conducted because the substance showed no propagation of detonation in the UN-Test A.1 and also no propagation of deflagration in the UN-test 1(c)(ii). 2.2 UN-Test 6 (c) Sample conditions: 1 × 30 kg fibre drum Observations: Only slow burning with black smoke and soot occurred. 2.3 Exit : Go to box 26
3. Box 26	Is the result a mass explosion? 3.1 Answer from Test Series 6 : **No** 3.2 Exit : Go to box 28
4. Box 28	Is the major hazard that from dangerous projections? 4.1 Answer from Test Series 6 : **No** 4.2 Exit : Go to box 30
5. Box 30	Is the major hazard radiant heat and/or violent burning but with no dangerous blast or projection hazard? 5.1 Answer from Test Series 6 : **No** 5.2 Exit : Go to box 32
6. Box 32	Is there nevertheless a small hazard in the event of ignition or initiation? 6.1 Answer from Test Series 6 : **No** 6.2 Exit : Go to box 35
7. Box 35	Is the substance or article manufactured with the view to producing a practical explosive or pyrotechnic effect? 7.1 Answer : **No**

	7.2 Exit : Go to box 38
8. Conclusion	**NOT CLASS 1** 8.1 Exit : Consider for another class/division

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on explosiveness can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.14	VII 7.11	Explosiveness	E.4.15	3.13

Further information / references

Further information about classification and testing for explosives can be found in the *Guidance on the Application of the CLP criteria*, section 2.1.

Gharagheizi F. Quantitative structure-property relationship for prediction of the lower flammability limit of pure compounds. Energy & Fuels 22 (2008) 3037-3039.

Gharagheizi F. A new group contribution-based model for estimation of lower flammability limit of pure compounds. J. Haz. Mat. 170 (2009a) 595-604.

R.7.1.11.2 Self-reactive substances and mixtures

Self-reactive substances are primariliy assigned to the headline ‘Flammability’ therefore please also refer to chapter R.7.1.10.4.

The sensitivity of self-reactive substances to impact (solids and liquids) and friction (solids only) may be of importance for the safe handling of the substances, in the event that these substances have pronounced explosive properties. If data according to EU test method A.14 as described in Regulation (EC) No 440/ 2008 are available, then this information should be part of the hazard communication in the registration dossier (REACH Annex VII, 7.11).

R.7.1.11.3 Organic peroxides

Organic peroxides are primariliy assigned to the headline ‘Flammability’ therefore please also refer to chapter R.7.1.10.9.

The sensitivity of organic peroxides to impact (solids and liquids) and friction (solids only) may be of importance for the safe handling of the substances, in the event that these substances have pronounced explosive properties. If data according to EU test method A.14 as described in Regulation (EC) No 440/ 2008 are available, then this information should be part of the hazard communication in the registration dossier (REACH Annex VII, 7.11).

R.7.1.12 Self-ignition temperature

The terminology used in Annex VII of REACH is not very precise. Therefore, some guidance in interpretation appears necessary:

For liquids and gases, the term '**auto-ignition**' instead of 'self-ignition' is generally used. Auto-ignitability is of high importance for the assignment of temperature classes in explosion protection (i. e. ATEX in Europe) of plants and equipment.

For solids and liquids adsorbed on a large surface, **self-heating** may occur by reaction with air with subsequent ignition. According to the CLP Regulation, Annex I, section 2.11, a self-heating substance or mixture is a liquid or solid substance or mixture, other than a pyrophoric liquid or solid, which, by reaction with air and without energy supply, is liable to self-heat; this substance or mixture differs from a pyrophoric liquid or solid in that it will ignite only when in large amounts (kilograms) and after long periods of time (hours or days). Therefore solids are considered under self heating substances in the chapter below.

Table R.7.1–9 Assignment of CLP hazard classes to the information requirement 'Self ignition temperature' according to REACH, Annex VII and the Test Method Regulation.

Information requirement according to Art. 10 (a) (vi) of the REACH Regulation (EC) No. 1907/2006 (the no. in brackets is the respective no. in the table in Annexes VII to IX to REACH)	**CLP Regulation (EC) No. 1272/2008** (the no. in brackets is the respective chapter no. in Annex I to CLP)	**Chapter in revised R.7(a) guidance**	**Corresponding test method according to The Test Method Regulation (EC) No. 440/2008**	**Corresponding test method according to CLP Regulation**	**Chapter in the *Guidance on the Application of the CLP Criteria*** (ex RIP 3.6)
Self ignition temperature (7.12)	For gases and liquids*	R.7.1.12.1	A.15 Auto-ignition temperature (liquids and gases)	n.a.	n.a.
	For solids * Note: the UN Test N.4 is preferable to generate the information for this endpoint. Refer to R.7.1.10.7	R.7.1.12.2, R.7.1.10.7	A.16 Relative self-ignition temperature for solids	n.a.	Section 2.11

* Note that regardless of whether the hazard class or category is listed in Article 14 (4) (a) of REACH, the chemical safety assessment (when required) must be performed in accordance with Article 14 (3) of REACH. Furthermore, according to Article 10 (a) (iv) of REACH the technical dossier of a registration for a substance under the REACH Regulation must include information on classification and labelling of the substance as specified in section 4 of Annex VI to the REACH Regulation.

R.7.1.12.1 Auto-ignition

Type of property

For liquids and gases, the term '**auto-ignition**' instead of 'self-ignition' is generally used. Auto-ignitability is of high importance for the assignment of temperature classes in explosion protection (i. e. ATEX in Europe) of plants and equipment. In this chapter, only the **auto-ignition** phenomena will be discussed.

Definition

The degree of auto-ignitability is expressed in terms of the auto-ignition temperature. The auto-ignition temperature is the lowest temperature at which the test substance will ignite when mixed with air under the conditions defined in the test method.

Test method(s)

For testing Auto-ignition temperature, method A.15 of Regulation (EC) 440/2008 should be used, which references several national and international standards (e.g. EN 14522, etc.). The test procedure is applicable to gases, liquids and vapours which, in the presence of air, can be ignited by a hot surface.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for self-ignition temperature.

'The study does not need to be conducted:

- *if the substance is explosive or ignites spontaneously with air at room temperature; or*
- *for liquids non flammable in air, e.g. no flash point up to 200°C; or*
- *for gases having no flammable range, or*
- *for solids,if the substance has a melting point ≤ 160°C, or if preliminary results exclude self-heating of the substance up to 400°C.'*

This means:

For gases:

Only gases classified as flammable according to the CLP Regulation have to be considered.

For liquids:

The auto-ignition temperature should be determined according to Directive EC 440/2008, method A.15. No data are required for liquids classified as:

- pyrophoric; or
- explosive, unstable or division 1.1 to 1.6; or
- organic peroxide; or
- self-reactive substance.

Further, the auto-ignition temperature does not have to be determined for liquids having no flash point up to 200°C. In practice, liquids with a boiling point above 350°C will not have a flash point below 200°C. Therefore, determination of the auto-ignition temperature is not necessary in such cases if the flash point is not known.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

For the determination of the auto-ignition temperature, the weight of evidence approach is not possible. Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

For the determination of the auto-ignition temperature, (Q)SAR approaches are strongly discouraged for the purpose of classification/ risk assessment.

- Grouping of substances and read-across approach

For the determination of the auto-ignition temperature read-across is usually not possible. However interpolation may still be possible within homologous series.

However, it is not possible to read across from methyl compounds to ethyl and propyl compounds and vice versa.

- Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies. Substances which decompose below room temperature or which react vigorously with moisture may be difficult to test. In such cases, the test may be waived due to technical reasons.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

The auto-ignition temperature can be considerably reduced by the presence of catalytic impurities. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

Endpoint specific information in the registration dossier / in IUCLID

Material and methods

- description of the apparatus or reference to the standard or the test method applied;
- quantity of sample used.

Results and discussion

- the value or the range of the auto-ignition temperature;
- if testing is waived, the reasons for waiving must be documented in the dossier.

For liquids/gases: observations (e.g decomposition with air, reactions with moisture, etc.)

For solids see the below chapter R.7.1.12.2.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on auto flammability (self-ignition temperature) can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.12	VII 7.12	Auto flammability	E.4.13	3.11

R.7.1.12.2 Self-heating substances

For solids and liquids adsorbed on a large surface, **self-heating** may occur by reaction with air with subsequent ignition. According to the CLP Regulation Annex I, section 2.11 the following definition is provided:

'A self-heating substance or mixture is a liquid or solid substance or mixture, other than a pyrophoric liquid or solid, which, by reaction with air and without energy supply, is liable to self-heat; this substance or mixture differs from a pyrophoric liquid or solid in that it will ignite only when in large amounts (kilograms) and after long periods of time (hours or days).'

The ECHA *Guidance on the Application of the CLP criteria* gives in Chapter 2.11 detailed information on the self-heating property, the CLP-classification, the relevant test method and the relation to the DSD and the transport of dangerous goods regulations.

See Section R.7.1.10.7 of this guidance document for further details and information.

R.7.1.13 Oxidising properties

Some of the information requirements according to REACH Annex VII were phrased such that they correspond to 'indications of danger' as given in Annex II of DSD. For substances, classification and labelling according to the CLP Regulation has been mandatory since 1 December 2010 (and will become mandatory for mixtures (preparations) from 1 June 2015, when the DSD and DPD will be repealed). Consequently, information requirements on oxidising properties are inherently covered by classification of the substance according to the CLP Regulation. However, the physical hazards according to CLP Regulation are structured completely differently from the physicochemical properties according to DSD (and therefore also REACH, Annex VII). This means that for some of the CLP hazard classes an unambiguous assignment to one of the headlines (information requirements) in Annex VII to REACH is not possible. The assignment of hazard classes to the headline 'oxidising properties' as shown in the table below (Table R.7.1–10) must therefore only be understood as a means to structure this document in accordance with Annex VII to REACH.

Table R.7.1–10 Assignment of CLP hazard classes to the information requirement 'Oxidising properties' according to REACH, Annex VII and correlation between the Test method Regulation and the test method according to CLP and supporting link with the *Guidance on the Application of the CLP criteria*.

Information requirement according to Art. 10 (a) (vi) of the REACH Regulation (EC) No. 1907/2006 (the no. in brackets is the respective no. in the table in Annexes VII to IX to REACH)	**CLP Regulation (EC) No. 1272/2008** (the no. in brackets is the respective chapter no. in Annex I to CLP)	**Chapter in revised R.7(a) guidance**	**Corresponding test method according to The Test Method Regulation , Regulation (EC) No. 440/2008**	**Corresponding test method according to CLP Regulation**	**Chapter in the Guidance on the Application of the CLP Criteria** (ex RIP 3.6)
Oxidising properties (7.13)	Oxidising gases (2.4) *	R.7.1.13.1	n.a.	**ISO 10156**	**2.4**
	Oxidising liquids (2.13) *	R.7.1.13.2	A.21 Oxidising properties (liquids)	**UN Test O.2**	**2.13**
	Oxidising solids (2.14) *	R.7.1.13.3	A.17 Oxidising properties (solids)	**UN Test O.1**	**2.14**

* Note that regardless of whether the hazard class or category is listed in Article 14 (4)(a) of REACH the chemical safety assessment (when required) must be performed in accordance with Article 14 (3) REACH. Furthermore, according to Article 10(a)(iv) of REACH the technical dossier of a registration of a substance under the REACH Regulation must include information on classification and labelling of the substance as specified in section 4 of Annex VI to the REACH Regulation.

R.7.1.13.1 Oxidising gases

Definition

The following definition of oxidising gases is provided in CLP Annex I, section 2.4.1.:

'Oxidising gas means any gas or gas mixture which may, generally by providing oxygen, cause or contribute to the combustion of other material more than air does.'

The criteria '*more than air does'* is further defined in a Note under Table 2.4.1 in Section 2.4.1 as '*having an oxidising power greater than 23.5 % as determined by a method specified in ISO 10156 as amended'.*

Classification criteria and relevant information

All oxidising gases are classified as oxidising gas, Category 1 (Ox. Gas 1, H270). Detailed guidance on the classification criteria and the test method(s) can be found in the *Guidance on the Application of the CLP criteria*, section 2.4.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for oxidising properties.

'The study does not need to be conducted if:

- *the substance is explosive, or*
- *the substance is highly flammable, or*
- *the substance is an organic peroxide, or*
- *the substance is incapable of reacting exothermically with combustible materials, for example on the basis of the chemical structure (e.g. organic substances not containing oxygen or halogen atoms and these elements are not chemically bonded to nitrogen or oxygen, or inorganic substances not containing oxygen or halogen atoms).*

The full test does not need to be conducted for solids if the preliminary test clearly indicates that the test substance has oxidising properties.

Note that as there is no test method to determine the oxidising properties of gaseous mixtures, the evaluation of these properties must be realised by an estimation method based on the comparison of the oxidising potential of gases in a mixture with that of the oxidising potential of oxygen in air.'

According to above indents, the study therefore does not need to be conducted if the gas:

- is classified as highly flammable; or
- does not contain oxygen, fluorine and/or chlorine which are chemically bonded to elements other than carbon or hydrogen.

The other above cited indents are not relevant for this endpoint.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

There are not many gases that are oxidising. Most oxidising gases are identified as such in the UN-RTDG and in ISO 10156: 2010 *Gas cylinders - Gases and gas mixtures: - Determination of fire potential and oxidizing ability for the selection of cylinder valve outlets.*

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

There is no known scientific literature that refers to test results for gases that are not classified in ISO 10156 nor in the UN-RTDG.

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

At present (Q)SAR is generally not applicable for the determination of oxidising limits of gases. Application of (Q)SAR is not possible. However, assessment of the chemical structure may be used to exclude oxidising behaviour of a substance. Possibly, this relation could be exploited in the development of future QSPR methods.

- Grouping of substances and read-across approach

For the determination of the oxidising gases read-across is usually not possible. However interpolation may still be possible within homologous series.

- Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

The normal level of impurities in the technical grade of oxidising gases does not impact the result of the test. Tests should be performed with the lowest concentration of impurities in the gas encountered in the normal manufacturing process and the moisture content should be less than or equal to 0.01 mol%. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

How to conclude on the DSD classification

All gases with a positive test result according to the test method described in ISO 10156 are classified 'Oxidising O, R8'.

Endpoint specific information in the registration dosser/in IUCLID

Material and methods

- reference to the standard applied.

Results and discussion

- if the test is positive indicate that the gas is 'oxidising';
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on oxidising properties can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.15	VII 7.13	Oxidising properties	E.4.16	3.14

Further information / references

Further information about classification and testing for oxidising gases can be found in the ECHA *Guidance on the application of CLP criteria*, section 2.4.

The test method is described in ISO 10156. The test is qualitative. If reaction is observed during the test, the gas to be evaluated is oxidizing.

For several gases, a 'coefficient of oxygen equivalency' (Ci) has been deduced from the explosion ranges observed during the tests. The Ci factors are listed in ISO 10156 along with the list of oxidising gases.

R.7.1.13.2 Oxidising liquids

Definition

The following definition of oxidising liquids is provided in CLP Annex I, section 2.13.1.:

'Oxidising liquid means a liquid substance or mixture which, while in itself not necessarily combustible, may, generally by yielding oxygen, cause, or contribute to, the combustion of other material.'

Classification criteria and relevant information

According to the CLP Regulation, a liquid is classified as an oxidising liquid if, in testing according to the UN Test O.2 of the UN-MTC (Part III, Section 34) it is at least as oxidising as a 65 % aqueous solution of nitric acid. The CLP Regulation has three categories for Oxidising Liquids. The category is also determined through the UN Test O.2, by comparison to various reference oxidisers.

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for oxidising properties.

'The study does not need to be conducted if:

- *the substance is explosive, or*
- *the substance is highly flammable, or*
- *the substance is an organic peroxide, or*
- *the substance is incapable of reacting exothermically with combustible materials, for example on the basis of the chemical structure (e.g. organic substances not containing oxygen or halogen atoms and these elements are not chemically bonded to nitrogen or oxygen, or inorganic substances not containing oxygen or halogen atoms).*

The full test does not need to be conducted for solids if the preliminary test clearly indicates that the test substance has oxidising properties.

Note that as there is no test method to determine the oxidising properties of gaseous mixtures, the evaluation of these properties must be realised by an estimation method based on the comparison of the oxidising potential of gases in a mixture with that of the oxidising potential of oxygen in air.'

The first indent states that explosive substances should not be tested for oxidising properties. For instance, organic substances with oxidising functional groups may be explosive and should first undergo the screening procedures for explosive properties in Annex 6 of the UN-MTC to rule out possible explosive behaviour. Such substances may also be thermally unstable and show self-reactive behaviour. Substances that have been classified as Explosives according to the CLP Regulation or have been assigned risk phrases R2 or R3 according the DSD, should normally not be tested for oxidising properties, since they are known to be explosive.

The second indent states that highly flammable substances do not have to be tested for oxidising properties. While it is not very clear what 'highly flammable' means in this case (whether it is or is not intended to mean 'extremely flammable' and 'flammable'), liquids that have a low flash point, or which are pyrophoric, are rarely oxidising. This implies that liquids classified as Flammable Liquids category 1 or 2, or as Pyrophoric Liquids, according to the CLP Regulation, normally do not need to be tested for oxidising properties. This corresponds to classification with risk phrases R12, R11 or R17 according to the DSD. If they contain oxidising functional groups, such substances may instead show self-reactive or explosive behaviour.

The third indent states that organic peroxides should not be tested for oxidising properties. Organic peroxides are distinguished by their chemical structure, and should be treated according to the procedures for the hazard class Organic Peroxides of the CLP Regulation, see Section R.7.1.10.9 of this document.

Waiving according to the fourth indent relies on the absence of particular molecular structural features. The wording is more precise in section 2.13.4 of Annex I to the CLP Regulation, which is in principle the same as the wording as in section 6 of Appendix 6 to the UN-MTC.

The last two paragraphs above quoted from Column 2 Specific rules for adaptation from Column 1 are not applicable for this endpoint.

According to Section 2.13.4.1 of Annex I to the CLP Regulation, an organic liquid does not have to be assessed for oxidising properties if:

a. *'the substance does not contain oxygen, fluorine or chlorine; or*

b. *the substance contains oxygen, fluorine or chlorine and these elements are chemically bonded only to carbon or hydrogen.'*

For inorganic liquids, assessment of oxidising properties does not have to be done if the substance does not contain any oxygen or halogen atoms, according to section 2.13.4.2 of Annex I to the CLP Regulation.

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

The UN Test O.2 of the UN-MCT is also used for classification according to the UN-RTDG,and consequently also in the various regulations on transport of dangerous goods e.g. ADR and RID. A liquid that has been classified as belonging to Division 5.1 according to the regulations on transport of dangerous goods on the basis of results from the UN Test O.2, is an Oxidising Liquid according to the criteria of the CLP Regulation.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

For the determination of whether a liquid is an oxidising liquid, weight of evidence is not possible. Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

At the time of writing, no reliable (Q)SAR-methods exist for sufficiently accurate predictions of oxidising properties. As explained above, however, assessment of the chemical structure may be used to exclude oxidising behaviour of a substance. Possibly, this relation could be exploited in the development of future QSPR-methods. Such an assessment of chemical structure would formally form part of a Column 2 adaptation justification.

- Grouping of substances and read-across approach

For the determination of the whether a liquid is an oxidising liquid, read-across is usually not possible. However interpolation may still be possible within homologous series.

- Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies. Some oxidising substances may decompose when heated. Substances may occasionally react with cellulose in other ways than through oxidation of the cellulose (e.g. through breaking chemical bonds within the cellulose). See also section 2.13.4.4 of Annex I to the CLP Regulation.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

Minor impurities will usually not influence the test, unless they are very strong oxidisers. Expert judgement should be used to determine whether impurities may have an effect. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

A few substances may show other reactions than pure oxidation of the cellulose, or may decompose. If this is suspected, expert judgement should be sought. See also section 2.13.4.4 of Annex I to the CLP Regulation.

How to conclude on the DSD classification

Any substance classified as an oxidising liquid according to the CLP-criteria should normally be classified with risk phrase R8 or R9 according to the DSD. The DSD-criteria for classification with risk phrase R9 are not very precise, but if the CLP classification is Category 1, the substance should be classified with risk phrase R9 if the reaction with cellulose is violent, e.g. if spontaneous ignition occurs in the test.

In the DSD, the A.21 test method of Regulation (EC) 440/2008 is used for the assessment of oxidising properties of liquids. This method is in principle identical to the UN Test O.2 of the UN-MTC used in the CLP Regulation. However, the DSD does not make any division corresponding to the categories of the CLP, and therefore only one reference substance is used in the A.21 test method. Since the CLP Regulation method is used for classification of substances, it is strongly advisable to use the UN Test O.2 instead of the A.21 test method. This is because the O.2 test method will also give more detailed information on the oxidising behaviour of a substance (or mixture), since more reference mixtures are used.

Endpoint specific information in the registration dosser/in IUCLID

Material and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied.

Results and discussion

- indicate the results of the spontaneous ignition test;
- indicate the mean pressure rise time for the test substance;
- indicate the mean pressure rise time for the reference substance(s);
- interpretation of results;
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on oxidising properties can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.15	VII 7.13	Oxidising properties	E.4.16	3.14

Further information / references

The ECHA guidance document *Guidance on the Application of the CLP criteria* gives in Chapter 2.13 detailed information on the oxidising property, the CLP-classification, the UN Test O.2 and the relation to the DSD and the transport of dangerous goods regulations.

R.7.1.13.3 Oxidising solids

Definition

The following definition of oxidising solids is provided in CLP Annex I, section 2.14.1:

'Oxidising solid means a solid substance or mixture which, while in itself is not necessarily combustible, may, generally by yielding oxygen, cause, or contribute to, the combustion of other material.'

Classification criteria and relevant information

According to the CLP Regulation, a solid is classified as an oxidising solid if in testing according to the UN Test O.1 of the UN-MTC (Part III, Section 34), it is at least as oxidising as potassium bromate in a 3:7 mixture with cellulose. The test is based on the burning behaviour of a mixture of cellulose and the tested solid. The CLP Regulation has three categories for oxidising solids. The category is also determined through the UN Test O.1 in the UN-MTC by comparison to reference mixtures of cellulose and potassium bromate[20].

Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for oxidising properties.

'The study does not need to be conducted if:

- *the substance is explosive, or*

[20] At the time of writing, work is in progress at the UN-level to modify Test O.1: Test for oxidising solids. This includes changing the reference substance and introducing a gravimetric method for the measurement. For further information, see document UN/SCEGHS/23/INF.17 available at the following link: http://www.unece.org/fileadmin/DAM/trans/doc/2012/dgac10c4/UN-SCEGHS-23-INF17.doc-UN-SCETDG-41-INF.43e.pdf .

- *the substance is highly flammable, or*
- *the substance is an organic peroxide, or*
- *the substance is incapable of reacting exothermically with combustible materials, for example on the basis of the chemical structure (e.g. organic substances not containing oxygen or halogen atoms and these elements are not chemically bonded to nitrogen or oxygen, or inorganic substances not containing oxygen or halogen atoms).*

The full test does not need to be conducted for solids if the preliminary test clearly indicates that the test substance has oxidising properties.

Note that as there is no test method to determine the oxidising properties of gaseous mixtures, the evaluation of these properties must be realised by an estimation method based on the comparison of the oxidising potential of gases in a mixture with that of the oxidising potential of oxygen in air.'

The first indent states that explosive substances should not be tested for oxidising properties. For instance, organic substances with oxidising functional groups may be explosive and should first undergo the screening procedures for explosive properties in Annex 6 of the UN-MTC to rule out possible explosive behaviour. Such substances may also be thermally unstable and show self-reactive behaviour. Substances that have been classified as Explosives according to the CLP-regulation or have been assigned risk phrases R2 or R3 according the DSD, should normally not be tested for oxidising properties, since they are known to be explosive.

The second indent states that highly flammable substances do not have to be tested for oxidising properties. While it is not very clear what 'highly flammable' means in this case (whether it is or is not intended to mean 'extremly flammable' and 'flammable'), solids classified as Flammable Solids or as Pyrophoric Solids according to the CLP-regulation are rarely oxidising. This corresponds to classification with risk phrases R11 or R17 according to the DSD. If they contain oxidising functional groups, such substances may instead show self-reactive or explosive behaviour.

The third indent states that organic peroxides should not be tested for oxidising properties. Organic peroxides are distinguished by their chemical structure, and should be treated according to the procedures for the hazard class Organic Peroxides of the CLP-regulation, see Section R.7.1.10.9 of this document.

Waiving according to the fourth indent relies on the absence of particular molecular structural features. The wording is more precise in section 2.14.4 of Annex I to the CLP-regulation, which is in principle the same as the wording as in Section 6 of Appendix 6 to the UN-MTC.

The first note under last indent from REACH Annex VII, which allows waiving of further testing, namely *'[...] if the preliminary test clearly indicates that the test substance has oxidising properties'* is relevant only when using the A.17 test method of Regulation (EC) 440/2008, which is **not** the preferred test method since it belongs to the DSD classification system. The UN Test O.1 used for classification according to the CLP Regulation does not include any preliminary test.

The last note taken from Column 2 'Specific rules for adaptation from Column 1' is not applicable for this endpoint. For inorganic solids, assessment of oxidising properties does not have to be done if the substance does not contain any oxygen or halogen atoms, according to section 2.14.4.2 of Annex I to the CLP Regulation.

According to section 2.14.4.1 of Annex I to the CLP-regulation, an organic solid does not have to be assessed for oxidising properties if:

a. *'the substance does not contain oxygen, fluorine or chlorine; or*

b. *the substance contains oxygen, fluorine or chlorine and these elements are chemically bonded only to carbon or hydrogen.'*

Adaptation possibilities according to Annex XI to REACH

- Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

The UN Test O.1 of the UN-MTC is also used for classification according to the UN-RTDG, and consequently also in the various regulations on transport of dangerous goods e.g. ADR and RID. A solid that has been classified as belonging to Division 5.1 according to the regulations on transport of dangerous goods on the basis of results from the UN Test O.1, is an oxidising solid according to the criteria of the CLP Regulation.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

- Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

- (Q)SAR

At the time of writing, no reliable (Q)SAR-methods exist for sufficiently accurate predictions of oxidising properties. As explained above, however, assessment of the chemical structure may be used to exclude oxidising behaviour of a substance. Possibly, this relation could be exploited in the development of future (Q)SPR-methods. Such an assessment of chemical structure would formally form part of a Column 2 adaptation argument.

- Grouping of substances and read-across approach

For the determination of the oxidising solids read-across is usually not possible. However interpolation may still be possible within homologous series.

- Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies. Some substances may decompose upon heating. Substances may occasionally react with cellulose in other ways than through oxidation of the cellulose.

Further adaptation possibilities

Not foreseen.

Impurities; uncertainties

The UN Test O.1 is (currently) performed using the unaided eye as measuring instrument. Only by expert judgement and thorough experience can the result of the test be correctly judged, and even then uncertainties may arise.

Minor impurities will usually not influence the test, unless they are very strong oxidisers. Expert judgement should be used to determine whether impurities may have an effect.

A few substances may show other reactions than pure oxidation of the cellulose, or may decompose. If this is suspected, expert judgement should be sought. Particle size and size distribution can have an influence on the test results.

How to conclude on the DSD classification

Any substance classified as an oxidising solid according to the CLP Regulation criteria should normally be classified with risk phrase R8 or R9 according to the DSD. The DSD-criteria for classification with risk phrase R9 are not very precise, but if the CLP Regulation classification is Category 1, the substance should be classified with risk phrase R9 if the reaction with cellulose is violent.

In the DSD, the A.17 test method of Regulation (EC) 440/2008 is used for the assessment of oxidising properties of solids. Although the principle of this method is to a large extent the same as that of the UN Test O.1 of the UN-MTC, the experimental set-up, reference substance (barium nitrate) and measured quantity differ. Furthermore, the DSD does not make any division corresponding to the categories of the CLP. Since the CLP Regulation is used for classification of substances, it is not advisable to use the A.17 method (which belongs to the DSD classification system). Instead, the UN Test O.1 should be used, which will also give more detailed information on the oxidising behaviour of a substance (or mixture), since more reference mixtures are used.

Endpoints specific information in the registration dosser/in IUCLID

Material and methods

- description of the apparatus and dimensions or reference to the standard or the test method applied;
- particle size and distribution.

Results and discussion

- if testing is waived, the reasons for waiving must be documented in the dossier.

If the UN test O.1 was used:

- indicate if a vigorous reaction was observed;
- indicate the maximum burning time for the test mixture;
- indicate the maximum burning time for the reference mixtures;
- interpretation of results, including any relevant special observations;
- estimated accuracy of the result (including bias and precision).

If A.17 test method was used:

- indicate if in the preliminary test, a vigorous reaction was observed;
- indicate the maximum burning rate for the test mixture;
- indicate the maximum burning rate for the reference mixture;
- interpretation of results, including any relevant special observations;
- estimated accuracy of the result (including bias and precision).

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on oxidising properties can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.15	VII 7.13	Oxidising properties	E.4.16	3.14

Further information / references

The ECHA *Guidance on the Application of the CLP criteria* gives in Chapter 2.14 detailed information on the oxidising property, the CLP-classification, the UN Test O.1 and the relation to the DSD and the transport of dangerous goods regulations.

R.7.1.14 Granulometry

Advice to registrants with regard to nanomaterials characterisation of granulometry can be found in *Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the *Guidance on IR&CSA*, section 2.2.3 Granulometry.

R.7.1.14.1 Type of property

Granulometry is not a specific physico-chemical property of a substance. The original particle size distribution is highly dependent on the industrial processing methods used and can also be affected by subsequent environmental or human transformations. Particle size is usually measured in micrometers (= 10^{-6} m; µm; 'microns').

Granulometry is of considerable importance for the toxic properties of a substance as it influences aspects such as:

- the route of exposure of humans and toxicity by inhalation;
- the choice of route of administration for animal testing;
- the efficiency of uptake in an organism;
- the distribution in the environment.

Granulometry is of importance for combustible dusts as it influences aspects such as the likelihood to form combustible/explosive dust - air mixtures.

In general all powder materials have a range of particle sizes (particle size distribution), a presentation of the particle size distribution (e.g. using a histogram of the particle size vs. mass, particle size vs. number of particles, etc.) is therefore necessary to interpret the data.

For inhalation exposure it is well know that the human toxicity will be related with the place of deposition into the respiratory tract. The location of deposition mainly depends on the properties of the particle (size, shape, density etc) that are commonly taken into account considering the aerodynamic diameter of the particle (see definition below). Thus, the general approach has been to use mass fractions (e.g. health related fractions as defined by EN 481 or the EPA PM Fractions). For instance, in Europe, from the publication of the EN 481 the OELs for powder materials have been defined for one or several fractions (inhalable, thoracic or respirable).

Photocentrifuge method - the method of determining the particle size distribution, which is described in ISO 13318-2:2007, is applicable to powders that can be dispersed in liquids, powders that are present in slurry form and some emulsions. Typical particle size range for analysis is from about 0.1 µm to 5 µm. The method is applicable to powders in which all particles have the same density and comparable shapes and do not undergo chemical or physical change in the suspension liquid. It is usually necessary that the particles have a density higher than that of the liquid.

Light extinction liquid-borne particle counter – in ISO 21501-3:2007 a calibration and verification method for a light extinction liquid-borne particle counter (LSLPC) is described, which is used to measure the size and particle number concentration of particles suspended in liquid. The light extinction method is based on single particle measurements and the typical size range of particles measured by this method is between 1 µm and 100 µm.

Light scattering liquid-borne particle counter - in ISO 21501-2:2007 a calibration and verification method for a light scattering liquid-borne particle counter (LSLPC) is described, which is used to measure the size and particle number concentration of particles suspended in liquid. The light scattering method is based on single particle measurements and the typical size range of particles measured by this method is between 0.1 µm and 10 µm.

Centrifugal X-ray method - the method of determining the particles size distribution described in ISO 13318-3:2004 is applicable to powders which can be dispersed in liquids or powders which are present in slurry form. The typical particle size range for analysis is from 0.1 µm to 5 µm. The method is applicable to powders in which all particles have the same effective density, chemical composition and comparable shapes.

The CEN document, EN 481 'Workplace Atmospheres – size fraction definitions for measurement of airborne particles' (CEN 1993) provides definitions of the inhalable, thoracic and respirable size fractions, and target specifications (conventions) for sampling instruments to measure these fractions. The current standard defines sampling conventions for particle size fractions which are to be used in assessing the possible health effects resulting from inhalation of airborne particles in the workplace. The different particle sizes defined in EN 481 are:

- inhalable fraction (the mass fraction of particles that can be inhaled by nose and mouth. Particles >100 µm are not included in the inhalable convention;
- thoracic fraction (the mass fraction of the inhaled particles that passes the larynx). The convention for thoracic fraction sets that 50% of the particles in air with an aerodynamic diameter of 10 µm belong to the thoracic fraction;
- respirable fraction (the mass fraction of the inhaled particles that reaches the alveoli) The convention for respirable fraction sets that 50% of particles with an aerodynamic diameter of 4 µm belong to the respirable fraction.

R.7.1.14.2 Definitions

Aerodynamic diameter: the diameter of a sphere of density 1 $g.cm^{-3}$ with the same terminal velocity (falling speed) due to gravitational force in calm air as the particle under the prevailing conditions of temperature, pressure and relative humidity (CEN, 1993). The aerodynamic diameter is used to compare partcles of different sizes, shapes and densities and it is a useful parameter to predict where in the respiratory tract such particles may be deposited. It is used in contrast to 'optical', 'measured' or 'geometric' diameters which are representations of actual diameters which in themselves cannot be related with the deposition within the respiratory tract.

Particle diffusion diameter: for particles of aerodynamic diameter less than 0.5 µm, the particle diffusion diameter should be used instead of the particle aerodynamic diameter. For diffusion, the appropriate *equivalent diameter* is the diffusion (mobility) diameter. This is defined as the diameter of a sphere with the same diffusion coefficient as the particle under the prevailing conditions of temperature, pressure and relative humidity.

The parameter of interest is the effective hydrodynamic radius, or effective Stoke's radius R_s. Particle size distribution (effective hydrodynamic radius) requires information on water insolubility. Fibre length and diameter distributions require information on the fibrous nature of the product and on stability of the fibrous shape under electron microscope conditions.

A fibre: is a water insoluble particle with an aspect ratio (length/diameter > 3) and diameter < 100 µm. Fibres of length < 5 µm need not be considered.

Particle: Minute piece of matter with defined physical boundaries. (ISO/TS 27687:2008)

Agglomerate: A collection of weakly bound particles of aggregates or mixtures of the two where the resulting external surface area is similar to the sum of the surface areas of the individual components (ISO/TS 27687:2008).

Aggreggate: Particle comprising strongly bonded or fused particles where the resulting external surface area may be significantly smaller than the sum of calculated surface areas of the individual components (ISO/TS 27687:2008).

R.7.1.14.3 Test methods

Many methods are available for particle size measurements, but none of them is applicable to the entire size range (see Table R.7.1–11). Sieving, microscopic sedimentation and elutriation techniques are most commonly employed. Methods for determining particle size distribution are designed to provide information on the transportation and sedimentation of insoluble particles in water and air. An integrated testing strategy (ITS) detailing the appropriate methods for determination of particle size distribution of respirable and inhalable particles is shown in Figure R.7.1–6.

Details of methods for determining particle size distribution and for fibre length and diameter distributions are outlined in OECD TG 110 and in the 'Guidance Document on the Determination of Particle Size Distribution, Fibre Length and Diameter Distribution of Chemical Substances' (JRC, 2002).

The particle size distribution is carried out on the material under investigation and not as airborne dust.

The measurement principle of the method used will determine what kind of diameter of the particle can be determined: for instance, optical diameter when using light scattering or aerodynamic diameter when using impactors. Methods which determine the mass median aerodynamic diameter (MMAD) need the generation of representative test atmospheres using suitable generation equipment and correct sampling techniques. They can be used in case of airborne particles (dusts, smokes, fumes), nebulised particles (wet aerosol) or dispersed particles (dry aerosol).

Figure R.7.1–6 Integrated testing strategy for granulometry

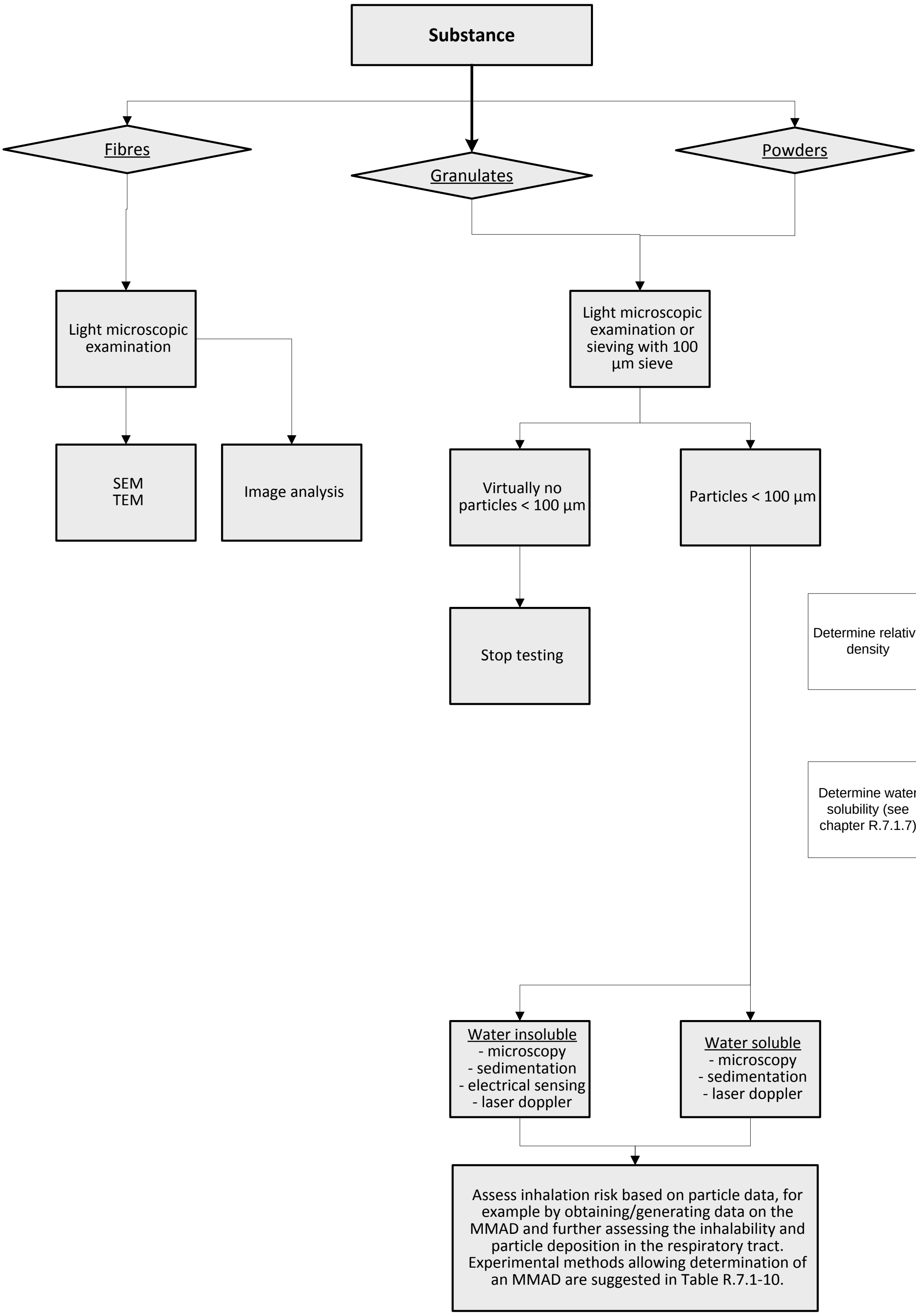


Table R.7.1–11 Methods to determine particle size distribution of a material

Method and details	Material and size range	MMAD
Microscopic examination It is preferable to prepare samples directly in order not to influence shape and size of the particles. This method determines size distribution of particles.	Particles of all kinds Size range: 0.5–5000 microns (light microscope) and <0.1–10 microns (SEM/TEM)	MMAD cannot be determined
Sieving Sieving using wire-mesh sieves and perforated sheet metal sieves is not suitable to determine the distribution of particles of respirable and inhalable size since their range is only 100-10,000 microns. Micro mesh sieves (range 5-100 micron) may give better results. However, since these sieves are generally operated in combination with mechanical or ultrasonic vibration, modification of median size and form may result. Sieving not suitable to determine distribution of particles of respirable size, but might be suitable to determine bigger particles.	Dry powders/granulates Size range: 100–10,000 microns (wire mesh/metal sieves) and 5-100 (micromesh)	MMAD cannot be determined
Sedimentation (gravitational settling) Method is based on gravitational settling of particles in liquid and the effective hydrodynamic radius is determined. Effective hydrodynamic radius distribution should be measured 3x with no two values differing by >20%. Requires sufficient numbers of radius intervals be used to resolve the radius distribution curve. Binary or ternary mixtures of latex spheres (2-100 microns) are recommended as calibration material. Method might be suitable to determine the distribution of particles of respirable and inhalable size.	Dry powders/granulates Size range: 2-200 microns	MMAD cannot be determined
Electrical Sensing Zone (e.g. Coulter) method Samples are suspended in an electrolytic solution. As the particle is drawn through an aperture, the change in conductance gives a measure of particle size. The important parameter is the settling velocity in the liquid phase, which depends on both density and diameter. Particles having a density of several g/cm^3 can be determined. Applicable to particles that are complete electrical isolators in the fluid. Difference in density between particles and fluid must not be too large. Method might be suitable to determine the distribution of particles of respirable and inhalable size.	Dry powders/granulates (non-conducting) Size range: 1-1000 microns	MMAD cannot be determined
Phase Doppler Anemometry Expensive technique. Particle size distribution can be measured either in air or in liquid. The method presupposes that the particles are spherical with	Dry powders/granulates Size range: 0.5-80 microns (in	MMAD cannot be determined

known refractive index. Method might be suitable to determine the distribution of particles of respirable and inhalable size.	air); 0.5-1000 microns (in liquid)	
Determination of fibre length and diameter distributions Light microscopy used to examine likelihood of fibres present by comparing similarities to known fibrous or fibre releasing substances or other data. Extreme care required during sample preparation to avoid fibre breaking and clumping. Care should also be taken to avoid contamination by airborne fibres. Samples might be prepared by (a) producing suspensions in water by gentle hand agitation or vortex mixing or (b) transfer of dry material onto copper tape either directly or by spraying of the dry fibres by use of atomiser or pipette. Length and diameter distributions should be measured independently at least twice and at least 70 fibres counted. No two values in a given histogram interval should differ by > 50% or 3 fibres, whichever is larger. The presence of long thin fibres would indicate a need for further, more precise measurements.	Fibrous products Size range: diameters as small as 0.1 micron and as large as 100 micron and lengths as small as 5 micron and as large as 300 micron	

It is advantageous to have accurate information about the propensity of materials to produce airborne dust (the *dustiness* of the material). No single method of dustiness testing is likely to represent and reproduce the various types of processing and handling used in industry. The measurement of dustiness depends on the test apparatus used, the properties of the dust and various environmental variables (i.e the dustiness is not a measurement of the 'dust as it is') . There are a number of methods for measuring the dustiness of bulk materials, based on the health related aerosol fractions defined in EN 481. Two methods (the rotating drum method and the continuous drop method) are detailed in EN 15051 'Workplace atmospheres – Measurement of the dustiness of bulk materials – Requirements and reference test methods' (CEN, 2006).

Dustiness is a relative term (derived from the amount of dust emitted during a standard test procedure). This is dependent on the method chosen, the condition and properties of the tested bulk material, and various environmental variables in which the tests are carried out. Thus, the two methods in EN 15051 may provide different results (the methods are intended to simulate handling processes) The standard is currently under revision (draft of European standard available) and the final publication is expected for 2013. The standard has been divided in 3 parts (a general part and one part for each of the methods). The methods (Table R.7.1–12) as described in the standard are used to determine dustiness in terms of the health related fractions defined by EN 481. Further analyse (e.g. analysing the contents on the dust collection stages) can be used to obtain the particle sizedistribution. These methods require the generation of representative test atmospheres using suitable generation equipment and correct sampling techniques.

Table R.7.1–12 Methods to generate/sample airborne dispersed or nebulised particles

Method and details	Material and size range	MMAD
Cascade impaction Cascade impactors can be used to obtain the size distribution of an aerosol (i.e in this context a dust cloud). Air samples are drawn through a device which consists of several stages on which particles are deposited on an impactation substrate. Particles will impact on a certain stage depending on their aerodynamic diameter . The cut- off size can be calculated from the jet velocities at each stage by weighing each stage before and after sampling and the MMAD derived from these calculations. This is a well established technique to measure the size distribution of particles (allowing calculating any mass fraction). Some models are specifically designed to give the 3 health related fractions defined by the EN 481. Please also check ISO/TR 27628:2007, which contains specific information on methods for bulk aerosol characterization and single particle analysis while using cascade impaction method.	Particles in an aerosol Size range: 0.1-20 and 0.5-80 microns	MMAD can be determined *via* an appropriate coupled analytical technique.
Laser scattering/diffraction In general, the scattering of the incident light gives distinct pattern which are measured by a detector. This technique is particle property dependent – i.e. material has unique scattering and diffraction properties which are also particle size dependent. It is important to calibrate the instrument with similar material (of the same size range as the material to be measured). Laser scattering techniques are suitable for geometric particles, viz spheres, cubes and monocrystals. Particle size will be established optically. The MMAD can be calculated by means of a calculation correction. Further information about corrections and limitations of the methods can be found in CEN/TR 16013-1 and CEN/TR 16013-2. Please also check ISO 13320:2009 Particle size analysis – Laser diffraction methods taking into account the possible limitations of the method, a the technique assumes a spherical particle shape in its optical model. The resulting particle size distribution is different from that obtained by methods based on other physical principles (e.g. sedimentation, sieving).	Particles of all kind Size range: 0.1 um to 3 mm (with special instrumentation and conditions, the size range can be extended above 3 mm and below 0.1 mm)	MMAD can be determined.
Rotating drum method (prEN 15051-2) This method is based on size selective sampling of an airborne dust cloud produced by the repeated lifting and dropping of a material in a rotating drum. Air drawn through the drum passes through a specially designed outlet and a 3-stage fractionating system consisting of two	Dry powders/granulates/friable products Size range: 0.5-10,000 microns	MMAD cannot be determined.

porous polyurethane foams and a membrane filter. The mass of dust collected on each collection stage is determined gravimetrically to give a direct measure of the biologically relevant size fractions. This method simulates a wide range of material handling processes in industry and determines the biologically relevant size functions of a material in the airborne state. This method is suitable to determine the respirable thoracic or inhalable fractions.		
Continuous drop method (prEN 15051-3) This method is based on the size selective sampling of an airborne dust cloud produced by the continuous single dropping of material in a slow vertical air current. The dust released by dropping material is conducted by the airflow to a sampling section where it is separated into the inhalable and respirable fractions. This method is suitable to determine the respirable and inhalable fractions.	Dry powders/granulates/friable products Size range: 0.5-10,000 microns	MMAD can be determined.

Table R.7.1–13 Methods that measure inhalable fractions only or that give no detailed distributions

Method and details	Material and size range	MMAD
Elutriation Particles are drawn out on a column at varying velocity. The velocity is used to calculate particle size and the weight of the remaining sample at a particular velocity is used to calculate the distribution. The method is limited to particles >15 microns. The method is not suitable to determine the distribution of particles of respirable size, but might be suitable to determine the distribution of particles of inhalable size	Dry powders/granulates Size range: 15-115 microns	MMAD cannot be determined.
Air jet sieve Air is aspirated through a weighted sample on a fine sieve and the weight loss measured. The method is capable of estimating the non-floatable fraction of the material under investigation. Aggregation of the particles will result in unreliable values. In addition, since the lower detection limit is only 10 micron, this method is not suitable to determine the distribution of particles of respirable size. The method is not suitable to determine the distribution of particles of the respirable fraction, but might be suitable to determine the distribution of particles between 10 and 10,000 microns..	Particles of all kind Size range: 10-10,000 microns	MMAD cannot be determined.
Cyclons The use of a cyclone is a simple approach to determining whether respirable and/or inhalable particles are present in the test atmospheres by constructing the cyclone cut off points at 4.25 and 100 microns. By measuring the weight of particles which pass through the cyclone it can be decided whether more sophisticated methods have to be applied to determine the size distribution of the particles smaller than 10 micron. This method is suitable to determine the respirable, thoracic or inhalable fraction.	Particles of all kind Size range: 0.1-200 microns	MMAD cannot be determined.

R.7.1.14.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VII to REACH

Column 2 of REACH Annex VII provides the following specific rules for adaptation of the standard information requirement for granulometry.

‘*The study does not need to be conducted if the substance is marketed or used in a non solid or granular form.*’

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

As the granulometry of a substance is highly dependent on the industrial processing methods and possibly also on handling of the material, any published data on granulometry will be pertinent only to the particular sample or process.

There are a number of web sites and electronic databases that include compilations of and evaluations of data on particle properties. However, there appear to be a limited number of reference books that provide particle size data.

The equivalence of the various national and international standard methods for particle size distribution has not been tested and is not known.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

In some situations where data is available from multiple sources (e.g. information on particle size distribution of different batches, or information from different methods), a weight of evidence approach may be used. Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

There are no QSPR/(Q)SAR tools available for predicting particle size and the data will therefore need to be experimentally determined. Application of (Q)SAR is not possible.

Grouping of substances and read-across approach

At present grouping and read across are not applicable.

Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies. Testing should always be possible for solids or granular substances.

Further adaptation possibilities

Not foreseen.

R.7.1.14.5 Impurities; uncertainties

There is a particular problem in relation to sedimentation and Coulter counter measurements. The effect of impurities on particle shape should be considered when measuring fibre length and diameter distributions.

The small quantities used as samples must be representative of product batches comprising many kilograms; therefore sampling and sample handling require great care.

Great care should also be taken due to the fact that non-conducting particles in a non-conducting liquid may be electrically charged resulting in non-representative settling of particles of a certain size. In addition, in the process of particle size distribution determination, it is very important to take the electrostatic charge of the particles into account. Electrostatically charged particles behave differently and may influence sampling.

It is useful to distinguish between aggregates and agglomerates. While an aggregate is held together by strong forces and may be considered to be permanent, agglomerates are held together with weak forces and may break up under certain circumstances. As small particles often form agglomerates, sample pre-treatment (e.g. the addition of dispersing agents, agitation or low-level ultrasonic treatment) may be required before the primary particle size can be determined. However, great care must be taken to avoid changing the particle size distribution.

R.7.1.14.6 Endpoint specific information in the registration dossier / in IUCLID

Material and methods

- sample preparation, such as any sonication, grinding, or addition of dispersion agents (if any);
- if a suspending medium is used (e.g. sedimentation test): indicate type of medium, temperature, pH, concentration and solubility of the substance in the suspending medium;
- the type of method used.

Results and discussion

- in the particle size field: mean and standard deviation;
- in the particle size distribution at different passages field: size and distribution;
- approximate information on particle shape (e.g. spherical, platelike, needle shaped) if available;
- for fibres: indicate both length and diameter of fibres.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on particle size distribution (Granulometry) can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.5	VII 7.14	Particle size distribution (Granulometry)	E.4.6	3.5

R.7.1.14.7 Further information / references

CEN 1993	EN 481: Workplace atmospheres. Size fraction definitions for measurement of airborne particles
CEN 2006	EN 15051: Workplace atmospheres. Measurement of the dustiness of bulk materials – Requirements and reference test methods
JRC (2002)	"Guidance Document on the Determination of Particle Size Distribution, Fibre Length and Diameter Distribution of Chemical Substances", ISBN 92-894-3704-9, EUR 20268 EN, http://publications.jrc.ec.europa.eu/repository/handle/111111111/5555
OECD TG 110	Test No. 110: Particle size distribution/fibre length and diameter distributions
prEN 15051-1 rev	Workplace exposure - Measurement of dustiness of bulk materials - Part 1: Requirements and choice of test methods
prEN 15051-2	Workplace exposure - Measurement of the dustiness of bulk materials - Part 2: Rotating drum method
prEN 15051-3	Workplace exposure - Measurement of the dustiness of bulk materials - Part 3: Continuous drop method
(ISO/TS 27687:2008)	Nanotechnologies-Terminology and definitions for nano-objects-Nanoparticle, nanofibre, and nanoplate
CEN/TR 16013-1:2010	Workplace exposure. Guide for the use of direct-reading instruments for aerosol monitoring. Choice of monitor for specific applications
CEN/TR 16013-2:2010	Workplace exposure. Guide for the use of direct-reading instruments for aerosol monitoring. Evaluation of airborne particle concentrations using optical particle counters

R.7.1.15 Adsorption/Desorption

Advice to registrants with regard to nanomaterials characterisation of adsorption/desorption can be found in *Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the Guidance on IR&CSA, section 2.2.4 Adsorption/desorption.

R.7.1.15.1 Type of property

Adsorption/desorption is not a specific physicochemical property of a substance. This property indicates the binding capacity (or 'stickiness') of a substance to solid surfaces, and so is essential for understanding environmental partitioning behaviour.

Information on adsorption/desorption is an essential input to environmental exposure models, because:

- adsorption to suspended matter can be an important physical elimination process from water in sewage treatment plants (STPs). This in turn may mean that sewage sludge, if spread to land, is a major source of the substance in soil;
- adsorption to suspended matter in receiving waters affects both the concentration in surface water and the concentration in sediment;
- desorption of a substance from soil directly influences its mobility and potential to reach surface or groundwaters.

Consequently, information on adsorption/desorption is also an important factor in test strategies for assessing toxicity to sediment- or soil-dwelling organisms.

Substances that adsorb strongly to biological surfaces (e.g., gills, skin, etc.) may lead to toxic effects in higher organisms after biomagnification.

The information is also relevant for assessing environmental persistence. For example: degradation rates in sediment and soil are also assumed to be reduced by default if a substance is highly sorptive (since it is less bioavailable to microorganisms). This may lead to consideration of soil/sediment simulation testing in some cases.

Finally, there may be practical implications for test performance: Substances that adsorb strongly to surfaces can be difficult to test in aquatic systems.

R.7.1.15.2 Definition

Adsorption is caused by temporary (reversible) or permanent bonding between the substance and a surface (e.g. due to van der Waals interactions, hydrogen bonding to hydroxyl groups, ionic interactions, covalent bonding, etc.). The OECD guidances offer further information (OECD 2000a, OECD 2000b, OECD 2001, OECD 2002).

The organic carbon normalized adsorption coefficient (K_{oc}) is the ratio of a substance concentration sorbed in the organic matter component of soil or sediment to that in the aqueous phase at equilibrium. In other words, $K_{oc} = K_d/f_{oc}$, where K_d is the distribution coefficient for adsorption, and f_{oc} the organic carbon content – the fraction organic carbon present in the soil or sediment. In turn, K_d is the experimental ratio of a substance's concentration in the soil (C_s) to that in the aqueous phase (C_{aq}) at equilibrium; namely $K_d = C_s/C_{aq}$. The organic matter normalized distribution coefficient (K_{om}) is similarly defined, but refers to the organic matter content of soil rather than the organic carbon content (OECD, 2000a).

R.7.1.15.3 Test method(s)

The adsorption of a substance to sewage sludge, sediment and/or soil can be measured or estimated using a variety of methods, which are tabulated in

Table R.7.1–14 in order of increasing complexity. The dissociation constant (if appropriate) should be known before testing. Information on vapour pressure, solubility in water and organic solvents, octanol-water partition coefficient and stability/degradability is also useful.

Table R.7.1–14 Methods for the measurement of adsorption

Method and Description	Applicability / Notes
Adsorption control within an inherent biodegradability test (OECD TG 302B) Estimate of the extent of adsorption to STP sludge made from the elimination level in a Zahn-Wellens inherent biodegradation test. (e.g. OECD TG 302B). 3-hour value recommended. Values beyond 24 hours not normally used. Where data are not available for adsorption up to 24 hours, data from time scales beyond this can only be used if adsorption is the only removal mechanism, with an upper limit of 7 days.	Highly adsorptive substances that are water soluble
HPLC method: OECD TG 121; EU C.19: Estimation of the Adsorption Coefficient (K_{oc}) on Soil and on Sewage Sludge using High Performance Liquid Chromatography (HPLC) (Original Guideline, adopted 22 January 2001) Calibration with reference substances (preferably structurally related to the test substance) of known K_{oc} allows the K_{oc} of the test substance to be estimated. Test substance K_{oc} value should lie within the calibration range of the reference substances.	Measurement of log K_{oc} in the range 1.5 to 5.0. Validated for several chemical types, see test guideline for details. Poorly soluble and volatile substances as well as mixtures. Ionisable substances: test both ionised and unionised forms in appropriate buffer solutions where at least 10 % of the test compound will be dissociated within pH range 5.5 to 7.5. May not be suitable for: substances that react with the column, solvent or other test system components; surface active substances; substances that interact in a specific way with inorganic soil components such as clay minerals; inorganic compounds; moderate to strong acids and bases.
Batch test of adsorption of substances on activated sludge (ISO 18749) Screening method to determine the degree of adsorption of substances on activated or primary sludge in sewage treatment plants (ISO, 2004). The method does not differentiate between adsorption and other elimination methods (such as complex formation, flocculation, precipitation, sedimentation or biodegradation).	Suitable for substances that: are water soluble, or allow for stable suspensions/dispersions/emulsions, are not significantly removed by abiotic processes (e.g. stripping/foaming), do not de-flocculate activated sludge, are not readily biodegradable, and have a sufficiently sensitive analytical method.

Sediment and soil adsorption/desorption isotherm (OPPTS 835.1220) Screening method according to US-EPA guideline (OPPTS, 1996) using three soil types.	
Batch equilibrium method (OECD TG 106; EU C.18: Absorption – Desorption Using a Batch Equilibrium Method (Updated Guideline, adopted 21 January 2000) Test uses a range of actual soils and so represents a more realistic scenario than the HPLC (OECD 121) method.	Used for substances with K_{oc} values that cannot be reliably determined using other techniques (e.g. surfactants). Requires a quantitative analytical method for the substance, reliable over the range of test concentrations. For ionisable substances, soil types should cover a wide range of pH. Adjustments for poorly soluble substances given in the test guideline.
OECD TG 312: Leaching in Soil Columns (Original Guideline, adopted 13 April 2004) K_d values can be derived from column leaching studies.	Appropriate study design to estimate K_d values particularly for unstable test substances that degrade significantly during the equilibrium time of 'shake flask' sorption studies
Simulation tests and direct field measurement: including OECD guidance document no. 22 (OECD, 2000b). Monolith lysimeters can be used to study the fate and behaviour of substances in an undisturbed soil profile under outdoor conditions. They allow for monitoring of the volume of leaching/drainage water as well as the concentrations of a substance and its transformation products. They are mainly used in pesticide studies. Field leaching studies can also be carried out where hydrodynamically isolated soil layers are analysed *in situ*. Although such studies are the most realistic, their reproducibility and representativity may be limited (e.g. due to the effects of large-scale soil structure, weather events, the soil conditions at the time of application, etc.). Since data from these methods are unlikely to be encountered for the vast majority of industrial substances, they are not considered further here. Further information can be found in guidance for pesticide registration.	

R.7.1.15.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex VIII and IX to REACH

Screening information on adsorption (and desorption) is required for substances manufactured or imported in quantities of 10 t/y or more. Depending on the results, further information (for example, a test) may be required for substances manufactured or imported in quantities of 100 t/y or more.

Column 2 of REACH Annexes VIII and IX provides two exemptions.

'The study does not need to be conducted if:

- *based on the physicochemical properties the substance can be expected to have a low potential for adsorption (e.g. the substance has a low octanol water partition coefficient), or*
- *the substance and its relevant degradation products decompose rapidly.'*

Or in other words, the substance and its relevant degradation products decompose rapidly. Therefore, if a substance hydrolyses, it might be more appropriate to also determine the degree of adsorption of the hydrolysis products.

In practice, a cutoff value of log K_{ow} = 3 can be applied for adsorption potential. However, caution should be exercised in using this criterion, as substances that are water soluble and have a low octanol-water partition coefficient do not necessarily always have a low adsorption potential. A *measured* adsorption coefficient is usually needed for ionising substances, since it is important to have information on pH-dependence (cationic substances in particular generally adsorb strongly). Similarly, measured values will normally be needed for surface active substances (e.g. surfactants), because K_{ow} values (predicted or measured) are likely to be poor predictors of adsorption for these types of substance. For ionisable substances, partition coefficients should also be corrected according to the pH of the environment being assessed (see Annex 2). For complex mixtures (e.g. UVCBs), a single value of K_{oc} will not be definitive. In such cases a range of values or a representative value can be given, depending on the substance.

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

For all organic substances manufactured or supplied in quantities of 10 tonnes per year or more, the K_{oc} should be estimated using read-across or QSPR methods as a first step. If the property is likely to be a significant determinant in the calculation of risk (e.g. following a sensitivity analysis), then a test should be conducted to provide a more reliable value for substances manufactured or supplied in quantities of 100 t/y or more. In general, confirmatory testing would not be expected for non-ionising substances with a log K_{ow} value below 3, or for substances that degrade rapidly (in which case the degradation products may be more relevant). The HPLC method may be used as a first step in testing, with the batch equilibrium method being considered only if more definitive data become necessary for the Chemical Safety Assessment. Column leaching studies might be an option under some circumstances (e.g. for unstable test substances that degrade significantly during the equilibrium time of shake flask sorption studies).

If estimation methods are not appropriate (e.g. because the substance is a surfactant or ionisable at environmentally-relevant pH), then a batch equilibrium test may need to be considered at the 10 tonnes per year band, and would be essential at the 100 tonnes per year band.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

Soil sorption (K_{oc}) of organic non-ionic substances can often be estimated from their octanol-water partition coefficient (K_{ow}), as well as from other properties such as aqueous solubility.

Such methods, including QSPR, are useful in the first instance to indicate the qualitative/quantitative adsorption coefficient of a substance. In some instances an estimated value may be sufficient for this endpoint. In all such cases the estimated method must be proven to be valid for the type of substance considered (see the general guidance for use and applicability of QSPR), and if possible a sensitivity analysis should be conducted with values generated from different models. Using a range of values in the CSA will help to highlight if the adsorption coefficient is an important factor for environmental behaviour of the substance. In general an estimated value will be sufficient if it is indicated that the adsorption coefficient will not affect the CSA, i.e. no risk is identified for the sediment/soil compartments. Estimated values are essential for substances for which experimental measurement is not feasible i.e. for *difficult substances*. Estimated values are also useful for comparing screening tests [e.g. HPLC method (OECD 121; EC C19)]. A number of reviews of K_{oc} prediction have been published recently (Lyman 1990, Reinhard & Drefahl 1999, Doucette 2000, Delle Site 2001, Doucette 2003, Dearden 2004). That of Doucette (2000) contains a number of worked examples of the estimation of log K_{oc} values. Additional information on the K_{oc} can be found in Gerstl (1990), Briggs (1981) and Nendza (1998).

Grouping of substances and read-across approach

Read-across and/or QSPR prediction for K_{oc} are important predictive tools and should be the first method used to predict Koc if reliable measured data do not exist and the model is valid for the substance. However if these options do not give meaningful and valid information or if K_{oc} is an important factor in the CSA (i.e. risks are indicated for sediment/soil compartments based on a predicted value and log $K_{ow} > 3$), then an experimental value should be measured.

Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies. In general, partition coefficients that are measured with a suitable standard method are preferred (and they are usually essential for surfactants and ionic substances that dissociate at environmentally relevant pH).

Further adaptation possibilities

Not foreseen. The K_{oc} is not directly relevant for environmental classification or the PBT assessment. However, it is a key property for exposure assessment so the information requirement should not be waived.

R.7.1.15.5 Impurities;uncertainties

Impurities can have an impact on the measurement of adsorption/desorption. Expert judgement should be used when considering whether impurities may affect the determination of the adsorption/desorption. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

R.7.1.15.6 Endpoint specific information in the registration dossier/ in IUCLID

HPLC method (OECD TG 121, EU C.19)

Materials and methods

- description of the HPLC equipment and operating conditions (column, mobile phase, means of detection, temperature);
- dead time and method used for its determination;

- reference substances (identity, purity, Koc, retention times) with results of at least 6 measurements with at least one of them above and one below the expected value for the test substance;
- quantities of test and reference substances introduced in the column.

Results and discussion

- average retention data and estimated d log Koc value for test compound;
- all values of log Koc derived from individual measurements.

Batch equilibrium method (OECD TG 106, EU C.18)

Materials and methods

- details on soil types (nature and sampling site(s), organic C, clay content and soil texture, and pH, if relevant Cation Exchange Capacity);
- information on the test substance (nominal and analytical test concentrations, stability and adsorption on the surface of the test vessel, solubilising agent if relevant (and justification for its use), radiochemical purity if relevant);
- details on test conditions (e.g. soil/solution ratio, number of replicates and controls, sterility, test temperature, and pH of the aqueous phase before and after contact with the soil);
- details on sampling (e.g. frequency, method);
- details on the analytical methods used for determination of the substance (detection limit, recovery %).

Results and discussion

- soil dry mass, total volume of aqueous phase, concentration of test substance in solution and/or soil after agitation and centrifugation, equilibration time, Koc, if appropriate mass balance;
- explanations of corrections made in the calculations, if relevant (e.g. blank run).

Leaching in soil columns (OECD TG 312)

Materials and methods

- details on soil types (nature and sampling site(s), organic C, clay content and soil texture, Cation Exchange Capacity, bulk density (for disturbed soil), water holding capacity and pH;
- information on the test substance (amount of test substance and, if appropriate, reference substance applied, solubilising agent if relevant (and justification for its use), radiochemical purity if relevant);
- details on test conditions (number of replicates and controls, test temperature, amount, frequency and duration of application of artificial rain);
- details on the analytical methods used for determination of the substance (detection limit, recovery %);
- reference substance used.

Results and discussion

- Koc, tables of results expressed as concentrations and as % of applied dose for soil segments and leachates;
- mass balance, if appropriate;
- leachate volumes;
- leaching distances and, where appropriate, relative mobility factors.

Adsorption control within an inherent biodegradability test (OECD TG 302B)

Materials and methods

- details on inoculum;
- information on the test substance (toxicity to bacteria, test concentration);
- details on test conditions (blank controls used, inoculum and test compound ratio (as DOC));
- details on sampling (frequency);
- details on the analytical methods used for determination of the DOC or COD;
- reference substance.

Results and discussion

- estimate of the extent of adsorption to STP sludge made from the elimination level in this Zahn-Wellens inherent biodegradation test, based on the 3-hour value if possible;
- values beyond 24 hours should not normally be used but where data is not available for adsorption up to 24 hours, data from time scales beyond this can only be used if adsorption is the only removal mechanism, with an upper limit of 7 days;
- if relevant results of testing of inhibition of biodegradation.

Simulation test/field measurement (OECD TG 22)

Materials and methods

- details on soil types (nature and sampling site(s); if relevant: organic C, clay content and soil texture, Cation Exchange Capacity and pH;
- details on lysimeter;
- information on the test substance (nominal and analytical test concentrations, solubilising agent if relevant (and justification for its use), radiochemical purity if relevant);
- details on test climate conditions (e.g. air temperature, solar radiation, humidity, potential evaporation or rate of artificial rainfall), soil temperature and soil moisture and duration of the study;
- details on sampling (frequency, method);
- details on the analytical methods used for determination of the test substance (detection limit, recovery %).

Results and discussion

- concentration of test substance in soil layers; Koc, if appropriate mass balance and concentrations and as % of applied dose for soil segments and leachates;
- explanations of corrections made in the calculations, if relevant (e.g. blank run).

Distribution modelling

Materials and methods

- model name and version;
- date of the model development;
- model type description e.g. steady-state, dynamic, fugacity, Gaussian, Level I-IV, etc.;
- environmental compartments which the model covers;
- information on model segmentation and environmental properties;
- input parameters (minimum information required for assessing the partitioning and degradation behaviour):

    - o vapour pressure;

- water solubility;
- molecular weight;
- octanol-water partition coefficient;
- information on ready biodegradability;
- for inorganic substances: it is recommended to have information on the partition coefficients and possible abiotic transformation products;

- temperature effect.

Results and discussion

- key exposure routes and distribution of the substance among them.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on adsorption/desorption can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
5.4.1	VIII 9.3.1	Adsorption / desorption	E.5.5.2	4.1.4
5.4.2	/	Henry's Law constant	E.5.5.3	4.1.4
5.4.3	X 9.3.4	Distribution modelling	E.5.5.4	4.1.4
5.4.4	X 9.3.4	Other distribution data	E.5.5.5	4.1.4

R.7.1.15.7 Further information/references

Briggs G.G. (1981) Theoretical and experimental relationships between soil adsorption, octanol-water partition coefficients, water solubilities, bioconcentration factors and the parachor. *J. Agric. Food Chem.* 29, 1050-1059.

Dearden J.C. (2004) QSAR modelling of soil sorption. In Cronin M.T.D. and Livingstone D.J. (Eds.), *Predicting Chemical Toxicity and Fate*, CRC Press, Boca Raton, FL, pp. 357-371.

Delle Site, A., (2001) Factors affecting sorption of organic compounds in natural sorbent/water systems and sorption coefficients for selected pollutants. *J. Phys. Chem. Ref. Data* 30, 187-439.

Doucette W.J. (2000) Soil and sediment sorption coefficients. In Boethling R.S. and Mackay D. (Eds.), *Handbook of Property Estimation Methods for Chemicals: Environmental and Health Sciences.* Lewis, Boca Raton, FL, pp. 141-188.

Doucette W.J. (2003) Quantitative structure-activity relationships for predicting soil/sediment sorption coefficients for organic chemicals. *Environ. Toxicol. Chem.* 22, 1771-1788

ECETOC (1998). Technical Report No. 74: QSARs in the Assessment of the Environmental Fate and Effects of Chemicals. European Centre for Ecotoxicology and Toxicology of Chemicals. Brussels.

EU C.18 Adsorption – desorption using a batch equilibrium method.

EU C.19 Estimation of the adsorption co-efficient (K_{oc}) on soil and on sewage sludge using high performance liquid chromatography (HPLC).

Gerstl Z. Estimation of organic chemical sorption by soils. *J. Contaminant Hydrology* (1990) 6, 357-375.

ISO (2004). *Water quality: adsorption of substances on activated sludge – batch test using specific analytical methods.* International Standard ISO 18749. First edition February 2004.

Lyman W.J. Adsorption coefficient for soils and sediments. In Lyman W.J., Reehl W.F. and Rosenblatt D.H. (Eds.), *Handbook of Chemical Property Estimation Methods*, American Chemical Society, Washington DC, 1990, pp. 4.1-4.33.

Mueller, M. and Kordell, W. (1996). Comparison of screening methods for the estimation of adsorption coefficients on soil. *Chemosphere* 32(12), 2493-2504.

Nendza M. *Structure-Activity Relationships in Environmental Sciences.* Chapman & Hall, London, 1998.

OECD (2000a). Adsorption – desorption using a batch equilibrium method. Organisation for Economic Co-operation and Development (OECD) Guideline for the testing of chemicals 106.

OECD (2000b) Guidance Document No. 22: Performance of Outdoor Monolith Lysimeter Studies. Organisation for Economic Co-operation and Development (OECD), Paris.

OECD (2001). Estimation of the adsorption co-efficient (K_{oc}) on soil and on sewage sludge using high performance liquid chromatography (HPLC). Organisation for Economic Co-operation and Development (OECD) Guideline for the testing of chemicals 121.

OECD (2002) Guidelines for Testing of Chemicals (Draft): Leaching in Soil Columns. Organisation for Economic Co-operation and Development (OECD), Paris.

OPPTS (1996). Sediment and soil adsorption/desorption isotherm. United States Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances Fate, Transport and Transformation Test Guideline 835.1220. Draft of April 1996.

Poole S.K. and Poole C.F. (1999) Chromatographic models for the sorption of neutral organic compounds by soil from air and water. *J. Chromatogr. A* 845, 381-400.

Reinhard M. and Drefahl A. (1999). *Handbook for Estimating Physico-chemical Properties of Organic Compounds.* Wiley, New York.

SETAC (1993). Guidance Document on Sediment Toxicity Tests and Bioassays for Freshwater and Marine Environments. In *Workshop on Sediment Toxicity Assessment at Renesse,* *Netherlands on 8-10 November 1993*. Hill I, Mathiessen P, Heimbach F (eds). Society of Environmental Toxicology and Chemistry – Europe, Brussels.

R.7.1.16 Stability in organic solvents and identity of relevant degradation products

R.7.1.16.1 Type of property

The stability in organic solvents is required for substances manufactured or imported in quantities of ≥ 100 t/a only if their stability in organic solvent is considered critical (REACH Annex IX, section 7.15).

There are rare occasions when it is important to have information on the stability of a compound in an organic solvent, to ensure confidence in the test results. However, for many substances, the stability in organic solvents will not be critical and testing need not be conducted.

Examples of when stability in organic solvents could be important are:

- for certain solubility measurements (e.g. octanol–water partition coefficient);
- to check on the stability of reagent solutions, fortification standards or calibration standards;
- when a test substance is dosed as a solution in an organic solvent (e.g. ecotoxicity studies);
- when a test substance is extracted from an environmental sample, plant or animal tissue or diet matrix (arising from a variety of physicochemical property, ecotoxicity and animal toxicity studies) into an organic solvent and stored pending analytical measurement.

R.7.1.16.2 Definition

A study of the stability of a test compound in an organic solvent is normally undertaken for a specific time period to confirm whether the test compound is stable under these conditions for the duration of the storage of the organic solvent or extract containing the test substance. Often several time periods are selected to check whether there is any particular downward trend in stability over time.

The stability of the test substance at a particular time period during the study is normally expressed as a percentage of the concentration of the test substance in the solvent extract, at that time period compared with the initial starting concentration of the test substance at t = 0, namely:

$$\frac{C_t}{C_0} \times 100\ \%$$

where C_t is the concentration of test substance in solvent extract at $t = t_1, t_2, t_3..., t_n$; and C_0 is the concentration of test substance in solvent extract at t = 0.

R.7.1.16.3 Test method(s)

A number of physical, chemical and biological processes can result in a decline in the actual concentration of a test substance in an organic solvent over time. Information on the stability of a test substance in a solvent is desirable, particularly when samples are to be stored. However, there does not appear to be any generally accepted methodology for performing such stability studies. Factors affecting the rate of degradation include rates of hydrolysis, of photolysis and of oxidation.

Typically, one or more concentrations of the test substance in the solvent are made up and analysed immediately after preparation (i.e. $t = 0$). They are then stored in appropriate vessels under the required test conditions (e.g. temperature, absence of light) and analysed, along with a freshly prepared solution of the test substance at the original test concentration(s), at regular intervals during the period of interest.

At each time of analysis, a sample is withdrawn from storage and mixed thoroughly before taking any aliquot for analysis. The analysis is carried out using the recommended method to determine whether any significant loss of the test substance has occurred during storage. It is important to analyse freshly made standards of the test substance in the organic solvent at the same time as analysing stored samples, so that any losses that may occur of the test substance during sampling, sample treatment and analysis are taken into consideration.

It is important to be able to have a check on the temperature to ensure that the temperature regime has been maintained throughout the period of the stability study.

Unlabelled reference material of suitable known purity may be used where a reliable method of analysis is available. Where an analytical method is still under development or is unlikely to be sufficiently sensitive, radio-labelled compounds should be used if available. Use of radio-labelled compounds can shorten the analysis time and help facilitate identification of any degradation products, should the test substance not be stable in the organic solvent.

Recovery or spiking experiments should normally be run. The number of spiking levels or the range of concentrations tested within a project should be left to the judgement of the analyst.

Further information should be obtained by checks on the stability of standards of the test substance in organic solvents as part of routine analytical protocols, to confirm whether the test substance is unstable under normal storage conditions.

Further tests may be necessary to identify storage conditions which minimise any degradation of the test substance not only in organic solvents, but also during the conducting of other tests, such as water solubility, surface tension and in the preparation of test media for ecotoxicity studies (OECD, 2000). Identification of the degradation product(s) will allow an assessment of whether they are likely to be more toxic than the parent material in subsequent ecotoxicity studies.

R.7.1.16.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex IX to REACH

Column 2 of REACH Annex IX provides the following specific rules for adaptation of the standard information requirement for stability in organic solvents and identity of relevant degradation products:

‘The study does not need to be conducted if the substance is inorganic.’

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

Stability data of substances in organic solvents are not normally reported in standard published sources of physicochemical data. Relevant sources of basic information regarding stability and storage conditions of substances are the Hazardous Substances Data Base (HSDB) and Sax's 'Dangerous Properties of Industrial Materials'.

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

At present (Q)SAR is generally not applicable for determination of stability in organic solvent and degradation products. Application of (Q)SAR is not possible.

Grouping of substances and read-across approach

At present grouping and read across are not applicable.

Testing is technically not possible

Testing should always be considered, if none of the waiving possibilities applies.

Further adaptation possibilities

Not foreseen.

R.7.1.16.5 Impurities; uncertainties

Impurities can have an impact on the measurement of stability in organic solvent and degradation products. Expert judgement should be used when considering whether impurities may affect the determination of the stability in organic solvent and degradation products. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

R.7.1.16.6 Endpoint specific information in the registration dossier / in IUCLID

This endpoint needs to be fulfilled on a case by case basis. As several different methods can be used to document this intrinsic property, we recommend the same strategy for drafting robust study summaries as described for the other endpoints. The general aspects described in section 2 of the Practical guide 3: How to report robust study summaries should also be applied for this endpoint. All endpoint specific characteristics should be described in such a way that the robust study summary allows an independent assessment of the endpoints reliability and completeness. The objectives, methods, results and conclusions of the full study report should be reported in a transparent manner as described for all other endpoints in this guidance.

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on stability in organic solvents can be found in:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.17	IX 7.15	Stability in organic solvents and identity of relevant degradation products	E.4.18	3.15

R.7.1.16.7 Further information / references

OECD Series on Testing and Assessment Number 23 Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures, ENV/JM/MONO(2000)6 (http://www.oecd.org/officialdocuments/displaydocumentpdf?cote=env/jm/mono(2000)6&doclanguage=en).

R.7.1.17 Dissociation constant

R.7.1.17.1 Type of property

Information on the dissociation constant is **supplementary data** for hazard assessment (OECD TG 112, 1981). The dissociation of a substance in water is of importance in assessing its impact upon the environment and may also influence the ADME of a substance and consequently its effects on human health. It governs the form of the substance which in turn determines its behaviour and transport. It may affect the adsorption of the substance on soils and sediments and absorption into biological cells.

The dissociation constant may also be an important factor in deciding which method or conditions should be used to determine the octanol-water partition coefficient (K_{ow}) and soil adsorption partition coefficient (K_{oc}). Slight changes in pH can considerably affect the form in which the substance is present in solution, especially if the pK_a value is within the environmentally-relevant pH range[21]. The dissociated and non-dissociated species may have significantly different water solubilities and partition coefficients. Therefore, significantly different bioavailability and toxicity may result. It is important to note that the dissolution of

[21] Fresh surface waters have pH values in the range 4-9, whereas marine environments have a stable pH of about 8. pH normally varies between 5.5 and 7.5 for agricultural soils and sewage treatment plant tanks.

salts from their crystal lattice into individual ions is not intended to be covered by the endpoint dissociation constant. Therefore this section refers only to acid dissociation (pK_a).

R.7.1.17.2 Definition

Dissociation is the reversible splitting of a substance into two or more chemical species, which may be ionic (OECD TG 112, 1981). The process can be represented as:

$$RX \leftrightarrow R^+ + X^-$$

The dissociation constant (K) for this process is expressed as the ratio of concentrations of the species on either side of the equation in water at equilibrium:

$$K = \frac{[R^+][X^-]}{[RX]}$$

Where the cation R+ is hydrogen, the substance can be considered an acid, and so this constant becomes an acid dissociation constant (K_a).

$$K_a = \frac{[H^+][X^-]}{[HX]}$$

A substance can have more than one acidic (or basic[22]) group, and the dissociation constant can be derived for each dissociation step in a similar way.

The K_a is related to pH as follows (where p is –log10):

$$pK_a = pH - \log_{10}\left(\frac{[X^-]}{[HX]}\right)$$

In practice for a simple substance having one dissociating group, the pK_a is equivalent to the pH at which the ionised and non-ionised forms are present in equal concentration (i.e. the substance has undergone 50% dissociation).

It is important to differentiate between dissociation and hydrolysis as hydrolysis is a separate standard information requirement according to Annex VIII of the REACH regulation. Hydrolysis is defined as reaction of a substance RX with water, with the net exchange of the group X with OH at the reaction centre (OECD TG 111, 2004).

$$RX + H_2O \rightarrow ROH + HX$$

22 Base strength is expressed as the acidity of the conjugate acid. The term pK_b was once used to express basicity so that the same scale could be used alongside acidity – care should be taken when citing older sources to check which term has been used. For consistency, dissociation of bases should preferably be expressed using the pK_a of the conjugate acid.

R.7.1.17.3 Test method(s)

OECD test guideline 112 (Dissociation constants in water, adopted May 1981) describes three laboratory methods to determine the pK_a of a substance. The three methods are appropriate for particular types of substances as described in the test guideline[23].

R.7.1.17.4 Adaptation of the standard testing regime

Adaptation possibilities according to column 2 of Annex IX to REACH

Column 2 of REACH Annex IX provides the following specific rules for adaptation of the standard information requirement for dissociation constant:

'A study does not need to be conducted if:

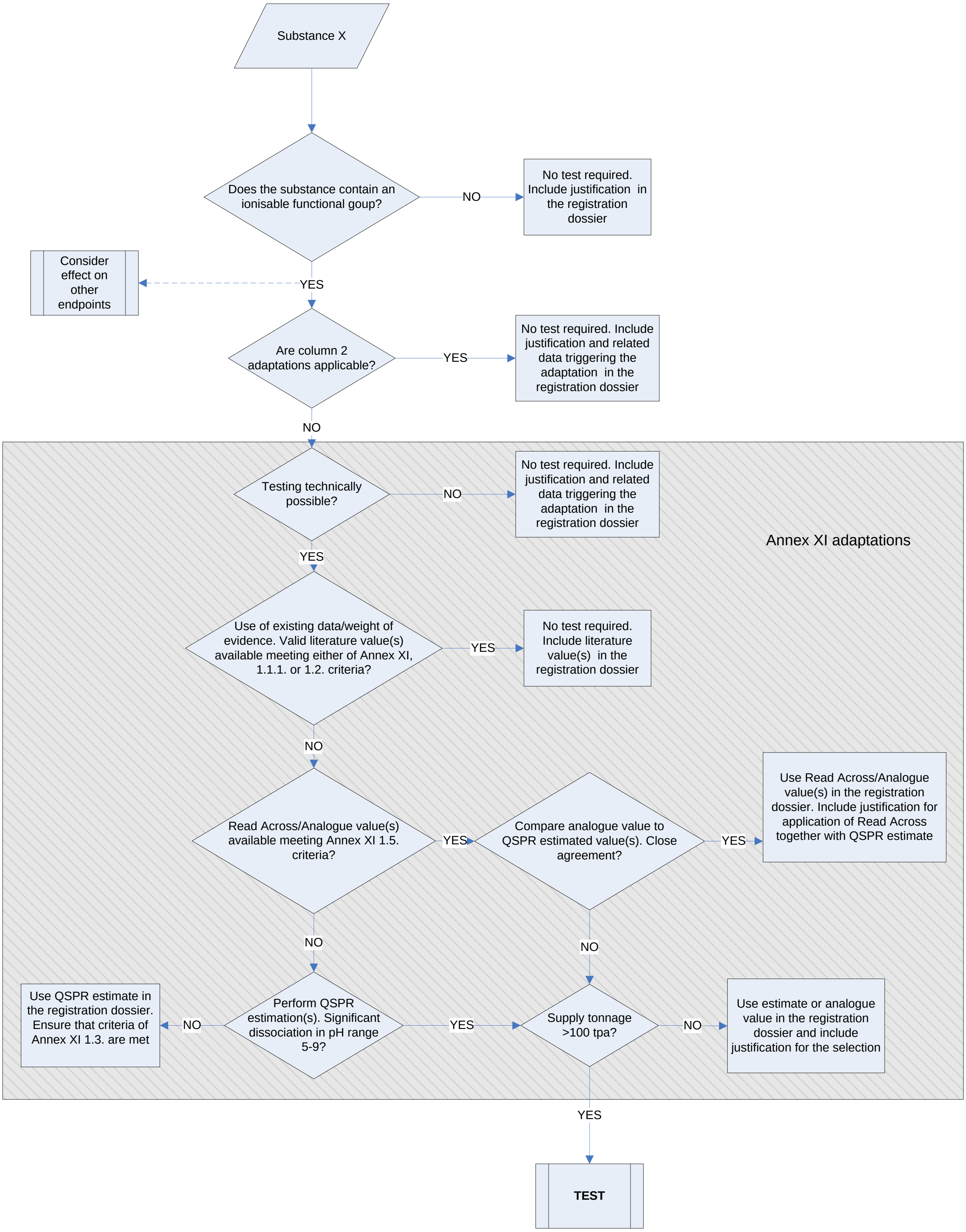
- *the substance is hydrolytically unstable (half-life less than 12 hours) or is readily oxidisable in water; or*
- *it is scientifically not possible to perform the test (e.g. because the analytical method is not sensitive enough).'*

In all cases where the above specific rules for adaptation are used to waive testing, evidence demonstrating the existence of that property of the substance which triggers the adaptation rule should be provided in the IUCLID dossier, e.g. if the test is not performed because the substance is hydrolytically unstable (half life < 12 hours) then the dossier must contain valid data on the hydrolysis clearly indicating a half life < 12 hours.

It is important to note that OECD TG 112 allows the use of a small amount of a water-miscible solvent to aid dissolution of sparingly soluble substances. Therefore low solubility will only prevent performance of the test in the context of the column 2 rules above for substances which remain highly insoluble and undetectable by analytical techniques in the presence of water miscible solvents.

23 The test method is available at the following link: http://www.oecd-ilibrary.org/environment/test-no-112-dissociation-constants-in-water_9789264069725-en

Figure R.7.1–7 Integrated testing strategy for dissociation constant



Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

Many literature sources for dissociation constant exist; some reference textbooks and on-line sources are listed in Section R.7.1.17.7. These should be searched for published, valid data. As mentioned in section 1.1.1 of Annex XI to REACH a number of conditions need to be met before any such data can be used. Namely:

- *'adequacy for the purpose of classification and/or risk assessment;*
- *sufficient documentation is provided to assess the adequacy of the study; and*
- *the data are valid for the endpoint being investigated and the study is performed using an acceptable level of quality assurance.'*

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used provided that data from a number of distinct sources indicate a similar value for the dissociation constant which is supported by one or more relevant QSPR predictions.The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

Estimated pK_a data can be generated by valid QSPR methods. In general, pK_a values that are measured with a suitable method are preferred to QSPR predictions. If an estimated pK_a value suggests that the substance will dissociate significantly at environmentally relevant pH, a test may be required to confirm the result.

There have been a few attempts to model pK_a values of diverse sets of substances. Klopman and Fercu (1994) used their MCASE methodology to model the pK_a values of a set of 2464 organic acids, and obtained good predictions; a test set of about 600 organic acids yielded a standard error of 0.5 pK_a unit. Klamt *et al.* (2003) employed their COSMO-RS methodology to predict pK_a values of 64 organic and inorganic acids, with a standard error of 0.49 pK_a unit. A comparison of commercially available software for the prediction of pK_a was done by Dearden *et al.* (2007).

Grouping of substances and read-across approach

For most ionisable substances supplied at greater than 100 t/y that are predicted to dissociate at environmentally relevant pHs, a test will typically be required for dissociation constant. Similar substances (analogues) for which measured pK_a data according to a reliable method are available may be considered for read-across. Such values should be reinforced by estimated methods for pK_a (e.g. the result of a QSPR prediction; see section above). In some instances it may be acceptable to read-across dissociation constant from an analogue. However if there is significant variation between the analogue read-across and the predicted pK_a then a test should be conducted.

Testing is technically not possible

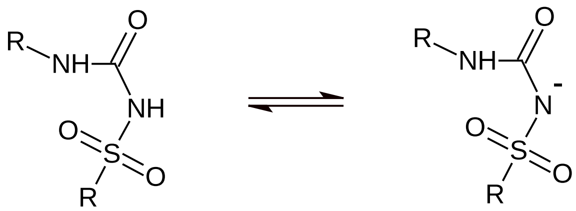
Testing should always be considered if none of the waiving possibilities applies. Instances where testing is technically not possible as a consequence of the properties of the substance are expected to be limited to highly reactive or unstable substances, and substances which in contact with water emit flammable gases.

Further adaptation possibilities

As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

No dissociating groups

If the substance cannot dissociate due to a lack of relevant functional groups, the dissociation constant is irrelevant and testing information does not need to be provided. However, ionisable groups might not always be obvious (e.g. in sulphonyl urea herbicides, which contain the function -$S(=O)_2NH.C(=O)NH$-, the acid group is $S(=O)_2NH$).



If a substance is much more soluble in water than expected, this may be an indication that dissociation has occurred.

UVCBs

For complex mixtures (e.g. UVCBs) containing ionisable components the assessment of pK_a is clearly complicated. Estimation of the representative constituent's pK_a values, if appropriate, should be considered.

R.7.1.17.5 Impurities; uncertainties

Impurities can have an impact on the measurement of dissociation constant. Expert judgement should be used when considering whether impurities may affect the determination of the dissociation constant. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

The presence of multiple dissociation/equilibrium reactions can complicate determination of the dissociation constant(s). In cases where multiple dissociation reactions can take place due to the presence of numerous dissociating groups and/or the presence of tautomerisation and/or zwitterionic forms, care should be taken in the interpretation of experimental results. QSPR predictions for such substances should also be carefully analysed as the models may not account for concurrent equilibria/dissociations. Additionally QSPR predictions may not account for intramolecular hydrogen bonding effects which can have a pronounced effect on the observed dissociation constant. In some cases, formation of intramolecular hydrogen bonding depends on the cis/trans isomerism of the substance, as is the case for the isomers fumaric and maleic acid. Care should be taken when using QSPR predictions for such molecules, as cis/trans isomerism is typically not taken into account.

The extent of ionisation may vary according to pH, ionic strength and/or the level of common ions in the test medium (common ion effect), and relatively small changes may significantly alter the equilibrium between dissociated and non-dissociated species.

R.7.1.17.6 Endpoint specific information in the registration dossier / in IUCLID

Knowledge of an ionisable substance's pK_a is important for all such substances. For substances supplied at levels below 100 tonnes per annum dissociation constant is not a testing requirement. Ideally however, a literature value, analogue value and/or QSPR prediction can be obtained and provided for such substances, especially if dissociation is relevant for interpreting the results of other physicochemical or fate and (eco)toxicological tests and for chemical safety assessment. For ionisable substances supplied at tonnages greater than 100 t/y, dissociation constant is a standard information requirement.

For substances which contain multiple ionisable functionalities, all measured macro pK_a values should be reported and preferably assigned to specific micro-reactions.

With regard to study summaries of experimental data, the IUCLID dossier should contain all relevant information regarding the endpoint and as a minimum the items listed below:

Materials and methods

- type of method;
- test guideline followed.

Test Materials

- test material identity.

Results and discussion

- concentration of the substance;
- test results as pK_a-value(s);
- temperature of the test medium (ºC);
- if testing is waived, the reasons for waiving must be documented in the dossier.

Any deviation from the guideline method used (and reasons for it) or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on dissociation constant can be found in:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.21	IX 7.16	Dissociation constant	E.4.22	3.16

R.7.1.17.7 References on dissociation constant

Balogh G.T., Gyarmati B., Nagy B., Molnar L. and Keseru G.M. Comparative evaluation of in silico pK_a prediction tools on the Gold Standard dataset. *QSAR Comb Sci* (2009) 28:1148-1155.

Dearden J.C., Cronin M.T.D., and Lappin D.C. A comparison of commercially available software for the prediction of pK_a. *J. Pharm. Pharmacol.* (2007) 59, Suppl. 1, A-7.

Klamt A., Eckert F., Diedenhofen M. and Beck M.E. (2003) First principles calculations of aqueous pK(a) values for organic and inorganic acids using COSMO-RS reveal an inconsistency in the slope of the pK(a) scale. *J. Phys. Chem. A* 107, 9380-9386.

Klopman G. and Fercu D. (1994) Application of the multiple computer automated structure evaluation methodology to a quantitative structure-activity relationship study of acidity. *J. Comput. Chem.* (1994) 15, 1041-1050.

Liao C. and Nicklaus M.C. Comparison of nine programs predicting pK_a values of pharmaceutical substances. *J. Chem. Inf. Model.* (2009) 49, 2801-2812.

Manchester J, Walkup G, Rivin O. and You Z.P. Evaluation of pK_a estimation methods on 211 druglike compounds. *J Chem Inf Model* (2010) 50, 565-571.

Meloun M. and Bordovská S. Benchmarking and validating algorithms that estimate pK_a values of drugs based on their molecular structure. *Anal. Bioanal. Chem.* (2007) 389, 1267-1281.

OECD (1981). Dissociation constants in water (titration method – spectrophotometric method – conductometric method). Organisation for Economic Co-operation and Development (OECD) Guideline for the testing of chemicals no 112.

R.7.1.18 Viscosity

R.7.1.18.1 Type of property

Viscosity is a property:

- needed for substance characterization;
- needed for the classification of aspiration hazard of liquids;
- which gives an indication of the penetration of the substance within soil.

R.7.1.18.2 Definition

Viscosity: viscosity is the (inner) resistance of a substance (gas, liquid) to a shift caused by laminar flow.

Dynamic viscosity (= dynamic viscosity coefficient) η:

Quantifies the property 'viscosity' by the quotient shear stress τ / shear rate $\dot{\gamma}$ $(\eta = \tau / \dot{\gamma})$

Kinematic viscosity (= kinematic viscosity coefficient) ν:

is given by the quotient dynamic viscosity to density $(\nu = \eta / \rho)$.

R.7.1.18.3 Test method(s)

Five different types of test methods are standardized for liquid substances:

- capillary viscometer;
- flow cup;
- rotational viscometer;
- rolling ball viscometer;
- drawn-shear viscometer.

There exist a lot of standardized determination methods with sometimes very specialised application ranges with respect to products, especially mixtures. For substances (within the scope of the REACH Regulation) the following standardised determination methods are recommended:

- Capillary viscometer: EN ISO 3104, EN ISO 3105, DIN 51562, BS 188, NF 60-100, ASTM D445, ASTM D4486;
- Flowcup: EN ISO 2431;
- Rotational viscometer: EN ISO 3219, DIN 53019;
- Rolling ball viscometer: DIN 53015.

For newtonian liquids (liquids for which the viscosity is independent of the shear stress and shear rate) any determination method may be used within the scope and applicability specifications. For non-newtonian liquids (liquids for which the viscosity depends on the shear rate) only the use of rotational viscometers is possible. Because the viscosity is remarkably temperature dependent each determination must be accompanied by the temperature at which the measurement was made. It is recommended to use the mean of two test runs. It is also recommended to determine the viscosity at at least two different temperatures. The classification criteria for aspiration hazard refer to kinematic viscosity at 40°C.

If explosives, pyrophorics or self-reactives are to be characterized, determination of the viscosity may not be practicable. For pyrophorics and self-reactives testing under inert gas should be considered. In any case the determination method has to be chosen carefully.

The use of the most recent update of the standard is advised; they are accessible *via* numerous websites, see R.7.1.1.3.

R.7.1.18.4 Adaptation of the standart testing regime

Within the REACH Regulation requirements testing of viscosity is only of interest for liquid substances.

Adaptation possibilities according to column 2 of Annex IX to REACH

Column 2 of REACH Annex IX does not provide any specific rules for adaptation from column 1.

Adaptation possibilities according to Annex XI to REACH

Use of existing data: Data on physical-chemical properties from experiments not carried out according to GLP or the test methods referred to in Article 13 (3) of REACH

If experimental data are available (study reports or literature data) meeting the criteria in section 1.1.1 of Annex XI to REACH, these could be used to meet the endpoint data requirements. If an estimation method is used as a source of information according to Column 2 of Annex VII, the QSAR model must meet the criteria set out in section 1.3 of Annex XI to REACH.

Weight of evidence

Where no single source of existing data (study reports, QSAR, literature data) is considered sufficiently reliable, thus not fully meeting the criteria in section 1.1.1 of Annex XI to REACH, or where several sources of similar reliability with deviating results exist, a weight of evidence approach may be used. The criteria in section 1.2 of Annex XI to REACH must then be met.

(Q)SAR

For the determination of the viscosity, (Q)SAR approaches are discouraged for the purpose of classification / risk assessment, except when the mean absolute error of the (Q)SAR is less than 5 %.

Grouping of substances and read-across approach

For the determination of the viscosity read across is not possible.

Testing is technically not possible

Testing should always be considered if none of the waiving possibilities applies. But the testing is technically not possible:

- if the substance is a solid;

- if liquid explosives, pyrophorics or self-reactives are to be characterized, determination of the viscosity may not be practicable (see above section Test method(s)).

Further adaptation possibilities

- the viscosity does not have to be determined experimentally if conclusive and consistent literature data are available;

- data for viscosity generated with the same tests and classification principles as specified in the CLP Regulation generated in conjunction with transport classification can satisfy the REACH requirements on a case-by-case basis.

As stated in Annex IX of REACH, when for certain endpoints, it is proposed to not provide information for other reasons than those mentioned in column 2 of that Annex or in Annex XI of REACH, this fact and the reasons must also be clearly stated. Such an approach may then be used.

R.7.1.18.5 Impurities; uncertainties

The influence of impurities is negligible if their concentration is below 1 %. The influence of higher concentrations may be significant. There exists no generalised tendency of the influence on the viscosity. Therefore utmost care should be taken in the selection of the key study(ies), or weight-of-evidence approaches, that the data selected is representative of the substance being registered by the respective companies.

R.7.1.18.6 Endpoint specific information in the registration dossier / in IUCLID

Materials and methods

- type of method;
- test guideline followed.

Results and discussion

- viscosity value and unit according to the used test method;
- preferred units are m Pa·s (for dynamic viscosity) and mm^2/s (for static viscosity) but other units are also accepted;
- each measured value should be accompanied with temperature (in °C). Usually two values are needed. Preferably one value is measured at approximately 20°C and another at an approximately 20°C higher temperature. Two determinations of viscosity should be measured for each temperature;
- for non-Newtonian liquids, the results obtained are preferably in the form of flow curves, which should be interpreted;
- individual and mean values should be provided at each temperature (from OECD Guideline 114 'Viscosity of liquids').

Any deviation from the guideline method used or any other special consideration should be reported. In cases where there is more than one source of data, the endpoint summary under results and discussion should provide a justification for the selection of the key study chapter.

Reference to other ECHA Guidance Documents

Further detailed guidance on viscosity can be found in the following chapters:

IUCLID Section	REACH Annex	Endpoint title	IUCLID 5 End User Manual Chapter	ECHA Practical Guide 3
4.22	IX 7.17	Viscosity	E.4.23	3.17

R.7.1.19 Shape

Please *check Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the *Guidance on IR&CSA*, section 2.2.3.3 Recommendations for shape.

R.7.1.20 Surface area

Please check *Appendix R7-1 Recommendations for nanomaterials applicable to: Chapter R7a Endpoint specific guidance* of the *Guidance on IR&CSA*, section 2.2.3.4 Recommendations for surface area.

R.7.1.21 Further information to be submitted for classification and labelling in hazard classes of the substance in accordance with article 10 (a) (iv) REACH

The criteria listed in the table below (Table R.7.1–15) should be provided for general registration purposes according to Article 10 (a) (iv) and section 4 of Annex VI to REACH. The assignment of hazard classes to relevant subchapters in R.7.1.21.1 to R.7.1.21.3 should therefore only be understood as a means to structure this document in accordance with Annexes VII to XI to the REACH Regulation.

Table R.7.1–15 Information to be submitted for general registration purposes according to Article 10 (a) (iv) REACH, CLP hazards classes and corresponding tests methods according to the Test Method Regulation and CLP[24]

CLP Regulation (EC) No. 1272/2008 (the no. in brackets is the respective chapter no. in Annex I to CLP)	**Corresponding test method according to the Test Method Regulation, Regulation (EC) No. 440/2008**	**Chapter in revised R.7(a) guidance**	**Information requirement according to REACH Regulation (EC) No. 1907/2006**	**Corresponding test method according to CLP Regulation**
Flammable aerosols (2.3)[25]	n.a.	R.7.1.21.1	See Article 10 (a) (iv) REACH requirements	Test methods according to 75/324/EC amended by 2008/47/EC (harmonised with **UN-MTC Section 31)**
Gases under pressure (2.5)	n.a.	R.7.1.21.2	See Article 10 (a) (iv) REACH requirements	n.a.
Corrosive to metals (2.16)	n.a.	R.7.1.21.3	See Article 10 (a) (iv) REACH requirements	**UN Test C.1 (UN-MTC Section 37.4)**

24 Please note that REACH information requirements regarding classification and labelling in accordance with Article 10(a)(iv) of the REACH Regulation are not limited to the items listed in this table. This table stresses that, while the REACH Regulation does not require the generation of information regarding the following hazard classes (Article 10(a)(vi) of the REACH Regulation, see Table R.7.1–1), any information available on these hazard classes must be included in a REACH registration dossier for a substance pursuant to Article 10(a)(iv) of the REACH Regulation.

25 The 4th ATP to the CLP Regulation amends the criteria in the CLP Annex I, Section 2.3 Flammable aerosols by changing the scope and title to Section 2.3 Aerosols.

R.7.1.21.1 Flammable aerosols

For further guidance on these please check the *Guidance on the Application of the CLP criteria*, chapter 2.3[26].

R.7.1.21.2 Gases under pressure

For further guidance please check the *Guidance on the Application of the CLP criteria* chapter 2.5.

R.7.1.21.3 Corrosive to metals

For further guidance please check the *Guidance on the Application of the CLP criteria* chapter 2.16.

[26] The 4th ATP to the CLP Regulation amends the criteria in the CLP Annex I, Section 2.3 Flammable aerosols by changing the scope and title to Section 2.3 Aerosols. Consequently the *Guidance on the Application of the CLP criteria*, Part 2: Physical hazards has been restructured to take account of the 4th ATP, which applies to substances from 1 December 2014 and to mixtures from 1 June 2015. Once the 4th ATP is applied a Guidance corrigendum will be made to delete the outdated sub-chapter 2.3.1 Flammable aerosols in the *Guidance on the Application of the CLP criteria*.

Appendix R.7.1-1 to Section R.7.1

Appendix R.7.1–1 Henry's law constant and evaporation rate

The Henry's law constant (HLC) is one of the most important factors in determining the environmental fate of chemicals. Henry's law states that the mass of gas dissolved by a given volume of solvent is proportional to the pressure of the gas with which it is in equilibrium. HLC is the ratio of the equilibrium concentration of the chemical in the gas phase (C_G) and that in the liquid phase (C_L):

$$\mathrm{HLC} = \frac{C_\mathrm{G}}{C_\mathrm{L}}$$

Therefore, HLC quantifies the partitioning of substances between the aqueous phase and the gas phase such as rivers, lakes and seas with respect to the atmosphere (gas phase). Indeed, this constant is a fundamental input for fugacity models that estimate the multimedia partitioning of chemicals (Mackay, 1991). As HLC is a ratio of two concentrations, it is without unit if both concentrations are expressed in the same unit. Some prefer to express the gas concentration in pascals and the liquid concentration in mol/m^3, thus giving the unit Pa·m^3/mol for the HLC.

For many chemicals, volatilisation can be an extremely important removal process, with half lives as low as several hours. HLCs can give qualitative indications of the importance of volatilisation. For substances with HLC values less than 0.01 Pa·m^3/mol, the substance is less volatile than water and as water evaporates the concentration of the substance in the aqueous phase will increase; for substances with HLC values around 100 Pa·m^3/mol, volatilisation will be rapid.

However, the degree of volatilisation of substances from the aquatic environment is highly dependent on the environmental parameters for the specific water bodies in question, such as the depth and the gas exchange coefficient (influenced e.g. by wind speed and water flow rate). The HLC cannot be used for evaluation of the removal of a substance from the water phase without considering these factors. As the n-octanol/water partition coefficient (K_{ow}) is used to predict bioaccumulation potential in air-breathing organisms, this aspect is especially important in a PBT context.

For example, where a substance has both a low vapour pressure and low water solubility, HLC can be relatively large if calculated using the ratio of vapour pressure and water solubility, which might imply that volatilisation is an important fate process. In practice, adsorption to dissolved organic carbon is likely to be much more relevant, and volatilisation will be lower than the HLC value suggests.

Experimental determination of Henry's law constant

The experimental approaches can be classified into two major groups: dynamic equilibration approach (often referred to as the *gas purge* approach) and the static equilibration approach. The following table (Table R.7.1–16) briefly summarises the reviewing work done by Staudinger and Roberts (1996).

Table R.7.1–16 Experimental approaches for the determination of HLC

Approach	Average Relative Standard Deviations (RSDs) / Notes
Dynamic approach	
Batch air stripping (bubble column) Henry's law constant (HLC) values are determined by measuring the rate of loss of the substance of interest from water by isothermally stripping with a gas (typically air) in a suitable bubble column apparatus.	Average RSDs determined from different literature sources ranged from 2.8 to 21
Concurrent flow (wetted wall column) Values are determined based on the use of a wetted wall (desorption) column. The wetted wall column equilibrates an organic solute between a thin film of water and a concurrent flow of gas. Substance-laden water is introduced into the wetted wall column where it comes in contact with a substance-free gas stream flowing concurrently. HLC: The knowledge of flow rates and compound masses present in the separated phase streams enables the direct calculation of HLC.	Average RSDs determined from different literature sources ranged from 19 to 52 Preliminary work must be performed to ensure that phase equilibrium is reached.
Static approach	
Single equilibration A known mass of a substance is introduced into an air-tight vessel with a known volume of water and air. When the equilibrium is attained the substance concentration is determined in one or both phases.	Average RSDs determined from different literature sources ranged from 2.8 to 30
Multiple Equilibration A liquid sample containing a known quantity of solute is allowed to equilibrate with a known volume of solute-free air. The air is the expelled and a new equilibration with the same amount of solute-free air is started. This process can be repeated until the number of equilibrations exhausts the mass of solute remaining in the system.	RSDs ranged from 0.7 to 3.5 This method is applicable for substances with $0.1 \leq$ HLC ≤ 2 The experimental error is reduced with a larger number of equilibrations.
EPICS Technique HLC is determined by measuring the gas headspace concentration ratios from pairs of sealed bottles. Relative rather than absolute air-phase concentrations are required.	Average RSDs determined from different literature sources ranged from 2.9 to 19
Variable Headspace The method is based upon the measurement of the relative equilibrium air-phase concentration (gas chromatography peak areas) from aliquots of the same solution in multiple containers having different headspace-to-liquid volume ratios.	Average RSDs determined from different literature sources ranged from 0.5 to 7.9

A data-analysis of reviewed experimental studies for HLC can be found in Staudinger and Roberts (1996). HLC values can also be found in one or more of the following references: Sander (1999), CRC Handbook of Chemistry and Physics (2000), the NIST Chemistry WebBook

(1998), and 'The Handbook of Environmental Data on Organic Chemicals' (Verschueren K, 2001).

Main factors affecting Henry's Law Constant values

Staudinger and Roberts (1996) thoroughly explain all the factors affecting HLC values and report equations that quantify the effect of temperature and pH. According to their work, in a majority of cases temperature is the main parameter affecting HLC values for natural waters with moderate contamination (1 mg/ml or less). Other conditions that have influence on HLC values are listed in Table R.7.1–17 (Staudinger and Roberts, 1996):

Table R.7.1–17 Conditions that have influence on HLC values

pH	Important for compound (substance) classes that dissociate to a significant extent in water because only nondissociated species undergo air-water exchange. For most natural waters (6 < pH < 8) the apparent HLC will be significantly less than the intrinsic HLC.
Compound Hydration	Important for aldehydes, which hydrate nearly completely in water, resulting in HLC apparent being several orders of magnitude lower than the intrinsic constant.
Compound concentration/ Complex mixtures effects	If a solution cannot be regarded as diluted (e.g. concentration approaching 10.0 mg/ml) HLC apparent will be lower than HLC values determined at lower concentrations.
Dissolved salts	If the ionic strength of a solution is high (e.g. seawater) the apparent HLC will be higher than the HLC determined in pure water.
Suspended solids /Dissolved Organic Matter (DOM)	If a compound is easily adsorbed (e.g. pesticides) the apparent HLC will be higher than the HLC determined in pure water.
Surfactants	Compounds with high Kow are expected to have an effect on HLC by lowering its value. Recorded effects increase in direct proportion with Kow.

It is worth noting that because of the complex nature of the water matrix the net effect of a possible combination of the parameters listed above may be more than the simple sum of individual effects (Staudinger and Roberts, 1996).

QSPR prediction of Henry's law constant

The prediction of HLC has been reviewed by Schwarzenbach *et al.* (1993), Reinhard and Drefahl (1999), Mackay *et al.* (2000) and Dearden and Schüürmann (2003). The most important approaches are:

- Ratio of water solubility (c_w) to vapour pressure (vp);
- Estimation using connectivity indices;
- Estimation using group and bond contribution methods.

The first method for estimating HLC is not strictly a QSAR method as it uses the water solubility (c_w) and vapour pressure (vp). It is not a highly accurate method, but neither is the measurement of HLC, especially for substances with very high or very low HLC values. vp/c_w can be converted to the dimensionless form of HLC (ratio of concentrations in air and water, c_a/c_w) or K_{aw} by the following equation, which is valid at 25°C:

$$c_a/c_w = 40.874\ vp/c_w$$

Since both water solubility and vapour pressure can be calculated by QSAR methods, then this approach might in some circumstances be a QSAR based method. The method is limited to substances of low water solubility (< 1.0 mol/L). If QSAR calculated values are used for vp and/or c_w, then the respective uncertainties must be considered. For miscible compounds or compounds with water solubility > 1 mol/L the vp/c_w method is not valid.

The second method is based on a combination of connectivity indices and calculated polarisability (Nirmalakhandan and Speece, 1988). A relatively narrow range of chemical types was used to develop the model, so it is not widely applicable. Moreover, Schüürmann and Rothenbacher (1992) found it to have poor predictive power.

Most prediction methods for HLC use a group or bond contribution approach, although some have used physicochemical properties (Dearden *et al.* 2000). The group and bond contribution methods were first used by Hine and Mookerjee (1974), who obtained, for a set of 263 diverse simple organic chemicals, a standard deviation of 0.41 log unit for the group contribution method and one of 0.42 for the bond contribution method. Cabani *et al.* (1981) claimed an improvement in the group contribution method over that of Hine and Mookerjee, whilst Meylan and Howard (1991) extended the bond contribution method and obtained, for a set of 345 diverse chemicals, a standard error of 0.34 log unit.

Evaporation rate

Evaporation rates generally have an inverse relationship to boiling points, i.e. the higher the boiling point, the lower the rate of evaporation. Knowledge of the evaporation rate of spills of volatile liquids can be useful in several respects. If it is known that a spill of a high vapour pressure liquid will evaporate completely in a short period of time, it may be preferable to isolate the area and avoid any intervention or clean-up. The evaporation rate also controls the atmospheric concentration of the vapour and hence the threat of explosion or fire. Data on the volatility properties of the liquid, its temperature, the wind speed, and the spill dimensions are used to calculate the evaporation rate and hence the fraction evaporated at any time.

The substance's tendency to partition into the atmosphere is controlled by the vapour pressure, which is essentially the maximum vapour pressure that a pure substance can exert in the atmosphere. This can be viewed as a kind of *solubility* of the chemical in the atmosphere. Using the ideal gas law ($PV=nRT$), the vapour pressure P in the pressure unit pascal (Pa) can be converted into a solubility (mol/m^3), where the gas constant R is 8.314 Pa.m^3/mol·K and T is absolute temperature (K).

Conversion from vapour pressure into concentration in air under ambient temperature:

% volume = vapour pressure (Pa)/101 325 x 100

or ppm = vapour pressure (Pa)/101 325 x 1 000 000

Since the molar volume is the same for all ideal gases (equal volumes of all gases under the same conditions of temperature and pressure contain the same number of molecule) ppm ≡ volume (i.e. ml/m^3). To convert to weight per unit volume:

X ppm = X x MW/24.041 mg/m^3, 1 mg/m^3 = 24.041/MW ppm

In the formulation of paints and related products, solvents are chosen based on their evaporation characteristics appropriate to the application technique and the curing temperature. To a large extent the evaporation rate of a solvent determines where and how it can be used. In determining the evaporation rate of solvents, n-butyl acetate is used as the standard and is assigned an evaporation rate value of 1. Other solvents are assigned evaporation rate values that indicate how fast they evaporate in relation to n-butyl acetate. For instance, a solvent that evaporates three times as fast as n-butyl acetate would be assigned a value of 3, whereas a solvent that evaporates half as fast as n-butyl acetate would be assigned a value of 0.5.

The rate of evaporation is determined using ASTM D3539-87. A known volume of liquid is spread on a known area of filter paper that is suspended from a sensitive balance in a cabinet. Dry air or nitrogen at 25 ºC is passed through the cabinet at a known rate. The loss of weight is determined and plotted against time.

References for Appendix R.7.1-1

ASTM D 3539-87 (Reapproved 2004) "Standard Test Methods for Evaporation Rates of Volatile Liquids by Shell Thin-Film Evapourometer".

Cabani S., Gianni P., Mollica V. and Lepori L. Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solut. Chem.* (1981) **10**, 563-595.

CRC (2000) CRC Handbook of Chemistry and Physics 81st Edition. Editor in Chief, D. Lide. CRC Press.

Dearden, J.C., Cronin, M.T.D., Ahmed, S.A. and Sharra, J.A. QSPR prediction of Henry's law constant: improved correlation with new parameters. In Gundertofte, K. and Jørgensen, F.S. (Eds), *Molecular Modelling and Prediction of Bioactivity.* Kluwer Academic/Plenum Publishers, New York, 2000, pp. 273-274.

Dearden J.C. and Schüürmann G. Quantitative structure-property relationships for predicting Henry's law constant from molecular structure. *Environ. Toxicol. Chem.* (2003) 22, 1755-1770.

Hine J. and Mookerjee P.K. The intrinsic hydrophilic character of organic compounds. Correlations in terms of structural contributions. *J. Org. Chem.* (1974) **40**, 292-298.

Mackay D. (1991) Multimedia environmental models. Lewish Publishers. Chelsea, MI

Mackay D., Shiu W.Y. and Ma K.C. Henry's law constant. In Boethling R.S. and Mackay D. (Eds.), *Handbook of Property Estimation Methods for Chemicals.* Lewis, Boca Raton, FL, 2000, pp. 69-87.

Meylan W.M. and Howard P.H. Bond contribution method for estimating Henry's law constants. *Environ. Toxicol. Chem.* (1991) **10**, 1283-1293.

W.G. Mallard and P.J. Linstrom, Eds., NIST Chemistry WebBook, NIST Standard Reference Database Number 69. November 1998. National Institute of Standards and Technology, Gaithersburg, MD, 20899.

Nirmalakhandan NN and Speece (1988). QSAR model for predicting Henry's Law Constant. Environ. Sci. Technol. **22**, 1349-1357.

Reinhard M. and Drefahl A. *Handbook for Estimating Physico-chemical Properties of Organic Compounds.* Wiley, New York, 1999.

Russell C.J., Dixon S.L. and Jurs P.C. Computer-assisted study of the relationship between molecular structure and Henry's law constant. *Anal. Chem.* (1992) **64**, 1350-1355.

Sander, R. (1999) Compilation of Henry's Law Constants for Inorganic and Organic Species of Potential Importance in Environmental Chemistry, Version 3 (http://www.henrys-law.org/henry-3.0.pdf).

Schwarzenbach R.P., Gschwend P.M. and Imboden D.M. *Environmental Organic Chemistry*. Wiley, New York, 1993.

Schüürmann G. and Rothenbacher C. Evaluation of estimation methods for the air-water partition coefficient. *Fresenius Environ. Bull.* (1992) **1**, 10-15.

Staudinger, J. and Roberts, P.V. (1996) A critical review of Henry's Law Constants for Environmental Applications*. Crit. Rev. Environ. Sci. Technol*. 26(3):205-297.

Verschueren K (2001) Handbook of Environmental Data on Organic Chemicals 4th Edition (2 volume set). John Wiley & Sons Inc, New York, USA.

R.7.2 Skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation

R.7.2.1 Introduction

Irrespective of whether a substance can become systemically available, changes at the site of first contact (skin, eye, mucous membrane/ gastro-intestinal tract, or mucous membrane/ respiratory tract) can be caused by exposure to a substance. These changes are considered local effects. A distinction in local effects can be made between those observed after single and those after repeated exposure. In this guidance document, the focus will be on local effects after single ocular, dermal or inhalatory exposure. However, wherever possible, use should also be made of existing repeated dose data insofar as they may contain valuable information for the purpose of assessing and classifying effects after single ocular, dermal or inhalatory exposure.

Substances causing local effects after single exposure can be further distinguished as irritant or corrosive substances, depending on the severity, reversibility or irreversibility of the effects observed. Corrosive substances are those which may destroy living tissues with which they come into contact. Irritant substances are non-corrosive substances which, through immediate contact with the tissue under consideration may cause inflammation (see Section R.7.2.1.1 for complete definitions). These tissues are in the present context skin, eye (cornea, iris and conjunctiva) and mucous epithelia such as the respiratory tract. Criteria for classification of irritant and corrosive substances are given in Annex I to Regulation (EC) No 1272/2008 on Classification, Labelling and Packaging of substances and mixtures (CLP Regulation).

Certain substances may also cause irritant effects only after repeated exposure, for example organic solvents. This type of substance may have defatting properties (Ad-hoc Working group on Defatting substances, 1997). Substances that have a similar mode of action need to be considered for labelling with the supplemental statement EUH066 *"Repeated exposure may cause skin dryness or cracking"*.

Information on the mechanisms underlying corrosion and irritation of skin, eye and respiratory tract is given in Appendix R.7.2–1 *Mechanisms of local toxicities: skin corrosion/irritation*, *serious eye damage/eye irritation and respiratory tract corrosion/irritation*.

R.7.2.1.1 Definitions of skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation

Definitions of skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation can be found in the CLP Regulation [27].

Skin irritation: Defined in Section 3.2.1.1 of Annex I to the CLP Regulation as *"[…] the production of reversible damage of the skin following the application of a test substance for up to 4 hours."*.

Dermal concern after repeated exposure: Used for a substance which may cause skin dryness, flaking or cracking upon repeated exposure but which cannot be considered as skin irritant (see Section 1.2.4 of Annex II to the CLP Regulation).

---

[27] Please note that the 8th Adaptation to Technical and Scientific Progress (ATP) of the CLP Regulation is currently under discussion. The 8th ATP will take into account the 5th Revision of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which was adopted in 2012 and contains in particular refined criteria for skin corrosion/irritation and serious eye damage/eye irritation.

Skin corrosion: Defined in Section 3.2.1.1 of Annex I to the CLP Regulation as *"[...] the production of irreversible damage to skin; namely, visible necrosis through the epidermis and into the dermis, following the application of a test substance for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of observation at 14 days, by discolouration due to blanching of the skin, complete areas of alopecia, and scars. [...]"*.

Eye irritation: Defined in Section 3.3.1.1 of Annex I to the CLP Regulation as *"[...] the production of changes in the eye following application of a test substance to the anterior surface of the eye, which are fully reversible within 21 days of application."*.

Serious eye damage: Defined in Section 3.3.1.1 of Annex I to the CLP Regulation as *"[...] the production of tissue damage in the eye, or serious physical decay of vision, following application of a test substance to the anterior surface of the eye, which is not fully reversible within 21 days of application. [...]".*

Respiratory tract irritation: There is no EU or OECD TG for respiratory tract irritation and testing for respiratory tract irritation is not a standard information requirement under REACH. Respiratory tract irritation is considered under the CLP Regulation (Table 3.8.1 of Annex I) as a transient target organ effect, i.e. an *"[...] effect which adversely alter[s] human function for a short duration after exposure and from which humans may recover in a reasonable period without leaving significant alteration of structure or function. [...]*". More specifically, respiratory tract irritation is often used to describe either or both of two different toxicological effects, *sensory irritation* and *local cytotoxic effects*. However, classification in STOT-SE Category 3 for respiratory tract irritation is generally limited to local cytotoxic effects. *"[...] Respiratory irritant effects [are] characterised [by] localised redness, oedema, pruritis and/or pain and they impair function with symptoms such as cough, pain, choking, and breathing difficulties [...]"* (see Section 3.8.2.2.1 of Annex I to the CLP Regulation).

Respiratory tract corrosion: There is no EU or OECD TG for respiratory tract corrosion and testing for respiratory tract corrosion is not a standard information requirement under REACH. Respiratory tract corrosion is defined in Section 3.1.2.3.3 of Annex I to the CLP Regulation as *"[...] destruction of the respiratory tract tissue after a single, limited period of exposure analogous to skin corrosion; this includes destruction of the mucosa. [...]".*

Classification and labelling under the CLP Regulation:

Substances and mixtures causing skin corrosion/irritation, serious eye damage/eye irritation and/or respiratory tract corrosion/irritation can be further characterised by their classification under the CLP Regulation [28].

Detailed information on the classification and labelling of substances and mixtures can be found in the *Guidance on the Application of the CLP criteria*.

28 Please note that the 8th Adaptation to Technical and Scientific Progress (ATP) of the CLP Regulation is currently under discussion. The 8th ATP will take into account the 5th Revision of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which was adopted in 2012 and contains in particular refined criteria for skin corrosion/irritation and serious eye damage/eye irritation.

a) For Skin effects

- **Skin corrosives** are classified in Category 1 with the Hazard statement H314 *"Causes severe skin burns and eye damage"*. Further subcategorisation is defined based on the Draize skin corrosion *in vivo* test:
    - Subcategory 1A: Destruction of skin tissue occurs after exposure times ≤ 3 minutes and is observed within a period ≤ 1 hour after exposure,
    - subcategory 1B: Destruction of skin tissue occurs after exposure times > 3 minutes and ≤ 1 hour and is observed within a period ≤ 14 days after exposure,
    - subcategory 1C: Destruction of skin tissue occurs after exposure times > 1 hour and ≤ 4 hours and is observed within a period ≤ 14 days after exposure.
- **Skin irritants** are classified in Category 2 with the Hazard statement H315 *"Causes skin irritation".*

b) For Eye effects

- **Substances or mixtures causing serious eye damage** are classified in Category 1 with the Hazard statement H318 *"Causes serious eye damage".*
- **Substances or mixtures causing eye irritation** are classified in Category 2 with the Hazard statement H319 *"Causes serious eye irritation"*.

c) For Specific Target Organ Toxicity with relevance to the respiratory tract

- **Substances or mixtures causing respiratory tract corrosion** are classified for Acute Toxicity by inhalation and labelled as EUH071 "*Corrosive to the respiratory tract*" if the corrosive effect causes the death of the animals within the criteria for Acute toxicity, or in Specific Target Organ Toxicity after Single Exposure (STOT-SE) Category 1 (with the Hazard statement H370 *"Causes damage to the respiratory tract"*) or Category 2 (with the Hazard statement H371 *"May cause damage to the respiratory tract"*), depending on the dose level required to cause the toxic effects.
- **Substances or mixtures causing respiratory tract irritation** *via* a local cytotoxic effect are classified in Specific Target Organ Toxicity after Single Exposure (STOT-SE) Category 3 with the Hazard statement H335 *"May cause respiratory irritation".*

According to Section 1.2.6 of Annex II to the CLP Regulation, the Hazard statement EUH071 must also be applied to inhaled substances or mixtures classified for skin corrosion and not tested for acute inhalation toxicity.

Note that dermal and respiratory tract irritation following repeated exposure are not discussed in the present context, since this Guidance focuses on acute effects after single exposure. However, data from repeated exposure studies may be useful in certain cases (e.g. if the substance was identified as a corrosive or strong irritant after the first application or for deriving quantitative information). Nevertheless, for the sake of completeness, both the definition of dermal irritation after repeated exposure as well as the related Hazard Statement

EUH066 ("Repeated exposure may cause skin dryness or cracking") are given here. More guidance on local effects after repeated exposure can be found in Section R.7.5 on repeated dose toxicity.

R.7.2.1.2 Objective of the guidance on skin corrosion / irritation, serious eye damage / eye irritation and respiratory tract corrosion / irritation

The general objectives are:

a. to establish whether information from physical/chemical data, from non-testing methods (grouping, QSARs and expert systems), from *in vitro* studies, from animal studies or human experience provides evidence that the substance is, or is likely to be, corrosive.

b. to establish whether information from physical/chemical data, from non-testing methods (grouping, QSARs and expert systems), from *in vitro* studies, from animal studies or human experience provides evidence of significant skin, eye or respiratory tract irritation.

c. to establish if possible the time of onset and the extent and severity of the responses and information on reversibility.

d. If possible to gather, in the process of hazard identification, any quantitative data on dose-response relationships that might allow the derivation of DNELs essential for a complete risk assessment.

If a risk assessment is necessary, both the severity of the identified hazard (in so far as it can be judged from the test data) and the probability of the occurrence of an acute corrosive or irritant response in humans must be assessed based on the likelihood of any exposure to the substance and in relation to the route, pattern and extent of the expected exposure.

Please note that there are currently no standard tests and no OECD TGs available for acute respiratory tract irritation and there is no testing requirement for respiratory tract irritation under the REACH Regulation. Consequently no testing and assessment strategy for respiratory tract corrosion/irritation is included in this guidance. Nevertheless, account should be taken of any existing and available data that provide evidence of the respiratory tract corrosion/irritation potential of a substance. For instance, acute inhalation studies including histopathological evaluation of the respiratory tract and/or examinations of nasal or bronchioalveolar lavage as well as repeated inhalation studies may provide important information for classification and labelling (See Section R.7.2.12 for further details).

SKIN CORROSION/IRRITATION

R.7.2.2 Information requirements on skin corrosion/irritation

The information on skin corrosion/irritation that is required to be submitted for registration and evaluation purposes is specified in Annexes VI to XI to the REACH Regulation. According to Annex VI, the registrant should gather and evaluate all existing available information before considering further testing. This includes physico-chemical properties, (Q)SAR ((Quantitative) Structure-Activity Relationship), grouping, in vitro data, animal studies, and human data. For classified substances, information on exposure, use and risk management measures should also be collected and evaluated in order to ensure safe use of the substance.

If these data are inadequate for hazard and risk assessment, further testing should be carried out in accordance with the requirements of Annexes VII (≥1 tpa) and VIII (≥10 tpa) to the REACH Regulation.

R.7.2.2.1 Information requirements for quantities of ≥1 tpa (Annex VII to the REACH Regulation) [29]

If new testing data are necessary, these must be derived from *in vitro* methods only. Annex VII does not foresee *in vivo* testing for skin corrosion/irritation.

The standard information requirements at this tonnage level for skin corrosion/irritation (see Section 8.1 in Column 1 of Annex VII) can be fulfilled by following four steps: (1) assessment of the available human and animal data, (2) assessment of the acid or alkaline reserve, (3) *in vitro* skin corrosivity study, (4) an *in vitro* skin irritation study.

Section 8.1 in Column 2 of Annex VII lists specific rules for adaptation according to which steps 3 and 4 do not need to be conducted. These rules are applicable when:

1. the available information already indicates that the criteria are met for classification as corrosive to the skin or irritating to eyes, or

2. the substance is flammable in air at room temperature (Please note that this rule should actually read: “the substance is **spontaneously** flammable in air at room temperature”), or

3. the substance is classified as very toxic in contact with skin (i.e. the “Substance is fatal in contact with skin” and classified in Category 1 for Acute toxicity according to current CLP terminology), or

[29] Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

an acute toxicity study *via* the dermal route does not indicate skin irritation up to the limit dose level (2000 mg / kg body weight) (Please see footnote d to Figure R.7.2–2 for further information).

R.7.2.2.2 Information requirements for quantities of ≥10 tpa (Annex VIII to the REACH Regulation) [30]

For substances manufactured or imported in quantities of ≥10 tpa *in vivo* testing is the standard information requirement of Annex VIII (Column 1) for skin corrosion/irritation, in case the information requirement cannot be met with the information obtained as specified in section 8.1 of Annex VII.

Before new tests are carried out to determine the properties listed in Annex VIII, all available *in vitro* data, *in vivo* data, historical human data, data from valid (Q)SARs and data from structurally related substances (read-across approach) must be assessed first. Due to the sequential nature of the REACH standard information requirements, the reader is reminded that at quantities of ≥10 tpa, the information requirements of Annex VII to the REACH Regulation also apply. This means that before a new *in vivo* test is performed, the appropriate *in vitro* testing must be undertaken according to the rules set out in section 8.1 of Annex VII and must be documented in the technical dossier (IUCLID). Finally, the information generated at Annex VII level must be taken into account in determining whether an *in vivo* test at Annex VIII level is really needed.

Column 2 of Annex VIII lists the following specific rules that allow deviation from the standard information required by Annex VIII for skin corrosion/irritation:

- the substance is classified as corrosive to the skin or as a skin irritant, or
- the substance is a strong acid (pH ≤ 2.0) or base (pH ≥ 11.5), or
- the substance is flammable in air at room temperature (Please note that this rule should actually read: “the substance is **spontaneously** flammable in air at room temperature”), or
- the substance is classified as very toxic in contact with skin (i.e. the “Substance is fatal in contact with skin” and classified in Category 1 for Acute toxicity according to current CLP terminology), or
- an acute toxicity study by the dermal route does not indicate skin irritation up to the limit dose level (2000 mg/kg body weight) (Please see footnote d to Figure R.7.2–2 for further information).

The *in vitro* methods that can be used to adapt the standard information requirements are detailed in Sections R.7.2.3.1 and R.7.2.4.1 of this Guidance, under “*In vitro* data”. In case results of the *in vitro* testing are used to adapt the standard information requirement of *in vivo*

---

30 Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

testing at Annex VIII level, an adaptation e.g. according to Annex XI to the REACH Regulation will need to be submitted in order to successfully submit a compliant dossier.[30]

Guidance on the application of these rules is given in the testing and assessment strategies described in Sections R.7.2.6 and R.7.2.11 of this Guidance.

It should be noted that the conditions of acceptance by ECHA of implementation of any of the adaptation rules laid down in Annex XI are strict, and whenever an adaptation argument is being used (e.g. use of (Q)SARs, read-across or non-validated *in vitro* test methods), scientific justification, solid documentation and readiness for risk assessment and Classification and Labelling must be provided by registrants. For detailed information on these rules, see Annex XI to the REACH Regulation.

R.7.2.3 Information sources on skin corrosion/irritation

R.7.2.3.1 Non-human data on irritation/corrosion

Non-testing data on irritation/corrosion

Physico-chemical properties

Relevant information can be inferred from basic physico-chemical characteristics of a substance (e.g. extreme pH). Extreme pH values may indicate the potential of a substance to cause skin corrosion:

If the pH ≤ 2 or pH ≥ 11.5, then consider the substance to be corrosive to the skin (Category 1) when the pH is used as the sole basis for the classification decision (see also Section R.7.2.4.1 of this Guidance).

Grouping, (Q)SARs and expert systems [31]

In REACH Annex XI two types of non-testing methods are mentioned which can be used for adaptation of standard information requirements, either as standalone (where possible) or in combination with other information (in the context of a *Weight-of-Evidence* assessment):

- qualitative and quantitative Structure-Activity-Relationships (SARs/QSARs, section 1.2, including expert systems, generally incorporating multiple (Q)SARs, expert rules and data) on the one hand, and
- grouping of substances and read-across approaches [32] on the other.

---

[31] Further information can be found in *Chapter R.6 QSAR and grouping of chemicals* of *the Guidance on IR&CSA*, the OECD Guidance on Grouping of Chemicals, Second Edition (OECD, 2014a), the new OECD Guidance on an Integrated Approach for Testing and Assessment (IATA) for skin corrosion and irritation (OECD, 2014b) and the JRC report on Alternative methods for regulatory toxicology (Worth, 2014).

[32] The relevant terminology is not always used consistently. With reference to the ECHA Guidance on QSAR and grouping, the terms category approach and analogue approach are used to describe techniques for grouping of substances, whilst the term read-across is reserved for a technique to fill data gaps, i.e. to transfer knowledge from one or more substances called source(s) to another substance with data gaps, named target substance.

The adaptation of standard information requirements can be applied for the assessment of skin corrosion/irritation, if it provides relevant and reliable data for the substance of interest. As specified in Annex XI of the REACH regulation, the use of non-testing methods needs to be justified and sufficiently documented. In the case of QSARs and expert systems, registrants need to prepare property predictions by completion of a QSAR Prediction Reporting Format (QPRF). The QPRF is a harmonised template for summarising and reporting substance-specific predictions generated by (Q)SAR models. For filling a data gap under REACH, it is also necessary to provide information on the prediction model employed following a QSAR Model Reporting Format (QMRF) document. The QMRF is a harmonised template for summarising and reporting key information on (Q)SAR model validity, including the results of any validation studies. The information is structured according to the OECD (Q)SAR validation principles (for further information see http://www.oecd.org/env/ehs/risk-assessment/validationofqsarmodels.htm). The JRC QSAR Model Database is an inventory of information on available QMRFs, freely accessible online (https://eurl-ecvam.jrc.ec.europa.eu/databases/jrc-qsar-model-database). More detailed guidance on QSAR models, their use and reporting formats, including the QMRF, is provided in Section R.6.1 of Chapter R.6 of the *Guidance on IR&CSA*.

In general, there are several different ways in which non-testing methods can be used in the context of an Integrated Approach to Testing and Assessment (IATA) (OECD, 2014b), e.g.:

- for direct prediction of corrosion/irritation potential or the absence thereof,
- as part of a *Weight-of-Evidence* scheme (where the information from non-testing methods alone is not sufficient for a decision), or
- in order to decide how best to proceed with further (*in vitro*) testing (i.e. *via* a top-down or bottom-up approach). For further information see Section R.7.2.6.2.

In the case of skin corrosion and irritation, many of the models have a mechanistic basis, which provides additional information on the relevance of the model.

- SAR and read-across on skin irritation and corrosion:

SARs and read-across are treated together in this section because the existence of a SAR (structural alert or set of fragments) provides one means of justifying read-across. In fact, structural alerts are substructures in the substance that are considered to reflect some kind of chemical or biochemical reactivity that underlies the toxicological effect. The occurence of a structural alert for a substance suggests the presence of an effect, based on the notion that structural analogues that have exhibited corrosion (or irritation) potential can be used to predict a corrosive or irritant effect for the substance of interest, or to tailor further testing and assessment, as indicated in the OECD IATA for skin corrosion/irritation (OECD, 2014b).

Knowledge on structural alerts for skin irritation/corrosion is always evolving (in particular where new classes of substances are introduced into the market). Therefore predictions based on read-across may also be possible for chemically similar substances if it can be shown that their similarity reflects reactive substructures able to react with skin tissue, even if that substructure has so far not been coded into a structural alert in any of the available literature or software models.

Negative data from structural analogues may also be used to make predictions in certain cases. The absence of one of the known structural alerts for irritation and corrosion alone does not prove absence of effect, as knowledge of structural alerts for irritation and corrosion might be incomplete. For instance, other substructures (not yet identified as structural alerts) or other properties of the substance may be responsible for a corrosive or irritant effect. As an example, irritant contact dermatitis may occur indirectly, such as in the case of exposure to organic solvents with defatting properties. Substances that have a similar mechanism of action need to be considered for the supplemental labelling *'Repeated exposure may cause skin dryness or cracking'* (EUH066) (Ad-hoc Working group on Defatting Substances, 1997).

An example of a simple SAR is the use of the hydroperoxide group (R-O-O-H) as an alert for corrosivity, which is mechanistically based on the fact that hydroperoxides are both acidic and oxidisers. Another SAR is the peroxide group (R_1-O-O-R_2), based on the fact that peroxides decompose easily and thus have a low thermal stability. The radicals formed by breaking the O-O bond are reactive and may be the cause of irritation or corrosion.

A variety of SARs (including hydroperoxides) for predicting the presence of irritation or corrosion have been described by Hulzebos *et al*. (2001, 2003, 2005), and some of these have been incorporated into the BfR (German Federal Institute for Risk Assessment) rule-base, and the SICRET tool (Walker *et al*., 2005, see Appendix R.7.2–2). The BfR alerts ("inclusion rules") for corrosion and irritation have more recently been incorporated into the Toxtree software (https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/qsar_tools/toxtree) and into the OECD QSAR Toolbox (http://www.qsartoolbox.org/).

- QSARs and expert systems for skin corrosion and irritation:

An overview of available (Q)SARs for skin corrosion and irritation is provided in Table R.7.2–1. QSARs and expert systems for skin corrosion and irritation have been described in several reviews (Hulzebos *et al*., 2001, 2003, 2005; Patlewicz *et al*., 2003; Gallegos Saliner *et al*., 2006, 2008). A comparison of the predictive capacities of three popular commercial tools is also available (Mombelli 2008). A few examples are presented in Appendix R.7.2–2, including literature-based QSAR models, and expert systems.

Most of the QSARs reported in the literature have been developed from small data sets of specific groups of substances, although in some cases more diverse and larger datasets were also examined. In general, it has been suggested that basic physico-chemical parameters such as acidity, basicity, hydrophobicity, and molecular size as well as electrophilic reactivity, are useful to predict the toxic potential of homologous substances. In contrast, models intended to predict the toxic potential of heterogeneous groups of substances emphasise the commonality of structural features.

A number of models are coded into expert systems, which are computer programs that guide hazard assessment by predicting toxicity endpoints of certain chemical structures based on the available information. Expert systems can be based on an automated rule-induction system (e.g. TOPKAT, HazardExpert and MultiCASE), or on a knowledge-based system (e.g. DEREK Nexus or the BfR- former DSS [33]). More details on available expert systems are reported in Appendix R.7.2–2.

Not all of the models were developed with EU regulatory purposes in mind, so it is important to assess in each case whether the endpoint or effect being predicted corresponds to the regulatory endpoint of interest. The rule-base at the heart of the former BfR DSS has been developed to predict EU regulatory endpoints, however predictions refer to the former Dangerous Substance Directive (DSD) classification/labelling system used in the EU before the CLP regulation came into force, and in borderline cases the results of the prediction may not

---

[33] Distribution of the BfR expert system "Decision Support System for Local Lesions" (DSS) mentioned in previous versions of this guidance has been discontinued. However, the rule-base for skin and eye irritation/corrosion included in this system has been incorporated into software tools such as the OECD QSAR Toolbox or Toxtree (cf. below).

fully reflect the correct CLP classification. More details on this model are reported in Appendix R.7.2–2.

It should also be noted that the criteria for classification as skin irritant Category 2 based on the mean score for erythema/eschar or for oedema in the *in vivo* test have changed from ≥2 under DSD to ≥2.3 under CLP. Consequently predictions as skin irritant Cat 2 from models developed based on the DSD criteria should be interpreted with caution since they may lead to overprediction and should not be used for direct classification under CLP. These models can however be argued to be "conservative" and therefore acceptable for predicting no classification under CLP.

Based on the BfR rule-base, the freely downloadable OECD QSAR Toolbox software contains two profilers relevant for corrosion/irritation, which encode both the “inclusion rules” (structural alerts predicting corrosion/irritation potential) and the “exclusion rules” (“IF…THEN NOT…” rules predicting the absence of irritation/corrosion potential) due to certain physico-chemical properties. The use in combination with other profilers (e.g. for skin metabolism) and data for analogues allows for the prediction of skin corrosion/irritation for new chemicals through read-across or category approaches. More details on the Toolbox specific contents for skin corrosion and irritation are reported in Appendix R.7.2–2.

In the case of classification models for skin corrosion, where it is not indicated in the supporting documentation whether the predicted classification should be Skin Corrosive Category 1A, 1B or 1C, Category 1 prediction without further sub-categorisation should be used. Very few models are available (see Gallegos Saliner *et al*., 2006, 2008 for review). Available models tend to focus on defined chemical classes (e.g. acids, bases, phenols) and might be useful as an alternative to *in vitro* testing for such classes. For classification and labelling, the BfR rule-base provides information that is the closest to the regulatory goal, since the system was designed to predict former EU Risk Phrases for skin irritation (R38) and corrosion (R34, R35) under the Dangerous Substance Directive (DSD). However, in borderline cases and as highlighted above, the prediction may not fully reflect the correct classification under CLP.

Table R.7.2–1 Overview of available (Q)SARs for skin corrosion/irritation. See Appendix R.7.2–2 for more information on these models.

Category of model or source	Reference or name of the model	Applicability domain
Literature models	Barratt (*et al.*) (1995a, 1996 a,b,c); Whittle *et al.* (1996)	Diverse local models for acids, bases , phenols, neutral organic and electrophiles
	Hayashi et al. (1999)	Phenols
	Kodithala *et al.* (1999)	Phenols, esters, and alcohols
	Nangia *et al*. (1996)	Bases
	Smith *et al*. (2000 a,b)	Esters
	Gerner *et al*. (2004); Hulzebos *et al.* (2005); Walker *et al*. (2004)	New Chemicals Database, organic chemicals with no significant hydrolysis potential and purity > 95%
	Golla *et al*. (2009)	Organic chemicals from diverse classes

Data repositories	Danish QSAR database (http://qsar.food.dtu.dk/, also included in the OECD QSAR Toolbox)	Industrial chemicals, pesticides, etc.
Computerised models	PaDEL-DDPredictor (http://padel.nus.edu.sg/software/padelddpredictor/) (Liew and Yap, 2013)	Calculated by the model based on the range of descriptors
	BfR rule-base, free (included in the OECD QSAR Toolbox and Toxmatch, Toxtree, ToxPredict and Ambit)	EU New chemicals (NONS) database, organic chemicals with no significant hydrolysis potential and purity > 95%
	ACD/Percepta	Organic chemicals
	Derek Nexus, commercial	Organic chemicals and some metals
	HazardExpert, commercial	Organic chemicals
	MolCode, commercial	Organic chemicals
	MultiCASE, commercial	Organic chemicals
	TOPKAT, commercial	Organic chemicals
Review papers	Hulzebos *et al*. (2001, 2003, 2005)	N.A.
	Patlewicz *et al*. (2003)	N.A.
	Gallegos Saliner *et al*. (2006, 2008)	N.A.
	Mombelli (2008)	N.A.

Abbreviation: N.A. = not applicable.

Testing data on skin corrosion/irritation

The internationally accepted test methods for skin corrosion/irritation as described in the Annex to the EU Test Methods (TM) Regulation (Council Regulation (EC) No 440/2008) and in OECD TGs (available at http://www.oecd.org/env/ehs/testing/oecdguidelinesforthetestingofchemicals.htm#Test_Guidelines) are: EU method B.4 (OECD TG 404), EU B.40 (OECD TG 430), EU B.40bis (OECD TG 431), OECD TG 435 and EU B.46 (OECD TG 439).

Please note that the latest version of an adopted test guideline should always be used when generating new data, independently of whether it is published by EU or OECD.

The testing strategy developed for skin corrosion/irritation (see Section R.7.2.6 of this Guidance) emphasises the need to evaluate all available information (including physico-chemical properties) before undertaking any *in vivo* testing. This strategy employs screening elements designed to avoid, as far as possible, *in vivo* testing of corrosive and severely irritating substances. In particular, *in vitro* tests should usually be performed first, and it should be assessed whether *in vivo* testing can be completely avoided.

In vitro data

Accepted *in vitro* test methods to detect skin corrosion/irritation (i.e. Category 1 and 2 under CLP) and/or absence of effects (i.e. not classified under CLP) are listed in Table R.7.2–2. More information on the specific scope and limitations of these tests is provided in Section R.7.2.4.1 of this Guidance, under "Testing data on skin corrosion/irritation".

In Table R.7.2–2, when the classification outcome in the column "Classification according to the CLP Regulation" is indicated as "Cat. 1B/1C" or "Cat. 1/Cat. 2", this means that the test method alone cannot differentiate between those (sub-)categories and more information is needed to conclude on the exact classification. For instance if the result of an *in vitro* skin irritation study according to B.46/OECD TG 439 is positive, it cannot be concluded whether the substance is either corrosive (Cat. 1) or irritant (Cat. 2) to the skin and therefore additional information on skin corrosion potential is needed e.g. by performing an *in vitro* skin corrosion study.

Table R.7.2–2 Accepted *in vitro* test methods for skin corrosion/irritation

	Test method	Validation status, regulatory acceptance	EU Test Methods/ OECD test guideline	Classification according to CLP Regulation	EURL ECVAM DB-ALM protocol Nr.
Skin corrosion					
	TER	Validated and regulatory acceptance	B.40/TG 430	Cat. 1 or non corrosive	115
	EpiDerm ™ SCT	Validated and regulatory acceptance	B.40 bis/TG 431	Cat. 1, 1A, 1B/1C or non-corrosive	119
	EpiSkin ™	Validated and regulatory acceptance	B.40 bis/TG 431	Cat. 1, 1A, 1B and 1C or non-corrosive[34]	118
	SkinEthic ™ RHE	Validated and regulatory acceptance	B.40 bis/TG 431	Cat. 1, 1A, 1B/1C or non-corrosive	-
	epiCS®	Validated and regulatory acceptance	B.40 bis/TG 431	Cat. 1, 1A, 1B/1C or non-corrosive	-

34 The EpiSkin SOP allows for differentiating between the 3 sub-categories and OECD GD 203 suggests the use of this method to distinguish 1B from 1C before *in vivo* testing is considered. However, OECD TG 431 currently only permits the use of EpiSkin to distinguish 1A from 1B/1C.

	Corrositex (*in vitro* membrane barrier test method)	Validated and regulatory acceptance	N.A./TG 435	Cat. 1, 1A, 1B and 1C or non-corrosive	116
Skin irritation					
	EpiDerm ™ SIT	Validated and regulatory acceptance	B.46/TG 439	Cat. 1/Cat. 2 or NC	138
	EpiSkin ™	Validated and regulatory acceptance	B.46/TG 439	Cat. 1/Cat. 2 or NC	131
	SkinEthic ™ RHE	Validated and regulatory acceptance	B.46/TG 439	Cat. 1/Cat. 2 or NC	135
	LabCyte EPI-MODEL24 SIT	Validated and regulatory acceptance	B.46/TG 439	Cat. 1/Cat. 2 or NC	-

Abbreviations: N.A. = not available; NC = not classified; RHE = Reconstructed Human Epidermis; SCT = Skin Corrosion Test; SIT = Skin Irritation Test; TER=Transcutaneous electrical resistance.

Further test method developments may occur and the registrants are advised to follow the latest updates through e.g. the EURL ECVAM website (https://eurl-ecvam.jrc.ec.europa.eu/) and ECHA's test methods webpage (http://echa.europa.eu/support/testing-methods-and-alternatives) for potential new test guidelines and test guideline updates.

Animal data

Annex I to the CLP Regulation defines skin corrosion/irritation as local toxic effects, and, as such, an assessment of skin corrosion/irritation is normally part of the acute testing phase of a toxicity programme and it is an early requirement of all regulatory programmes. Testing for skin corrosion/irritation has, historically, used animal models and a variety of test methodologies depending upon, for example, the laboratory undertaking the test and the area and intended application. An IATA, which aims at minimisation of animal testing and instead largely relies on internationally approved *in vitro* tests, has been adopted by the OECD in 2014 as Guidance Document 203 (OECD, 2014b). Thereby, animal models have become unnecessary in most cases when testing for this endpoint. This is in line with one of the objectives of the REACH Regulation, as described in Articles 13(1) and 25(1), on that animal testing should be undertaken only as a last resort, i.e., where a substance falls outside of the applicability domain of the available *in vitro* methods or the results are not conclusive.

In cases in which *in vivo* testing may be necessary, current approaches for skin corrosion/irritation testing *in vivo* are covered by the Acute Dermal Irritation/Corrosion test method (EU B.4/OECD TG 404). This guideline requires a tiered approach, whereby existing and relevant data are evaluated first. The guideline also recommends that testing in animals should only be conducted if determined to be necessary after consideration of available alternative methods. The *in vivo* test uses one animal (the rabbit is the preferred species); in the absence of severe effects this is followed by a further testing of up to two animals (in a sequential manner up to a total maximum of three animals). When two animals are used, if both exhibit the same response, no further testing is needed.

Both EU and OECD methods use the scoring system developed by Draize (1944). The EU criteria for classification are based on the mean tissue scores obtained over the first 24-72 hour period after exposure and on the reversibility or irreversibility of the effects observed. Skin irritants (Category 2) cause significant inflammation of the skin (erythema and/or oedema) but this effect is transient, i.e. the affected sites are repaired within the observation period of the test.

A corrosive substance causes full thickness destruction of the skin tissue and is classified as Skin corrosive (Category 1) and sub-classified in subcategory 1A, 1B or 1C depending upon the exposure time (3 min, 1 hour, and 4 hours, respectively) and observation time (1 hour, 14 days, and 14 days, respectively).

For existing animal data, the use of methods other than those specified in the Annex to the EU Test Methods Regulation, or corresponding OECD methods may be accepted on a case-by-case basis.

In addition to the EU B.4/OECD TG 404 mentioned above, further animal data may be available e.g. from:

- o Acute dermal toxicity test (EU B.3/OECD TG 402)
- o Skin sensitisation tests (EU B.6/OECD TG 406, EU B.42/OECD TG 429, and OECD TG 442A and 442B)

Section R.7.2.6 of this Guidance provides comments on how to use information from these test in a testing and assessment strategy for skin corrosion/irritation. Additional *in vivo* tests may also provide relevant information (see paragraph 37 of the OECD Guidance Document 203 (OECD, 2014b)) although the reporting and scoring of the irritation in these tests may not be sufficient in all cases to allow a final conclusion to be drawn.

R.7.2.3.2 Human data on skin corrosion/irritation

Existing human data include historical data that should be taken into account when evaluating intrinsic hazards of substances. *New* testing in humans for hazard identification purposes is not acceptable for ethical reasons.

Existing data can be obtained from case reports, poison information centres, medical clinics, occupational experience, epidemiological studies and volunteer studies. Their quality and relevance for hazard assessment should be critically reviewed. However, in general, human data can be used to determine a corrosive or irritating potential of a substance. Good quality and relevant human data have precedence over other data. However, absence of incidence in humans does not necessarily overrule *in vitro* data or existing animal data of good quality that are positive.

R.7.2.4 Evaluation of information on skin corrosion/irritation

R.7.2.4.1 Non-human data on skin corrosion/irritation

Non-testing data on skin corrosion/irritation

In 2014, the OECD approved an IATA for skin corrosion/irritation. The IATA includes description of various types of data that can be used in the assessement of these hazards, including the types of infomation presented below. The IATA has a modular approach, whereby the *domain, role in IATA, strengths, weaknesses and limitations* of each type of data are given in a tabular form. It is also explained with flow diagrams how the data can then be integrated. Detailed guidance is given on the *Weight-of-Evidence* approach and on how quality, adequacy, coverage and consistency of data is assessed within a *Weight-of-Evidence* approach (OECD, 2014b).

Physico-chemical properties

According to the current EU and OECD guidelines, substances should not be tested on animals for skin corrosion/irritation if they can be predicted to be corrosive to the skin (Category 1) from their physico-chemical properties. In particular, substances exhibiting strong acidity (pH ≤2.0) or alkalinity (pH ≥11.5) in solution are predicted to be corrosive to the skin and should not be tested on animals. Testing with *in vitro* methods can nevertheless be performed, especially if skin corrosion sub-categorisation is required. It should also be noted that although prediction of skin corrosion based on pH extremes shows a very high specificity (> 90%), and therefore a low number of false positives (Worth *et al*., 1998), it cannot be ruled out that some substances may be overpredicted if classification is based solely on pH data. However, substances that have other pH values will need to be considered further for their potential for skin corrosion/irritation.This model is included in the OECD IATA for skin corrosion and irritation (OECD, 2014b). Several studies have investigated and confirmed the usefulness of pH as a predictor of corrosion (Worth and Cronin, 2001) and as an element in tiered testing strategies (Worth, 2004).

Where extreme pH is the only basis for classification as corrosive, it may also be important to take into consideration the acid/alkaline reserve, i.e. a measure of the buffering capacity of a substance (Young *et al*., 1988; Botham *et al*., 1998; Young and How, 1994). However, it should be noted that for pure substances the sensitivity of pH for identifying skin corrosivity may actually be significantly reduced when combined with acid/alkaline reserve information (Worth *et al*., 1998). The buffering capacity should not be used alone to exonerate from classification of the substance as corrosive. Indeed, when the acid/alkaline reserve suggests that the substance might be non-corrosive, further *in vitro* testing should be considered (see Section 3.2.2.2 of Annex I to the CLP Regulation).

Grouping, (Q)SARs and expert systems

Guidance has been developed by the former ECB (Worth *et al*., 2005) on how to apply (Q)SARs for regulatory use. Guidance on how to assess the validity and suitability of (Q)SAR models and adequacy of their predictions is given in Section R.6.1 of Chapter R.6 of the Guidance on IR&CSA and in the OECD Guidance document on the validation of (Q)SAR models (OECD, 2007). Essentially, the determination of whether a (Q)SAR result may be used to replace a test result can be broken down into three main steps:

1. an evaluation of the scientific validity (relevance and reliability) of the model,
2. an assessment of the applicability of the model to the chemical of interest and the reliability of the individual model prediction,
3. an assessment of the adequacy of the information for making the regulatory decision, including an assessment of completeness, i.e. whether the information is sufficient to make the regulatory decision, and if not, what additional (experimental) information is needed.

The assessment of model validity needs to be performed along the lines of the OECD principles for (Q)SAR validation (OECD, 2007), e.g. in terms of a defined endpoint, an unambiguous algorithm, a defined applicability domain, the statistical characteristics ("goodness-of-fit"), and mechanistic interpretation.

The following questions, *inter alia*, should be addressed when assessing the reliability of an individual prediction:

i. Is the chemical of interest within the scope of the model, according to the defined applicability domain of the model?

ii. Is the defined applicability domain suitable for the regulatory purpose?

iii. How well does the model predict chemicals that are similar to the chemical of interest?

iv. Is the model estimate reasonable, taking into account other information?

The mechanism of skin corrosion and irritation involves toxicodynamic and toxicokinetic parameters. Some models predict skin corrosion and irritation based on toxicodynamic properties only (e.g. acidity or basicity, electrophilicity, other reactivity, surfactant activity, solving membranes). Such models have to be additionally evaluated to check whether they also take account of toxicokinetic parameters related to the potential of a substance to cross relevant outer membranes of the skin (stratum corneum) and to be active in the living tissue underneath; alternatively these models have to be used in combination with data covering such toxicokinetic parameters. Conversely some models predict (the absence of) corrosion and irritation solely from e.g. physico-chemical properties considered to illustrate the toxicokinetic behaviour of a substance. Such models should be evaluated to check whether they also take account of the activity of the substance (toxicodynamics), in particular for its potential corrosivity (whereby the corrosive action itself may lead to membrane destruction and subsequent tissue damage).

For example, the BfR rule-base implemented in Toxtree and the OECD QSAR Toolbox contains both physico-chemical exclusion rules and structure-based inclusion rules (structural alerts). Evaluations of these rules for the prediction/exclusion of skin corrosion/irritation (Rorije and Hulzebos, 2005, on the physico- chemical exclusion rules; Gallegos Saliner *et al.*, 2007, on the structural alerts) have been carried out in accordance with the OECD principles for (Q)SAR validation (see Appendix R.7.2–2). However, inclusion and exclusion rules were evaluated separately, and not used in combination in these works.

When applied, these two sets of rules may sometimes provide contradictory information, i.e. a structural alert might indicate corrosion/irritation potential, while at the same time, based on physico-chemical properties, absence of effect is predicted. In such cases, it is recommended to consider additional information (e.g. on skin permeability or on the behaviour of chemically similar substances). In other cases, applicability of one (or more) of the physico-chemical exclusion rules might indicate absence of a corrosion/irritation potential of the target substance, while no structural alert for corrosion/ irritation is triggered. Given that the absence of any known structural alert is not equivalent to the absence of a potential effect, in such a situation the substance should still be examined for potentially reactive substructures (and examining the behaviour of chemical analogues would still be beneficial).

While these considerations apply to the use of the BfR rule-base for direct classification/non-classification, less certainty might be required e.g. for a decision on further *in vitro* testing: where the exclusion rules suggest the absence of an effect, a bottom-up approach could be followed, i.e. a test for irritation and not one for corrosion might be initiated (see Section R.7.2.6.2).

There is no other model available which sufficiently describes the absence of effects. Neutral organics[35] are expected not to be irritants; however their defatting potential should be discussed. Predicted absence of reactivity needs to be described in sufficient detail or be substantiated with other information.

[35] By definition a neutral organic is a chemical which does not have potential reaction centres, even after skin metabolism.

Testing data on skin corrosion/irritation

In vitro data

There are EU and OECD adopted test guidelines (see Section R.7.2.3.1), according to which substances can be classified as skin corrosives, skin irritants, or not classified.

Annex VII to the REACH Regulation requires information from the *in vitro* tests specified below for skin corrosion/irritation, and not from animal tests. Guidance on how *in vitro* data can also be used to fulfil Annex VIII requirements, is given in Section R.7.2.6 of this document [36].

Data from the following types of test can be used for Annex VII requirements and may be accepted for Annex VIII requirements for skin corrosion/irritation when general rules for adaptation specified in Annex XI are used:

- For skin irritation:
  - **Reconstructed human epidermis (RHE) tests** (EU B.46/OECD TG 439): These tests are considered scientifically valid for the prediction of irritant (Category 2) and non-irritant (no category) substances for Annex VII purposes, and also Annex VIII according to the rules laid down in Annex XI (see Section R.7.2.6 of this Guidance).

  The specific scope and limitations of these tests are:

  - They discriminate skin irritants (Category 2) from substances not classified for skin irritation (no Category) under CLP. However, they cannot discriminate skin irritants (Category 2) from skin corrosives (Category 1). The latter discrimination needs to be addressed with an *in vitro* skin corrosion test;
  - Cell viability in these models is measured by the MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, Thiazolyl blue) assay. If a test substance acts directly on the MTT (e.g. is a direct MTT-reducer), is naturally coloured, or becomes coloured during tissue treatment, additional controls should be used to detect and correct for test substance interference with the viability measurement technique. Detailed description of how to correct for direct MTT reduction and interferences by colouring agents is available in the Standard Operating Procedures (SOPs) for the four validated test methods and referenced in the OECD and EU TGs [37];
  - This test method may not be applicable to all groups of chemical classes. For example metals or inorganic metal compounds were not included in the

---

36 Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

37 A revision of OECD TG 439 including the use of HPLC/UPLC-spectrophotometry as an alternative way to measure MTT formazan is currently under discussion at the OECD with a high probability of adoption in April 2015. If this revision is accepted, it will reduce the limitation of these test methods towards strongly coloured substances.

validation study and there is experience that some metals (e.g. cobalt) may give a false positive result;

- They do not allow testing of gases and aerosols.

- For skin corrosion:

  - **Transcutaneous electrical resistance (TER) test method** (EU B.40/OECD TG 430)
  - **Reconstructed Human Epidermis (RHE) test method** (includes more than one protocol) (EU B.40 bis/OECD TG 431)
  - ***In vitro* membrane barrier test method** (OECD TG 435)

  All the above-mentioned tests allow for the discrimination of skin corrosives (Category 1) from non-corrosive substances.

  The specific scope and limitations of these tests are:

  - None of them allows testing of gases and aerosols;
  - Only the ***in vitro* Membrane Barrier** test method for skin corrosion is accepted to discriminate between skin corrosive subcategories 1A, 1B and 1C and non-corrosives;
  - The ***in vitro* Membrane Barrier** test method has a limited applicability domain (only acids, bases and acid derivatives). In addition, test materials not causing detectable changes in the detection system (e.g. typically 4.5 < pH < 8.5) cannot be tested;
  - The **RHE** test method can be used to distinguish subcategory 1A from subcategories 1B and 1C. The protocol of **EpiSkin,** which is one of the four validated methods included in the RHE test guideline, also allows for the discrimination of subcategory 1B from subcategory 1C and, according to the OECD IATA (OECD, 2014b), this information may be used in a *Weight-of-Evidence* assessment;
  - TER cannot be used to subcategorise skin corrosive substances;
  - The use of the **RHE** test method may not be applicable to all groups of chemical classes. For example there is reasonable doubt on the adequacy of this model for certain groups of fatty amine derivatives where RhE assays did not predict corrosivity, whereas these substances were corrosive in *in vivo* rabbit studies (Houthoff *et al.*, 2014). Furthermore, metals or inorganic metal compounds were not included in the validation study and there is experience that some metals (e.g. cobalt) may give a false positive result;

In relation to cell viability measurement by the MTT assay in **RHE** models, the same limitations as those specified above for the *in vitro* skin irritation test (EU method B.46/OECD TG 439) apply.

- Quality aspects of existing *in vitro* data:

For quality assessment of existing *in vitro* data that will form the basis for later possible *Weight-of-Evidence* considerations, see Section R.4.4 of Chapter R.4 of the Guidance on IR&CSA, and for aspects that need to be taken into account in such a *Weight of Evidence* see Section R.5.2.1.2 of Chapter R.5 of the Guidance on IR&CSA.

Animal data

Well-reported studies, particularly if conducted in accordance with the principles of GLP, can be used to identify substances which would be considered to cause, or not to cause, skin corrosion or skin irritation. There may be a number of skin corrosion/irritation studies already available for an existing substance, none of which are fully equivalent to an OECD TG or an EU test method such as those in the Annex to the EU Test Methods Regulation. If the results from such a batch of studies are consistent, they may, together, provide sufficient information on the skin corrosion/irritation potential of the substance.

If the results from a variety of studies are unclear, based on the criteria given below for evaluation of the data, the registrant will need to decide which of the studies is/are most reliable, relevant for the endpoint in question and adequate for classification purposes.

Particular attention should be given to the persistence of irritation effects, even those which do not lead to classification. Effects such as erythema, oedema, fissuring, scaling, desquamation, hyperplasia and opacity which do not reverse within the test period may indicate that a substance will cause persistent damage to the human skin.

Data from studies other than skin corrosion/irritation ones (e.g. other toxicological studies on the substance in which local responses of skin have been reported) may provide useful information though they may not be well reported in relation to, for example, the basic requirements for information on skin irritation. However, it should be noted that skin reactions and symptoms are not systematically scored in e.g. acute and sub-acute dermal toxicity studies since these studies are not specifically designed to address skin corrosion/irritation.

- Quality Aspects of existing *in vivo* data:

Data from **existing** irritation studies in animals must be taken into account before further testing is considered. A quality assessment of any such studies should be done using, for example, the system developed by Klimisch *et al*. (1997), as described in Section R.4.2 of Chapter R.4 of the *Guidance on IR&CSA*, and a judgement will need to be made as to whether any further testing is required. Some examples to note are:

i. Was the animal species used the rabbit or was it another species such as the rat or the mouse? The rat and the mouse are not as sensitive as the rabbit for irritation testing.

ii. How many animals were used? Current methodology requires a maximum of 3 animals tested in a sequential manner (with 1 animal being sufficient if skin corrosion effects are observed in the first tested animal, or 2 animals being sufficient if consistent effects are observed in the first and the second tested animals) but 6 were frequently used in the past (See Section 3.2.2.3.2.2 of the *Guidance on the Application of the CLP criteria* for the evaluation of results from tests that have been conducted with more than 3 animals).

iii. How many dose levels were used? If dilutions were included, what solvent was used (as this may have influenced absorption)? Which dose volume was used?

iv. Which exposure period was used? Single or repeated exposure?

v. The method used to apply the substance to the skin should be noted i.e. whether occluded or semi-occluded and whether the application site was washed after treatment.

vi. Check the observation period used post-exposure. Shorter periods than those in the current guideline may be adequate for non-irritants but may require a more severe classification for irritants when the observation period is too short to measure full recovery.

Irritation scores from old reports, reports produced for regulatory submission in the USA or in publications may be expressed as a Primary Irritation Score. Without the original data it is not always possible to convert these scores accurately into the scoring system used in the EU. For extremes, i.e. where there is either no irritation or severe irritation, it may not be necessary to look further, but average irritation scores pose a problem and expert judgement may be required to avoid repeat testing.

Observations such as those above can all be used to assess whether the existing animal test report available can be used to reliably predict the irritation potential of a substance, thus avoiding further testing.

R.7.2.4.2 Human data on skin corrosion/irritation

Well-documented *existing* human data from different sources can often provide very useful information on skin corrosion/irritation , sometimes for a range of exposure levels. Often the only useful information available on irritation is obtained from human experience (e.g. occupational settings). The usefulness of all human data on irritation will depend on the extent to which the effect, and its magnitude, can be reliably attributed to the substance of interest.

The quality and relevance of existing human data for hazard assessment should be critically reviewed. For example, in occupational studies with mixed exposure it is important that the substance causing skin corrosion or skin irritation has been accurately identified. There may also be a significant level of uncertainty in human data due to poor reporting and lack of specific information on exposure.

Examples of how existing human data can be used in hazard classification for irritation are provided in an ECETOC monograph (ECETOC, 2002).

Human data on local skin effects may be obtained from existing data on single or repeated exposure. The exposure could be of accidental nature or prolonged, for example in occupational settings. The exposure is usually difficult to quantify. When looking at the effects, corrosivity is characterised by destruction of skin tissue, namely visible necrosis through the epidermis and into the dermis. Corrosive reactions are typified by ulcers, bleeding and bloody scabs. After recovery the skin will be discoloured due to blanching of the skin and will present complete areas of alopecia and scars.

In addition to human data on local skin effects (which originate from clinical and occupational studies, poison information centres, case reports and retrospective epidemiological studies) existing human data from skin irritation human patch testing (HPT) might also be available. HPT is a controlled study involving the exposure of small patches of skin of human volunteers to substances for which skin corrosion and other unacceptable toxicological hazards can be excluded. HPT data have been compiled for example by Jírová *et al.* (2010), Basketter *et al.* (2012), as well as Ishii *et al.* (2013). Testing with human volunteers to obtain primary hazard data on skin corrosion/irritation for regulatory purposes is discouraged. Available good quality data should nevertheless be considered as appropriate and used for Classification and Labelling decision making. It should however be noted that the CLP Regulation does not contain clear criteria for classification for skin irritation based on human data.

R.7.2.4.3 Exposure considerations for skin corrosion/irritation

Exposure-based waiving from testing is not applicable to the endpoints of skin corrosion/irritation. Exposure-based waiving from testing as specified in Annex XI (3) of the

REACH Regulation only applies to tests listed in Sections 8.6 and 8.7 of Annex VIII, Annex IX and Annex X according to the REACH text.

R.7.2.4.4 Remaining uncertainty on skin corrosion/irritation

Usually it is possible to unequivocally identify (or accept) a substance as being corrosive, whatever type of study provides the information.

There may be a significant level of uncertainty in human data on irritant effects (e.g. because of poor reporting, lack of specific information on exposure, subjective or anecdotal reporting of effects, small number of subjects).

Data from studies in animals and from *in vitro* tests performed according to internationally accepted test methods will usually give relevant information on the skin corrosion/irritation potential of a substance. In general, it is assumed that substances which cause skin corrosion/irritation in EU or OECD TG-compliant studies in animals or *in vitro* will cause skin corrosion/irritation in humans, and those which are not irritant in EU or OECD TG-compliant studies will not be irritant in humans (Please note that in general test animals are considered to be more sensitive to skin corrosion/irritation effects than humans (e.g. OECD, 2014b)). It should be borne in mind that one of the limitations of the *in vivo* corrosion/irritation studies is the subjective grading of the lesions. Moreover, inconsistent results from a number of similar studies increase the uncertainty in deriving data from animal or *in vitro* studies.

The scope of the *in vitro* tests for corrosion/irritation has also some limitations, as explained in Section R.7.2.4.1 under “Testing data on corrosion/irritation”. In addition inconsistent results from two or more *in vitro* tests could add to the overall uncertainty in interpretation of the data.

R.7.2.5 Conclusions on skin corrosion/irritation

R.7.2.5.1 Concluding on suitability for Classification and Labelling

In order to conclude on Classification and Labelling according to the CLP Regulation, all the available information needs to be taken into account and consideration should be given to both the *Guidance on the Application of the CLP criteria* and the various remarks (related to Classification and Labelling) made throughout this guidance document [38].

R.7.2.5.2 Concluding on suitability for Chemical Safety Assessment

A dose-response assessment is difficult to make for skin corrosion/irritation simply because up to now most data have been generated for undiluted substances in accordance with test guidelines and traditional practice (which continues today). From a risk characterisation perspective it is therefore advisable to use the outcome of the classification procedure, i.e. a substance that is classified is assumed to be sufficiently characterised. However, a complete risk assessment requires both hazard and dose-response data and for local effects the concentrations is often the determinative dose metric. Consequently, if dose-response data are available, they must be taken into account (Figure R.7.2–1). For instance, dose-response information might be available from sub-acute or sub-chronic dermal toxicity studies (as such

[38] Please note that the 8th Adaptation to Technical and Scientific Progress (ATP) of the CLP Regulation is currently under discussion. The 8th ATP will take into account the 5th Revision of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which was adopted in 2012 and contains in particular refined criteria for skin corrosion/irritation and serious eye damage/eye irritation.

studies require a determination of a non-irritant dose in the dose selection), or from human experience and may in certain cases be determined using *in vitro* studies. However, when information is used from existing dermal toxicity studies (e.g. repeated dose), it should be noted that the test conditions do not reflect the test conditions used in the *in vivo* skin corrosion/irritation study: e.g. test material is applied in dilution *vs.* neat, vehicles/solvents are often used, exposure duration is different and test material application areas differ (see Module 5 of the OECD IATA (OECD, 2014b)).

Guidance on the possibilities for derivation of DNELs for skin corrosion/irritation is given in Appendix R.8-9 of Chapter R.8 of the *Guidance on IR&CSA*.

R.7.2.5.3 Information not adequate

A *Weight-of-Evidence* approach comparing available adequate information with the tonnage-triggered information requirements under REACH may result in the conclusion that the requirements are not fulfilled. In order to proceed to further information gathering, the testing and assessment strategy described in Section R.7.2.6 below is recommended.

R.7.2.6 Testing and assessment strategy for skin corrosion/irritation

The OECD has approved an IATA for skin corrosion/irritation (OECD, 2014b), which includes a description of various types of data that can be used in the assessement of these hazards. The IATA has a modular approach, whereby the *domain, role in IATA, strengths, weaknesses and limitations* of each type of data are given in a tabular form. Some parts of the IATA provide more detailed scientific background than the present document. Furthermore, the IATA gives detailed guidance on the *Weight-of-Evidence* approach. At the *Weight-of-Evidence* step, all existing information is integrated and assessed in order to decide whether further *in vitro* testing of the substance (or *in vivo* testing as a last option if *in vitro* testing is not possible or not conclusive) is necessary. While the OECD IATA provides slightly more detailed guidance than the testing and assessment strategy below, there is no conceptual difference between the two.

R.7.2.6.1 Objective / General principles

The following testing and assessment strategy is recommended for developing adequate and scientifically sound data for assessment/evaluation and classification of the skin corrosive and skin irritating properties of substances. For existing substances with insufficient data, this strategy can also be used to decide which additional data, beside those already available, are needed. The testing and assessment strategy is aimed at the identification of skin corrosion/irritation by using different elements where appropriate, depending on the information available. A basic principle of the strategy is that the results of one study or from an information source are evaluated before another study is initiated. The strategy seeks to ensure that the data requirements are met in the most efficient and humane manner so that animal usage and costs are minimised.

The different elements provided in Figure R.7.2–1 describe information sources that can be used to conclude on a substance's hazard potential towards skin. The elements described in Figure R.7.2–2 can be rearranged as appropriate, especially those in Part 1. This may be particularly helpful in cases where a conclusion can be drawn from certain elements without having to consider all of them. If judged relevant, elements in Part 1 can be omitted and *in vitro* testing can be performed immediately.

Figure R.7.2–2 is divided into three parts whereby Part 1 aims at evaluating existing information that may be available on the substance. In Part 2 existing information and relevant data should be assessed in order to consider whether there is enough information available to conclude on the substance hazard properties within a *Weight-of-Evidence* analysis, in case it is not possible to make a conclusion based on single elements described in Part 1. In case no conclusion can be drawn from Parts 1 and 2, new data should be generated in Part 3 by first performing relevant *in vitro* testing. Only in case no conclusion can be drawn based on the *in vitro* testing, can *in vivo* testing be performed (for substances at or above 10 tonnes per annum only).

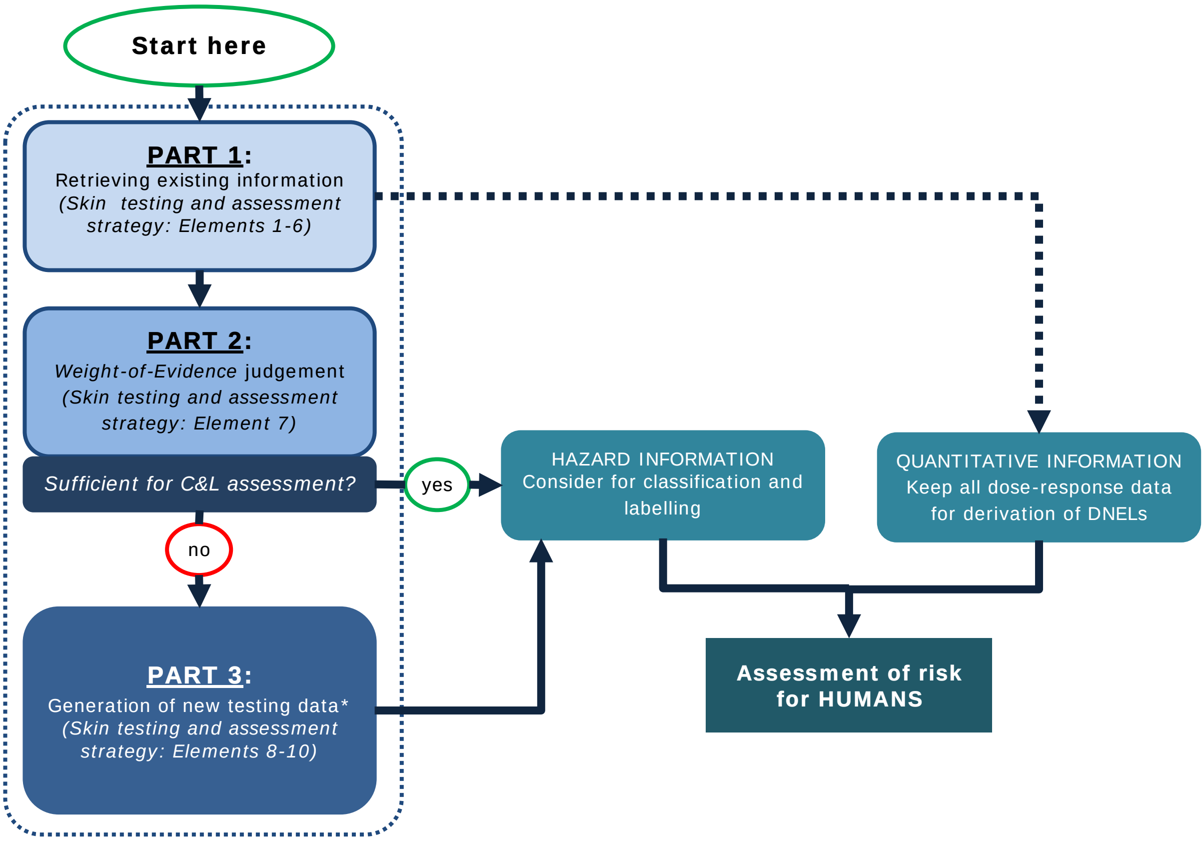
Some guidance for testing is provided by the specific rules for adaptation from standard information requirements, as described in column 2 of Annexes VII-X to the REACH Regulation, together with some general rules for adaptation from standard information requirements in Annex XI.

Risk assessment of the skin corrosion/irritation potential of a substance is normally made in a qualitative way provided that the substance has been classified as being corrosive or irritant to the skin. Existing test guidelines do not contain dose-response assessment, consequently a quantitative analysis will often not be possible. Therefore, hazard identification and appropriate classification is the key determinant in the information gathering strategy below. As a consequence, the use of *Assessment Factors* is of limited use in order to take into account uncertainty of data. However, the registrant is encouraged to keep and use all quantitative data that might be encountered in the process of retrieving hazard information in the context of the present testing strategy and to perform a complete risk assessment, comprising assessment of qualitative hazard as well as quantitative information.

It is recommended that the testing and assessment strategy be followed until element 6 (Figure R.7.2–1 and Figure R.7.2–2 in all cases and thereafter the *Weight-of-Evidence* analysis be performed. Clearly, all information sources/elements can be rearranged as appropriate, i.e. not all elements will necessarily be accompanied by data but it is important that all potential data sources are explored prior to starting the *Weight-of-Evidence* analysis. While it is recommended that this approach be followed, other approaches may be deemed more appropriate and efficient on a case-by-case basis. For example, in case there is no existing data and it is anticipated that generation of “pre-testing data” would be non-conclusive, it may be appropriate to directly proceed to the information generation part. Furthermore, prior to performing any new *in vivo* test, the use of *in vitro* methods should be fully exploited (see Articles 13(1) and 25(1) of the REACH Regulation) by using the general rules of Annex XI for adaptation of the standard testing regime set out in Annexes VII to X.

If the substance is not classified for skin corrosion/irritation, no risk assessment for this endpoint is performed, regardless of the exposure. Please note that there are no options for exposure-based waiving for these endpoints in the REACH Regulation.

The following flow chart (Figure R.7.2–1) gives an overview of a possible approach for defining a testing and assessment strategy for skin corrosion and irritation.



*Generation of new testing data according to Annex VII to VIII to the REACH Regulation and with due observation of the rules for adaptation of the standard testing regime laid down in Annex XI.[39]

Figure R.7.2–1 Overview of the testing and assessment strategy for skin corrosion/irritation

R.7.2.6.2 Testing and assesment strategy for skin corrosion/irritation

Recommended approach

The testing and assessment strategy presented here comprises three parts (see Figure R.7.2–2): Part 1 (elements 1 to 6) is about retrieving existing information, Part 2 (element 7) represents a *Weight-of-Evidence* analysis and expert judgement, and Part 3 (elements 8 to 10) is about the generation of new information by testing.

In Part 1, existing and available information from the literature and databases is gathered and considered in the strategy approach. The order of the different elements, i.e. 1 to 6, is only indicative and they may be arranged as appropriate. This may be especially helpful in cases

[39] Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

where a reliable conclusion can be drawn from certain elements without having to consider all of them. For instance, if the substance has an extreme pH (≤ 2.0 or ≥ 11.5) skin corrosivity is considered implicit (element 1c) and therefore the substance should be classified as skin corrosive (Category 1) according to CLP and further testing is not required. At the end of Part 1, and if no final conclusion could be derived directly from one or several of the available pieces of information, all the information collected should be analysed using a *Weight-of-Evidence* approach (element 7).

In the information generation part (elements 8 to 10), new information on the corrosion/irritation potential of substances is produced by means of *in vitro* (elements 8 and 9) or, as a last resort (see Articles 13(1) and 25(1) of the REACH Regulation), *in vivo* testing (element 10). Therefore, before concluding the *Weight-of-Evidence* analysis in element 7 and *in vitro* testing (elements 8 and 9), new *in vivo* tests should not be conducted. More information on how to use the *in vitro* methods for skin corrosion/irritation within the testing strategy can be found in the following paragraphs.

While it is recommended that this approach be followed, other approaches may be more appropriate and efficient on a case-by-case basis. For example, in case there is no existing data and it is anticipated that compilation of data at elements 1-7 would be non-conclusive, it may be appropriate to directly proceed to the information generation part.

Figure R.7.2–2 Testing and assessment strategy for evaluating the skin corrosion/irritation potential of substances (footnotes a to h are detailed below the figure).

Element	Information	Conclusion [40]
Existing data on physico-chemical properties		
1a	Is the substance spontaneously flammable in contact with air (pyrophoric) or water at room temperature? →	**YES**: No testing required (Column 2 adaptation in section 8.1 of Annexes VII and VIII)
1b	Is the substance an organic hydroperoxide or an organic peroxide? →	**YES**: Consider classifying as: ▪ corrosive (Skin Corrosive Cat. 1B) if the substance is a hydroperoxide, or ▪ irritant (Skin Irritant Cat. 2) if the substance is a peroxide. **OR** Provide evidence supporting deviating classification or non-classification [41].
1c	Is the pH of the substance ≤ 2.0 or ≥ 11.5? [a] →	**YES**: Consider classifying as corrosive (column 2, section 8.1. of Annex VIII) if pH is used as the sole basis for classification decision. Where classification is based upon consideration of pH alone, subcategorisation is not possible and therefore Skin Corrosive Cat.1 should be applied.
1d	Are there other physical or chemical properties that indicate that the substance is corrosive/irritant? →	**YES**: Use this information for *Weight-of-Evidence* analysis (Element 7).
Existing human data		
2	Are there adequate existing human data [b] which provide evidence that the substance is a corrosive or irritant? →	**YES**: Consider classifying accordingly.

[40] Please note that the 8th Adaptation to Technical and Scientific Progress (ATP) of the CLP Regulation is currently under discussion. The 8th ATP will take into account the 5th Revision of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which was adopted in 2012 and contains in particular refined criteria for skin corrosion/irritation and serious eye damage/eye irritation.

[41] Information on e.g. *in vitro* testing may provide evidence on a more suitable classification, if there is some doubt on the correct classification.

Existing animal data from corrosion/irritation studies		
3	Are there data from existing studies *on corrosion and irritation* in laboratory animals, which provide sound conclusive evidence that the substance is a corrosive, irritant or non-irritant? →	**YES**: Consider classifying accordingly (either Skin Corrosive Cat. 1, 1A, 1B, 1C or Skin Irritant Cat. 2) or consider no classification.
Existing data from general toxicity studies via the dermal route and from sensitisation studies		
4a	Is the substance classified as fatal in contact with skin ($LD_{50} \leq$ 50 mg/kg bw, CLP hazard statement H310)? [c] →	**YES**: The substance will be classified for acute dermal toxicity (column 2 adaptation in section 8.1 of Annexes VII and VIII). No new testing for skin irritation/corrosion is needed in this case.
4b	Has the substance proven to be a corrosive, irritant or non-irritant in a suitable acute dermal toxicity test? [d] →	**YES**: If test conditions are consistent with OECD TG 404, consider classifying accordingly (Skin Corrosive Cat. 1, 1A, 1B, 1C or Skin Irritant Cat. 2) or consider no classification.
4c	Has the substance proven to be a corrosive or an irritant in sensitisation studies or after repeated exposure? [e] →	**YES**: This information cannot be used for considering a concrete classification conclusion but must be used exclusively within the integrated *Weight-of-Evidence* judgement.
Existing/new (Q)SAR data and read-across		
5a	Are there structurally related substances (suitable “read-across” or grouping), which are classified as corrosive to the skin (Skin Corrosive Cat. 1), or do suitable (Q)SAR methods indicate corrosion potential of the substance? [f] →	**YES**: Consider classifying as Skin Corrosive Cat. 1.
5b	Are there structurally related substances (suitable “read-across” or grouping), which are classified as irritant to the skin (Skin Irritant Cat. 2), or indicating that the substance is non-irritant, or do suitable (Q)SAR methods indicate irritant or non-irritant potential of the substance? [f] →	**YES**: Consider classifying accordingly.
Existing in vitro data		
6a	Has the substance demonstrated corrosive properties in an EU/OECD adopted *in vitro* test? Data from *in vitro* test methods that have been validated and are considered scientifically valid but are not yet adopted by EU and/or OECD may also be used if the provisions defined in Annex XI are met. →	**YES**: Consider classifying as corrosive. If discrimination between Skin Corrosive Cat. 1A, 1B and 1C is not possible, Cat. 1 must be chosen. If a negative result is obtained and there is no existing data from (an) *in*

		vitro skin irritation study(ies), the irritation potential must be determined, e.g. with an *in vitro* skin irritation test.
6b	Has the substance demonstrated irritant or non-irritant properties in an EU/OECD adopted *in vitro* test? Data from *in vitro* test methods that have been validated and are considered scientifically valid but are not yet adopted by EU and/or OECD may also be used if the provisions defined in Annex XI are met. →	**YES**: Consider classifying accordingly (Skin Irritant Cat. 2) or consider no classification. If a positive result is obtained and there is no exisiting data from (an) *in vitro* skin corrosion study(ies), the corrosion potential must be determined e.g. with an *in vitro* skin corrosion test (Element 8).
6c	Are there data from (a) non-validated suitable *in vitro* test(s), which provide sound conclusive evidence that the substance is corrosive/ irritant? [g] →	**YES**: Consider classifying accordingly (Skin Corrosive Cat 1, 1A, 1B, 1C or Skin Irritant Cat. 2).
Weight-of-Evidence analysis		
7	The "elements" described above may be arranged as appropriate. Taking all available existing and relevant data mentioned above (Elements 1-6) into account, is there sufficient information to make a decision on whether classification/labelling is necessary, and – if so – how to classify and label? →	**YES**: Classify accordingly (Skin Corrosive Cat. 1, 1A, 1B, 1C or Skin Irritant Cat. 2) or consider no classification. If discrimination between Skin Corrosive Cat 1A, 1B and 1C is not possible, Cat. 1 must be chosen.
New in vitro tests for corrosivity [g]		
8	Does the substance demonstrate corrosive properties in (an) EU/OECD adopted *in vitro* test(s) for skin corrosion? → Data from *in vitro* test methods that have been validated and are considered scientifically valid but are not yet adopted by EU and/or OECD may also be used if the provisions defined in Annex XI are met.	**YES**: Classify accordingly (Skin Corrosive Cat. 1A, 1B or 1C). If discrimination between Cat. 1A, 1B and 1C is not possible, Cat. 1 must be chosen. If a negative result is obtained, the irritation potential of the substance must be determined, e.g. with an *in vitro* skin irritation test (Element 9), in order to determine if the substance should be classified as Skin Irritant Cat. 2 or not classified.
New in vitro tests for irritation [g]		
9	Does the substance demonstrate irritating or non-irritating properties in (an) EU/OECD adopted *in vitro* test(s) for skin irritation? Data from *in vitro* test methods that have been validated and are considered scientifically valid but are not yet adopted by EU and/or OECD may also be used if the provisions defined in Annex XI are met. →	**YES**: Classify accordingly (Skin Irritant Cat. 2) or consider no classification. If a positive result is obtained and there is no existing data from (an) *in vitro* skin corrosion study(ies), the corrosion potential must be determined e.g. with an *in vitro* skin corrosion test

		(Element 8). If a conclusion on skin corrosion/irritation cannot be drawn by using *in vitro* testing, *in vivo* testing should be performed (at Annex VIII level only).
New in vivo test for corrosion/irritation as a last resort (Annex VIII to the REACH Regulation) [h]		
10	Does the substance demonstrate corrosive or irritant properties in an EU/OECD adopted *in vivo* test? →	**YES**: Classify accordingly (Skin Corrosive Cat. 1, 1A, 1B, 1C or Skin Irritant Cat. 2). **NO:** No classification needed.

Notes to the information scheme on skin corrosion/irritation:

a) Note that if the buffering capacity suggests that the substance may not be corrosive, further data are needed to confirm this, preferably using an appropriate *in vitro* test method.

b) Data from case reports, occupational experience, poison information centres, HPTs or from clinical studies.

c) If the substance is classified as fatal in contact with skin ($LD_{50} \leq 50$ mg/kg bw), further testing for skin corrosion/irritation would result in severe suffering or death of the animal. Thus, further testing is not required and sufficient labelling (warning) is provided by the Hazard statement H310 *"Fatal in contact with skin"* and the GHS Pictogram GHS06 with the signal word "Danger". The classification as fatal in contact with skin requires strict risk management measures and hence, since all contact with the skin must be avoided, there is no need to investigate the skin corrosion/irritation potential further. In case existing information on skin corrosion/irritation is available, it should be included in the registration dossier and used for classification and labelling for skin corrosion/irritation.

d) Has the substance proven to be either an irritant or a corrosive in an acute dermal toxicity test carried out with **rabbits** with the undiluted test substance (liquids) or with a suitable suspension (solids)? In case of signs of skin corrosion, classify as Skin Corrosive (subcategorisation as 1A, 1B or 1C, where possible). In all other cases: calculate or estimate the amount of test substance per cm^2 and compare this to the test substance concentration of 80 µl or 80 mg/cm^2 employed in the EU B.4/OECD TG 404 for dermal corrosion/irritation test with rabbits. If in the same range and adequate scoring of skin effects is provided, classify or not as Skin Irritant Category 2. In case conclusive negative data was obtained in rabbits, stop. If not in the same range and inadequate scoring of skin effects, use for *Weight-of-Evidence* analysis and proceed.

In case the test was performed in other species, which may be less sensitive (e.g. **rat**), evaluation must be made with caution. Usually, the rat is the preferred species for toxicity studies within the EU. The limit dose level of 2000 mg/kg bw of a solid is normally applied as a 50% suspension in a dose volume of 4 ml/kg bw onto a skin surface area of about 5x5 cm. Assuming a mean body weight of 250 g, a dose of 1 ml of the suspension will be applied to an area of 25 cm^2, i.e 20 mg test substance per cm^2. In case of an undiluted liquid, 0.5 ml is applied to 25 cm^2, i.e. 20 µl/cm^2. Considering the fact that the rat skin is less sensitive compared to rabbit skin, much lower exposures are employed and, in general, the scoring of dermal effects is performed less accurately, the results of dermal toxicity testing in rats will not be adequate for classification with respect to skin irritation. Only in case of evidence of skin corrosivity in the rat dermal toxicity test can the test substance be classified as Skin Corrosive Category 1. All other data should be used for *Weight of Evidence*.

e) Regarding data from skin sensitisation studies, the skin of guinea pigs is less sensitive than that of rats which is, in turn, less sensitive than that of rabbits. Only in case of evidence of skin corrosivity in the sensitisation test (Maximisation or Buhler) with the neat material or dilutions of solids in water, physiological saline or vegetable oil, should the test substance be classified as Skin Corrosive Category 1. However, care should be exercised when interpreting findings from guinea pig studies, particularly from maximisation protocols, as intradermal injection with adjuvant readily causes necrosis. All other data should be used for *Weight of Evidence* only. Information on irritant properties from skin sensitisation tests cannot be used to conclude on a specific classification regarding acute skin irritation but may be used in a *Weight-of-Evidence* analysis. In general, irritation data from the Local Lymph Node Assay are not usable. The test substance is applied to the dorsum of the ear by open topical application, and specific vehicles for enhancement of skin penetration are used.

f) Conclusion on no classification can be made if the *in silico* model has been shown to predict adequately the absence of the classified effect and if it also fulfils the requirements of Annex XI to the REACH Regulation. Prediction of the absence of the classified effect can be made either by triggering an exclusion rule in the BfR system (to be checked on a case-by-case basis), or based on a negative prediction in a classification QSAR that was trained on both positive and negative substances. The suitability of the model (reliability, relevance) should be very carefully checked to make sure that the prediction is fit for purpose, and the applicability of the model to the substance should also be justified (e.g. fulfilment of the conditions of Section 1.3 of Annex XI to the REACH Regulation should be checked). For read-across, generation of new *in vivo* data should be avoided.

g) New *in vitro* testing should be performed following a top-down or bottom-up approach. Please see the following paragraph "How to use the *in vitro* methods for skin corrosion/irritation within the strategy". While it may be appropriate to use information from non-validated *in vitro* tests if already existing, it is highly recommended to adhere to the test protocols whose scientific validity has been established by formal validation and which, ideally, have been officially adopted by the European Commission and/or by the OECD. Data obtained from suitable non-validated suitable *in vitro* tests can only be used according to the criteria set out in Annex XI, section 1.4 of the REACH Regulation, i.e. only positive results can be accepted.

h) *In vivo* testing should not be conducted in case the substance falls under the scope of the specific *in vitro* tests performed, and there are no substance-specific limitations on use of those tests and the Registrant formulates an adaptation according to Annex XI to the REACH Regulation. Due to the current standard *in vivo* information requirement at Annex VIII level and above, an adaptation needs to be built up in a registration dossier in order to successfully submit a compliant dossier [42].

---

42 Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

How to use the *in vitro* methods for skin corrosion/irritation within the testing and assessment strategy

For skin corrosion and irritation no single *in vitro* test method can fully replace the *in vivo* test (EU TM B.4 / OECD TG 404) across the full range of skin responses. However, the *in vitro* methods specified in Section R.7.2.3.1 and R.7.2.4.1 may replace the *in vivo* test depending on the outcome of the study or when combined within a tiered testing strategy.

New *in vitro* testing should be performed following a top-down or bottom-up approach (Figure R.7.2–3). The top-down approach should be used when available information suggests that the substance may be irritant or corrosive to the skin. The bottom-up approach, on the other hand, should be followed only when available information suggests that the substance may not be irritant to the skin.

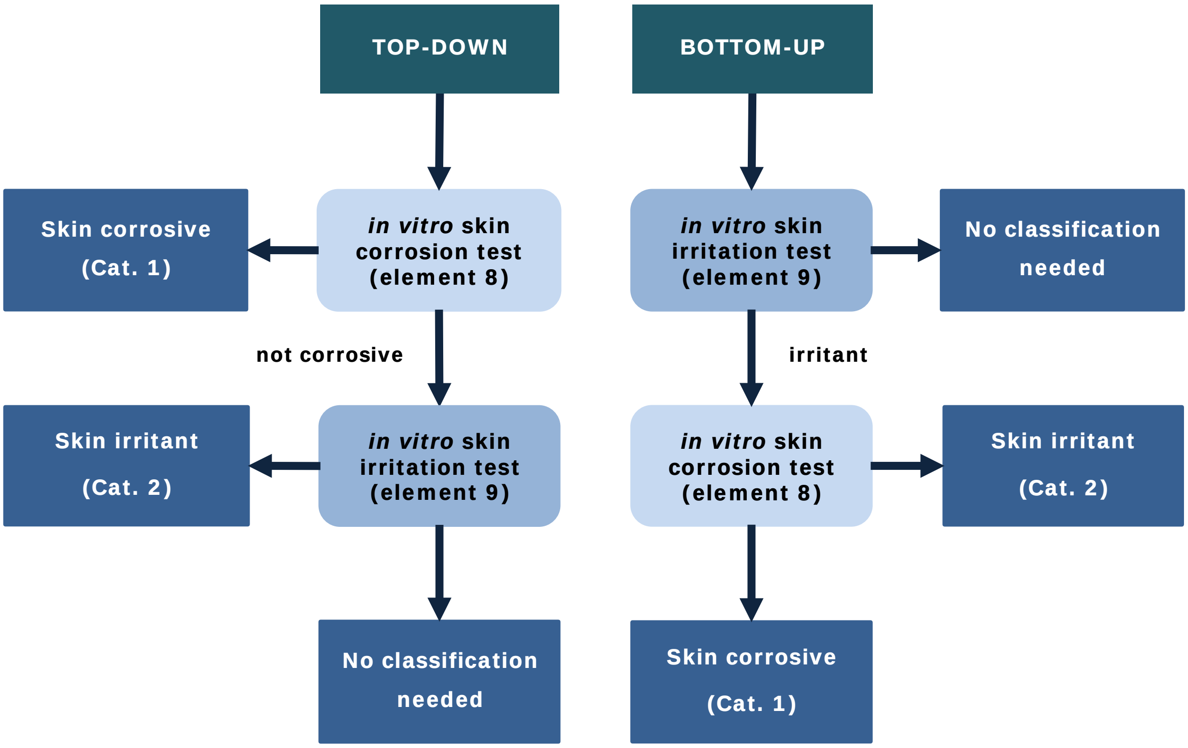


Figure R.7.2–3 Schematic presentation of Top-down and Bottom-up approaches for Skin Corrosion/irritation.

There are certain steps to be considered before any testing (*in vitro* or *in vivo*) is conducted. These steps are specified in Section 8.1 in column 1 of Annexes VII and VIII to the REACH Regulation and also in Figure R.7.2–2.

After these steps, no *in vivo* testing, as specified in section 8.1 of Annex VIII, is necessary in cases where:

a) the substance falls under the scope of the specific *in vitro* tests performed, and there are no substance-specific limitations to using those tests, and

b) the Registrant formulates an adaptation according to Annex XI to the REACH Regulation.

If an *in vivo* study for skin irritation is a standard information requirement (i.e. for substances registered at or above 10 tonnes per annum) and the steps above have been followed, the Registrant should choose to adapt the standard information requirement for the *in vivo* study by using Annex XI adaptation possibilities. Due to the current standard *in vivo* information requirement at Annex VIII level and above, an adaptation needs to be built up in a registration dossier in order to successfully submit it [43].

Instructions on how to submit *in vitro* information instead of *in vivo* can be found e.g. in Section 3.7 of *Practical Guide 1: How to report in vitro data* (available at http://echa.europa.eu/practical-guides).

It is important to note that it is the responsibility of the registrant to ensure that the chosen test method is suitable for the substance in order to obtain adequate information from the *in vitro* studies. For most substances, the use of adopted EU or OECD TGs for skin corrosion/irritation purposes will provide results that will have regulatory acceptance under REACH.

43 Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

SERIOUS EYE DAMAGE/EYE IRRITATION

R.7.2.7 Information requirements for serious eye damage/eye irritation

The information on serious eye damage/eye irritation that is required to be submitted for registration and evaluation purposes is specified in Annexes VI to XI to the REACH Regulation. According to Annex VI, the registrant should gather and evaluate all existing available information before considering further testing. This includes physico-chemical properties, (Q)SAR ((Quantitative) Structure-Activity Relationship), grouping, *in vitro* data, animal studies, and human data. For classified substances, information on exposure, use and risk management measures should also be collected and evaluated in order to ensure safe use of the substance.

If these data are inadequate for hazard and risk assessment, further testing should be carried out in accordance with the requirements of Annexes VII (≥1 tpa) and VIII (≥10 tpa) to the REACH Regulation.

R.7.2.7.1 Information requirements for quantities of ≥1 tpa (Annex VII to the REACH Regulation) [44]

If new testing data are necessary, these must be derived from *in vitro* methods only. Annex VII does not foresee *in vivo* testing for serious eye damage/eye irritation.

The standard information requirements at this tonnage level for serious eye damage/eye irritation (see Section 8.2 in Column 1 of Annex VII) can be satisfied by following three steps: (1) assessment of the available human and animal data, (2) assessment of the acid or alkaline reserve, (3) *in vitro* eye irritation study (Please note that when the REACH Regulation refers to the “eye irritation” endpoint, this covers both serious eye damage and eye irritation).

Section 8.2 in Column 2 of Annex VII lists specific rules for adaptation according to which step 3 is not necessary. These rules are applicable when:

4. the available information indicates that the criteria are met for classification as corrosive to the skin or irritating to eyes (Please note that when a substance is classified as Skin corrosive Category 1 under the CLP Regulation, the risk of severe damage to eyes is considered implicit and the substance is also classified in Category 1 for Serious eye damage), or

   the substance is flammable in air at room temperature (Please note that this rule should actually read: “the substance is **spontaneously** flammable in air at room temperature”).

[44] Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

R.7.2.7.2 Information requirements for quantities of ≥10 tpa (Annex VIII to the REACH Regulation) [45]

For substances manufactured or imported in quantities of ≥10 tpa *in vivo* testing is the standard information requirements of Annex VIII (Column 1) for serious eye damage/eye irritation, in case the information requirement cannot be met with the information obtained as specified in section 8.2 of Annex VII.

Before new tests are carried out to determine the properties listed in Annex VIII, all available *in vitro* data, *in vivo* data, historical human data, data from valid (Q)SARs and data from structurally related substances (read-across approach) must be assessed first. Due to the sequential nature of the REACH standard information requirements, it is reminded that at quantities of ≥10 tpa, the information requirements of Annex VII to the REACH Regulation also apply. This means that before a new *in vivo* test is performed, the appropriate *in vitro* testing must be undertaken according to the rules set out in section 8.2 of Annex VII and must be documented in the technical dossier (IUCLID). Finally, the information generated at Annex VII level must be taken into account in determining whether an *in vivo* test at Annex VIII level is really needed.

Column 2 of Annex VIII lists the following specific rules that allow deviation from the standard information required by Annex VIII for serious eye damage/eye irritation:

- the substance is classified as irritating to eyes with risk of serious damage to eyes (Please note that the reference to irritating to eyes with risk of serious damage to eyes here means Category 1 for Serious eye damage, according to the CLP Regulation), or

- the substance is classified as corrosive to the skin and provided that the registrant classified the substance as eye irritant (Please note that the reference to eye irritation here means Category 1 for Serious eye damage, according to the CLP Regulation), or

- the substance is a strong acid (pH ≤ 2.0) or base (pH ≥ 11.5), or

- the substance is flammable in air at room temperature (Please note that this rule should actually read: "the substance is **spontaneously** flammable in air at room temperature").

The *in vitro* methods that can be used to adapt the standard information requirements are detailed in Sections R.7.2.8.1 and R.7.2.9.1 of this Guidance, under "*In vitro* data". It should be noted that the use of an EU or OECD adopted *in vitro* test methods (one or several in combination) on serious eye damage/eye irritation may provide adequate information for the **replacement** of the regulatory *in vivo* test (EU TM B.5 / OECD TG 405). The standard testing requirement of Annex VIII should be adapted according to the general rules laid down in Annex XI, allowing avoidance of unnecessary animal testing as required in Annex VIII by the

---

[45] Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

use of non-testing data or *in vitro* testing (see Section R.7.2.8.1 of this Guidance for possible alternatives to animal testing), and in order to successfully submit a compliant dossier.[45]

Guidance on the application of these rules is given in the testing and assessment strategy for serious eye damage/eye irritation described in Section R.7.2.11 of this Guidance.

It should be noted that the conditions of acceptance by ECHA of implementation of any of the adaptation rules laid down in Annex XI are strict, and whenever an adaptation argument is being used (e.g. use of (Q)SARs, SARS, read-across or non-validated *in vitro* test methods), scientific justification, solid documentation and readiness for risk assessment and Classification and Labelling must be provided by registrants. For detailed information on these rules, see Annex XI to the REACH Regulation.

R.7.2.8 Information sources on serious eye damage/eye irritation

R.7.2.8.1 Non-human data on serious eye damage/eye irritation

Non-testing data on serious eye damage/eye irritation

Physico-chemical properties

Relevant information can be inferred from basic physico-chemical characteristics of a substance (e.g. extreme pH). *Extreme* pH values may indicate the potential of a substance to cause skin corrosion or serious eye damage:

If the pH is ≤ 2 or pH ≥ 11.5, then consider the substance to be corrosive to the skin (Category 1) and to cause serious eye damage (Category 1) when pH is used as the sole basis for the classification decision (See also Sections R.7.2.4.1 and R.7.2.9.1 of this Guidance).

Grouping, (Q)SARs and expert systems [46]

In REACH Annex XI two types of non-testing methods are mentioned which can be used for adaptation of standard information requirements, either as standalone (where possible) or in combination with other information (in the context of a *Weight-of-Evidence* assessment):

- qualitative and quantitative Structure-Activity-Relationships (SARs/QSARs, section 1.2, including expert systems, generally incorporating multiple (Q)SARs, expert rules and data) on the one hand, and
- grouping of substances and read-across approaches [47].

---

[46] Further information can be found in *Chapter R.6 QSAR and grouping of chemicals* of *the Guidance on IR&CSA*, the OECD Guidance on Grouping of Chemicals, Second Edition (OECD, 2014a), the new OECD Guidance on an Integrated Approach for Testing and Assessment (IATA) for skin corrosion and irritation (OECD, 2014b) and the JRC report on Alternative methods for regulatory toxicology (Worth, 2014).

[47] The relevant terminology is not always used consistently. With reference to the ECHA Guidance on QSAR and grouping, the terms category approach and analogue approach are used to describe techniques for grouping of substances, whilst the term read-across is reserved for a technique to fill data gaps, i.e. to transfer knowledge from one or more substances called source(s) to another substance with data gap, named target substance.

The adaptation of standard information requirements can be used for the assessment of serious eye damage/eye irritation, if it provides relevant and reliable data for the substance of interest. As specified in Annex XI to the REACH Regulation, the use of non-testing methods needs to be justified and sufficiently documented. In the case of QSARs and expert systems, registrants need to prepare property predictions by completion of a QSAR Prediction Reporting Format (QPRF). The QPRF is a harmonised template for summarising and reporting substance-specific predictions generated by (Q)SAR models. For filling a data gap under REACH, it is also necessary to provide information on the prediction model employed following a QSAR Model Reporting Format (QMRF) document. The QMRF is a harmonised template for summarising and reporting key information on (Q)SAR model validity, including the results of any validation studies. The information is structured according to the OECD (Q)SAR validation principles (for further information see http://www.oecd.org/env/ehs/risk-assessment/validationofqsarmodels.htm). The JRC QSAR Model Database is an inventory of information on available QMRFs, freely accessible online (https://eurl-ecvam.jrc.ec.europa.eu/databases/jrc-qsar-model-database). More detailed guidance on QSAR models, their use and reporting formats, including the QMRF, is provided in Section R.6.1 of Chapter R.6 of the *Guidance on IR&CSA*.

In general, there are several different ways in which non-testing methods can be used in the context of an IATA (an IATA for serious eye damage and eye irritation is currently under development by the OECD), e.g.:

- for direct prediction of serious eye damage/eye irritation potential or the absence thereof,
- as part of a *Weight-of-Evidence* scheme (where the information from non-testing methods alone is not sufficient for a decision), or
- in order to decide how best to proceed with further (*in vitro*) testing (i.e. *via* a top-down or bottom-up approach). For further information see Section R.7.2.11.2.

- SARs and read-across for serious eye damage and eye irritation:

In principle, the same considerations apply as with the use of SARs and read-across for skin corrosion/irritation (see Section R.7.2.3.1). Structural alerts for serious eye damage/eye irritation have been described in the literature, e.g. in Gerner *et al.* (2005).

The occurrence of structural analogues that exhibit serious eye damage (or eye irritation) potential can also be used to predict the effect in the substance of interest and adapt the respective information requirements. Negative data from structural analogues may also be used to make predictions in certain cases, however, absence of one of the known structural alerts for irritation and corrosion alone does not prove absence of effect, as knowledge of structural alerts for irritation and corrosion might be incomplete. For instance, other substructures (not yet identified as structural alerts) or other properties of the substance may be responsible for a corrosive or irritant effect.

- QSARs and expert systems for serious eye damage and eye irritation:

An overview of available (Q)SARs for serious eye damage/eye irritation is provided in Table R.7.2–3. An extensive review of the state-of-the-art was published by the former ECB (Gallegos Saliner *et al*. 2006, 2008). In Appendix R.7.2–3 some examples are given to illustrate currently available models and the techniques that have been used to develop them. Examples of models based on classical regression and classification techniques, together with more innovative approaches, are collected in Appendix R.7.2–3.

The most widely used expert systems for assessing eye irritation are the same as those used for assessing skin corrosion and irritation. Details on automated rule-induction systems (e.g. TOPKAT and MultiCASE), and on knowledge-based systems (e.g. DEREK Nexus, and the BfR rule-base) are reported in Appendix R.7.2–3.

The freely downloadable OECD QSAR Toolbox software contains two profilers relevant for serious eye damage/eye irritation based on the BfR rule-base, which encode “inclusion rules” (structural alerts predicting serious eye damage/eye irritation potential) with a suggestion that exclusion of serious eye damage/eye irritation potential might be possible based on certain physico-chemical properties. The use in combination of profilers and data for analogues could allow for the prediction of serious eye damage/eye irritation for new substances through a read-across or category approach. More details on the OECD QSAR Toolbox specific contents for skin irritation and corrosion are reported in Appendix R.7.2–3.

Not all of the models were developed with EU regulatory purposes in mind, so it is important to assess in each case whether the endpoint or effect being predicted corresponds to the regulatory endpoint of interest. The BfR model for the prediction of serious eye damage/eye irritation has been developed to predict EU regulatory endpoints, however predictions refer to the former DSD classification/labelling system used in the EU before the CLP Regulation came into force, and in borderline cases the results of the prediction may not fully reflect the correct CLP classification. More details on this model are reported in Appendix R.7.2–3.

It should also be noted that the criteria for classification as eye irritant Category 2 based on the mean score for corneal opacity and conjunctival redness in the *in vivo* test have changed from ≥2 and ≥2.5, respectively, under DSD to ≥1 and ≥2.0, respectively, under CLP. Consequently predictions as eye irritant Cat 2 from models developed based on the DSD criteria should be interpreted with caution since they may lead to underprediction and should not be used for direct classification under CLP.

In the case of classification models for serious eye damage/eye irritation, the classification criteria used in model development should be compared with the EU classification criteria, to assess the relevance of the model. Where it is not indicated in the supporting literature whether the predicted classification should be Category 1 (Serious eye damage) or Category 2 (Eye irritation), the category chosen should be supported with expert judgement.

Table R.7.2–3 Overview of available (Q)SARs for serious eye damage/eye irritation. See Appendix R.7.2–3 for more information on these models.

Category of model or source	Reference or name of the model	Applicability domain
Literature models	Solimeo *et al.* (2012)	Not available
	Abraham *et al.* (2003)	Pure bulk liquids
	Gerner *et al.* (2005)	Based on Physico-chemical values
	Barratt (1995b, 1997)	Neutral organic chemicals
Computerised models	PaDEL-DDPredictor (http://padel.nus.edu.sg/software/padelddpredictor/) (Liew and Yap, 2013)	Calculated by the model based on the range of descriptors
	BfR rule-base, free	EU New chemicals (NONS) database, organic chemicals with

	(included in the OECD QSAR Toolbox and Toxmatch, Toxtree, ToxPredict and Ambit)	no significant hydrolysis potential and purity >95%
	ACD/Percepta, commercial	Organic chemicals
	Derek Nexus, commercial	Organic chemicals and some metals
	HazardExpert, commercial	Organic chemicals
	MolCode, commercial	Organic chemicals
	MultiCASE, commercial	Organic chemicals
	TOPKAT, commercial	Organic chemicals
Review papers	Patlewicz *et al*., 2003	N.A.
	Gallegos Saliner *et al*. (2006, 2008)	N.A.

Abbreviation: N.A. = not applicable.

Testing data on serious eye damage/eye irritation

The internationally accepted test methods for serious eye damage/eye irritation as described in the Annex to the EU Test Methods (TM) Regulation (Council Regulation (EC) No 440/2008) and in OECD TGs (available at http://www.oecd.org/env/ehs/testing/oecdguidelinesforthetestingofchemicals.htm#Test_Guidelines) are: EU B.5 (OECD TG 405), EU B.47 (OECD TG 437), EU B.48 (OECD TG 438) and OECD TG 460.

At the OECD there are currently three additional draft TGs under discussion regarding the eye hazard, i.e. EpiOcular™ EIT, Short-time exposure (STE) test method and Cytosensor® microphysiometer (CM) test method. Additional test methods may become available for addressing the eye hazard, therefore the reader is advised to check the OECD website and ECHA's test methods webpage (http://echa.europa.eu/support/testing-methods-and-alternatives) to check the current status of these test methods.

Please note that the latest version of an adopted test guideline should always be used when generating new data, independently of whether it is published by the EU or OECD.

The testing and assessment strategy developed for serious eye damage/eye irritation (see Section R.7.2.11 of this Guidance) emphasises the need to evaluate all available information (including physico-chemical properties) before undertaking any *in vivo* testing. This strategy employs screening elements designed to avoid, as far as possible, *in vivo* testing of corrosive and severely irritating substances. In particular, *in vitro* tests should usually be performed first, and it should be assessed whether *in vivo* testing can be completely avoided.

In vitro data

Accepted *in vitro* test methods to detect serious eye damage (Category 1 under CLP) and/or absence of effects requiring classification for serious eye damage/eye irritation (i.e. not classified under CLP) are listed in Table R.7.2–4. More information on the specific scope and limitations of these tests is provided in Section R.7.2.9.1 under "Testing data on serious eye damage/eye irritation".

Table R.7.2–4 Accepted *in vitro* test methods for serious eye damage / eye irritation

Test method	Validation status, regulatory acceptance	EU Test Method / OECD test guideline	Classification according to CLP Regulation	EURL ECVAM DB-ALM protocol Nr.
Serious eye damage / eye irritation				
BCOP	Validated and regulatory acceptance	B.47 / OECD TG 437	Cat 1 or NC	98, 124
ICE	Validated and regulatory acceptance	B.48 / OECD TG 438	Cat 1 or NC	80
FL	Validated and regulatory acceptance	N.A. / OECD TG 460	Cat 1	71
CM [48]	Validated and considered to be scientifically valid	N.A. / OECD draft TG available and being considered for adoption	Cat 1 or NC	130
STE [49]	Validated and considered to be scientifically valid	N.A. / OECD draft TG available and being considered for adoption	Cat 1 or NC	N.A.
EpiOcular ™ EIT [50]	Validated and considered to be scientifically valid	N.A. / OECD draft TG available and being considered for adoption	NC	N.A.
Ocular Irritection® Assay [51]	Validated	N.A. / N.A.	Cat 1	157

48 The CM test method was validated by EURL ECVAM and considered to be scientifically valid (https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/topical-toxicity/eye-irritation; section 1.2) and was also reviewed by ICCVAM (http://ntp.niehs.nih.gov/?objectid=807EF83B-92CC-9A6C-3FFE8725DF1F9F5D); A draft OECD Test Guideline is available at: http://www.oecd.org/env/ehs/testing/section4healtheffects.htm.

49 The STE test method was validated by JaCVAM and peer-reviewed by ICCVAM and considered to be scientifically valid (http://ntp.niehs.nih.gov/?objectid=2D70C7A2-CCDB-D782-06CB38302BD7D10E); A draft OECD Test Guideline is available at: http://www.oecd.org/env/ehs/testing/section4healtheffects.htm.

50 The EpiOcular ™ EIT test method was validated by EURL ECVAM and considered to be scientifically valid (https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations); A draft OECD Test Guideline is available at: http://www.oecd.org/env/ehs/testing/section4healtheffects.htm.

Test methods currently with limited application under REACH					
	IRE [52]	validated	N.A. / N.A.	Cat 1	85
	HET-CAM [52]	Validated	N.A. / N.A.	Cat 1	47, 96

NOTE: During the validation exercise EURL ECVAM concluded that the SkinEthic ™ Human Corneal Epithelium (HCE) is not sufficiently sensitive for identifying substances not classified for serious eye damage/eye irritation (the test method produced an unacceptable number of false negative results in the validation study) and recommended optimisation and further validation of the test method by the developer (EURL ECVAM, 2015).

---

51 The Ocular Irritection® Assay has undergone an external prospective and retrospective validation study co-sponsored by In Vitro International (the method developer) and INT.E.G.RA (Eskes *et al.*, 2014) and appears to be a suitable test method for the identification of substances causing serious eye damage (CLP Category 1) and not requiring classification for the eye hazard. The test method is also proposed by the developer to be suitable for the identification of substances not classified for serious eye damage/eye irritation based on the outcome of a validation study. However, an independent peer-review of the validation study is still pending and therefore the final applicability of the test method still needs to be confirmed. Therefore conclusions on classification cannot be drawn from negative results before the scientific validity of the test method to correctly identify substance not requiring classification for serious eye damage/eye irritation has been confirmed.

52 Concerning the IRE and HET-CAM test methods, ICCVAM validation assessments in 2007 and 2010 that these test methods were not sufficiently accurate for regulatory use or that there was not sufficient data, especially for Category 2 chemicals, to make a final conclusion on their validity and recommended additional studies (http://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/ocular/in-vitro/index.html & http://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/ocular/in-vitro-test-methods/index.html). The Manual of Decisions of the Competent Authorities (EC, 2009) concluded that there is enough evidence available to conclude that the test methods are able to detect substances causing severe damage to eyes. Positive results can therefore be used for classification purposes i.e. leading to a classification of Category 1 for serious eye damage and labelling with H318 *“Causes serious eye damage”* according to CLP.

Abbreviations: BCOP = Bovine Corneal Opacity and Permeability; CM = Cytosensor Microphysiometer; EpiOcular ™ EIT = EpiOcular™ Eye Irritation Test; FL = Fluorescein Leakage; HET-CAM = Hen's Egg Test on Chorioallantoic Membrane; ICE = Isolated Chicken Eye; IRE = Isolated Rabbit Eye; N.A. = not available; STE = Short-Time Exposure.

The test methods indicated in Table R.7.2–4 above are either organotypic assays (BCOP, ICE, IRE and HET-CAM), cytotoxicity and cell function based assays (CM, FL and STE), reconstructed human cornea-like epithelium assays (EpiOcular™ EIT), or *in chemico* assays (Ocular Irritection®). These test methods are mainly concerned with modelling the immediate effects of substances on the cornea**.** *In vivo* eye irritation endpoints which may not be covered by the above-mentioned optimised protocols are the following:

i. persistence/reversibility of effects
ii. discolouration on the cornea [53]

Concerning persistence and reversibility of effects, the OECD TGs for BCOP (OECD TG 437) and ICE (OECD TG 438) and the OECD GD 160 (OECD, 2011) state that histopathological examination of the corneas may be potentially useful when a more complete characterization of corneal damage is needed. Some evidence has been published showing that histopathology may support the identification of irreversible effects produced by non-extreme pH detergent and cleaning products when used in combination with the ICE test method (Cazelle *et al.*, 2014). However, more work is still needed to assess the usefulness of the histopathological evaluation concerning identification of irreversible effects.

There are currently no validated *in vitro* eye irritation test methods available that could be used for the direct identification of Eye irritants Category 2 under CLP.

Additional test methods currently under development to assess different ranges of eye irritation potential are e.g. the *Ex Vivo* Eye Irritation Test (EVEIT) and the Porcine Cornea Reversibility Assay (PorCORA). The EVEIT and PorCORA test methods are organotypic assays which use either isolated rabbit or porcine corneas, respectively, and have been proposed to be able to discriminate between reversible and irreversible (persistent) effects by directly monitoring the recovery process in excised corneas kept in culture for several days following chemical exposure (Frentz *et al.*, 2008; Spöler *et al.,* 2010; Piehl *et al.*, 2010, 2011).

Testing and Assessment strategies combining different test methods according to their applicability domain and capacity to classify in the different ranges of serious eye damage/eye irritation (from those listed in table R.7.2-4 and those mentioned in the previous paragraphs) still need to be developed to facilitate the identification of Category 2 substances on the basis of methods that currently can only be used to directly identify Category 1 and/or not classified substances.

Further test method developments may occur and the registrants are advised to follow the latest updates through e.g. EURL ECVAM website (https://eurl-ecvam.jrc.ec.europa.eu/) and ECHA's test methods webpage (http://echa.europa.eu/support/testing-methods-and-alternatives) for potential new test guidelines and test guideline updates.

[53] Current *in vitro* TGs (listed in table R.7.2-4 above) do not cover discoloration of the cornea, but some test methods may give indications about this effect.

Animal data

Annex I to the CLP Regulation defines serious eye damage/eye irritation as local toxic effects, and, as such, an assessment of serious eye damage/eye irritation is normally part of the acute testing phase of a toxicity programme and it is an early requirement of all regulatory programmes. Testing for serious eye damage/eye irritation has, historically, used animal models and a variety of test methodologies depending upon, for example, the laboratory undertaking the test, the area and intended application. However, in line with one of the objectives of the REACH Regulation, as described in Articles 13(1) and 25(1) and Annex VI, animal testing should be undertaken only as a last resort after i) considering all existing available test data and ii) generating information whenever possible by means of alternative methods to animal testing such as *in vitro* methods, QSAR models, grouping or read-across.

In cases in which *in vivo* testing is necessary, current approaches for serious eye damage/eye irritation testing *in vivo* are covered by the Acute Eye Irritation/Corrosion test method (EU B.5/OECD TG 405). This guideline recommends a tiered approach, whereby existing and relevant data are evaluated first. The guideline also recommends that testing in animals should only be conducted if determined to be necessary after consideration of available alternative methods. The *in vivo* test uses one animal (the rabbit is the preferred species); in the absence of severe effects this is followed by a further testing of up to two animals (a total maximum of three animals).

Both EU and OECD methods use the scoring system developed by Draize (1944). The EU criteria for classification are based on the mean tissue scores obtained over the first 24-72 hour period after exposure and on the reversibility or irreversibility of the effects observed. Currently, *irritants* (Category 2 Eye irritants) cause significant inflammation of the eye (conjunctiva redness/oedema, cornea and/or iris) but this effect is transient, i.e. the affected sites are repaired within the observation period of the test. A substance causing considerable damage to the cornea and/or iris is classified in Category 1 for Serious Eye Damage. The criteria for classification in Category 1 for Serious Eye Damage include persistence of effects (effects on the cornea, iris or conjunctiva that are not expected to be reversed or have not fully reversed within an observation period of normally 21 days, i.e. with a score >0), irreversible staining of the eye and/or criteria for the degree of severity.

For existing data, the use of methods other than those specified in the Annex to the EU Test Methods Regulation, or corresponding OECD methods, such as the rabbit Low Volume Eye Test (LVET) (Griffith *et al*., 1980) may be accepted on a case-by-case basis (see also ESAC, 2009).

R.7.2.8.2 Human data on serious eye damage/eye irritation

Existing human data include historical data that should be taken into account when evaluating intrinsic hazards of substances. *New* testing in humans for hazard identification purposes is not acceptable for ethical reasons.

Existing data can be obtained from case reports, poison information centres, medical clinics, occupational experience, epidemiological studies and volunteer studies. Their quality and relevance for hazard assessment should be critically reviewed. However, in general, human data can be used to determine a corrosive or irritating potential of a substance. Good quality and relevant human data have precedence over other data. However, absence of incidence in humans does not necessarily overrule *in vitro* data or existing animal data of good quality that are positive.

R.7.2.9 Evaluation of information on serious eye damage / eye irritation

R.7.2.9.1 Non-human data for serious eye damage / eye irritation

Non-testing data on serious eye damage / eye irritation

Physico-chemical properties

According to the current EU and OECD guidelines, substances should not be tested on animals for serious eye damage/eye irritation if they can be predicted to be corrosive to the skin (Category 1 of CLP) or cause serious eye damage (Category 1 of CLP) from their physico-chemical properties. In particular, substances exhibiting strong acidity (pH ≤2.0) or alkalinity (pH ≥11.5) in solution are predicted to be corrosive to the skin or cause serious eye damage, and should not be tested on animals. Testing with *in vitro* methods can nevertheless be performed to confirm classification decisions (see Section 3.3.2.3 of Annex I to the CLP Regulation).

A substance known or predicted to be corrosive to the skin can be considered to cause Serious Eye Damage (Category 1). However, no conclusion can be made regarding serious eye damage/eye irritation potential when the pH has an intermediate value (when 2.0< pH <11.5). Where extreme pH is the only basis for classification as serious eye damage, it may also be important to take into consideration the acid/alkaline reserve, i.e. a measure of the buffering capacity (Young *et al.*, 1988,; Young and How, 1994). However, the buffering capacity should not be used alone to exonerate from classification of the substance as corrosive. Indeed, when the acid/alkaline reserve suggests that the substance may not cause serious eye damage, further *in vitro* testing should be considered (see Section 3.3.2.3 of Annex I to the CLP Regulation).

Grouping, (Q)SARs and expert systems

Guidance has been developed by the former ECB (Worth *et al.*, 2005) on how to apply (Q)SARs for regulatory use. Guidance on how to assess the validity and suitability of (Q)SAR models and adequacy of their predictions is given in Section R.6.1 of Chapter R.6 of the Guidance on IR&CSA. Essentially, the determination of whether a (Q)SAR result may be used to replace a test result can be broken down into three main steps:

1. evaluation of the scientific validity (relevance and reliability) of the model,
2. assessment of the applicability of the model to the chemical of interest and the reliability of the individual model prediction,
3. assessment of the adequacy of the information for making the regulatory decision, including an assessment of completeness, i.e. whether the information is sufficient to make the regulatory decision, and if not, what additional (experimental) information is needed.

The assessment of model validity needs to be performed along the lines of the OECD principles for (Q)SAR validation (OECD, 2007), e.g. in terms of a defined endpoint, an unambiguous algorithm, a defined applicability domain, the statistical characteristics (“goodness-of-fit”), and mechanistic interpretation.

The following questions, *inter alia*, should be addressed when assessing the reliability of an individual prediction:

i. Is the chemical of interest within the scope of the model, according to the defined applicability domain of the model?

ii. Is the defined applicability domain suitable for the regulatory purpose?

iii. How well does the model predict chemicals that are similar to the chemical of interest?

iv. Is the model estimate reasonable, taking into account other information?

The mechanism of serious eye damage/eye irritation involves toxicodynamic and toxicokinetic parameters. Some models predict serious eye damage and eye irritation based on toxicodynamic properties only (e.g. acidity or basicity, electrophilicity, other reactivity, surfactant activity, solving membranes). Such models have to be additionally evaluated to check whether they also take account of toxicokinetic parameters related to the potential of a substance to cross relevant outer membranes of the eye (cornea) and to be active in the living tissue underneath; alternatively, these models have to be used in combination with data covering such toxicokinetic parameters. Conversely models predict (the absence of) serious eye damage/eye irritation solely from e.g. physico-chemical properties considered to illustrate the toxicokinetic behaviour of a substance. Such models should be evaluated to check whether they also take account of the activity of the substance (toxicodynamics), in particular for its potential to cause serious eye damage (whereby the corrosive action itself may lead to membrane destruction and subsequent tissue damage).

For example, the BfR rule-base implemented in Toxtree and the OECD QSAR Toolbox contains both physico-chemical exclusion rules and structure-based inclusion rules (structural alerts). Evaluations of these rules for the prediction/exclusion of eye irritation (Tsakovska *et al.*, 2005, on structural alerts; Tsakovska *et al.*, 2007, on physico-chemical exclusion rules) have been carried out in accordance with the OECD principles for (Q)SAR validation (see Appendix R.7.2–3). However, inclusion and exclusion rules were evaluated separately, and not used in combination in these works.

When applied, these two sets of rules may sometimes provide contradictory information, i.e. a structural alert might indicate serious eye damage/eye irritation potential, while at the same time, based on physico-chemical properties, absence of effect is predicted. In such cases, it is recommended to consider additional information (e.g. on the behaviour of chemically similar substances). In other cases, applicability of one (or more) of the physico-chemical exclusion rules might indicate absence of serious eye damage/eye irritation potential of the target substance, while no structural alert for serious eye damage/eye irritation is triggered. Given that the absence of any known structural alert is not equivalent to the absence of a potential effect, in such a situation the substance should still be examined for potentially reactive substructures (and examining the behaviour of chemical analogues would still be beneficial).

While these considerations apply to the use of the BfR rule-base for direct classification/non-classification, less certainty might be required e.g. for a decision on further *in vitro* testing: where the exclusion rules suggest the absence of an effect, a bottom-up approach could be followed, i.e. a test for eye irritation and not one for serious eye damage might be initiated (see Section R.7.2.11.2).

There is no other model available which sufficiently describes the absence of effects. Neutral organics[54] are expected not to be irritants. Predicted absence of reactivity needs to be described in sufficient detail or be substantiated with other information.

Testing data on serious eye damage/eye irritation

In vitro data

There are EU and OECD adopted test guidelines (see Section R.7.2.8.1), according to which substances can be classified as causing serious eye damage or not classified.

Annex VII to the REACH Regulation requires information from *in vitro* tests for serious eye damage/eye irritation, and not from animal tests. Guidance on how *in vitro* data can also be used to fulfil Annex VIII requirements, is given in Section R.7.2.11 of this document [55].

Data from the following types of tests can be used for Annex VII requirements and may be accepted for Annex VIII requirements for serious eye damage/eye irritation when the general rules for adaptation specified in Annex XI are used:

- **Bovine Corneal Opacity and Permeability (BCOP) test method** (EU B.47/OECD TG 437): The specific scope and limitations are:
  - This test is recommended to identify substances inducing serious eye damage, i.e. substances to be classified in Eye Damage Category 1 under CLP, without further testing, and also recommended to identify substances that do not require classification for eye irritation or serious eye damage i.e. leading to non-classification under CLP, without further testing;
  - If, as a result of testing, the substance is neither classified as Eye Damage Category 1 nor identified as not requiring classification under CLP, further testing/evaluation is required;
  - This test may result in false positive Category 1 predictions (serious eye damage) for alcohols and ketones and false negative predictions (underpredicted Category 1 substances) for substances that would be classified as Category 1 (serious eye damage) *in vivo* based on persistence of effects only (i.e., that do not meet Category 1 classification criteria based on the mean scores obtained from the first 3 observation days but show persistent effects at the 21st observation day) (Adriaens *et al.*, 2014; OCED, 2013). See also Section R.7.2.8.1 for "*In vitro* test methods for serious eye damage/eye irritation";
  - This test does not allow testing of gases and aerosols.

---

54 By definition a neutral organic is a chemical which does not have potential reaction centres, even after skin metabolism.

55 Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

- **Isolated Chicken Eye (ICE) test method** (EU B.48/OECD TG 438): The specific scope and limitations are:
    - This test is recommended to identify substances inducing serious eye damage, i.e. substances to be classified in Eye Damage Category 1 under CLP, without further testing, and also recommended to identify substances that do not require classification for eye irritation or serious eye damage i.e. leading to non-classification under CLP, without further testing;
    - If, as a result of testing, the substance is neither classified as Eye Damage Category 1 nor identified as not requiring classification under CLP, further testing/evaluation is required;
    - Similar limitations in relation to false positive and false negative predictions, as specified for the BCOP assay above, apply to this test method as well;
    - This test does not allow testing of gases and aerosols.

- **Fluorescein leakage (FL) test method** (OECD TG 460): The specific scope and limitations are:
    - This test is recommended to identify substances inducing serious eye damage, i.e. substances to be classified in Eye Damage Category 1 under CLP, without further testing;
    - This test is not recommended for the identification of substances which should be classified as Eye irritants Category 2 or of substances which should not be classified for serious eye damage and eye irritation;
    - This test is only applicable to water soluble substances and/or where the toxic effect is not affected by dilution;
    - Its applicability domain does not include strong acids and bases, cell fixatives and highly volatile substances;
    - If, as a result of testing, the substance is not classified as Eye Damage Category 1 under CLP, further testing/evaluation is required.

In the case of REACH Annex VIII information requirement, and beside results from the test methods mentioned above, a positive outcome (Serious Eye Damage Category 1) from one of five *in vitro* assays (i.e. the IRE, HET-CAM, CM, STE, Ocular Irritection^{TM} assay) is also accepted in the EU to classify a substance as Eye Damage Category 1 under CLP using the adaptations of the standard testing regime specified in REACH Annex XI. A negative outcome, i.e. leading to non-classification according to CLP, can also be accepted for fulfilling the Annex VIII information requirement on the basis of test data obtained with the CM, EpiOcular EIT and STE test methods, in case the substance falls into the applicability domain of the test method(s) and Annex XI adaptations are used.

Currently, there are no validated *in vitro* methods available for the direct identification of Category 2 Eye irritants.

- Quality Aspects of exisiting *in vitro* data:

For quality assessment of existing *in vitro* data that will form the basis for later possible *Weight-of-Evidence* considerations, see Section R.4.4 of Chapter R.4 of the *Guidance on IR&CSA*, and for aspects that need to be taken into account in such a *Weight of Evidence* see Section R.5.2.1.2 of Chapter R.5 of the *Guidance on IR&CSA*.

Animal data

Well-reported studies, particularly if conducted in accordance with the principles of GLP, can be used to identify substances which would be considered to cause, or not to cause serious eye damage or eye irritation. There may be a number of serious eye damage/eye irritation studies already available for an existing substance, none of which are fully equivalent to an OECD TG or an EU test method such as those in the Annex to the EU Test Methods Regulation. If the results from such a batch of studies are consistent, they may, together, provide sufficient information on the serious eye damage/eye irritation potential of the substance.

If the results from a variety of studies are unclear, based on the criteria given below for evaluation of the data, the registrant will need to decide which of the studies are most reliable, relevant for the endpoint in question and adequate for classification purposes.

Particular attention should be given to the persistence of irritation effects, even those which do not lead to classification. Effects such as persistent corneal opacity, discolouration of the cornea by a dye substance, adhesion, pannus, and interference with the function of the iris or other effects that impair sight which do not reverse within the test period may indicate that a substance will cause persistent damage to the human eye.

Data from studies other than skin corrosion/irritation studies (e.g. other toxicological studies on the substance in which local responses of skin have been reported) may provide useful information though they may not be well reported in relation to, for example, the basic requirements for information on skin irritation.

Data from studies other than serious eye damage/eye irritation studies (e.g. other toxicological studies on the substance in which local responses of the eye have been reported) may provide useful information though they may not be well reported in relation to, for example, the basic requirements for information on eye irritation. More notably, eye reactions and symptoms are not systematically scored in studies not specifically designed to address serious eye damage/eye irritation.

- Quality Aspects of existing *in vivo* data:

Data from **existing** irritation studies in animals must be taken into account before further testing is considered. A quality assessment of any such reports should be done using, for example, the system developed by Klimisch *et al*. (1997), as described in Section R.4.2 of Chapter R.4 of the *Guidance on IR&CSA*, and a judgement will need to be made as to whether any further testing is required. Some examples to note are:

i. Was the animal species used the rabbit or was it another species such as the rat or the mouse? Normally the rabbit is used for eye irritation testing.

ii. How many animals were used? Current methodology requires a maximum of 3 animals tested in a sequential manner (with 1 or 2 animals being sufficient if serious eye damage/irreversible effects are observed in the first or second tested animal, respectively) but 6 were frequently used in the past (see Section 3.3.2.3.2.2 of the *Guidance on the Application of the CLP criteria* for the evaluation of results from tests that have been conducted with more than 3 animals).

iii. How many dose levels were used? If dilutions were included, what solvent was used (as this may have influenced absorption)? Which dose volume was used?

iv. Check the observation period used post exposure. Shorter periods than in the current guideline may be adequate for non-irritants but may require a more severe classification for irritants when the observation period is too short to measure full recovery.

v. Was initial pain noted after instillation of the test substance onto the eye? Was the substance washed out from the eye? Was fluorescent staining used?

vi. How was the test material applied onto the eye?

Irritation scores from old reports, reports produced for regulatory submission in the USA or in publications may be expressed as a Maximum Average Score (MAS). Without the original data it is not always possible to convert these scores accurately into the scoring system used in the EU. For extremes, i.e. where there is either no irritation or severe irritation, it may not be necessary to look further, but average irritation scores pose a problem and expert judgement may be required to avoid repeat testing.

Observations such as those above can all be used to assess whether the existing animal test report available can be used reliably to predict the irritation potential of a substance, thus avoiding further testing.

- Specific considerations:

A refinement of the classical Draize test is the rabbit low volume eye test (LVET). The test protocol deviates from OECD TG 405 in that in the LVET, 10 µl is directly applied onto the cornea. The grading scale and the data interpretation in the LVET is exactly the same as those used in OECD TG 405. The validity of the LVET was reviewed by EURL ECVAM between 2006 and 2009 *via* retrospective validation for the detergent and cleaning products applicability domain (for further details, see https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/topical-toxicity/eye-irritation). Anatomical and physiological considerations for rabbit and human eyes indicate that a dose volume of 10 µl is appropriate (A.I.S.E. 2006): the tear volume in both rabbit and man is approximately the same (~ 7-8 µl), and after blinking, the volume capacity in the human eye is ~10 µl after blinking. Furthermore the use of direct cornea exposure mimics human exposure scenarios that can be reasonably expected (e.g. accidental ocular exposure during household use) and for the specific use domain of household detergents and cleaning products as well as their main ingredients (i.e. surfactants) as used in these products. These considerations suggest that the LVET is also potentially a suitable test to demonstrate toxicological effects on man of potential eye hazards of substances. The LVET has been used in industry for the safety evaluation of single substances (Griffith *et al*., 1980) and detergent and cleaning products (Freeberg *et al*., 1984; Freeberg *et al*. 1986a,b; Cormier *et al*., 1995; Roggeband *et al*., 2000), and has shown to be a very good predictor of the effects in man. It still overpredicts, but less than the classical Draize test of OECD TG 405.

After peer review, the LVET was not recommended for prospective use, i.e. to generate new data but it was acknowledged that existing LVET data of the limited use domain mentioned above may be used for purposes of classification and labeling decisions. Moreover, it was recognised that existing LVET data of this limited use domain may be used as supplementary data for future validation studies. No additional testing should however be performed to further develop or validate the LVET test. It was also pointed out that the LVET has a tendency to classify in lower hazard categories when compared to OECD TG 405. Nevertheless, it was acknowledged that these data may still be useful on a case-by-case basis, with respect to test data for household detergents, cleaning products and surfactants used in such products (ESAC, 2009).

In summary, available data from the LVET on substances should be considered and must be carefully evaluated. For the classification of substances it must be taken into account that the test has a limited applicability domain (detergent and cleaning products). Consequently, within the applicability domain of household detergents, cleaning products and their main ingredients, positive LVET data (be it Category 2 or Category 1) can be used for the appropriate classification for either serious eye damage or eye irritation, but negative data from LVET as a *stand alone method* (in the absence of any other information) are not conclusive for *no classification*.

R.7.2.9.2 Human data on serious eye damage / eye irritation

Well-documented *existing* human data of different sources can often provide very useful information on serious eye damage/eye irritation, sometimes for a range of exposure levels. Often the only useful information available on irritation is obtained from human experience (e.g. occupational settings). The usefulness of all human data on irritation will depend on the extent to which the effect, and its magnitude, can be reliably attributed to the substance of interest. Experience has shown that it is difficult to obtain useful data on substance-induced eye irritation, but data may be available on human ocular responses to certain types of mixtures (e.g. Freeberg *et al.*, 1986a).

The quality and relevance of existing human data for hazard assessment should be critically reviewed. For example, in occupational studies with mixed exposure it is important that the substance causing serious eye damage or eye irritation has been accurately identified. There may also be a significant level of uncertainty in human data due to poor reporting and lack of specific information on exposure.

Examples of how existing human data can be used in hazard classification for irritation are provided in an ECETOC monograph (ECETOC, 2002).

Substances causing Serious eye damage Category 1 give more severe corneal opacity and iritis than Eye irritants Category 2. Category 1 substances induce considerable tissue damage which can result in serious physical decay of vision. It is recognised that such severe lesions usually do not reverse within 21 days (relates to animals) (see Section 3.3 of Annex I to the CLP Regulation). In contrast, the effects of Category 2 substances are reversible within 21 days. In humans, an ophthalmic examination by a physician would reveal a decay of vision. If it is not transient but persistent it implies classification in Category 1. If the discrimination between Category 1 and Category 2 is not obvious, then Category 1 might be chosen, however other types of information may be generated e.g. by performing *in vitro* testing, to support the conclusion (for further information, see Section 3.3 of the *Guidance on the Application of the CLP criteria*).

R.7.2.9.3 Exposure considerations for serious eye damage / eye irritation

Exposure-based waiving from testing is not applicable to the endpoint of serious eye damage/eye irritation. Exposure-based waiving from testing as specified in Annex XI (3) of the REACH Regulation only applies to tests listed in Sections 8.6 and 8.7 of Annex VIII, Annex IX and Annex X according to the REACH text.

R.7.2.9.4 Remaining uncertainty on serious eye damage / eye irritation

Usually it is possible to unequivocally identify (or accept) a substance as causing serious eye damage, whatever type of study provides the information.

There may be a significant level of uncertainty in human data on irritant effects (e.g. because of poor reporting, lack of specific information on exposure, subjective or anecdotal reporting of effects, small numbers of subjects).

Data from studies in animals and from *in vitro* tests performed according to internationally accepted test methods will usually give relevant information on the serious eye damage/eye irritation potential of a substance. In general, it is assumed that substances which cause serious eye damage/eye irritation in EU or OECD TG-compliant studies in animals or *in vitro* will cause serious eye damage/eye irritation in humans, and those which are not irritant in EU or OECD TG-compliant studies will not be irritant in humans (Please note that in general test animals are considered to be more sensitive to serious eye damage/eye irritation than humans (e.g. Adriaens *et al.*, 2014)). It should be borne in mind that some of the limitations of the *in vivo* serious eye damage/eye irritation study include its high variability, the variable exposure being dependent on the physico-chemical properties of the test substance, and the subjective grading of the lesions (Adriaens *et al.*, 2014; Cormier *et al.*, 1996; Prinsen, 2006; Marzulli and Ruggles, 1973; Weil and Scala, 1971). Moreover, inconsistent results from a number of similar studies increases the uncertainty in deriving data from animal or *in vitro* studies.

The scope of the *in vitro* tests for serious eye damage/eye irritation has also some limitations, as explained in Section R.7.2.9.1 under "Testing data on serious eye damage/eye irritation". In addition inconsistent results from two or more *in vitro* tests could add to the overall uncertainty in interpreting the data.

R.7.2.10 Conclusions on serious eye damage/eye irritation

R.7.2.10.1 Concluding on suitability for Classification and Labelling

In order to conclude on Classification and Labelling according to the CLP Regulation, all the available information needs to be taken into account and consideration should be given to both the *Guidance on the Application of the CLP criteria* and the various remarks (related to Classification and Labelling) made throughout this guidance document [56].

R.7.2.10.2 Concluding on suitability for Chemical Safety Assessment

A dose-response assessment is difficult to make for serious eye damage/eye irritation simply because up to now most data have been generated for undiluted substances in accordance with test guidelines and traditional practice (which continues today). From a risk characterisation perspective it is therefore advisable to use the outcome of the classification procedure, i.e. a substance that is classified is assumed to be sufficiently characterised. However, a complete risk assessment requires both hazard and dose-response data and for local effects the concentration is often the determinative dose metric. Consequently, if dose-response data are available, they must be taken into account (see Figure R.7.2–4).

Guidance on the possibilities for derivation of DNELs for serious eye damage/eye irritation is given in Appendix R.8-9 of Chapter R.8 of the *Guidance on IR&CSA*.

R.7.2.10.3 Information not adequate

A *Weight-of-Evidence* approach comparing available adequate information with the tonnage-triggered information requirements under REACH may result in the conclusion that the requirements are not fulfilled. In order to proceed to further information gathering the testing and assessment strategy described in Section R.7.2.11 below is recommended.

---

56 Please note that the 8th Adaptation to Technical and Scientific Progress (ATP) of the CLP Regulation is currently under discussion. The 8th ATP will take into account the 5th Revision of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which was adopted in 2012 and contains in particular refined criteria for skin corrosion/irritation and serious eye damage/eye irritation.

R.7.2.11 Testing and assessment strategy for serious eye damage/eye irritation

R.7.2.11.1 Objective / General principles

The following testing and assessment strategy is recommended for developing adequate and scientifically sound data for assessment/evaluation and classification of the serious eye damage and eye irritation properties of substances. For existing substances with insufficient data, this strategy can also be used to decide which additional data, beside those already available, are needed. The testing and assessment strategy is aimed at the identification of serious eye damage/eye irritation by using different elements where appropriate, depending on the information available. A basic principle of the strategy is that the results of one study or from an information source are evaluated before another study is initiated. The strategy seeks to ensure that the data requirements are met in the most efficient and humane manner so that animal usage and costs are minimised.

The different elements provided in the Figure R.7.2–4 describe information sources that can be used to conclude on a substance's hazard potential towards the eye. The elements described in Figure R.7.2–5 can be rearranged as appropriate, especially those in Part 1. This may be particularly helpful in cases where a conclusion can be drawn from certain elements without having to consider all of them. If judged relevant, elements in Part 1 can be omitted and *in vitro* testing can be performed immediately.

Figure R.7.2–5 is divided into three parts whereby Part 1 aims at evaluating existing information that may be available on the substance. In Part 2 existing information and relevant data should be assessed in order to consider whether there is enough information available to conclude on the substance's hazard properties within a *Weight-of-Evidence* analysis, in case it is not possible to make a conclusion based on single elements described in Part 1. In case no conclusion can be drawn from Parts 1 and 2, new data should be generated in Part 3 by first performing relevant *in vitro* testing. Only in case no conclusion can be drawn based on the *in vitro* testing, can *in vivo* testing be performed (for substances at or above 10 tonnes per annum only).

Some guidance for testing is provided by the specific rules for adaptation from standard information requirements, as described in column 2 of Annexes VII-X to the REACH Regulation, together with some general rules for adaptation from standard information requirements in Annex XI.

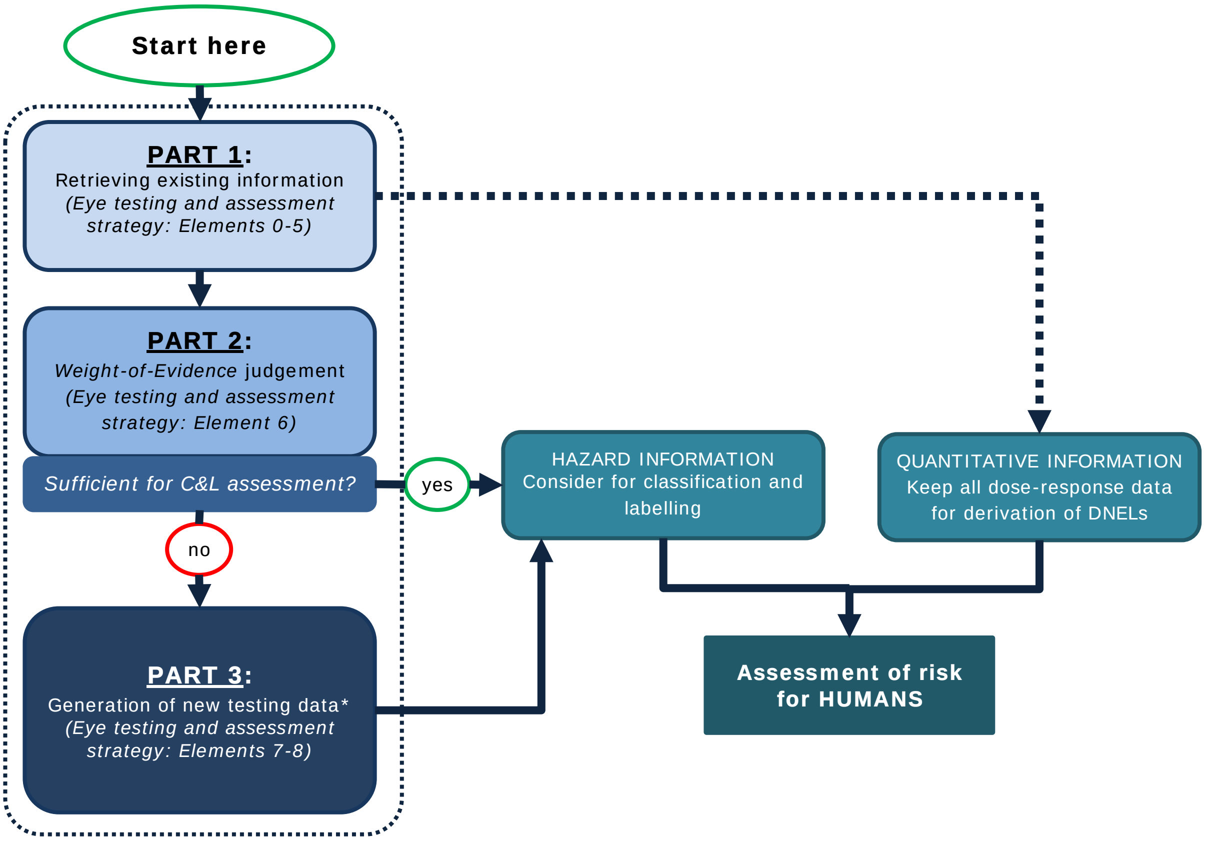
Risk assessment of the serious eye damage/eye irritation potential of a substance is normally made in a qualitative way provided that the substance has been classified as causing serious eye damage/eye irritation. Existing test guidelines do not contain dose-response assessment, consequently a quantitative analysis will often not be possible. Therefore, hazard identification and appropriate classification is the key determinant in the information gathering strategy below. As a consequence, the use of *Assessment Factors* is of limited use in order to take into account uncertainty of data. However, the registrant is encouraged to keep and use all quantitative data that might be encountered in the process of retrieving hazard information in the context of the present testing strategy and to perform a complete risk assessment, comprising assessment of qualitative hazard as well as quantitative information.

It is recommended that the testing and assessment strategy be followed until element 5 (Figure R.7.2–4 and Figure R.7.2–5) in all cases and thereafter the *Weight-of-Evidence* analysis be performed. Clearly, all information sources/elements can be rearranged as appropriate, i.e. not all elements will necessarily be accompanied by data but it is important that all potential data sources are explored prior to starting the *Weight-of-Evidence* analysis.

While it is recommended that this approach be followed, other approaches may be more appropriate and efficient on a case-by-case basis. For example, in case there is no existing data and it is anticipated that generation of "pre-testing data" would be non-conclusive, it may be deemed appropriate to directly proceed to the information generation part. Furthermore, prior to performing any new *in vivo* test, the use of *in vitro* methods should be fully exploited (see Articles 13(1) and 25(1) of the REACH Regulation) by using the general rules of Annex XI for adaptation of the standard testing regime set out in Annexes VII to X.

If the substance is not classified for serious eye damage/eye irritation, no risk assessment for this endpoint is performed, regardless of the exposure. Please note that there are no options for exposure-based waiving for these endpoints in the REACH Regulation.

The following flow chart (Figure R.7.2–4) gives an overview of a possible approach for defining a testing and assessment strategy for serious eye damage/eye irritation.



*Generation of new testing data according to Annex VII to VIII to the REACH Regulation and with due observation of the rules for adaptation of the standard testing regime laid down in Annex XI.

Figure R.7.2–4 Overview of the testing and assessment strategy for serious eye damage/eye irritation

R.7.2.11.2 Testing and assessment strategy for serious eye damage/eye irritation

Recommended approach

The testing and assessment strategy for serious eye damage/eye irritation (see Figure R.7.2–5) is completely analogous in structure to that for skin corrosion/irritation. The testing and assessment strategy consists of three parts: Part 1 (elements 0 to 5) is about retrieving exisiting information, Part 2 (element 6) represents a *Weight-of-Evidence* analysis and expert judgement (element 6), and Part 3 is about generation of new information by testing (elements 7 to 8).

In Part 1, existing and available information from the literature and databases is gathered and considered in the strategy approach. The order of the different elements, i.e. 0 to 5, is only indicative and they may be arranged as appropriate. This may be particularly helpful in cases where a reliable conclusion can be drawn from certain elements without having to consider all of them. For instance, if the substance is classified as corrosive to the skin or has an extreme pH ($\leq$ 2.0 or $\geq$ 11.5) serious eye damage is considered implicit (element 1c) and therefore the substance should be classified as causing serious eye damage (Category 1) according to CLP and further testing is not required. At the end of Part 1 and if no final conclusion could be derived directly from one or several of the available pieces of information, all the information collected should be analysed using a *Weight-of-Evidence* approach (element 6).

In the information generation part (elements 7 to 8), new information on the serious eye damage/eye irritation potential of substances is generated by means of *in vitro* (element 7) or, as a last resort (see Articles 13(1) and 25(1) of the REACH Regulation), *in vivo* testing (element 9). Therefore, before concluding the *Weight-of-Evidence* analysis in element 6 *and in vitro* testing (elements 7a and 7b), new *in vivo* tests should not be conducted. More information on how to use the *in vitro* methods for serious eye damage/eye irritation within the testing strategy can be found in the following paragraphs.

While it is recommended that this approach be followed, other approaches may be more appropriate and efficient on a case-by-case basis. For example, in case there is no existing data and it is anticipated that compilation of data at elements 0-6 would be non-conclusive, it may be appropriate to directly proceed to the information generation part.

Figure R.7.2–5 Testing and assessment strategy for evaluating the serious eye damage/eye irritation potential of substances (footnotes a to f are detailed below the figure).

Element	Information	Conclusion [57]
Conclusion of the information strategy on skin corrosion/irritation		
0	Is the substance classified as a skin corrosive? →	**YES**: When assigned Skin Corrosive Cat. 1, 1A, 1B or 1C, the risk of severe damage to eyes is considered implicit (Serious Eye Damage Cat. 1) (Column 2 adaptation of Annexes VII and VIII).
Existing data on physico-chemical properties		
1a	Is the substance spontaneously flammable in contact with air (pyrophoric) or water at room temperature? →	**YES**: No testing required (Column 2 adaptation of Annexes VII and VIII).
1b	Is the substance an organic hydroperoxide or an organic peroxide? →	**YES**: Consider classifying for: ■ When assigning a Skin Corrosive Cat. 1B classification for a hydroperoxide, the risk of serious eye damage is considered implicit. Consider classifying as Serious Eye Damage Cat. 1, or ■ When assigning a Skin Irritant Cat. 2 classification for a peroxide, the risk of eye irritation is considered implicit. Consider classifying as Eye Irritant Cat. 2. **OR** Provide evidence supporting deviating classification or non-classification [58].
1c	Is the pH of the substance ≤ 2.0 or ≥ 11.5? [a] →	**YES**: Consider classifying as Serious Eye Damage Cat. 1 (column 2 adaptation in section 8.2 of Annex VIII) if pH is used as the sole basis for classification decision.
1d	Are there other physical or chemical properties that indicate that the substance causes serious eye damage or eye irritation? →	**YES**: Use this information for *Weight-of-Evidence* analysis (Element 6).

[57] Please note that the 8th Adaptation to Technical and Scientific Progress (ATP) of the CLP Regulation is currently under discussion. The 8th ATP will take into account the 5th Revision of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which was adopted in 2012 and contains in particular refined criteria for skin corrosion/irritation and serious eye damage/eye irritation.

[58] Information on e.g. *in vitro* testing may provide evidence on more suitable classification, if there is some doubt on the correct classification.

Existing human data		
2	Are there adequate existing human data [b] which provide evidence that the substance has the potential to cause serious eye damage or eye irritation? →	**YES**: Consider classifying accordingly (Serious Eye Damage Cat. 1 or Eye Irritant Cat. 2).
Existing animal data from serious eye damage/eye irritation studies		
3	Are there data from existing studies on serious eye damage/eye irritation in laboratory animals, which provide sound conclusive evidence that the substance is seriously damaging to the eye, eye irritant or non-irritant? →	**YES**: Consider classifying accordingly (Serious Eye Damage Cat. 1 or Eye Irritant Cat. 2) or consider no classification.
Existing/new (Q)SAR data and read-across		
4	Are there structurally related substances (suitable "read-across" or grouping), which are classified as causing serious eye damage/eye irritation, or indicating that the substance is non-irritant, or do valid (Q)SAR methods indicate serious eye damage/eye irritation or non-irritation of the substance? [c] →	**YES**: Consider classifying accordingly.
Existing in vitro data		
5a	Has the substance demonstrated serious eye damage, eye irritation or non-irritating properties in an EU/OECD adopted *in vitro* test? Data from *in vitro* test methods that have been validated and are considered scientifically valid but are not yet adopted by EU and/or OECD may also be used if the provisions defined in Annex XI are met. →	**YES**: Consider classifying accordingly (Serious Eye Damage Cat. 1 or Eye Irritant Cat. 2) or consider no classification. If discrimination between Serious Eye Damage Cat. 1 and Eye Irritant Cat. 2 is not possible, Serious Eye Damage Cat. 1 must be chosen.
5b	Are there acceptable data from (a) non-validated suitable *in vitro* test(s), which provide sound evidence that the substance causes serious eye damage/eye irritation? [d] →	**YES**: Consider classifying accordingly (SeriousEye Damage Cat. 1 or Eye Irritant Cat. 2). If discrimination between Serious Eye Damage Cat. 1 and Eye Irritant Cat. 2 is not possible, Serious Eye Damage Cat. 1 must be chosen.
Weight-of-evidence analysis		
6	The "elements" described above may be arranged as appropriate. Taking all available existing and relevant data mentioned above (Elements 0 – 5) into account, is there sufficient information to make a decision on whether classification/labelling is necessary, and – if so – how to classify and label? →	**YES**: Classify accordingly (Serious Eye Damage Cat. 1 or Eye Irritant Cat. 2) or consider no classification.

New in vitro tests for serious eye damage/eye irritation (Annex VII to the REACH Regulation) [e]		
7a	Does the substance demonstrate serious eye damage, eye irritation or non-irritant properties in (an) EU/OECD adopted *in vitro* test(s) for the eye hazard charaterisation?[e]→ Data from *in vitro* test methods that have been validated and are considered scientifically valid but are not yet adopted by EU and/or OECD may also be used if the provisions of Annex XI are met.	**YES**: Classify accordingly (Serious Eye Damage Cat. 1 or Eye Irritant Cat. 2) or consider no classification. If discrimination between Serious Eye Damage Cat. 1 and Eye Irritant Cat. 2 is not possible, Serious Eye Damage Cat. 1 must be chosen. If a conclusion on the eye hazard cannot be drawn by using *in vitro* testing, *in vivo* testing should be performed (at Annex VIII level only).
7b	Does the substance demonstrate serious eye damage or eye irritant properties in (a) non-validated suitable *in vitro* test(s) for serious eye damage/eye irritation? [d] →	**YES**: Classify as Serious Eye Damage Cat. 1 or Eye Irritant Cat. 2. If a conclusion on the eye hazard cannot be drawn by using *in vitro* testing, *in vivo* testing should be performed (at Annex VIII level only).
New in vivo test for serious eye damage/eye irritation as a last resort (Annex VIII to the REACH Regulation) [f]		
8	Does the substance demonstrate serious eye damage or eye irritation in an OECD adopted *in vivo* test? →	**YES**: Classify accordingly (Serious Eye Damage Cat. 1 or Eye Irritant Cat. 2). **NO:** No classification needed.

Notes to the information scheme on serious eye damage/eye irritation:

a) Note that if the buffering capacity suggests the substance may not cause serious eye damage, further data are needed to confirm this, preferably using an appropriate *in vitro* test method.

b) Data from case reports, occupational experience, poison information centres or from clinical studies.

c) Conclusion on no classification can be made if the model has been shown to adequately predict the absence of the classified effect and if it fulfils the requirements of Annex XI to the REACH Regulation. Prediction of the absence of the classified effect can be made either by triggering an exclusion rule in the BfR system (to be checked on a case-by-case basis), or based on a negative prediction in a classification QSAR that was trained on both positive and negative substances. The suitability of the model (reliability, relevance) should be very carefully checked to make sure that the prediction is fit for purpose, and the applicability of the model to the substance should also be justified (e.g fulfilment of the conditions of Section 1.3

of Annex XI to the REACH Regulation should be checked). For read-across, generation of new *in vivo* data should be avoided.

d) Data obtained from non-validated suitable *in vitro* tests can only be used according to the criteria set out in section 1.4 of Annex XI to the REACH Regulation, i.e. only positive results can be accepted. However, there are already several EU/OECD adopted test methods which should be primarily used (see Table R.7.2–4).

e) New *in vitro* testing should be performed following a top-down or bottom-up approach. Please see the following paragraph "How to use the *in vitro* methods serious eye damage/eye irritation within the strategy". It is highly recommended to adhere to the test protocols whose scientific validity has been established by validation and which, ideally, have been officially adopted by the European Commission and/or by the OECD.

f) *In vivo* testing should not be conducted in case the substance falls under the scope of the specific *in vitro* test(s) performed, and there are no substance-specific limitations to using those tests, and the Registrant formulates an adaptation according to Annex XI to the REACH Regulation. Due to the current standard *in vivo* information requirement at Annex VIII level and above, an adaptation needs to be built up in the registration dossier in order to successfully submit a compliant dossier [59].

How to use the *in vitro* methods for serious eye damage/eye irritation within the testing and assessment strategy

For serious eye damage/eye irritation no single *in vitro* test method is currently able to fully replace the regulatory *in vivo* test, known as the Draize eye test (EU B.5/OECD TG 405) across the full range of ocular responses for different chemical classes. However, the *in vitro* test methods specified in Sections R.7.2.8.1 and R.7.2.9.1 may be used for partial replacement within a tiered testing strategy or as stand-alone test methods depending on the outcome of the study. Moreover, combinations of several alternative test methods may be able to fully replace the Draize eye test. Testing strategies such as the top-down or bottom-up approaches provide a means of incorporating existing information, QSAR predictions, read-across and grouping and *in vitro* test results.

New *in vitro* testing should be performed following a top-down or bottom-up approach (Scott *et al*., 2010). The top-down approach (start with an *in vitro* test able to identify substances that are seriously damaging to the eye, i.e. classified as Serious eye damage Cat. 1) should be used when all available collected information and the *Weight-of-Evidence* assessment result in a high *a-priori* probability of the substance being seriously damaging to the eye. The bottom-up approach, on the other hand (start with an *in vitro* test able to identify substances not requiring classification for serious eye damage/eye irritation, i.e. not classified) should be followed when all available collected information and the *Weight-of-Evidence* assessment result in a high *a-priori* probability of the substance being non-irritant to the eyes.

[59] Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

There are steps to be considered before any testing (*in vitro* or *in vivo*) is conducted. These steps are specified in Section 8.2 in column 1 of Annexes VII and VIII to the REACH Regulation and include:

1. Assessment of all available human and animal data (e.g. animal data may be available from acute dermal toxicity studies or a substance may already be classified as a skin corrosive or as causing serious eye damage);

2. Assessment of acid and alkaline reserve. It should be noted that other substance properties e.g. pH or others could indicate that the substance is seriously damaging or irritant to eyes. Consideration of information obtained from the use of (Q)SARs or from similar substances may be useful in predicting the serious eye damage/eye irritation potential of the substance (see Figure R.7.2–5);

3. Assessment of existing data on physico-chemical properties. No further testing is required if the substance is spontaneously flammable in air (pyrophoric) or water at room temperature.

   After following steps 1, 2 and 3, if a conclusion on classification cannot be drawn, *in vitro* studies for serious eye damage/eye irritation should be conducted;

4. One or more *in vitro* studies for serious eye damage/eye irritation should be performed, and the outcome can be:

   a) In the case of a positive and definitive result from e.g. the BCOP, ICE, FL or other scientifically valid *in vitro* test methods, the substance can be classified as inducing “serious eye damage” (Serious Eye Damage Category 1 under CLP), and no further test *in vivo* is necessary;

   b) In addition, the BCOP and ICE, or other scientifically valid *in vitro* test methods can also provide information on whether the substance does not require any classification for the eye hazard. If no classification is needed, no further testing *in vivo* is necessary;

   c) For Annex VII information requirement, as no *in vivo* testing is foreseen, a *Weight-of-Evidence* approach may be needed in order to conclude on the eye hazard potential of the substance. The assessment should take all relevant pieces of information into account. This means that in case where the available *in vitro* test(s) for serious eye damage/eye irritation does (do) not enable a definitive conclusion on the classification for the eye hazard to be drawn, information obtained e.g. from skin irritation testing should be considered. Thus, in case inconsistent *in vitro* results for serious eye damage/eye irritation are obtained, the Weight of Evidence including information on skin irritation (Category 2) may support the classification for eye irritation (Category 2), as a precautionary principle. See also the *Guidance on the Application of the CLP criteria*;

   d) For Annex VIII information requirements, if a definitive conclusion on the serious eye damage/eye irritation potential of the substance cannot be reached from the use of one or several *in vitro* methods used in a testing strategy, a further test conducted *in vivo* to assess the eye hazard potential of the substance is needed.

Note: Registrants must make sure that the substance falls within the scope and applicability domain of the specific *in vitro* tests performed, and there are no substance-specific limitations to using those tests (see *in vitro* tests for serious eye damage/eye irritation and sections R.7.2.8.1 and R.7.2.9.1).

Registrants who must fulfil the Annex VIII information requirement for an *in vivo* eye irritation study and have completed the above steps, may be able to do so by using an adaptation according to Annex XI to the REACH Regulation and without testing on animals. Due to the current standard *in vivo* information requirement at Annex VIII level and above, an adaptation needs to be built up in a registration dossier in order to successfully submit it. However, an *in vivo* eye irritation test may still be necessary depending on the assessment of the available information and outcomes of *in vitro* studies [60].

Instructions on how to submit *in vitro* information instead of *in vivo* can be found e.g. in section 3.7 of *Practical Guide 1: How to report in vitro data* (available at http://echa.europa.eu/practical-guides).

It is important to note that it is the responsibility of the registrant to ensure that the chosen test method(s) is (are) suitable for the substance in order to obtain adequate information from the *in vitro* studies. For most substances, the use of EU- or OECD-adopted test methods for the eye hazard characterisation will provide results that will have regulatory acceptance under REACH.

60 Please note that the information requirements in REACH Annexes VII and VIII in relation to skin corrosion/irritation and serious eye damage/eye irritation are currently under revision. This revision is expected to strengthen the role of *in vitro* methods and to remove the standard information requirement for an *in vivo* study at the Annex VIII level. As a consequence, once the new REACH Annexes come into force, an *in vivo* study would only be required where a substance falls outside of the applicability domain of the available *in vitro* methods or the results obtained from such methods would not allow a conclusive decision on (non-)classification and risk assessment.

RESPIRATORY TRACT CORROSION/IRRITATION

R.7.2.12 Information sources on respiratory tract corrosion/irritation

The evaluation of respiratory tract corrosion/irritation potential can be based on expert judgement using evidence such as: human and animal experience, existing (*in vitro*) data, substance properties like pH values, volatility (Saturated Vapour Concentration (SVC)) or dustiness, information from similar substances or any other pertinent data.

R.7.2.12.1 Animal data

There are currently no EU or OECD adopted test guidelines that deal specifically with respiratory tract corrosion or irritation. Studies that could inform on the respiratory tract corrosion/irritation potential of the substance concerned are single or repeated inhalation exposure studies (information on (histo-)pathological changes).

Single inhalation exposure studies *in vivo* may provide information on nasal irritation such as rhinitis, whereas histopathological examination of respiratory tract tissues of animals repeatedly exposed by inhalation (28-day and 90-day inhalation studies) may provide information on inflammatory/cytotoxic effects such as hyperemia, edema, inflammation or mucosal thickening. Data from bronchoalveolar lavage may give additional information on the inflammatory response.

It is noteworthy that, while histopathology is not a standard element of the OECD TG 436 for Acute Inhalation Toxicity, TG 436 specifies that "*Additional examinations included a priori by design may be considered to extend the interpretive value of the study, such as… providing evidence of irritation by microscope examination of the respiratory tract. Examined organs may include those showing evidence of gross pathology in animals surviving 24 or more hours, and organs known or expected to be affected. Microscopic examination of the entire respiratory tract may provide useful information for test articles that are reactive with water, such as acids and hygroscopic test articles*".

Moreover, the data on local dermal or ocular corrosion/irritation might contain information that is relevant for the respiratory endpoint and this should be considered accordingly. It is for instance a reasonable precaution to assume that corrosive (and severely irritating) substances would also cause respiratory tract irritation or even corrosion when vaporised or in the form of an aerosol. Furthermore, information from cases where symptoms have been described associated with occupational exposures can be used on a case-by-case basis to characterise the respiratory tract corrosion/irritation potency of a substance. Existing and available information from acute and repeated dose inhalation toxicity studies may also be considered sufficient to show that the substance causes respiratory tract corrosion/irritation at a specific concentration level or range. The data need to be carefully evaluated with regard to the exposure conditions (sufficient documentation required). Possible confounding factors should be taken into account.

R.7.2.12.2 Human data

Existing human data include historical data that should be taken into account when evaluating intrinsic hazards of substances. *New* testing in humans for hazard identification purposes is not acceptable for ethical reasons.

Existing human data can be obtained from case reports, poison information centres, medical clinics, and occupational experience or from epidemiological studies or volunteer studies. Their quality and relevance for hazard assessment should be critically reviewed. However, in general, human data can be used to determine a corrosive or irritating potential of a substance. Good quality and relevant human data have precedence over other data. However, absence of incidence in humans does not necessarily overrule existing good quality animal data that are positive.

Specifically with regard to respiratory tract irritation, there is a view in the occupational health literature that sensory irritation may be a more sensitive effect than overt tissue-damaging irritation, given that its biological function is to serve as an immediate warning against substances inhaled during a short period of time which could damage the airways, and that it triggers physiological reflexes that limit inhalation volumes and protect the airways. However, there is a lack of documented evidence to indicate that this is a generic position that would necessarily apply to all inhaled irritants.

R.7.2.13 Evaluation of information on respiratory tract corrosion/irritation

All data available should be evaluated to estimate a substance's potential to induce respiratory tract corrosion or irritation.

R.7.2.13.1 Animal data

The evaluation is based on data from inhalation studies (acute, repeated exposure):

- Clinical symptoms of dyspnoea or breathing difficulties,
- Histomorphology of the respiratory tract,
- Lavage examination (nasal, bronchoalveolar).

Useful information may be obtained from the single and repeated inhalation toxicity studies for classification and labelling as well as for DNEL derivation.

For derivation of a DNEL (acute - inhalation, local effects) information from animal studies with acute and/or repeated inhalation exposure may be used. This usually requires that in the study several exposure concentrations had been used that allow derivation of a No Observed Adverse Effect Concentration (NOAEC) and/or a Low Observed Adverse Effect Concentration (LOAEC) or a benchmark concentration (BMC) as starting points for DNEL derivation (Section R.8.2.1 and Appendix R.8-8 of Chapter R.8 of the *Guidance on IR&CSA*). In case such information is only available from repeated dose inhalation studies, derivation of a long-term DNEL (long-term - inhalation, local effects) might be more appropriate.

For classification and labelling purposes, the severity of the effects (reversible versus irreversible) and the target within the respiratory tract (upper versus lower respiratory tract) need to be considered.

In case animal studies show reversible effects (usually in the upper respiratory tract), the studies can be used as part of a *Weight-of-Evidence* evaluation for classification for STOT-SE Category 3. Reversible respiratory tract effects may be clinical signs of toxicity like dyspnoea or rhinitis and histopathological effects like hyperemia, oedema, minimal inflammation or thickened mucous layer which may be reflective of the characteristic clinical symptoms described above.

In case the studies show significant changes, more than transient in nature, especially in the lower respiratory tract (bronchiolar and alveolar region), classification for STOT-SE Category 1 or 2 might be considered, depending on the concentration at which the effects occur. Significant changes to the respiratory tract may include necrosis, or other morphological changes that are potentially reversible but provide clear evidence of marked organ dysfunction. However, if such effects were only observed in inhalation studies with repeated exposure and the mode of action indicates that the significant damage to the respiratory tract is due to repeated exposure, classification for "Specific Target Organ Toxicity after Repeated Exposure (STOT-RE), Category 1 or 2 might be more appropriate (see Section 3.9 of the *Guidance on the Application of the CLP criteria*).

For corrosive substances that may be acutely toxic, the additional labelling with EUH071 *"Corrosive to the respiratory tract"* should be considered (see Section 3.1 of the *Guidance on the Application of the CLP criteria*). It is presumed that corrosive substances will cause toxicity by inhalation exposure. The Hazard statement EUH071 must be assigned for substances that may be inhaled in addition to classification for acute inhalation toxicity, if data are available that indicate that the mechanism of toxicity is corrosivity. In cases where no acute inhalation test has been performed and the substance may be inhaled, this hazard statement must also be assigned. However, if corrosive substances are used in mixtures in sub-corrosive concentrations, it needs to be ensured that an appropriate classification for potential respiratory tract irritation is applied. For liquids the volatiliy/SVC, and for solids dustiness, if applicable, should be taken into consideration.

R.7.2.13.2 Human data

The evaluation is based on:

- Experience from occupational exposure;
- Published data on volunteers (objective measurements, psychophysical methods, and subjective reporting);
- Other data (e.g. from nasal lavage).

Consideration should be given to real-life human observational experience, if this is properly collected and documented (Arts *et al.*, 2006), e.g. data from well-designed workplace surveys or worker health monitoring programmes. For substances with an array of industrial uses and with abundant human evidence, the symptoms of respiratory tract irritation can sometimes be associated with certain concentrations of the irritants in the workplace air and might thus allow derivation of DNELs. However, the exposure details need to be well documented and due consideration should be given to possible confounding factors.

Data on sensory irritation of the airways may be available from volunteer studies. These include objective measurements of respiratory tract irritation such as electrophysiological responses, data from lateralization threshold testing, biomarkers of inflammation in nasal or bronchoalveolar lavage fluids. Including anosmics as subjects could exclude odour as a bias. Good quality and relevant human data have precedence over other data. However, absence of positive findings in humans does not necessarily overrule good quality animal data that are positive.

Human data demonstrating respiratory tract irritation are used primarily for classification for Specific Target Organ Toxicity after Single Exposure (STOT-SE), Category 3 (H335: *"May cause respiratory irritation"*) under CLP (see Section 3.8 of the *Guidance on the Application of the CLP criteria*).

Such effects are characterised by localised redness, oedema, pruritis and/or pain and they impair function with symptoms such as cough, pain, choking, and breathing difficulties.

Subjective human observations could be supported by objective measurements of clear respiratory tract irritation (such as electrophysiological responses, biomarkers of inflammation in nasal or bronchoalveolar lavage fluids). Furthermore, the symptoms observed in humans should also be typical of those that would be produced in the exposed population rather than being an isolated idiosyncratic reaction or response triggered only in individuals with hypersensitive airways. Ambiguous reports simply of 'irritation' must be excluded as this term is commonly used to describe a wide range of sensations including those such as smell, unpleasant taste, a tickling sensation, and dryness, which are outside the scope of classification for respiratory tract irritation.

R.7.2.14 Conclusions on respiratory tract corrosion/irritation

R.7.2.14.1 Concluding on suitability for Classification and Labelling

In order to conclude on Classification and Labelling according to the CLP Regulation, all the available information needs to be taken into account, and consideration should be given to both the *Guidance on the Application of the CLP criteria* and the various remarks (related to Classification and Labelling) made throughout this guidance document.

R.7.2.14.2 Concluding on suitability for Chemical Safety Assessment

A dose-response assessment might be possible. Animal studies, especially those with repeated inhalation exposure and several exposure concentrations, may be available that allow derivation of a NOAEC and/or a LOAEC as starting points for DNEL derivation.

Human data indicative of respiratory tract irritation that provide reliable quantitative information on the threshold for the irritative effects may also be used to derive DNEL (acute - inhalation, local effects) (see Section R.8.2.1 and Appendix R.8-8 of Chapter R.8 of the *Guidance on IR&CSA*).

R.7.2.15 References on skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation

Abraham MH, Hassanisadi M, Jalali-Heravi M, Ghafourian T, Cain WS and Cometto-Muniz JE (2003) Draize rabbit eye test compatibility with eye irritation thresholds in humans: a quantitative structure-activity relationship analysis. Toxicol Sci 76:384-91.

Ad-hoc Working group on Defatting substances (1997). ECBI/22/96-Add. 10.

Adriaens E, Barroso J, Eskes C, Hoffmann S, McNamee P, Alépée N, Bessou-Touya S, De Smedt A, De Wever B, Pfannenbecker U, Tailhardat M and Zuang V (2014) Retrospective analysis of the Draize test for serious eye damage/eye irritation: importance of understanding the in vivo endpoints under UN GHS/EU CLP for the development and evaluation of *in vitro* test methods. Arch Toxicol 88:701-23.

A.I.S.E. (2006) Low Volume Eye Test (LVET) – A.I.S.E. Submission to ECVAM – January 2006.

Arts JHE, de Heer C and Woutersen R (2006) Local effects in the respitory tract: relevance of subjectively measured irritation for setting occupational exposure limits. Int Arch Occup Environ Health 79:283-98.

Barratt MD (1995a) Quantitative structure activity relationships for skin corrosivity of organic acids, bases and phenols. Toxicol Lett 75:169-176.

Barratt MD (1995b) The role of structure-activity relationships and expert systems in alternative strategies for the determination of skin sensitisation, skin corrosivity and eye irritation. Altern Lab Anim 23:111-22.

Barratt MD (1996a) Quantitative Structure-Activity Relationships for Skin Corrosivity of Organic Acids, Bases and Phenols: Principal Components and Neural Network Analysis of Extended Sets. Toxicol *in Vitro* 10:85-94.

Barratt MD (1996b) Quantitative Structure-Activity Relationships for Skin Irritation and Corrosivity of Neutral and Electrophilic Organic Chemicals. Toxicol *in Vitro* 10:247-56.

Barratt MD, Dixit MB and Jones PA (1996c) The use of *in vitro* cytotoxicity measurements in (Q)SAR methods for the prediction of the skin corrosivity potential of acids. Toxicol *in Vitro* 10:283-90.

Barratt MD (1997) QSARs for the Eye Irritation Potential of Neutral Organic Chemicals. Toxicol *in Vitro* 11:1-8.

Basketter D, Jírová D and Kandárová H (2012) Review of skin irritation/corrosion Hazards on the basis of human data: a regulatory perspective. Interdiscip Toxicol 5:98-104.

Botham PA, Earl LK, Fentem JH, Roguet R and van de Sandt JJM (1998) Alternative methods for skin irritation testing: the current status. ECVAM Skin Irritation Task Force Report 1. Altern Lab Anim 26:195-211.

Cazelle E, Eskes C, Hermann M, Jones P, McNamee P, Prinsen M, Taylor H and Wijnands M. (2014). Suitability of histopathology as an additional endpoint to the Isolated Chicken Eye Test for classification of non-extreme pH detergent and cleaning products. Toxicology in Vitro 28, 657–666.

Cormier EM, Hunter JE, Billhimer W, May J and Farage MA (1995) Use of clinical and consumer eye irritation data to evaluate the low-volume eye test. J Toxicol Cut Ocul Toxicol 14:197-205.

Cormier EM, Parker RD, Henson C, Cruze LW, Merritt AK, Bruce RD and Osborne R (1996) Determination of the intra- and inter-laboratory reproducibility of the Low Volume Eye Test and its statistical relationship to the Draize test. Regul Toxicol Pharmacol 23: 156-61.

Draize JH, Woodard G and Calvery HO (1944) Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J Pharmacol Exp Ther 82:377-390.

ECETOC (2002) ECETOC Monograph No.32, Use of Human Data in Hazard Classification for Irritation and Sensitisation. Available at: http://www.ecetoc.org/publications

Enslein K, Borgstedt HH, Blake BW and Hart JB (1987) Prediction of rabbit skin irritation severity by structure activity relationships. *In Vitro* Toxicol 1:129-47.

Enslein K, Blake BW, Tuzzeo TM, Borgstedt HH, Hart JB and Salem H (1988) Estimation of rabbit eye irritation scores by structure activity equations. *In vitro* Toxicol 2:1-14.

ESAC (2009) Statement on the use of existing Low Volume Eye Test (LVET) data for weight of evidence decisions on classification and labelling of cleaning products and their main ingredients. Available at: https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/docs-eye-irritation/esac-statement-lvet

Eskes C, Hoffmann S, Facchini D, Ulmer R, Wang A, Flego M, Vassallo M, Bufo M, van Vliet E, d'Abrosca F and Wilt N (2014) Validation study on the Ocular Irritection assay for eye irritation testing. Toxicol In Vitro 28:1046-65.

EURL ECVAM (2015) Recommendation on the use of the EpiOcular™ Eye Irritation Test (EIT) for identifying chemicals not requiring classification and labelling for serious eye damage/eye irritation according to UN GHS. Available at: https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations (in publication).

Freeberg FE, Griffith JF, Bruce RD and Bay PHS (1984) Correlation of animal test methods with human experience for household products. J Toxicol Cut Ocul Toxicol 1:53-64.

Freeberg FE, Hooker DT and Griffith JF (1986a) Correlation of animal eye test data with human experience for household products: an update. J Toxicol Cut Ocul Toxicol 5:115-23.

Freeberg FE, Nixon GA, Reer PJ, Weaver JE, Bruce RD, Griffith JF and Sanders III LW (1986b) Human and rabbit eye responses to chemical insult. Fund Appl Toxicol 7:626-34.

Frentz M, Goss M, Reim M, & Schrage NF (2008). Repeated exposure to Benzalkonium chloride in the Ex Vivo Eye Irritation Test (EVEIT): observation of isolated corneal damage and healing. ATLA 36: 25-32.

Gallegos Saliner A, Patlewicz G and Worth AP (2006). Review of literature-based models for skin and eye irritation and corrosion. JRC report EUR 22320 EN. European Chemicals Bureau, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/doc/QSAR_Review_Irritation.pdf

Gallegos Saliner A, Tsakovska I, Pavan M, Patlewicz G and Worth AP (2007) Evaluation of SARs for the prediction of skin irritation/corrosion potential - structural inclusion rules in the BfR decision support system. SAR QSAR Environ Res 18:331-42.

Gallegos Saliner A, Patlewicz G and Worth AP (2008) A Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion. QSAR Comb Sci 27:49-59.

Gerner I, Schlegel K, Walker JD and Hulzebos E (2004) Use of physico-chemical property limits to develop rules for identifying chemical substances with no skin irritation or corrosion potential. QSAR Comb Sci 23:726-33.

Gerner I, Liebsch M and Spielmann H (2005) Assessment of the Eye Irritating Properties of Chemicals by Applying Alternatives to the Draize Rabbit Eye Test: The Use of QSARs and *In Vitro* Tests for the Classification of Eye Irritation. Altern Lab Anim 33:215-37.

Golla S, Madihally S, Robinson RL Jr and Gasem KA (2009) Quantitative structure-property relationships modeling of skin irritation. Toxicol *in Vitro* 23:176-84.

Griffith JF, Nixon GA, Bruce RD, Reer PJ and Bannan EA (1980) Dose-response studies with chemical irritants in the albino rabbit eye as a basis for selecting optimum testing conditions for predicting hazard to the human eye. Toxicol Appl Pharmacol 55:501-13.

Hayashi M, Nakamura Y, Higashi K, Kato H, Kishida F and Kaneko H (1999) A Quantitative Structure-Activity Relationship Study of the Skin Irritation Potential of Phenols. Toxicol *in Vitro* 13:915-22.

Houthoff E, Rugen P, Hart D (2014) Predictability of *in vitro* dermal assays when evaluating fatty amine derivatives. Toxicol *in Vitro (in press).*

Hulzebos EM, Janssen PAH, Maślankiewicz L, Meijerink MCM, Muller JJA, Pelgrom SMG, Verdam L and Vermeire TG (2001) The application of structure-activity relationships in human hazard assessment: a first approach. RIVM report 601515 008, Bilthoven, The Netherlands. Available at: http://rivm.openrepository.com/rivm/bitstream/10029/9562/1/601516008.pdf

Hulzebos EM, Maslankiewicz L and Walker JD (2003) Verification of literature derived SARs for skin irritation and corrosion. QSAR Comb Sci 22:351-63.Hulzebos E, Walker JD, Gerner I and Schlegel K (2005) Use of structural alerts to develop rules for identifying chemical substances with skin irritation or skin corrosion potential. QSAR Comb Sci 24:332-42.

Ishii S, Ishii K, Nakadate M and Yamasaki K (2013) Correlation study in skin and eye irritation between rabbits and humans based on published literatures. Food Chem Toxicol 55:596-601.

Jírová D, Basketter D, Liebsch M, Bendová H, Kejlová K, Marriott M and Kandárová H (2010) Comparison of human skin irritation patch test data with *in vitro* skin irritation assays and animal data. Contact Dermatitis 62:109-16.

Klimisch HJ, Andreae M and Tillmann U (1997) Systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharm 25:1-5.

Kodithala K, Hopfinger AJ, Thompson ED and Robinson MK (2002) Prediction of skin irritation from organic chemicals using membrane-interaction QSAR Analysis. Toxicol Sci 66:336-46.

Liew CY and Yap CW (2013) QSAR and predictors of eye and skin effects. Mol Inform 32:281-90.

Marzulli FN and Ruggles DI (1973) Rabbit eye irritation test: collaborative study. J Assoc Off Analyt Chem 56:905-14.

Mombelli E (2008) An evaluation of the predictive ability of the QSAR software packages, DEREK, HAZARDEXPERT and TOPKAT, to describe chemically-induced skin irritation. Altern Lab Anim 36:15–24.

Nangia A, Anderson PH, Berner B, Maibach HI (1996) High dissociation constants (pKa) of baasic permeants are associated with in vivo skin irritation in man. Contact Dermatitis 34:237-42.

OECD (2007) Guidance Document On the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. Environment, Health and Safety Publications, Series on Testing and Assessment No. 69 (ENV/JM/MONO(2007)2). Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2007)2&doclanguage=en

OECD (2011) Guidance Document On "The Bovine Corneal Opacity And Permeability (BCOP) And Isolated Chicken Eye (ICE) Test Methods: Collection Of Tissues For Histological Evaluation And Collection Of Data On Non-Severe Irritants. OECD Environment, Health and Safety Publications, Series on Testing and Assessment No. 160 (ENV/JM/MONO(2011)45). Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2011)45&doclanguage=en

OECD (2013) Streamlined Summary Document Supporting OECD Guideline 437 On The Bovine Corneal Opacity And Permeability For Eye Irritation/Corrosion. Environment, Health and Safety Publications, Series on Testing and Assessment No. 189 (ENV/JM/MONO(2013)13). Available at:

http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2013)13&doclanguage=en

OECD (2014a) Guidance on Grouping of Chemicals, Second Edition. Environment, Health and Safety Publications, Series on Testing and Assessment No. 194 (ENV/JM/MONO(2014)4). Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2014)4&doclanguage=en

OECD (2014b) New Guidance Document On An Integrated Approach On Testing And Assessment (IATA) For Skin Corrosion And Irritation. Environment, Health and Safety Publications, Series on Testing and Assessment No. 203 (ENV/JM/MONO(2014)19). Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2014)19&doclanguage=en

Patlewicz G, Rodford R and Walker JD (2003) Quantitative structure-activity relationships for predicting skin and eye irritation. Environ Toxicol Chem 22:1862-9.

Piehl M, Gilotti A, Donovan A, DeGeorge G & Cerven D (2010). Novel cultured porcine corneal irritancy assay with reversibility endpoint. Toxicology in Vitro 24: 231-239.

Piehl M, Carathers M, Soda R, Cerven D & DeGeorge G (2011). Porcine Corneal Ocular Reversibility Assay (PorCORA) predicts ocular damage and recovery for global regulatory agency hazard categories. Toxicology in Vitro 25: 1912-1918.

Roggeband R, York M, Pericoi M and Braun W (2000) Eye irritation responses in rabbit and man after single applications of equal volumes of undiluted model liquid detergent products. Food Chem Toxicol 38:727-34.

Rorije E and Hulzebos E (2005) Evaluation of (Q)SARs for the prediction of Skin Irritation/Corrosion Potential. Physico-chemical exclusion rules. Final report for ECB contract IHCP.B430206. European Chemicals Bureau, Joint Research Centre, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/information-sources/qsar-document-area/Evaluation_of_Skin_Irritation_QSARs.pdf/view

Scott L, Eskes C, Hoffmann S, Adriaens E, Alepée N, Bufo M, Clothier R, Facchini D, Faller C, Guest R, Harbell J, Hartung T, Kamp H, Varlet BL, Meloni M, McNamee P, Osborne R, Pape W, Pfannenbecker U, Prinsen M, Seaman C, Spielmann H, Stokes W, Trouba K, Berghe CV, Goethem FV, Vassallo M, Vinardell P and Zuang V (2010) A proposed eye irritation testing strategy to reduce and replace *in vivo* studies using Bottom-Up and Top-Down approaches. Toxicol In Vitro 24:1-9.

Smith JS, Macina OT, Sussman NB, Luster MI and Karol MH (2000a) A robust structure-activity relationship (SAR) model for esters that cause skin irritation in humans. Toxicol Sci 55:215-22.

Smith JS, Macina OT, Sussman NB, Karol MH and Maibach HI (2000b) Experimental validation of a structure-activity relationship model of skin irritation by esters. Quant Struct-Act Rel 19:467–74.

Solimeo R, Zhang J, Kim M, Sedykh A and Zhu H (2012) Predicting Chemical Ocular Toxicity Using a Combinatorial QSAR Approach. Chem Res Toxicol 25: 2763-69.

Spöler F, Frentz M & Schrage NF (2010). Towards a new *in vitro* model of dry eye: the Ex Vivo Eye Irritation Test. Dev Ophthalmol. 45: 93-107.

Tsakovska I, Netzeva T and Worth AP (2005) Evaluation of (Q)SARs for the prediction of Eye Irritation/Corrosion Potential - physico-chemical exclusion rules. JRC Report EUR 21897 EN, 42pp. European Chemicals Bureau, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/doc/Evaluation_of_Eye_Irritation_QSARs.pdf

Tsakovska I, Gallegos Saliner A, Netzeva T, Pavan M, Worth AP (2007) Evaluation of SARs for the prediction of eye irritation/corrosion potential - structural inclusion rules in the BfR decision support system. SAR QSAR Environ Res 18:221-35.

Walker JD, Gerner I, Hulzebos E, Schlegel K (2005) The Skin Irritation Corrosion Rules Estimation Tool (SICRET). QSAR Comb Sci 24:378-84.

Weil CS and Scala A (1971) Study of intra- and inter- laboratory variability in the results of rabbit eye and skin irritation tests. Toxicol Appl Pharmacol 19:276-360.

Whittle E, Barratt MD, Carter JA, Basketter DA and Chamberlain M (1996) Skin corrosivity potential of fatty acids: *In vitro* rat and human testing and (Q)SAR studies. Toxicol *In vitro* 10:95-100.

Worth AP and Cronin MTD (2001) The use of pH measurements to predict the potential of chemicals to cause acute dermal and ocular toxicity. Toxicol 169:119-31.

Worth AP (2004) The tiered approach to toxicity assessment based on the integrated use of alternative (non-animal) tests. *In:* Predicting Chemical Toxicity and Fate (Cronin MTD and Livingstone DJ Eds.) CRC Press, Boca Raton, FL, USA, pp.389-410.

Worth AP, Bassan A, Gallegos A, Netzeva TI, Patlewicz G, Pavan M, Tsakovska I, Vracko . (2005) The Characterisation of (Quantitative) Structure-Activity Relationships: Preliminary Guidance. JRC report EUR 21866 EN. European Chemicals Bureau, Joint Research Centre, European Commission, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/information-sources/qsar-document-area/QSAR_characterisation_EUR_21866_EN.pdf/view

Worth A, Barroso J, Bremer S, Burton J, Casati S, Coecke S, Corvi R, Desprez B, Dumont C, Gouliarmou V, Goumenou M, Gräpel R, Griesinger C, Halder M, Janusch Roi A, Kienzler A, Madia F, Munn S, Nepelska M, Paini A, Price A, Prieto P, Rolaki A, Schäffer M, J. Triebe, Whelan M, Wittwehr C and Zuang V (2014) Alternative methods for regulatory toxicology – a state-of-the-art review. JRC Science and Policy Reports, Report EUR 26797 EN. European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Systems Toxicology Unit, Institute for Health and Consumer Protection, European Commission Joint Research Centre, Ispra, Italy. Available at: https://ec.europa.eu/jrc/sites/default/files/echa_jrc_sla_report_public_05-09-14_withcover_ipo.pdf

Young JR, How MJ, Walker AP and Worth WMH (1988) Classification as corrosive or irritant to skin of preparations containing acidic or alkaline substances, without testing on animals. Toxicol *In Vitro* 2:19-26.

Young JR and How MJ (1994) Product Classification as Corrosive or Irritant by measuring pH and Acid/Alkali Reserve. *In:* Alternative Methods in Toxicology, Vol 10 - *In Vitro* Skin Toxicology: Irritation, Phototoxicity, Sensitization (Rougier A, Goldberg AM and Maibach HI Eds.) Mary Ann Liebert, Inc., New York, USA, pp.23-27.

Appendices R.7.2-1 to 3 to Section R.7.2

Appendix R.7.2–1 Mechanisms of local toxicities: skin corrosion/irritation, serious eye damage/eye irritation and respiratory tract corrosion/irritation

Content of Appendix R.7.2-1:

- Mechanisms of skin corrosion and irritation
- Mechanisms of serious eye damage/eye irritation
- Mechanisms of respiratory tract corrosion and irritation

MECHANISMS OF SKIN CORROSION AND IRRITATION

Clinically, different types of irritant contact dermatitis (ICD) exist, and have been classified on the basis of differences in morphology and mode of onset, as: acute irritant dermatitis (primary irritation); irritant reaction; delayed, acute irritant contact dermatitis; cumulative irritant dermatitis; traumatic irritant dermatitis, pustular and acneiform irritant dermatitis; non-erythematuous irritant dermatitis; and subjective irritation (Lammintausta and Maibach, 1990).

Two different pathogenetic pathways may be involved in ICD. Acute ICD is characterised by an inflammatory reaction which mimics allergic contact dermatitis, with the release of inflammatory mediators and cytokines. Chronic ICD, on the other hand, is characterised by disturbed barrier function, associated with an increased epidermal turnover which leads clinically to lichenification (Berardesca and Distante, 1994).

The clinically relevant elements of skin irritation are a disturbance of the desquamation process, resulting in scaling or hyperkeratosis (chronic effects), i.e. epidermal events, and an inflammatory response with vasodilation and redness in combination with extravasation of water, which may be observed as papules, vesicles and/or bullae and oedema (acute effects), i.e. events essentially taking place in the dermis (Serup, 1995). The onset of irritation takes place at the stratum corneum level and later in the dermis, whereas early events in sensitisation occur in the dermis. Variations in the skin reactions are dependent on the degree of injury induced, as well as on the effects of an irritant substance on different cell populations. For example, pigmentary alterations are due to effects on melanocytes, whereas ulcerations are due to extensive keratinocyte necrosis (skin corrosion). The release of cytokines and mediators can be initiated by a number of cells, including living keratinocytes and those of the stratum corneum, which thus modulate inflammation and repair (Sondergard *et al*., 1974; Hawk *et al*., 1983; Barker *et al*., 1991; Baadsgaard and Wang, 1991; Hunziker *et al*., 1992; Berardesca and Distante, 1994).

The physico-chemical properties, concentration, volume and contact time of the irritant give rise to variations in the skin response. Furthermore, inter-individual differences exist, based on age, gender, race, skin colour and history of any previous skin disease. In the same individual, reactivity differs according to differences in skin thickness and skin sensitivity to irritation of the different body regions. Finally, a greater sensitivity to some irritants (DMSO, propylene glycol, SLS and soap) has been reported during winter, because of the reduced hydration state of the skin (Frosch and Pilz, 1995). Although clinically different types of irritant reactions can be observed, they are all based on cellular and biochemical mechanisms which induce the irritant response. It is not yet possible to conclude whether the observed clinical differences are actually due to differences in biochemical mechanisms, and further investigations are needed.

According to Barratt (1995a) and further elaborated by Walker *et al*. (2004), for organic substances, the mechanisms leading to skin irritation are normally described by a two-stage process where a substance first has to penetrate the *stratum corneum* and then trigger a biological response in deeper epidermal or dermal layers.

For strong inorganic acids and bases, no *stratum corneum* penetration is needed because they erode the *stratum corneum*. According to the Technical Guidance Document (TGD) supporting Commission Directive 93/67/EEC on risk assessment for new notified and existing substances (EC, 2003), the percutaneous absorption of acrylates, quaternary ammonium ions, heterocyclic ammonium ions and sulphonium salts is slow, since these substances are binding to macromolecules in the skin. As a result of binding, corrosion can occur as the *stratum corneum* is eroded. Reactivity can be caused by electrophiles and/or pro-electrophiles. Electrophiles contain atoms, such as N, O or halogens attached to a C-atom, which makes that specific C-atom positively charged and therefore reactive with electron-rich regions of peptides and proteins. This causes irritation *via* covalent binding to the skin.

At this time, the following mechanisms are proposed for inducing skin irritation or skin corrosion by affecting the structure and function of the *stratum corneum* :

1. Mechanisms of skin irritation:

    - Reaction with skin proteins and interference with lipids in the *stratum corneum* by surface-active agents (denaturation of proteins, disruption of plasma membrane lipids),

    - Dissolving of plasma membrane lipids and thus defatting and disintegration of skin by low molecular weight organic substances.

2. Mechanisms of skin corrosion:

    - Erosion of the *stratum corneum* by most inorganic acids and bases and by strong organic acids with pH ≤2.0 and bases with pH ≥11.5, and

    - Binding to skin components in the *stratum corneum* by cationic surfactants and percutaneous absorption of acrylates, quaternary ammonium ions, heterocyclic ammonium ions and sulphonium salts.

3. Mechanisms that may lead to both skin irritation and corrosion:

    - Penetration of the *stratum corneum* by anionic or non-surfactant organic substances with sufficient hydrophobic and hydrophilic properties, and

    - Elicitation of an inflammatory and/or cytotoxic response in the epidermis or dermis.

    The severity of these responses may determine whether irritation or corrosion occurs.

MECHANISMS OF SERIOUS EYE DAMAGE AND EYE IRRITATION

Eye injury can be caused by many insults. These can be physical such as puncture by sharp objects. Eye injury can be caused by substances such as systemic drugs that can enter into the eye through the blood stream (e.g. Cyclosporine, vaccines, intravenous immunoglobulines, intravenous streptokinase). Various degrees of eye injury can also be caused by direct (topical) contact with substances or mixtures such as acids, alkalis, solvents or surfactants. These materials may come into contact with the eye intentionally, e.g. through the use of eye drops, medications, products intended for use around the eyes, but also unintentionally, e.g. accidental spills and splashes of consumer products or accidental exposures in the workplace.

In general, substances or mixtures which come directly into contact with the eye may cause local effects on the frontal tissues and substructures of the eye, e.g. cornea, conjunctiva, iris, lachrymal system and eye lids. There are several modes of action by which topical substances and mixtures cause eye injury e.g. cell membrane lysis, saponification and coagulation (see Table R.7.2–5).

Table R.7.2–5 Categories of irritant substances and their typical mode of action in eye irritation.

Substance / mixtures	*Mode of Action*
Inert substances	May cause effect due to large size. Protrusions may cause direct puncture of the eye.
Acids	May react directly with cellular components e.g. eye proteins and cause coagulation, lysis or precipitation resulting in relatively localised injury.
Bases (Alkalis)	May actively disrupt the cell membrane lipids by alkaline action i.e. saponification. May penetrate to the deeper layers of the eye tissue. May react directly with cellular components and cause coagulation or lysis of the tissue.
Solvents	May cause membrane lysis by dissolving lipids in plasma membranes of epithelial and underlying cells resulting in loss of the cells affected and, as a result, tissue degradation, which might be transient, depending on the repair mechanisms (cell proliferation, tissue restoration). May also cause coagulation.
Lachrymators	May stimulate the sensory nerve endings in the corneal epithelium causing an increase in tearing.

The degree of eye injury is usually dependent on the characteristics (chemical category/class) and concentration of the substance or mixture. Acids and alkalis usually cause immediate irritation to the eyes. Other substances may cause eye injuries that start as mild but progress to be more severe at a later period e.g. substances that react with cellular constituents *via* alkylation or oxidative attack on macromolecules. An example of these types of substances are e.g. peroxides, mustards and bleaches (Scott *et al.*, 2010).

Upon exposure of the ocular surface to eye irritants, inflammation of the conjunctiva can be induced. This includes dilation of the blood vessels causing redness, increased effusion of water causing swelling (oedema/chemosis) and an increase in the secretion of mucus leading to an increase in discharge. Visual acuity can be impaired. Effects on the cornea may be more severe (e.g. destruction of the cornea, or persistent corneal opacity or discoloration of the cornea by a dye substance), or reversible where effects are limited to the epithelia. Irritants may also produce an increase in tear production and changes to the tear film integrity such as increased wetness. Iritis may result from direct irritation or become a secondary reaction to the corneal injury. Once the iris is inflamed, infiltration of fluids can follow which affects the ability to adjust the size of the pupil and decreases the reaction to light leading to decreased visual acuity. Due to the richness of nerves in the iris, irritation also causes subjective symptoms such as itching, burning and stinging.

Eye injury can be reversible or irreversible depending on the degree of damage and degree of repair. Damage to the corneal epithelium alone can repair quickly, often with no permanent eye damage. The cornea may still repair fairly well if the damage goes beyond the basement membrane into the superficial part of the stroma but the repair process may take days or even weeks to occur. Once the damage extends significantly into the stroma, corneal ulceration can occur due to the subsequent series of inflammatory processes. If damage extends to and beyond the endothelium, corneal perforation may occur which is irreversible and may cause permanent loss of vision. Eye injury can cause different degrees of functional loss e.g. increase of tear production, opacification of the cornea, oedema and so decrease visual acuity.

The body has its own defence mechanisms, e.g. sensing the pain, stinging and burning, and the eyelids will blink to avoid full exposure to the substance. Increased tear production and blinking of the eyes with the help of the drainage apparatus help to dilute or clear the causative agent. Such defence mechanisms are highly developed in man with rapid adversive blinking and profuse tear production resulting from exposure of the eye to a foreign material that is irritating. It is well reported in the literature that species differences occur in the rate of blinking and tear production mechanism that can influence how effectively foreign materials are removed from the eye.

MECHANISMS OF RESPIRATORY TRACT CORROSION AND IRRITATION

Corrosion of the respiratory tract includes destruction of the mucosa followed by proliferation of epithelial cells. Remodeling of tissue may occur with chronic injury if repair mechanisms are unable to keep pace. Mild epithelial or endothelial injury without basement membrane damage, severe inflammation, or persistence of the inciting agent may be resolved by simple cellular regeneration. With more severe damage, a significant inflammation component may be elicited which may be followed by tissue destruction or fibrosis. In some cases, persistence of the inciting agent within the tissue may lead to the development of a granulomatous disease, as observed with inhalation exposure to crystalline silica or carbon nanotubes (Harkema *et al*., 2013).

Corrosive effects in the respiratory tract may be non-specific, e.g. induced by highly acidic or basic substances like sulphuric acid. However, acute necrosis and loss of olfactory epithelium may also be observed following inhalation or bloodborne exposure to toxicants that require metabolic activation by the P450 system, such as 3-methylfuran. Once the basement membrane is exposed, cytokines are released and inflammation takes place (Harkema *et al*., 2013).

The term "respiratory tract irritation" is often used to indicate either or both of two different toxicological effects. These are i) cytotoxic effects in the affected tissue, and ii) sensory irritation.

Cytotoxic effects in the respiratory tract are comparable to dermal and eye irritation. These effects are characterised by inflammation (increased blood flow (hyperemia), local infiltration with white blood cells, swelling, oedema) and there may also be haemorrhage, and eventual necrosis and other pathological changes. The effects are in principle reversible. A recent publication has proposed the term “tissue irritation” for this kind of effects (Brüning *et al.,* 2014).

Chronic irritation can lead to repeated episodes of cell proliferation in the affected tissues, and this may increase the risk of tumour development. The nature of effects depends on the substance and its primarily targeted region; the severity of effects depends on the concentration and duration of exposure. In general, repeated exposure studies in animals focus on observing (histo)pathological evidence for tissue damage. In case overt tissue damage (mucosal erosion and ulceration) occurs, a non-specific cytotoxic action at the site of contact along the respiration route can be assumed. Depending on the concentration and duration of exposure a severity gradient of lesions from anterior to posterior regions can be observed (in contrast to effects in certain mucosa types depending on the metabolic activation of the test substance) and, depending on the severity and the extent of the lesions, adjacent submucosal tissues can also be affected (e.g. by cartilage destruction). Such lesions are not fully reversible due to scar formation or replacement of the original mucosa, or may induce other serious health effects such as marked bleeding or persistent airway obstruction.

"Sensory irritation" refers to the local and central reflex interaction of a substance with the autonomic nerve receptors, which are widely distributed in the mucosal tissues of the eyes and upper respiratory tract. Three substance or substance-group specific target sites of sensory irritation generating different responses can be identified: a) nasal (and eye) irritation, i.e. interaction with the trigeminal nerve, b) pharyngeal irritation, i.e. interaction with the glossopharyngeal nerve, and c) larynx and lower respiratory tract, i.e. interaction with the vagus nerve.

Sensory irritation leads to unpleasant sensations such as pain, burning, pungency, and tingling. The severity depends on the airborne concentration of the irritant rather than on the duration of exposure. Sensory irritation is a receptor-mediated effect, and usually occurs almost immediately upon exposure to the inhaled irritant. It leads to reflex involuntary responses such as sneezing, lacrimation, rhinorrhea, coughing, vasodilatation of blood vessels in the nasal passages, and changes in the rate and depth of respiration. In humans, protective behavioural responses such as covering the nose and mouth can also occur. Sensory irritation is distinct from odour sensation, which is mediated *via* different nerve pathways (olfactory). However, there is evidence that odour perception and other cognitive influences can affect the perception of sensory irritation in humans.

In rodents, sensory irritation leads to a reflex reduction in the respiratory rate (breath-holding). This reflex effect on respiration can be measured experimentally (determination of the RD_{50} value in the Alarie assay (Alarie, 1973)) although results may vary considerably depending on the species and strain of rodents, on the exposure duration (time should be long enough to induce changes), and results also show inter-laboratory variability. Investigations of the correlation between the results of the Alarie test and human data are difficult since the parameters examined in humans and mice are different and adequate human data to determine a human equivalent to the RD_{50} is not available at the moment. The results of a study by Cometto-Muniz *et al*. (1994) indicate that RD_{50} values in animals are not easily comparable with “nasal pungency thresholds” in humans.

As indicated, human data are mostly based on subjective experiences and need to be carefully controlled in order to prevent confounding by odour perception (Dalton, 2003; Doty *et al*., 2004). Validated questionnaires have been developed for the investigation of sensory irritation responses in human volunteers. Emphasis was given to developing a spectrum of objective measurements (see review by Arts *et al*., 2006). Compiling toxicological profiles for substances in the workplace demonstrates that sensory irritation often appears to be a very sensitive and relevant endpoint in human risk assessment. Accordingly, 40 % of the occupational exposure limit values (OELs) are based on the avoidance of sensory irritation (Dick and Ahlers 1998; Edling and Lundberg 2000; van Thriel *et al.,* 2006). This endpoint is related to the interaction of volatile substances with neuronal sensors located in mucous membranes of the respiratory tract and the eyes. In many cases, data from controlled human studies are either not available or inadequate, so OELs are predominantly derived from animal data investigating local effects in the respiratory tract. These effects are usually measured as tissue irritation. Comparison of human data on sensory irritation with data from subacute and subchronic inhalation studies in animals led to the proposal of a default assessment factor of 3 for extrapolating animal data concerning local irritating effects to humans (Brüning *et al*., 2014).

REFERENCES

Alarie Y (1973) Sensory irritation of the upper airways by airborne chemicals. Toxicol Appl Pharmacol 24:279-97.

Baadsgaard O and Wang T (1991) Immune regulation in allergic and irritant skin reactions. Int J Dermatol 30:161-72.

Barker JN, Mitra RS, Griffiths CEM, Dixi, VM and Nickoloff BJ (1991) Keratinocytes as initiators of inflammation. Lancet 337:211-4.

Berardesca E and Distante F (1994) The modulation of skin irritation. Contact Dermatitis 31:281-7.

Cometto-Muñiz JE and Cain WS (1994) Sensory reactions of nasal pungency and odor to volatile organic compounds: the alkylbenzenes. Am Ind Hyg Assoc J 55:811-7.

Dalton P (2003) Upper airway irritation, odor perception and health risk due to airborne chemicals. Toxicol Lett 140-141:239-48.

Doty RL, Cometto-Muñiz JE, Jalowayski AA, Dalton P, Kendal-Reed M and Hodgson M (2004) Assessment of upper respiratory tract and ocular irritative effects of volatile chemicals in humans. Crit Rev Toxicol 34:85-142.

EC (2003) Technical Guidance Document in support of Commission Directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Available at: https://echa.europa.eu/documents/10162/16960216/tgdpart1_2ed_en.pdf

Frosch PJ and Pilz B (1995) Irritant patch test techniques. *In:* Handbook of Non-invasive Methods and the Skin (Serup J and Jemec GBE Eds.) CRC Press, Boca Raton, FL, USA, pp.587-91.

Harkema JR, Nikula KJ and Haschek WM (2013) Chapter 51: Respiratory System. *In:* Haschek and Rousseaux's Handbook of Toxicologic Pathology, 3rd edition, volume III, (Haschek WM, Rousseaux CG, Wallig MA Eds.) Elsevier/Academic Press, Amsterdam, The Netherlands, pp.1935-2003.

Hawk JLM, Black AK, Jaenicke KF, Barr RM, Soter NA, Mallett AI, Gilchrest BA, Hensby CN, Parrish JA and Greaves MW (1983) Increased concentrations of arachidonic acid, prostaglandins E2, D2, and 6-oxo-F1 alpha, and histamine in human skin following UVA irradiation. J Invest Dermatol 80:496-9.

Hunziker T, Brand CU, Kapp A, Waelti ER and Braathen LR (1992) Increased levels of inflammatory cytokines in human skin lymph derived from sodium lauryl sulphate-induced contact dermatitis. Br J Dermatol 127:254-7.

Lammintausta KH and Maibach HI (1990) Contact dermatitis due to irritation. *In:* Occupational Skin Disease (Adams RM Ed.) 2nd edition, Saunders WB, Philadelphia, PA, USA, pp.1-15.

Serup J (1995) The spectrum of irritancy and application of bioengineering techniques. *In:* Irritant Dermatitis. New Clinical and Experimental Aspects. Current Problems in Dermatology, vol.23 (Elsner P and Maibach HI Eds.) Karger, Basel, Switzerland, pp.131-43.

Sondergard J, Graevers MW and Jorgenson HP (1974) Recovery of prostaglandins in human primary irritant dermatitis. Arch Dermatol 110:556-7.

Walker JD, Gerner I, Hulzebos E and Schlegel K (2004) (Q)SARs for Predicting Skin Irritation and Corrosion: Mechanisms, Transparency and Applicability of Predictions. QSAR Comb Sci 23:721-5.

Appendix R.7.2–2 (Q)SARs and expert systems for skin corrosion and irritation

Content of Appendix R.7.2-2:

- Literature-based QSAR models
- Commercial models
- BfR rule-base
- OECD QSAR Toolbox

In principle, Annex XI to the REACH Regulation allows for an adaptation of the standard information requirements by using (Q)SARs, including the prediction of non-irritancy. However, for the endpoint skin corrosion/irritation, only very few of the currently available models are suitable for this purpose if used as stand-alone methods. Nevertheless, such models can still have merit when used as supporting information or in *Weight-of-Evidence* approaches and for positive prediction of skin corrosion/irritation.

LITERATURE-BASED QSAR MODELS

In the open scientific literature, (Q)SARs have been based on continuous (e.g. Primary Irritation Indices) or categorical (e.g. EU classifications) measures of skin irritation.

For defined classes of substances, categorical QSARs have been reported for discriminating between corrosives and non-corrosives (Barratt, 1996a, 1996b), and between skin irritants and non-irritants (Smith *et al.*, 2000a, 2000b). These studies did not actually provide a transparent algorithm for classifying chemicals, so they are of limited value for regulatory use. However, they illustrate the feasibility of developing such models.

A linear discriminant model for distinguishing between irritant and non-irritant liquid esters in human volunteers was reported by Smith *et al*. (2000a). As mentioned above the exact algorithm is not clear. In addition the primary irritation index for human irritation may need translation when these scores are considered for classification. However, the results could be informative for future model development for esters, since they indicate that irritant esters can be distinguished from non-irritants on the basis of a limited number of physico-chemical parameters.

For defined classes of substances, continuous QSARs for predicting the Primary Irritation Index (PII) have also been published (Barratt, 1996b; Hayashi *et al.*, 1999; Kodithala *et al.*, 2002). For example, the application of stepwise regression analysis to a set of 52 neutral and electrophilic organic substances produced the following model:

$$PII = 1.047 \log P - 0.244\ MV + 0.888\ DM + 0.353$$

$$N=52,\ r^2 = 0.422,\ r_{cv}^2 = 0.201,\ s=1.376,\ F=11.70$$

This equation indicates that the PII has a positive dependence on log P (logarithm of the octanol-water partition coefficient) and DM (dipole moment), and a negative dependence on MV (molecular volume). This model has a low goodness-of-fit (r^2) and a poor predictivity (as reflected by r_{cv}^2), so is not recommended for regulatory use. More research is needed into the development of models for predicting PII and it should be considered whether the information generated could be used in the setting of DNELs.

Some limited evidence indicates that the reactive effects of acids and bases can be predicted by using the acid/base dissociation constant (pKa), which can itself be predicted by using commercially available software products, such as the SPARC program. Evidence for the usefulness of pKa as a predictor of skin irritation for acids has been provided by Berner *et al.* (1988, 1989, 1990), whereas evidence for the usefulness of pKa as a predictor of skin irritation for bases has been provided by Nangia *et al*. (1996). Barratt also used pKa for predicting the effects of acids and bases (Barratt, 1995a). These studies did not address the question of how to use pKa where there are multiple functional groups in the substance of interest, and therefore multiple ionisation constants. Based on current knowledge, no clear recommendations can be made on how to use pKa information.

An overview on the available literature-based models for skin corrosion/irritation is given in Table R.7.2–6.

Table R.7.2–6 Available literature-based models for skin corrosion/irritation.

Reference	***Content***
QSAR models	
Barratt (1996a) Quantitative structure-activity relationships (QSARs) for skin corrosivity of organic acids, bases and phenols: Principal components and neural network analysis of extended datasets.	This paper describes QSAR models relating skin corrosivity data of organic acids, bases and phenols to their log(octanol/water partition coefficient), molecular volume, melting point and pK(a).
Barratt (1996b) Quantitative structure-activity relationships for skin irritation and corrosivity of neutral and electrophilic organic chemicals.	This paper describes QSAR models derived by relating skin irritation and corrosivity data of neutral and electrophilic organic chemicals to their log(octanol/water partition coefficient) (logP), molecular volume, dipole moment and 1/molecular weight.
Barrat (1996c) The use of in vitro cytotoxicity measurements in (Q)SAR methods for the prediction of the skin corrosivity potential of acids.	This paper describes quantitative structure-activity relationships (QSAR) methods that relate the severity of skin corrosivity (designated by the EC risk phrases R34 and R35) of acids to parameters that model their skin permeability and cytotoxicity. Skin permeability was modelled by log(octanol/water partition coefficient), molecular volume and melting point, while the cytotoxicity of the acids was accounted for by their pKa values and the in vitro cytotoxicity of their sodium salts towards Swiss mouse embryo 3T3 cells.
Gerner *et al.* (2004) Quantitative structure-property relationships modeling of skin irritation.	This paper describes limit values for specific physico-chemical properties that are appropriate for identifying chemical substances that have no skin irritation or corrosion potential. These physicochemical properties include melting point, molecular weight, octanol/water partition coefficient, surface tension, vapour pressure, aqueous solubility and lipid solubility.
Golla *et al.* (2009) Quantitative structure-property relationships modeling of skin irritation.	This paper describes a skin irritation QSPR model based on rabbit Draize test data for 186 compounds, which included chemicals from diverse molecular classes. The effectiveness of using a combination of traditional, functional group and structural descriptors has been studied. The effects of molecular size, reactivity and skin penetration on skin irritation have been also analysed.

Hayashi *et al.* (1999) A quantitative structure-activity relationship study of the skin irritation potential of phenols.	This paper describes QSARs for skin irritation potential derived using twenty-four phenols, using the following descriptors: the absolute hardness calculated from HOMO and LUMO energy levels for reactivity, and log P for permeability. The selection of the descriptors was based on the hypothesis that skin irritation is induced by reaction of phenols with macromolecules present in epidermal and dermal levels of the skin.
Hulzebos *et al.* (2005) Use of structural alerts to develop rules for identifying chemical substances with skin irritation or skin corrosion potential.	This paper describes the identification and categorisation of structural alerts for acute skin lesions as irritation or corrosion or a combination of corrosion/irritation alerts.
Kodithala *et al.* (2002) Prediction of skin irritation from organic chemicals using membrane-interaction QSAR analysis.	This paper describes membrane-interaction QSAR analysis carried out for a training set of 22 hydroxy organic compounds for which the Draize skin irritation scores, PII, had been determined. Skin irritation potency is predicted to increase with (1) increasing effective concentration of the compound available for uptake into phospholipid-rich regions of a cellular membrane, (2) increasing binding of the compound to the phospholipid-rich regions of a cellular membrane, and (3) the chemical reactivity of the compound as reflected by the highest occupied molecular orbital (HOMO) and/or lowest unoccupied molecular orbital (LUMO) of the molecule.
Walker *et al.* (2004) (Q)SARs for Predicting Skin Irritation and Corrosion: Mechanisms, Transparency and Applicability of Predictions.	This paper describes previously-developed (Q)SARs for predicting skin irritation and corrosion, proposes mechanisms of skin irritation and corrosion, and discusses the transparency and applicability of predictions.
Walker *et al.* (2005) The Skin Irritation Corrosion Rules Estimation Tool (SICRET).	This paper describes the Skin Irritation Corrosion Rules Estimation Tool (SICRET) that was developed to allow others to estimate whether their chemicals are likely to cause skin irritation or skin corrosion. SICRET uses physicochemical property limits to identify chemicals with no skin corrosion or skin irritation potential.
Whittle (1996) Skin corrosivity potential of fatty acids: *In vitro* rat and human testing and (Q)SAR studies.	This paper investigates the corrosive potential of a series of fatty acids - propanoic acid (C3), butanoic acid (C4), hexanoic acid (C6), octanoic acid (C8), decanoic acid (C10) and dodecanoic acid (C12) - according to an *in vitro* skin corrosivity test (IVSCT) using both rat skin and human skin. The results are discussed in the context of a QSAR for the corrosivity of organic acids, with the putative mechanism that corrosivity is a function of the ability of the chemical to permeate the skin together with its cytotoxicity, expressed in this case as acidity (pK(a)).
Worth and Cronin (2001) The use of pH measurements to predict the potential of chemicals to cause acute dermal and ocular toxicity.	This paper presents a the development of classification models based on pH data for predicting the potential of chemicals to cause skin corrosion, skin irritation and eye irritation. The possible application of these models in the context of tiered testing strategies is discussed.

Reviews and evaluation of existing models	
Gallegos Saliner *et al.* (2006) Review of Literature-Based Models for Skin and Eye Irritation and Corrosion.	This report reviews the state-of-the-art of *in silico* and *in vitro* methods for assessing dermal and ocular irritation and corrosion. In this review, emphasis is placed on literature-based QSAR models for skin and eye irritation and corrosion as well as computer-based expert systems.
Gallegos Saliner *et al.* (2008) Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion.	This paper reviews the state-of-the-art of *in silico* methods for assessing dermal and ocular irritation and corrosion. It is based on an in-depth review performed by the European Chemicals Bureau of the European Commission: Joint Research Centre . The most widely used *in silico* approaches are classified into methods to assess (1) skin irritation, (2) skin corrosion and (3) eye irritation. In this review, emphasis is placed on literature-based (Q)SAR models.
Gallegos Saliner *et al.* (2007) Evaluation of SARs for the prediction of skin irritation/corrosion potential: structural inclusion rules in the BfR decision support system.	This work evaluates the structural inclusion rules implemented in the Decision Support System for skin irritation and corrosion developed at the German Bundesinstitut für Risikobewertung (BfR) for predicting the absence of skin irritation and/or corrosion. The following assessments were performed: (a) a confirmation of the structural rules by rederiving them from the original training set (1358 substances), and (b) an external validation by using a test set of 200 chemicals not used in the derivation of the rules.
Mombelli (2008) An evaluation of the predictive ability of the QSAR software packages, DEREK, HAZARDEXPERT and TOPKAT, to describe chemically-induced skin irritation.	This paper reports the performance of the skin irritation module of three commercially-available software packages: DEREK, HAZARDEXPERT and TOPKAT. Their performances were tested on the basis of data published in the literature for 116 chemicals.
Rorije and Hulzebos (2005) Evaluation of (Q)SARs for the prediction of Skin Irritation/Corrosion Potential. Physicochemical exclusion rules.	This work evaluates the physical-chemical rule-base incorporated in the Decision Support System for skin irritation and corrosion developed at the German Bundesinstitut für Risikobewertung (BfR) for predicting the absence of skin irritation and/or corrosion. This evaluation includes 1) the compliance of the rule-base with the OECD principles on (Q)SARs, 2) the derivation of the (Q)SAR rules, 3) the external validation of these rules, including an assessment of the suitability of the dataset used for validation.

Further details on these models can be found in Chapter 3 of the JRC report “Alternative methods for regulatory toxicology - a state-of-the-art review” (Worth *et al.*, 2014).

COMMERCIAL MODELS

There is a number of software tools available that provide access to QSARs for skin corrosion/ irritation.

TOPKAT, which is commercialised by Accelrys (http://accelrys.com/solutions/scientific-need/predictive-toxicology.html), incorporates models to discriminate severe irritants from non-severe irritants, as well as mild/moderate irritants from non-irritants. These models are based on work by Enslein *et al*. (1987). The algorithm of TOPKAT is not very transparent. The model predicts a probability of a weak/mild/moderate and severe irritation. It states that probabilities <0.3 and >0.7 give sufficient certainty of the prediction. The model gives the sensitivity and specificity values of the specific classes such as acyclic etc., which are mostly around or above 90%. It also shows similar structures from the TOPKAT perspective including the experimental result. The TOPKAT predictions of weak/mild/moderate and severe irritation need to be translated to consider them for classification. The models indicate whether the prediction is in the applicability domain of the model.

There is a rule-base for irritation in **Derek Nexus** (Sanderson and Earnshaw, 1991; Combes and Rodford, 2004), which is developed and regularly updated by LHASA Ltd (http://www.lhasalimited.org/products/derek-nexus.htm). To predict toxicity, the program checks whether any alerts within the query structure match previously characterised toxicophores (substructure with potential toxic effect) in the knowledge base. The reasoning engine then assesses the likelihood of a structure being toxic, and a message indicating the nature of the toxicological hazard is provided together with relevant literature references. There are nine levels of confidence: certain, probable, plausible, equivocal, doubted, improbable, impossible, open, contradicted. The Derek Nexus rule-base has 25 structural alerts for the prediction of skin corrosion/irritation. There are some combined alerts for respiratory tract irritation and irritation of the gastrointestinal tract but these are not specific to skin corrosion or irritation. If Derek Nexus does not make a prediction of corrosion or irritation, it cannot be concluded that there is no effect – it could mean that none of the known alerts was found to be present in the substance of interest or it was outside the applicability domain of that specific alert. The Derek Nexus model is transparent in its algorithm, when the model is fired showing the structural alert and its limitations. The alert is supported with literature references and sometimes with example substances. The example substances are supposed to support the mechanistic reasoning. The Derek Nexus model can be used for positive identification of skin irritation. The confidence levels have to be taken into account for the purpose of classification. The Derek Nexus model cannot be used to predict non-corrosion/irritation as the model only contains alerts that detect the presence of corrosion/irritation.

HazardExpert is a rule-based software tool developed and commericalised by CompuDrug Chemistry Ltd. (http://www.compudrug.com/hazardexpertpro) for predicting the toxicity of organic substances in humans and in animals (Smithing and Darvas, 1992). HazardExpert uses a fragment-based approach to predict toxicokinetic effects and various human health effects, including membrane irritation. Since this endpoint is not clearly defined in HazardExpert, it is recommended not to use it directly for the assessment of skin or eye irritation. However, it could be used as supplementary information in a *Weight-of-Evidence* approach for positive prediction.

The Multiple Computer Automated Structure Evaluation (**MultiCASE**) program, developed by MultiCASE Inc. (http://www.multicase.com/case-ultra-models#skin_eye_tox_bundle), is an automated rule induction tool that automatically identifies molecular fragments likely to be relevant to the activity of molecules (Klopman, 1992; Klopman *et al.*, 1993). It also provides an indication of the importance of these fragments in relation to the potency of the molecules containing them. MultiCASE can be used to predict various human health endpoints, including eye irritation (Klopman *et al.*, 1993; Rosenkranz *et al.*, 1998). However, it is not clear how to relate the MultiCASE scoring system to Draize scores or regulatory classifications. In principle, the MultiCASE model can be used for positive and negative indications of skin irritation. The structural alert is provided as well as information on its internal validation. The MultiCASE model also indicates whether it is in the applicability domain of the model. The MultiCASE predictions of weak/mild/moderate and severe irritation need to be translated to consider them for classification.

ACD/Labs Percepta Predictors (http://www.acdlabs.com/products/percepta/predictors.php), developed by ACD/Labs, includes a module for skin and eye irritation. It estimates the potential of a compound to cause eye or skin irritation in a standard rabbit Draize test. The predictions are reported as qualitative irritation categories (not irritating, slightly irritating, irritating, highly irritating, and corrosive). Probabilistic models are supplemented by an expert system that identifies Structural Alerts relevant to the irritation properties of compounds. Overall, 21 structural alerts were formulated for rabbit eye irritation, and 17 alerts for the rabbit skin irritation case. The categorisation of effect needs to be compared to the CLP cut-offs if application for REACH purposes is intended.

PaDEL-DDPredictor includes several models for skin and eye irritation and corrosion (http://padel.nus.edu.sg/software/padelddpredictor/). The models have been built on a training set of 1707 compounds using one and two dimensional descriptors. The final predictions rely on consensus models based on majority voting from base models predictions. The applicability domain is defined by the range of descriptors for compounds in the training set.

QSAR PREDICTION REPOSITORY

The Danish EPA (http://qsar.food.dtu.dk/) has developed an in-house MultiCASE model for predicting severe versus mild skin irritation based on 800 test results taken from RTECS (Registry of Toxic Effects of Chemical Substances), the HSDB (Hazardous Substances Data Bank) and the former official list of EU-classified substances (Annex I of Directive 67/548/EEC, now replaced by Annex VI to the CLP Regulation). It is not clear how the RTECS and HSDB classification criteria for irritation correlate with the EU criteria. Due to limitations in the information for assessing the reliability of the prediction, these predictions are difficult to use in the regulatory context.

BFR DECISION SUPPORT SYSTEM

A decision support system (DSS) developed by the German Federal Institute for Risk Assessment (BfR) uses physico-chemical exclusion rules to predict the absence of skin corrosion/irritation potential in combination with structural inclusion rules (SARs) to predict the presence of such potential (Gerner *et al.*, 2004; Hulzebos *et al.*, 2005; Walker *et al.*, 2004). The exclusion rules are based on physico-chemical properties such as molecular weight, aqueous solubility, and log K_{ow}, whereas the inclusion rules are based on sub-structural molecular features. The physico-chemical rules are assumed to implicitly take into account bioavailability (skin penetration) whereas the structural rules take reactivity into account. The physico-chemical and structural rule-bases are designed to predict the former EU risk phrases for skin irritation (R38) and skin corrosion (R34 and R35).

The exclusion rules have the following general form:

IF (physico-chemical property) A THEN predict the absence of toxic effect B

Example: IF Log K_{ow} < -3.1 THEN the substance does not need to be considered for classification

Some of the exclusion rules can be applied to all structures within the domain, whereas others only refer to a subset containing certain elements.

The structural inclusion rules take the following general form:

IF (substructure) A THEN predict the occurrence of toxic effect B.

Example: IF *Chlorosilane alert is present* THEN the substance needs to be considered for "corrosive" classification.

The performance of the BfR physico-chemical rule-base for predicting the absence of skin effects has been assessed by the RIVM (Rorije and Hulzebos, 2005), whereas the structural rule-base for predicting the occurrence of skin effects has been assessed by the ECB (Gallegos Saliner *et al.*, 2007). The endpoint is the former EU (DSD) classification and the algorithms and domain of applicability are transparent. However, the exact chemical structures of the training set are not disclosed to users of the model, due to the data originating from the confidential notification procedure at the time of the development of the system. Though the rules are empirically derived, a mechanism of action can be deduced. Thus, in principle, the resulting predictions can be used as the basis for classification by comparison with CLP criteria. It should be determined, on a case-by-case basis, whether the predictions for a given substance provide a sufficient basis for classification, or whether additional information is needed in a *Weight-of-Evidence* approach.

OECD QSAR TOOLBOX

The freely downloadable OECD QSAR Toolbox software (http://www.qsartoolbox.org/) covers the skin corrosion/irritation endpoint with one experimental database and two profilers.

In more detail, the database of experimental data (called "Skin irritation" in the software) refers to the endpoint primary irritation index and collects the data available in:

1. The RIVM Skin Irritation database, which contains Primary Skin Irritation Indices from skin irritation tests from the following sources: ECVAM Workshop 6 on Corrosivity (Barratt (1995b); Botham *et al.* (1995)), and ECETOC Technical Report No.66 on Skin Irritation and Corrosion Reference Chemicals Data Bank (ECETOC, 1995).

2. Experimental results for Primary Skin Irritation Indices from LJMU. Additional experimental results gathered from OECD SIDS Dossiers published between 1992 and 2009 were added in 2010.

The OECD QSAR Toolbox allows for the identification of analogues based on mechanistic and endpoint specific profilers, and for the prediction of skin irritation/corrosion through the use of profilers (BfR rule-base), readacross, trend analysis and QSAR models. Information about inclusion and exclusion rules, details on the performance of the exclusion rules, and applicable chemical class-specific rules for the results of the Skin irritation/corrosion profiler can be found by searching the context menu in the the OECD QSAR Toolbox software.

REFERENCES

Barratt MD (1995a) Quantitative structure activity relationships for skin corrosivity of organic acids, bases and phenols. Toxicol Lett 75:169-176.

Barratt MD (1995b) Appendix A: Quantitative Structure-activity Relationships for Skin Corrosivity - ECVAM Workshop 6: corrosivity. Alternatives to Laboratory Animals 23:243-55.

Barratt MD (1996a) Quantitative Structure-Activity Relationships for Skin Corrosivity of Organic Acids, Bases and Phenols: Principal Components and Neural Network Analysis of Extended Sets. Toxicol *in Vitro* 10:85-94.

Barratt MD (1996b) Quantitative Structure-Activity Relationships for Skin Irritation and Corrosivity of Neutral and Electrophilic Organic Chemicals. Toxicol *in Vitro* 10:247-56.

Barratt MD, Dixit MB and Jones PA (1996c) The use of *in vitro* cytotoxicity measurements in (Q)SAR methods for the prediction of the skin corrosivity potential of acids. Toxicol *in Vitro* 10:283-90.

Berner B, Wilson DR, Guy RH, Mazzenga GC, Clarke FH, Maibach HI (1988) The relationship of pKa and acute skin irritation in man. Pharm Res 5:660-3.

Berner B, Wilson DR, Guy RH, Mazzenga GC, Clarke FH, Maibach HI (1989) Relationship of pKa and acute skin irritation in humans. J Toxicol Cutan Ocul Toxicol 8:481–92.

Berner B, Wilson DR, Steffens RJ, Mazzenga GC, Hinz R, Guy RH, Maibach HI (1990) The relationship between pKa and acute skin irritation for a series of basic penetrants in man. Fundam Applied Toxicol 15:760–6.

Botham PA, Chamberlain M, Barratt MD, Curren RD, Esdaile DJ, Gardner JR, *et al.* (1995) A Prevalidation study on *in vitro* skin corrosivity testing: The Report and recommendations of ECVAM Workshop 6. Altern Lab Anim 23:219-42.

Combes RD and Rodford R (2004) The Use of Expert Systems for Toxicity Prediction - Illustrated With Reference to the DEREK Program. *In:* Predicting Toxicity and Fate (Cronin, MTD and Livingstone DJ Eds.) CRC Press, Boca Raton, FL, USA, pp.193-204.

ECETOC (1995) Technical Report No.66, Skin Irritation and Corrosion: Reference Chemicals Data Bank. Available at: http://www.ecetoc.org/publications

Gallegos Saliner A, Patlewicz G and Worth AP (2006). Review of literature-based models for skin and eye irritation and corrosion. JRC report EUR 22320 EN. European Chemicals Bureau, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/doc/QSAR_Review_Irritation.pdf

Gallegos Saliner A, Tsakovska I, Pavan M, Patlewicz G and Worth AP (2007) Evaluation of SARs for the prediction of skin irritation/corrosion potential - structural inclusion rules in the BfR decision support system. SAR QSAR Environ Res 18:331-42.

Gallegos Saliner A, Patlewicz G and Worth AP (2008) A Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion. QSAR Comb Sci 27:49-59.

Gerner I, Schlegel K, Walker JD and Hulzebos E (2004) Use of physico-chemical property limits to develop rules for identifying chemical substances with no skin irritation or corrosion potential. QSAR Comb Sci 23:726-33.

Golla S, Madihally S, Robinson RL Jr and Gasem KA (2009) Quantitative structure-property relationships modeling of skin irritation. Toxicol *in Vitro* 23:176-84.

Hayashi M, Nakamura Y, Higashi K, Kato H, Kishida F and Kaneko H (1999) A Quantitative Structure-Activity Relationship Study of the Skin Irritation Potential of Phenols. Toxicol *in Vitro* 13:915-22.

Hulzebos E, Walker JD, Gerner I and Schlegel K (2005) Use of structural alerts to develop rules for identifying chemical substances with skin irritation or skin corrosion potential. QSAR Comb Sci 24:332-42.

Klopman G (1992) MULTICASE.1. A Hierarchical Computer Automated Structure Evaluation Program. Quant Struct-Act Rel 11:176-84.

Klopman G, Ptchelintsev D, Frierson M, Pennisi S, Renskers K and Dickens M (1993) Multiple computer automated structure evaluation methodology as an alternative to *in vitro* eye irritation testing. Altern Lab Anim 21:14-27.

Kodithala K, Hopfinger AJ, Thompson ED and Robinson MK (2002) Prediction of skin irritation from organic chemicals using membrane-interaction QSAR Analysis. Toxicol Sci 66:336-46.

Mombelli E (2008) An evaluation of the predictive ability of the QSAR software packages, DEREK, HAZARDEXPERT and TOPKAT, to describe chemically-induced skin irritation. Altern Lab Anim 36:15–24.

Rorije E and Hulzebos E (2005) Evaluation of (Q)SARs for the prediction of Skin Irritation/Corrosion Potential. Physico-chemical exclusion rules. Final report for ECB contract IHCP.B430206. European Chemicals Bureau, Joint Research Centre, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/information-sources/qsar-document-area/Evaluation_of_Skin_Irritation_QSARs.pdf/view

Rosenkranz HS, Zhang YP and Klopman G (1998) The Development and Characterisation of a Structure-activity Relationship Model of the Draize Eye Irritation Test. Altern Lab Anim 26:779-809.

Sanderson DM and Earnshaw CG (1991) Computer Prediction of Possible Toxic Action from Chemical Structure; The DEREK System. Hum Exp Toxicol 10:261-73.

Smithing MP and Darvas F (1992) HazardExpert: An Expert System for Predicting Chemical Toxicity. *In:* Food Safety Assessment (Finlay SF, Robinson SF, Armstrong DJ Eds.) American Chemical Society, Washington, DC, USA, pp.191-200.

Solimeo R, Zhang J, Kim M, Sedykh A and Zhu H (2012) Predicting Chemical Ocular Toxicity Using a Combinatorial QSAR Approach. Chem Res Toxicol 25: 2763-69.

Walker JD, Gerner I, Hulzebos E and Schlegel K (2004) (Q)SARs for Predicting Skin Irritation and Corrosion: Mechanisms, Transparency and Applicability of Predictions. QSAR Comb Sci 23:721-5.

Walker JD, Gerner I, Hulzebos E, Schlegel K (2005) The Skin Irritation Corrosion Rules Estimation Tool (SICRET). QSAR Comb Sci 24:378-84.

Whittle E, Barratt MD, Carter JA, Basketter DA and Chamberlain M (1996) Skin corrosivity potential of fatty acids: *In vitro* rat and human testing and (Q)SAR studies. Toxicol *In vitro* 10:95-100.

Worth AP and Cronin MTD (2001) The use of pH measurements to predict the potential of chemicals to cause acute dermal and ocular toxicity. Toxicol 169:119-31.Worth A, Barroso J, Bremer S, Burton J, Casati S, Coecke S, Corvi R, Desprez B, Dumont C, Gouliarmou V, Goumenou M, Gräpel R, Griesinger C, Halder M, Janusch Roi A, Kienzler A, Madia F, Munn S, Nepelska M, Paini A, Price A, Prieto P, Rolaki A, Schäffer M, J. Triebe, Whelan M, Wittwehr C and Zuang V (2014) Alternative methods for regulatory toxicology – a state-of-the-art review. JRC Science and Policy Reports, Report EUR 26797 EN. European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Systems Toxicology Unit, Institute for Health and Consumer Protection, European Commission Joint Research Centre, Ispra, Italy. Available at: https://ec.europa.eu/jrc/sites/default/files/echa_jrc_sla_report_public_05-09-14_withcover_ipo.pdf

Appendix R.7.2–3 (Q)SARs and expert systems for serious eye damage and eye irritation

Content of Appendix R.7.2-3:

- Literature-based QSAR models
- Commercial models
- BfR decision support system
- OECD QSAR Toolbox

In principle, Annex XI to the REACH Regulation allows for an adaptation of the standard information requirements by using (Q)SARs, including the prediction of non-irritancy. However, for the endpoint serious eye damage/eye irritation, only very few of the currently available models are suitable for this purpose if used as stand-alone methods. Nevertheless, such models can still have merit when used as supporting information or in *Weight-of-Evidence* approaches and for positive prediction of serious eye damage/eye irritation.

LITERATURE-BASED QSAR MODELS

In the open scientific literature, (Q)SARs have been based on continuous (e.g. molar eye scores) or categorical (e.g. EU classifications) measures of eye irritation. Examples of mathematical (continuous) models have been published models by Sugai *et al.* (1991) and Cronin *et al.* (1994), whereas examples of categorical models have been published by Sugai *et al.* (1990) and by Barratt (1997).

Regression models based on solvatochromic parameters can be used for predicting the degree of eye irritation, as illustrated by Abraham and coworkers (Abraham, 1993; Abraham *et al.*, 1998). The mechanistic basis of these models is that a substance is transferred from a pure organic liquid to an organic solvent phase consisting of the tear film and cell membranes on the surface of the eye. The more soluble the organic liquid in the initial phase, the greater the degree of irritation is. These models are worthy of further characterisation. However, for routine regulatory use, information on a number of so-called Abraham descriptors would also need to be made available.

Neural network approaches can also be used to model eye irritation (e.g. Patlewicz *et al.*, 2000). At present, however, many of these models lack transparency, especially in the algorithm. However if the training sets are provided as well as validation information they could possibly be used in a *Weight-of-Evidence* approach. Mechanistic reasoning should also be provided.

An approach called Membrane-Interaction QSAR analysis, developed by Kulkarni *et al.* (2001), provides a means of incorporating molecular dynamic simulations to generate membrane-solute interaction properties. The development and application of models based on molecular simulations requires the use of specialised expertise and software. They could be used to increase understanding of the mechanisms of eye irritation.

A classification approach called Embedded Cluster Modelling (ECM) provides a means of generating *elliptic models* in two or more dimensions (Worth and Cronin, 2000), so that irritants can be transparently identified as those substances located within the boundaries of the ellipse. The statistical significance of these “embedded clusters” can be verified by cluster significance analysis (CSA), as illustrated for an eye irritation dataset by Cronin (1996).

Different methods were applied to a dataset of 119 organic liquids classified as I (irritant) or NI (non-irritant) according to former EU classification criteria. The classification models (CMs) were developed by applying linear discriminant analysis (LDA), binary logistic regression (BLR), and classification tree (CT) analyses, using a single predictor variable (molecular weight), and assigning equal probabilities for the two classes (I/NI). (Worth and Cronin, 2003).

All of these models are simple to apply and are associated with a transparent algorithm. The statistics illustrate the inevitable trade-offs that result from the selection of different cut-off values. Thus, the BLR model does not identify many irritants, but it does so with a high degree of confidence. Conversely, the CT does not identify many of the non-irritants, but it has a low false negative rate. Thus, the combined use of the BLR and CT models could be useful for distinguishing between eye irritants and non-irritants.

An overview on the available literature-based models for serious eye damage/eye irritation is provided in Table R.7.2–7.

Table R.7.2–7 Available literature-based models for serious eye damage / eye irritation.

Reference	***Content***
QSAR models	
Abraham *et al.* (2003) Draize rabbit eye test compatibility with eye irritation thresholds in humans: a quantitative structure-activity relationship analysis.	Draize rabbit eye test scores, as modified maximum average score (MMAS), for 68 pure bulk liquids were adjusted by the liquid-saturated vapor pressure P. These 68 adjusted scores, as log (MMAS/P), were shown to be completely equivalent to eye irritation thresholds (EIT), expressed as log (1/EIT), for 23 compounds in humans. Thus, for the first time the Draize eye test in rabbits for pure bulk liquids is shown to be perfectly compatible with eye irritation thresholds in humans.
Barratt (1995) The role of structure-activity relationships and expert systems in alternative strategies for the determination of skin sensitisation, skin corrosivity and eye irritation.	This paper describes the derivation of a set of structural alerts for skin sensitisation, which have been incorporated into the expert system DEREK, and of Quantitative structure-activity relationships (QSARs) derived for predicting the skin corrosivity (for organic acids and bases) and for the eye irritation potential (for neutral organic chemicals).
Gerner *et al.* (2005) Assessment of the Eye Irritating Properties of Chemicals by Applying Alternatives to the Draize Rabbit Eye Test: The Use of QSARs and *In Vitro* Tests for the Classification of Eye Irritation.	This paper evaluates and discusses the nature of eye lesions and their importance for classification and labelling of possible hazards to human eyes, with a view to promoting the development of specific *in vitro* assays which are able to discriminate between eye damage, moderate eye irritation, and minor irritation effects which are completely reversible within a few days. Structural alerts for the prediction of eye irritation/corrosion hazards to be classified and labelled according to international classification criteria, are presented, which should be validated in accordance with internationally agreed (OECD) principles for (Q)SAR system validation. Physicochemical limit values for prediction of the absence of any eye irritation potential relevant for human health can make available a definition of the applicability domains of alternative methods developed for the replacement of the Draize eye irritation test.
Solimeo *et al.* (2012) Predicting Chemical Ocular Toxicity Using a Combinatorial QSAR Approach.	This paper describes QSAR models for a set of small molecules with animal ocular toxicity data compiled by the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods.

Reviews and evaluation of existing models	
Gallegos Saliner *et al.* (2006) Review of Literature-Based Models for Skin and Eye Irritation and Corrosion.	This report reviews the state-of-the-art of *in silico* and *in vitro* methods for assessing dermal and ocular irritation and corrosion. In this review, emphasis is placed on literature-based QSAR models for skin and eye irritation and corrosion as well as computer-based expert systems.
Gallegos Saliner *et al.* (2008) Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion.	This paper reviews the state-of-the-art of *in silico* methods for assessing dermal and ocular irritation and corrosion. It is based on an in-depth review performed by the European Chemicals Bureau of the European Commission: Joint Research Centre . The most widely used *in silico* approaches are classified into methods to assess (1) skin irritation, (2) skin corrosion and (3) eye irritation. In this review, emphasis is placed on literature-based (Q)SAR models.
Tsakovska *et al.* (2005) Evaluation of (Q)SARs for the prediction of Eye Irritation/Corrosion Potential - physicochemical exclusion rules.	In this study, an evaluation was performed of the physicochemical BfR-DSS rule-base (comprising 31 physicochemical exclusion rules) for predicting the absence of eye irritation/corrosion. According to the results of this study: a) the physicochemical exclusion rules for eye irritation/corrosion comply well with the OECD validation principles; b) predictions of no adverse effect (NOT R34/R35/R36/R41) can be made for 20 out of the 199 chemicals in the test set; c) 3 of the 45 irritants/corrosives are falsely predicted as non-irritant or non corrosive; d) the probability of a negative prediction being correct (Negative Predictive Value) is 0.87; and e) approximately 10% of Draize rabbit eye tests could be avoided by relying on the predictions of no adverse effect.
Tsakovska *et al.* (2007) Evaluation of SARs for the prediction of eye irritation/corrosion potential - structural inclusion rules in the BfR decision support system.	This work summarises the results of a study carried out by the ECB to assess the performance of the BfR structural rule-base. The assessment included: (a) evaluation of the structural alerts by using the training set of 1341 substances with experimental data for eye irritation and corrosion; and (b) external validation by using an independent test set of 199 substances. The test set of 199 substances contained 154 (77%) non-labelled substances and 45 (23%) labelled as eye irritants/corrosives, subdivided as follows: (i) 10 R36 substances (5%); (ii) 28 R41 substances (14%); and (iii) 7 substances (4%) labeled R34 or R35.

Further details on these models can be found in Chapter 4 of the JRC report "Alternative methods for regulatory toxicology - a state-of-the-art review" (Worth *et al.*, 2014).

COMMERCIAL MODELS

There is a number of software tools available that provide access to QSARs for serious eye damage/eye irritation.

The **TOPKAT** software (http://accelrys.com/solutions/scientific-need/predictive-toxicology.html) includes models for eye irritation based on structural fragments. These models were originally developed by Enslein *et al.* (1988). The TOPKAT algorithm is not very transparent. The model predicts a probability of a weak/mild/moderate and severe irritation. It states that probabilities <0.3 and >0.7 give sufficient certainty of the prediction. The model gives the sensitivity and specificity values of the specific classes such as acyclic, which are mostly around or above 90%. It also shows similar structures from the TOPKAT perspective including the experimental result. The TOPKAT predictions weak/mild/moderate and severe irritation need to be translated to consider them for classification. The models indicate whether the prediction is in the applicability domain of the model.

There is a rulebase for irritation in **Derek Nexus** (Sanderson and Earnshaw, 1991; Combes and Rodford, 2004), which is developed and regularly updated by LHASA Ltd (http://www.lhasalimited.org/products/derek-nexus.htm). See for a general outline the skin irritation section on (Q)SARs. The Derek Nexus rule-base has five alerts that are specific to eye irritation, plus one for eye lachrymation. If Derek Nexus does not make a prediction of irritation or corrosivity, it cannot be concluded that there is no effect – it could mean that none of known alerts was found to be present in the substance of interest or it was outside the applicability domain of that specific alert. The Derek Nexus model is transparent in its algorithm, when the model is fired showing the structural alert and its limitations. The alert is underlined with literature references and sometimes with example substances, which is not sufficient to consider them internally validated. The example substances underline the mechanistic reasoning. The Derek Nexus model can be used for positive identification of skin irritation. The confidence levels have to be translated to consider them for classification. The Derek Nexus model cannot be used to predict non-serious eye damage/eye irritation as the model only contains alerts that detect the presence of serious eye damage/eye irritation.

The fragment-based **MultiCASE** approach (http://www.multicase.com/case-ultra-models#skin_eye_tox_bundle) has been used to model eye irritation (Klopman *et al.*, 1993; Enslein *et al.*, 1988; Rosenkranz *et al.*, 1998; Klopman, 1998). The publications on these models do not define the algorithms. In principle, the MultiCASE model can be used for positive and negative indication for eye irritation. The structural alert is provided as well as the internal validation. The MultiCASE model also indicates whether it is in the applicability domain of the model. The MultiCASE predictions of weak/mild/moderate and severe irritation need to be translated to consider them for classification. The prediction should be underlined with mechanistic reasoning using other models or expert judgement.

ACD/Labs Percepta Predictors (http://www.acdlabs.com/products/percepta/predictors.php), developed by ACD/Labs, includes a module for skin and eye irritation. It estimates the potential of a compound to cause eye or skin irritation in a standard rabbit Draize test. The predictions are reported as qualitative irritation categories (not irritating, slightly irritating, irritating, highly irritating, and corrosive). Probabilistic models are supplemented by an expert system that identifies Structural Alerts relevant in the irritational properties of compounds. Overall, 21 structural alerts were formulated for rabbit eye irritation, and 17 alerts for the rabbit skin irritation case.

PaDEL-DDPredictor includes several models for skin and eye irritation and corrosion (http://padel.nus.edu.sg/software/padelddpredictor/). The models have been built from a training set of 1707 compounds using one and two dimensional descriptors. The final predictions rely on consensus models based on majority voting from base models predictions. The applicability domain is defined by the range of descriptors for compounds in the training set.

BFR DECISION SUPPORT SYSTEM

The decision support system (DSS) developed by the German Federal Institute for Risk Assessment (BfR) uses physico-chemical exclusion rules to predict the absence of serious eye damage/eye irritation potential in combination with structural inclusion rules (SARs) to predict the presence of such potential (Gerner *et al.*, 2005). These rules are used analogously to those described in the skin corrosion and irritation section above. The physico-chemical and structural rule-bases are designed to predict the former EU risk phrases for eye irritation (R36) and severe eye irritation/corrosion (R41). Independent assessments by the ECB support the performance of the physico-chemical rule-base for predicting the absence of eye effects

(Tsakovska *et al.*, 2005), as well as the performance of the structural rulebase for predicting the occurrence of eye effects (Tsakovska *et al*., 2007).

OECD QSAR TOOLBOX

The freely downloadable OECD QSAR Toolbox software (http://www.qsartoolbox.org/) covers the serious eye damage/eye irritation endpoint with one experimental database and two profilers.

In more detail, the database of experimental data (called "Eye irritation ECETOC" in the software) refers to the endpoint Modified Maximum Average Score (MMAS) and collects experimental results on rabbit eye irritation described in,ECETOC Technical Report No.48 on Eye Irritation Reference Chemicals Data Bank (ECETOC, 1992).

The OECD QSAR Toolbox allows for the identification of analogues based on mechanistic and endpoint specific profilers, and for the prediction of skin irritation/corrosion through the use of read across, trend analysis and QSAR models. Information about inclusion and exclusion rules, details on the performance of the exclusion rules, and applicable chemical class-specific rules for the results of the Eye irritation/corrosion profiler can be found by searching the context menu in the the OECD QSAR Toolbox software.

REFERENCES

Abraham MH (1993) Scales of solute hydrogen-bonding: their construction and application to physico-chemical and biochemical processes. Chem Soc Rev 22:73-83.

Abraham MH, Kumarsingh R, Cometto-Muñiz JE and Cain WS (1998) A Quantitative Structure-Activity Relationship (QSAR) for a Draize Eye Irritation Database. Toxicol *in Vitro* 12:201-7.

Abraham MH, Hassanisadi M, Jalali-Heravi M, Ghafourian T, Cain WS and Cometto-Muñiz JE (2003) Draize rabbit eye test compatibility with eye irritation thresholds in humans: a quantitative structure-activity relationship analysis. Toxicol Sci 76:384-91.

Barratt MD (1995) The role of structure-activity relationships and expert systems in alternative strategies for the determination of skin sensitisation, skin corrosivity and eye irritation. Altern Lab Anim 23:111-22.

Combes RD and Rodford R (2004) The Use of Expert Systems for Toxicity Prediction - Illustrated With Reference to the DEREK Program. *In:* Predicting Toxicity and Fate (Cronin, MTD and Livingstone DJ Eds.) CRC Press, Boca Raton, FL, USA, pp.193-204.

Cronin MTD, Basketter DA, York M (1994) A quantitative structure- activity relationship (QSAR) investigation of a Draize eye irritation database. Toxicol *in Vitro* 8:21-8.

Cronin MTD (1996) The use of cluster significance analysis to identify asymmetric QSAR datasets in toxicology. An example with eye irritation data. SAR QSAR Environ Res 5:167-75.

ECETOC (1992) Technical Report No.48, Eye Irritation Reference Chemicals Data Bank. ISSN-0773-8072-48(2). 2nd edition of June 1998 available at: http://www.ecetoc.org/publications

Gallegos Saliner A, Patlewicz G and Worth AP (2006). Review of literature-based models for skin and eye irritation and corrosion. JRC report EUR 22320 EN. European Chemicals Bureau, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/doc/QSAR_Review_Irritation.pdf

Gallegos Saliner A, Patlewicz G and Worth AP (2008) A Review of (Q)SAR Models for Skin and Eye Irritation and Corrosion. QSAR Comb Sci 27:49-59.

Gerner I, Liebsch M and Spielmann H (2005) Assessment of the Eye Irritating Properties of Chemicals by Applying Alternatives to the Draize Rabbit Eye Test: The Use of QSARs and *In Vitro* Tests for the Classification of Eye Irritation. Altern Lab Anim 33:215-37.

Klopman G (1998) The MultiCASE Program II. Baseline Activity Identification Algorithm (BAIA). J Chem Inf Comput Sci 38:78-81.

Kulkarni A, Hopfinger AJ, Osborne R, Bruner LH and Thompson ED (2001) Prediction of Eye Irritation from Organic Chemicals Using Membrane-Interaction QSAR Analysis. Toxicol Sci 59:335-45.

Patlewicz GY, Rodford RA, Ellis G and Barratt MD (2000) A QSAR Model for the Eye Irritation of Cationic Surfactants. Toxicol *in Vitro* 14:79-84.

Rosenkranz HS, Zhang YP and Klopman G (1998) The Development and Characterisation of a Structure-activity Relationship Model of the Draize Eye Irritation Test. Altern Lab Anim 26:779-809.

Sanderson DM and Earnshaw CG (1991) Computer Prediction of Possible Toxic Action from Chemical Structure; The DEREK System. Hum Exp Toxicol 10:261-73.

Solimeo R, Zhang J, Kim M, Sedykh A and Zhu H (2012) Predicting Chemical Ocular Toxicity Using a Combinatorial QSAR Approach. Chem Res Toxicol 25: 2763-69.

Sugai S, Murata K, Kitagaki T and Tomita I (1990) Studies on eye irritation caused by chemicals in rabbits—1. A quantitative structure-activity relationships approach to primary eye irritation of chemicals in rabbits. J Toxicol Sci 15:245-62.

Sugai S, Murata K, Kitagaki T and Tomita I (1991) Studies on eye irritation caused by chemicals in rabbits—II. Structure-activity relationships and *in vitro* approach to primary eye irritation of salicylates in rabbits. J Toxicol Sci 16:111-30.

Tsakovska I, Netzeva T and Worth AP (2005) Evaluation of (Q)SARs for the prediction of Eye Irritation/Corrosion Potential - physico-chemical exclusion rules. JRC Report EUR 21897 EN, 42pp. European Chemicals Bureau, Ispra, Italy. Available at: https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/doc/Evaluation_of_Eye_Irritation_QSARs.pdf

Tsakovska I, Gallegos Saliner A, Netzeva T, Pavan M, Worth AP (2007) Evaluation of SARs for the prediction of eye irritation/corrosion potential - structural inclusion rules in the BfR decision support system. SAR QSAR Environ Res 18:221-35.

Worth AP and Cronin MTD (2000) Embedded cluster modelling: a novel quantitative structure-activity relationhip for generating elliptic models of biological activity. *In:* Progress in the Reduction, Refinement and Replacement of Animal Experimentation (Balls M, van Zeller A-M, Halder ME, Eds.) Elsevier Science, Amsterdam, The Netherlands, pp.479-491.

Worth AP and Cronin MTD (2003) The use of discriminant analysis, logistic regression and classification tree analysis in the development of classification models for human health effects. J Mol Struct (Theochem) 622:97-111.

Worth A, Barroso J, Bremer S, Burton J, Casati S, Coecke S, Corvi R, Desprez B, Dumont C, Gouliarmou V, Goumenou M, Gräpel R, Griesinger C, Halder M, Janusch Roi A, Kienzler A, Madia F, Munn S, Nepelska M, Paini A, Price A, Prieto P, Rolaki A, Schäffer M, J. Triebe, Whelan M, Wittwehr C and Zuang V (2014) Alternative methods for regulatory toxicology – a state-of-the-art review. JRC Science and Policy Reports, Report EUR 26797 EN. European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Systems Toxicology Unit, Institute for Health and Consumer Protection, European Commission Joint Research Centre, Ispra, Italy. Available at: https://ec.europa.eu/jrc/sites/default/files/echa_jrc_sla_report_public_05-09-14_withcover_ipo.pdf

R.7.3 Skin and respiratory sensitisation

R.7.3.1 Introduction

A number of diseases are recognised as being, or presumed to be, allergic in nature. These include asthma, rhinitis, conjunctivitis, allergic contact dermatitis, urticaria and food allergies (the latter is not discussed in this document). In this Section, the endpoints discussed are those traditionally associated with occupational and consumer exposure to chemicals (proteins are not discussed in this document). Photosensitisation is potentially important but its mechanism of action is poorly understood, and it is not discussed in this document.

R.7.3.1.1 Definition of skin and respiratory sensitisation

A sensitizer is an agent that is able to cause an allergic response in susceptible individuals. The consequence of this is that following subsequent exposure *via* the skin the characteristic adverse health effects of allergic contact dermatitis or atopic dermatitis may be provoked. After inhalation exposure, adverse health effects include asthma (and related respiratory symptoms such as rhinitis) or extrinsic allergic alveolitis.

Respiratory hypersensitivity is a term that is used to describe asthma and other related respiratory conditions, irrespective of the mechanism (immunological or non-immunological) by which they are caused. In contrast, dermal allergy is based on an immunological mechanism.

It is perhaps helpful to attempt to define the term chemical respiratory hypersensitivity. One approach taken by the UK Health and Safety Executive was to describe the induction phase as the process of rendering the airways unusually sensitive (hypersensitive) such that following subsequent inhalation exposure an asthmatic reaction might be elicited associated with classical symptoms of airway narrowing, chest-tightening and bronchial restriction (HSE, 1997). Other approaches to definition of relevant terms are available elsewhere. For instance, various definitions are provided for specific sensitising agents in the workplace – all of which imply a mechanism whereby hypersensitivity of the respiratory tract is induced as the result of workplace exposure – and that this may result later in the development of occupational asthma (Bernstein et al., 1993). Lists of chemicals cited here, by the HSE, and elsewhere, as causes of respiratory sensitisation and occupational asthma are very similar, and in some instances identical (Chan-Yeung et al., 1993). Among the chemicals populating these lists are: diisocyanates, acid anhydrides, certain platinum salts, some reactive dyes, cyanuric chloride, and plicatic acid (from Western Red Cedar).

When directly considering human data in this document, the clinical diagnostic terms asthma, rhinitis and extrinsic allergic alveolitis have been retained.

These definitions are reflected in the criteria for the classification of skin and respiratory sensitizers, which provide a useful tool against which the hazardous properties of a substance can be judged. These criteria are given in the 22nd Adaptation to Technical Progress to Directive 67/548/EEC [Directive 96/54/EC, Official Journal L248; pp 227-229]; Annex VI has been recast in the 28th Adaptation to Technical Progress (ATP) (Directive 2001/59, Official Journal L225; pp 1- 333).

R.7.3.1.2 Objective of the guidance on skin and respiratory sensitisation

The general objectives are to determine:

- whether there are (Q)SAR data, existing *in vitro* or *in vivo* data, or human evidence indicating that the agent has skin or respiratory sensitisation potential

- whether the agent has skin sensitisation potential based on new tests according to the strategy as presented in this document.

Therefore, in the sections on skin sensitisation and respiratory sensitisation firstly an overview of types of data is given that may provide information on sensitisation, followed by guidance on the process of judging the available data in terms of adequacy, completeness and remaining uncertainty. In Section R.7.3.7 guidance is given on application of the data to reach a conclusion on suitability for classification and labelling and possibly potency. Finally in Section R.7.3.8 an integrated testing strategy (ITS) for skin sensitisation and an integrated evaluation strategy (IES) for respiratory sensitisation is presented.

R.7.3.1.3 Mechanisms of immunologically-mediated hypersensitivity

Among the key steps required for a chemical to induce sensitisation *via* skin contact are gaining access to the viable epidermis, protein binding, metabolic activation (if required), internalization and processing by Langerhans cells (LC), transport of antigen by LC to draining lymph nodes, and presentation to and recognition by T lymphocytes. For chemicals that sensitise *via* the respiratory tract, the relevant mechanisms are believed to be essentially similar, although gaining access to the respiratory epithelium may be somewhat easier than at skin surfaces due to the lack of a stratum corneum. Moreover, because the lining of the respiratory tract, the professional antigen presenting cells, and regulatory mechanisms in the respiratory tract differ from those in the skin, they all may have an impact on the type of immune response evoked. Although the site of induction of an adaptive immune response to a chemical allergen may be influenced by local conditions and local immunoregulatory mechanisms, the fact remains that the inherent properties of the chemical itself play a major role in determining whether an immune responses is induced and the qualitative characteristics of that response.

Although it is sometimes assumed that immune responses induced following encounter with antigen in or on the skin are often of selective Th_1-type, this is not necessarily the case. It is clear that cutaneous immune responses can be of either Th_1- or Th_2-type according to the nature of the antigen.

In the respiratory tract, chemical respiratory allergens appear to preferentially elicit Th_2-immune responses (Maestrelli et al., 1997); observations that are consistent with experimental experience in mice (Dearman et al., 2002; Herrick et al., 2003; Farraj et al., 2004), and possibly also rats (Arts et al., 1998). Th2 type immune responses are characterised by the production of cytokines such as IL4 and IL5 and by the production of IgE antibodies. However, the mechanisms through which chemicals are able to induce sensitisation of the respiratory tract are not fully understood and there remains controversy about the roles played by IgE antibody-mediated mechanisms, and whether IgE represents a mandatory universal requirement for the induction by chemicals of allergic sensitisation of the respiratory tract. The area is complicated because although for all chemical respiratory allergens there are patients who display serum IgE antibodies of the appropriate specificity, in other instances (and particularly with respect to the diisocyanates) there are symptomatic subjects in whom it is not possible to detect IgE antibody. There are two, non-mutually exclusive, possibilities. The first is that IgE does play a central role but that for one or more of various reasons it is not being detected accurately in the serum of patients with occupational asthma. The second is that allergic sensitisation of the respiratory tract by chemicals can be effected through IgE antibody-independent immunological mechanisms (Kimber et al., 2002 and 2005). These may also include Th_1-type immune responses. In this context it has been reported, for instance, that inhalation challenge of sensitised rodents with contact allergens may elicit respiratory allergic reactions (Garssen et al., 1991; Garcia et al., 1992; Buckley et al., 1994; Zwart et al., 1994; Satoh et al., 1995; Arts et al., 1998). This comes as no surprise because it is clear that contact sensitisation is systemic in nature and that there is no reason to suppose that encounter of sensitised animals with the relevant contact allergen at respiratory epithelial surfaces will not cause an adverse immunologic reaction. However, it is important to note that in reality only a very few precedents for the elicitation of pulmonary reactions by skin sensitising chemicals in humans have been observed, and in practice it may not represent a significant health issue.

In addition, there is a growing body of evidence that effective sensitisation of the respiratory tract by chemicals defined as respiratory allergens (such as for instance the acid anhydrides, diisocyanates and others) can and does occur in response to dermal contact (reviewed by Kimber et al., 2002). There are also experimental animal data and human evidence for sensitisation by inhalation and skin effects following dermal challenge (Kimber et al., 2002, Baur et al., 1984, Ebino et al., 2001, Stadler et al., 1984). Therefore, it is not necessarily the case that chemicals that cause allergic dermal reactions require sensitisation *via* the skin, or that chemicals that cause allergic airway reactions require sensitisation *via* the respiratory tract.

R.7.3.2 Information requirements for skin and respiratory sensitisation

The information requirements for sensitisation are described in REACH Annexes VI to XI, where the information that shall be submitted for registration purposes is specified.

Column 1 of Annex VII clearly informs on the standard information requirement for skin sensitisation data for substances produced or imported in quantities of ≥1 t/y.

The assessment of skin sensitisation shall comprise the following consecutive steps:

1. *an assessment of the available human, animal and alternative data,*

2. *In vivo testing*

Column 2 of Annex VII lists specific rules according to which the required standard information may be omitted, replaced by other information, or adapted in another way. If the conditions are met under which column 2 of this Annex allows adaptations, the fact and the reasons for each adaptation should be clearly indicated in the registration. For skin sensitisation column 2 reads:

Step 2 does not need to be conducted if:

- *the available information indicates that the substance should be classified for skin sensitisation or corrosivity; or*

- *the substance is a strong acid (pH<2.0) or base (pH>11.5); or*

- *the substance is flammable in air at room temperature.*

The Murine Local Lymph Node Assay (LLNA) is the first-choice method for in vivo testing. Only in exceptional circumstances should another test be used. Justification for the use of another test shall be provided. This means that in certain cases other in vivo methods may be conducted. In such cases convincing scientific justification for the use of another test shall be provided.

No information requirements are present for respiratory sensitisation. Respiratory sensitizers are indicated for harmonised classification and labelling in REACH Article 115, and respiratory sensitisation is mentioned in Annex I and XV which deal with respectively chemical safety report and preparation of these dossiers.

In addition to these specific rules, the required standard information set may be adapted according to the general rules contained in Annex XI. In this case as well, the fact and the reasons for each adaptation should be clearly indicated in the registration.

General requirements for generation of information on intrinsic properties of substances are given in REACH Article 13 which states that this information may be generated by means other than tests, provided the conditions specified in Annex XI are met.

R.7.3.3 Information for skin sensitisation and its sources

R.7.3.3.1 Non-human data for skin sensitisation

Non-testing data for skin sensitisation

Non-testing methods for skin sensitisation cover a breadth of different approaches namely read-across/chemical categories, chemistry considerations and (Q)SARs. Read-across/chemical categories are described in Sections R.6.1 and R.6.2.

A compendium of available (Q)SARs is not in existence at the present time, work is being carried out by ECB to develop an inventory of evaluated (Q)SARs which will populate the (Q)SAR Application Toolbox, a larger project currently led by the OECD. The JRC QSAR Model Database is being designed to help a user determine the validity and applicability of a model for a specific chemical and purpose. This is relevant to the assessment of adequacy. The OECD principles (described on Website http://www.oecd.org/document/23) will help to characterise the validity of a given model. Preliminary practical guidance on their interpretation has been developed (Worth *et al.*, 2005). Evaluated (Q)SARs will be documented in (Q)SAR Reporting Formats (see Section R.6.1.9 in Chapter R.6 of the *Guidance on IR&CSA*). More generic information on evaluating QSARs, their predictions and reporting formats is provided in Section R.6.1.6 in Chapter R.6 of the *Guidance on IR&CSA*.

Exploring the reaction chemistry of compounds forms the basis of most read-across justifications and many of the available skin sensitisation (Q)SARs. The skin sensitisation potential of a chemical is related to its ability to react with skin proteins to form covalently linked conjugates and recognition of these by the immune system. In the vast majority of cases, this is dependent on electrophilic reactivity of the skin sensitizer or a derivative produced (usually by oxidation) *in vivo* or abiotically (Barratt *et al.*, 1997). There are various types of electrophile-nucleophile reactions in skin sensitisation, perhaps the most frequently encountered are: Michael-type reactions; S_N2 reactions; S_NAr reactions; acylation reactions and Schiff-base formation. These chemical reaction mechanisms can serve as a means of describing the domain of applicability (the scope) of a (Q)SAR or form the basis for grouping chemicals into chemical categories. Recent work in this area has been described in (Aptula *et al.*, 2005, Aptula and Roberts 2006, Roberts *et al.*, 2007).

There are relatively few (Q)SARs for skin sensitisation reported in the peer reviewed literature. Available models include local and global (Q)SARs as well as expert systems.

Local (Q)SAR models

The majority of local models available have been developed for direct-acting electrophiles using the relative alkylation index (RAI) approach. This is a mathematical model derived by Roberts and Williams (1982). It is based on the concept that the degree of sensitisation produced at induction, and the magnitude of the sensitisation response at challenge, depends on the degree of covalent binding (haptenation; alkylation) to carrier protein occurring at induction and challenge. The RAI is an index of the relative degree of carrier protein haptenation and was derived from differential equations modelling competition between the carrier haptenation reaction in a hydrophobic environment and removal of the sensitizer through partitioning into polar lymphatic fluid. In its most general form the RAI is expressed as:

$$RAI = \log D + a \log k + b \log P \qquad (1)$$

Thus the degree of haptenation increases with increasing dose D of sensitizer, with increasing reactivity (as quantified by the rate constant or relative rate constant *k* for the reaction of the sensitizer with a model nucleophile) and with increasing hydrophobicity (as quantified by log P, P being the octanol/water partition coefficient). This RAI model has been used to evaluate a wide range of different datasets of skin sensitising chemicals. Examples include sulfonate esters (Roberst and Basketter 2000), sulfones (Roberts and Williams 1982), primary alkyl bromides (Basketter et al., 1992), acrylates (Roberts 1987), aldehydes and diketones (Patlewicz et al., 2001, Patlewicz et al., 2002, Patlewicz et al., 2004, Roberts et al., 1999, Roberts and Patlewicz 2002, Patlewicz et al., 2003).

This approach has been shown to be mechanistically robust but the breadth of available models so far is still somewhat limited. These types of models assume a reasonable appreciation of chemistry.

The covalent hypothesis has served and continues to be the most promising way of developing mechanistically based robust QSARs. These are local in that their scope is characterised by a mechanistic reactivity domain as outlined in Aptula et al., 2005, Aptula and Roberts 2006, Roberts *et al.*, 2007. An example of this type of mechanistic model has been recently published (Roberts *et al.*, 2006). In the RAI model, log*k*, has been typically modelled by experimental rate constants, substituents' constants or molecular orbital parameters. More effort is needed to encode reactivity into descriptors, this could be achieved through the systematic generation of *in vitro* reactivity data as outlined in (Aptula and Roberts 2006, Aptula et al., 2006b, Schultz et al., 2006, Gerberick et al., 2004) and in the next section.

Global statistical models

Global Statistical models usually involve the development of empirical QSARs by application of statistical methods to sets of biological data and structural descriptors.

These are perceived to have the advantage of being able to make predictions for a wider range of chemicals. In some cases, the scope/domain of these models are well described, in most other cases a degree of judgement is required in determining whether the training set of the model is relevant for the chemical of interest. Criticism often levied at these types of models is that they lack mechanistic interpretability. The descriptors might appear to lack physical meaning or are difficult to interpret from a chemistry perspective. The sorts of descriptors used may encode chemical reactivity/electrophilicity e.g. LUMO (the energy of the lowest molecular orbital) and partitioning effects e.g. Log P, but more commonplace is that a large number of descriptors are calculated that encode structural, topological and/or geometrical information. A number have been reported in the recent literature, examples include those developed using LLNA data (Devillers 2000, Estrada et al., 2003, Fedorowicz et al., 2005, Fedorowicz et al., 2005, Li et al., 2005, Miller et al., 2005, Ren et al., 2006, Li et al., 2007).

Expert systems

There are several commercial (Q)SAR models for skin sensitisation available. Examples include TOPKAT, CASE, Derek for Windows and TIMES.

Statistical Models

TOPKAT (current version 6.2) marketed by Accelrys Inc (San Diego, USA) comprises two suites of models; one for aromatics (excluding chemicals with 1 benzene ring) and the other for aliphatics and chemicals with 1 benzene ring. The first set of models discriminate between non-sensitizers and sensitizers, a probability is calculated for the submitted chemical structure. If the probability is greater than or equal to 0.7, the chemical is predicted to be a sensitizer, a non-sensitizer would have a probability of less or equal to 0.30. The second set of models resolve the potency: weak/moderate vs. strong where a probability of 0.7 or more indicates a strong sensitizer and a probability below 0.30 indicates a weak or moderate sensitizer. Probability values between 0.30 and 0.70 are referred to as indeterminate. An optimum prediction space algorithm ensures that predictions are only made for chemicals within the model applicability domain (Enslein et al., 1997, http://www.accelrys.com/products/topkat/).

CASE methodology and all its variants were developed by Klopman and Rosenkranz. There are a multitude of models for a variety of endpoints and hardware platforms. The CASE approach uses a probability assessment to determine whether a structural fragment is associated with toxicity (Cronin et al., 2003). The MCASE models that have been developed for skin sensitisation are described further in primary articles (Gealy et al., 1996, Graham et al., 1996, Johnson et al., 1997). There are two sensitisation modules available for purchase from MultiCase Inc (Ohio, USA) (http://www.multicase.com/products/prod0911.htm). In addition the (Q)SAR estimates for one MCASE skin sensitisation model are included in the Danish Environmental Protection Agency (EPA) (Q)SAR database which is currently hosted on the European Chemicals Bureau (ECB) website http://ecb.jrc.it/QSAR/ .

Knowledge based systems

Derek for Windows (DfW) is a knowledge-based expert system created with knowledge of structure-toxicity relationships and an emphasis on the need to understand mechanisms of action and metabolism. It is marketed and developed by LHASA Ltd (Leeds, UK) a not-for-profit company and educational charity (http://www.lhasalimited.org/index.php).

Within DfW (version 9), there are 361 alerts covering a wide range of toxicological endpoints. An alert consists of a toxicophore, a substructure known or thought to be responsible for the toxicity alongside associated literature references, comments and examples. The skin sensitisation knowledge base in DfW was initially developed in collaboration with Unilever in 1993 using its historical database of guinea pig maximisation test (GPMT) data for 294 chemicals and contained approximately forty alerts (Barratt et al., 1994). Since that time, the knowledge base has undergone extensive improvements as more data have become available (Payne and Walsh 1994). The current version (version 9) contains seventy alerts for skin sensitisation and the closely-related endpoint of photoallergenicity (Barratt et al., 2000, Langton et al., 2006).

Hybrids

TIssue MEtabolism Simulator (TIMES) software has been developed to integrate a skin metabolism simulator with 3D-QSARs for evaluating reactivity of chemicals in order to predict their skin sensitisation potency (Dimitrov et al., 2005, Dimitrov et al., 2005). The simulator contains 236 hierarchically ordered spontaneous and enzyme controlled reactions. Covalent interactions of chemicals/metabolites with skin proteins are described by 47 alerting groups. 3D-QSARs (COREPA) are applied for some of these alerting groups.

Clearly there are a breadth of different (Q)SARs and expert systems available for the estimation of skin sensitisation hazard. The approaches are quite varied and each has been developed on different sets of *in vivo* data (principally GPMT and LLNA). Whilst efforts have been made to characterise a number of the literature based models in terms of the OECD principles for QSAR validation (see Roberts et al., 2007 as an example), further work is still required for some of the commercial systems (ECETOC 2003). In addition, in many cases these models have been demonstrated to be reasonable for predicting skin sensitizers correctly but are limited in predicting non-sensitizers correctly (Roberts et al., 2007, ECETOC 2003). For this reason, careful interpretation of model predictions needs to be considered in light of other information e.g. analogue read-across (other similar chemicals with respect to their mechanistic domain).

Further work should explore encoding more knowledge/rules for non-reactive chemicals as well as those chemicals likely to undergo chemical or metabolic transformation.

Consideration of which model(s) to apply will be dependent on the specific chemical of interest, the underlying training set data and the applicability domain. These issues are described more fully in Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*. An example is illustrated here; if the chemical falls into a chemistry reactivity domain that is well characterised, then a local (Q)SAR model developed for this domain (such as those previously described) will give rise to the most robust prediction of skin sensitisation. Where the mechanism is not understood or not known *a priori* one or more of the expert systems such as TOPKAT, Derek for Windows or the others already described will be best placed to provide an estimate. These systems whilst not wholly transparent do provide a reasonable amount of supporting information to enable the robustness of a prediction to be evaluated. This is discussed in more detail in Section R.7.3.4.1.

Testing data for skin sensitisation

In vitro data

At present, no officially adopted EU-OECD *in vitro* tests for skin sensitisation exist. However, several systems are in the course of development (Eskes et al., 2005), based on an improved understanding of the biochemical and immunological mechanisms underlying the process (Worth et al., 2002). Currently, *in vitro* assays to detect the sensitising properties of a chemical are under development for the following areas:

- **Epidermal bioavailability**: skin penetration is a prerequisite for skin sensitisation. Information about the skin penetration properties can help to evaluate the potential of a chemical to be identified as a skin sensitizer (ECVAM, 2007).

- **Chemical reactivity**: since the majority of chemical allergens is electrophilic and reacts with nucleophilic amino acids, peptide reactivity assays can give an indication of skin sensitisation potency or potential to form a complete antigen (Gerberick et al., 2004, Aptula et al., 2006b).

- **Cell-based assays**: the knowledge that changes occur in epidermal Langerhans cells as a result of exposure to chemical allergens (e.g. the expression of surface markers and/or cytokines release) and that Langerhans cells can be replaced by blood derived dendritic-like cells or cell lines have been applied to design in vitro alternative tests (Kimber et al., 2001, Tuschl et al., 2000, Casati et al., 2005, Ryan et al., 2005, Sakaguchi et al., 2006, Aeby et al., 2004, Azam et al., 2006, Python et al., 2007). These systems have been shown to selectively express various mediators and/or markers of activation following exposure to chemical sensitizers and attempts to develop robust assays have started. Beside Langerhans cells, keratinocytes play a prominent role in the sensitisation process (Corsini et al., 1998, van Och et al., 2005, Vandebriel et al., 2005). In addition to chemical processing, LC activation requires the binding of cytokines produced by keratinocytes as a result of initial chemical exposure. Moreover the assessment of keratinocytes cytokine expression as a function of the ability of chemicals to induce cutaneous sensitisation is also the object of several investigations (Aiba et al., 2000, Herouet et al., 2000). Keratinocytes have been tested both in primary cultures, in co-culture with dendritic cells and as reconstituted epidermis (Casati et al., 2005, Kubilus et al., 1986, Coquette et al., 2003). The use of reconstituted skin models for the assessment of contact allergens is under investigation.

Owing to the complexity of the mechanisms of skin sensitisation, a single test will probably not be able to replace the currently required animal procedures. Efforts are still needed to identify the most relevant endpoints in the optimisation of existing tests. However, a combination of several *in vitro* tests, covering the relevant mechanistic steps of skin sensitisation, into a test battery could possibly lead to replacement of *in vivo* tests (Eskes et al., 2005). How the outputs from these tests could be combined is not as yet determined, although a general strategy has been presented (Jowsey et al., 2006). Until that date, *in vitro* tests may be used as supportive evidence in combination with other types of data for the identification of allergens (see Section R.7.3.8.3 for an ITS based on a WoE approach).

Animal data

Guideline-compliant tests

For new *in vivo* testing of skin sensitisation potential, the murine local lymph node assay (LLNA) is the REACH Annex VII-endorsed method. This assay has been validated internationally and has been shown to have clear animal welfare benefits and scientific advantages compared with the guinea pig tests described below. The LLNA is designed to detect the potential of chemicals to induce sensitisation as a function of lymphocyte proliferative responses induced in regional lymph nodes. This method is described in OECD TG 429/EU B.42.

Two further animal test methods for skin sensitisation are described in OECD TG 406/EU B.6: the guinea pig maximisation test (GPMT) and the Buehler test. The GPMT is an adjuvant-type test in which the acquisition of sensitisation is potentiated by the use of Freund's Complete Adjuvant (FCA) and in which both intradermal and topical exposure are used during the induction phase. The Buehler test is a non-adjuvant method involving for the induction phase topical application only.

Both the GPMT and the Buehler test are able to detect chemicals with moderate to strong sensitisation potential, as well as those with relatively weak sensitisation potential. In such methods activity is measured as a function of challenge-induced dermal hypersensitivity reactions elicited in test animals compared with controls. Since the LLNA is the preferred method for new *in vivo* testing, the use of the standard guinea pig tests to obtain new data on skin sensitisation potential will be acceptable only in exceptional circumstances and will require scientific justification. However, existing data of good quality deriving from such tests will be acceptable and will, if providing clear results, preclude the need for further *in vivo* testing.

ECETOC Monograph 29 (2000) contains a useful discussion of these tests.

Non-guideline compliant tests and refinements to the standard assays

Existing data may be available from tests that do not have an OECD guideline, for example:

i. other guinea pig skin sensitisation test methods (such as the Draize test, optimisation test, split adjuvant test, open epicutaneous test);

ii. additional tests (such as the mouse ear swelling test);

Information may also be available from other endpoints, for example, repeated dose dermal studies that show effects indicative of an allergic response, such as persistent erythema and/or oedema.

For new testing, refinements to the existing guideline methods may also be possible. In such cases, care should be taken to ensure that any modifications or deviations from standard methodologies are scientifically justified. For example, it might be feasible to conduct a reduced version of the LLNA (rLLNA) in which assessments are made on the basis of results from a vehicle control and a single (highest) concentration of the test substance (Eskes et al., 2005). In such cases, it is recommended that expert advice be sought before commencing the tests.

R.7.3.3.2 Human data on skin sensitisation

Human data on cutaneous (allergic contact dermatitis and urticarial) reactions may come from a variety of sources:

- consumer experience and comments, preferably followed up by professionals (e.g. diagnostic patch tests)
- diagnostic clinical studies (e.g. patch tests, repeated open application tests)
- records of workers' experience, accidents, and exposure studies including medical surveillance
- case reports in the general scientific and medical literature
- consumer tests (monitoring by questionnaire and/or medical surveillance)
- epidemiological studies
- human experimental studies such as the human repeat insult patch test (Stotts, 1980) and the human maximisation test (Kligman, 1966), although it should be noted that *new* experimental testing for hazard identification in humans, including HRIPT and HMT, is not acceptable for ethical reasons.

R.7.3.4 Evaluation of available information on skin sensitisation

For both steps of the effects assessment, i.e. hazard identification and dose (concentration)-response (effect) assessment, it is very important to evaluate the data with regard to their adequacy and completeness. The evaluation of adequacy shall address the reliability and relevance of the data. The completeness of the data refers to the conclusion on the comparison between the available adequate information and the information that is required under the REACH proposal for the applicable tonnage level of the substance. Such a conclusion relies on WoE approaches, mentioned in REACH Annex XI Section 1.2, which categorise available information based on the methods used: *guideline tests, non-guideline tests*, and other types of information which may justify adaptation of the standard testing regime. Such a WoE approach also includes an evaluation of the available data as a whole, i.e. both over or across endpoints: i.e. for a sensitive evaluation of sensitisation effects, it is necessary to efficiently integrate the information gathered for sensitisation with that obtained from the study of skin and eye irritation (and acute dermal toxicity).

This approach provides a basis to decide whether further information is needed on endpoints for which specific data appear inadequate or not available, or whether the requirements are fulfilled.

For this specific endpoint some additional remarks are made on the adequacy of the various types of data that may be available.

R.7.3.4.1 Non-human data on skin sensitisation

Non-testing data on skin sensitisation

The evaluation and assessment of a chemical using (Q)SARs is dependent on both the chemical of interest and the (Q)SAR model(s) used to make a prediction. Here we attempt to provide some specific advice for skin sensitisation. More general advice on (Q)SARs including evaluation of OECD principles is described in Section R.6.1.3 in Chapter R.6 of the *Guidance on IR&CSA*).

One of the first steps to consider is what information already exists on chemicals *similar* to the one of interest. Chemical similarity is a widely used concept in toxicology, and is based on the hypothesis that similar compounds have similar biological activities. This forms the underlying basis for developing (Q)SARs. In the case of skin sensitisation, the most robust means of comparing two or more chemicals is through an evaluation of their likely chemical reactivity. Recent work in this area has been investigating means of encoding reactivity for the different mechanistic domains in form of rules (Aptula and Roberts 2006, Aptula et al., 2006). (Note: This approach might involve the systematic generation of *in vitro* reactivity data for these different mechanistic domains. (see Aptula et al., 2006 as an example) .If the chemical reactivity is not known, or can not be determined through experimentation then a pragmatic means of identifying similar chemicals can be done through a substructural/analogue search.

There are a number of available computational tools and databases that facilitate the search and retrieval of similar analogues. Some like Leadscope (http://www.leadscope.com) are commercial, others like Chemfinder (www.chemfinder.com), ChemID (http://chem.sis.nlm.nih.gov/chemidplus/) or DssTox (http://www.epa.gov/nheerl/dsstox/) are freely available to use on the internet.

Some of the available search engines are linked to databases (through hyperlinks and indexes) whereas other facilities such as DssTox provide a repository of available QSAR datasets which can be downloaded for subsequent use in appropriate QSAR /database software tools.

Many of currently available tools containing public data have focussed on endpoints such as carcinogenicity, mutagenicity or acute toxicity. This means that an additional search is needed to identify skin sensitisation data. Much of the available skin sensitisation experimental data resides in peer reviewed publications. Cronin and Basketter (1994) published the results of over 270 *in vivo* skin sensitisation tests (mainly from the guinea pig maximisation test). All data were obtained in the same laboratory and represent one of the few occasions when large amounts of information from corporate databases was released into the open literature. A larger database of animal and human studies for 1034 compounds is described by Graham et al. (1996), the MCASE database. A comparatively large number of data have been published for the local lymph node assay, examples include publications by Ashby et al. (1995) and Gerberick et al (2005).

These publications are invaluable to identify analogues with associated skin sensitisation test data.

The second step involves an assessment of the similarity of the analogues identified. Considerations will include whether:

- the same endpoint is considered
- there are any additional functional groups or additional substituents that might influence the reactivity and sensitising behaviour (applicability domain considerations)
- the physico-chemical parameters similar (e.g. LogP, applicability domain considerations)

- there are impurities that influence the sensitisation profile
- the likely chemical mechanism is the same

These considerations may help identify an available local (Q)SAR for that chemical class/mechanistic group.

If an appropriate local model can not be identified then a third step of evaluating a chemical using one of the available global models/expert systems is merited.

Here a prediction needs to be evaluated in the context of the likely chemistry and the available *like* chemicals available within the training set. i.e. is the compound of interest within the scope of the model and are similar chemicals in the training set of the model well predicted. This type of information provides additional weight to whether the estimate derived is meaningful and relevant. For global models available in the literature, the training sets and the algorithm(s) are usually available to allow such comparisons to be made.

For expert systems such as Derek for Windows, TOPKAT etc, the training sets and to an extent the algorithms or descriptors used are often kept latent within the software. Some supporting information is provided on the robustness and relevance for a given prediction. For example, within DfW it is possible to see representative example chemicals and explanations of the mechanistic basis for the SAR developed. Within TOPKAT, it is possible to obtain an assessment of whether the chemicals falls within the applicability domain of the model (both with respect to the fragment and descriptor space), whether it is an example chemical in the database as well as perform a similarity assessment to identify analogues. Similar functionalities and features are present in many of the other commercial expert systems available.

Although the main factors driving skin sensitisation (and therefore the (Q)SARs) is the underlying premise of the electrophilicity of a chemical, other factors such as hydrophobicity encoded in the octanol/water partition coefficient (log P) may also be considered as playing a role in the modifying the sensitisation response observed. Within DfW, an assessment of the likely skin penetration ability is made using the algorithm by Potts and Guy. This relates the Kp value to log P and MW (Potts and Guy 1992). It is then possible to rationalise the output in terms of bands of penetration potential. Some have been described in (Howes et al., 1996).

Specific model and prediction information can be described in more detail in reporting formats ((Q)SAR Reporting Format). This summarises the pertinent information to consider for given model when evaluating an estimate as well as the estimate itself. More details are provided in Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*.

Other information such as results in other assays such as the Ames test (a common feature of genotoxic substances is that they can bind covalently to DNA and cause direct DNA damage) or aquatic toxicity tests may provide supporting information about the electrophilicity of the chemical of interest and hence its likely sensitisation ability. Some of this work is still at an early stage but correlations have been explored between mutagens and sensitizers (Wolfreys and Basketter 2004) and between aquatic toxicants and sensitizers (Aptula et al., 2006).

Testing data on skin sensitisation

In vitro data

Even though a number of *in vitro* methods are currently under development, none of these methods has yet undergone a formal validation process. According to Annex XI, *in vitro* data obtained with non-validated methods can only be used in a WoE approach. If such data are considered for the evaluation, expert judgement is needed to assess their reliability. In particular, attention should be paid to the level of optimisation of the method that should meet at least the ECVAM criteria for entering pre-validation (Curren et al., 1995), including evidence of the reproducibility of the method, its mechanistic relevance and predictive capacity (Balls et al., 1995, Hartung et al., 2004, Worth et al., 2001).

In vitro assays only cover a (specific) part of the process of sensitisation that occurs *in vivo*, therefore it is unlikely that a single method will be able to substitute for the animal test.

Animal data

Well reported studies using internationally acceptable protocols, particularly if conducted in accordance with the principles of GLP, can be used for hazard identification. Other studies (see Section R.7.3.3.1 and below), not fully equivalent to OECD test protocols, can, in some circumstances, provide useful information. Particular attention should be paid to the quality of these tests and the use of appropriate positive and negative controls. The specificity and sensitivity of all animal tests should be monitored through the inclusion of appropriate positive and negative controls. In this context, positive controls are the 6-monthly sensitivity checks with an appropriate positive control substance, and negative controls are the vehicle-treated control animals included as part of each test.

Guideline-compliant tests

For the conduct and interpretation of the LLNA the following points should be considered:

i. the vehicle in which the test material and controls have been applied;

ii. the concentrations of test material that have been used;

iii. any evidence for local or systemic toxicity, or skin inflammation resulting from application of the test material;

iv. whether the data are consistent with a biological dose response;

v. the submitting laboratory should be able to demonstrate its competency to conduct the LLNA.

OECD TG 429/EU B.42 provides guidance on the recommended vehicles, number of animals per group, concentrations of test chemical to be applied and substances to be used as a positive control. A preliminary study or evaluation of existing acute toxicity/dermal irritation data is normally conducted to determine the highest concentration of test substance that is soluble in the vehicle but does not cause unacceptable local or systemic toxicity. The submission of historical control data will demonstrate the ability of the test laboratory to produce consistent responses. Based on the use of radioactive labelling, chemicals that result in a stimulation index (SI) of ≥3 at one or more test concentrations are considered to be positive for skin sensitisation. Both positive and negative responses in the LLNA conducted as described in OECD TG 429/EU B.42 meet the data requirements for classification of a substance as a skin sensitizer: no further testing is required.

Alternative vehicles to those listed in OECD TG 429/EU B.42 may be used in the LLNA if sufficient scientific justification is provided. OECD TG 429/EU B.42 also states that endpoints other than radioactive labelling may be used to assess proliferation, on condition that justification and scientific support, which will include full citations and a description of the methodology, are provided.

The guinea pig test methods described in OECD TG 406/EU B.6, the GPMT (Magnusson et al., 1969, Schlede et al., 1995) and the Buehler, can also be used for hazard identification. Recommendations on conducting and analysing these methods are provided by Steiling et al.,

2001. Particular attention should be paid to the quality of these tests with consideration given to the following points:

i. numbers of test and control guinea pigs;

ii. number or percentage of test and control animals displaying skin reactions;

iii. whether skin irritation was observed at the induction phase;

iv. whether the maximal non-irritating concentration was used at the challenge phase;

v. the choice of an appropriate vehicle (ideally, one that solubilises or gives a stable suspension or emulsion of the test material, is free of allergenic potential, is non-irritating, enhances delivery across the stratum corneum, and is relevant to the usage conditions of the test material, although it is recognised that it will not always be possible to meet all these conditions);

vi. whether there are signs of systemic toxicity (a sighting study should be performed to determine an appropriate induction dose that causes irritation but not systemic toxicity);

vii. staining of the skin by the test material that may obscure any skin reactions (other procedures, such as chemical depilation of the reaction site, histopathological examination or the measurement of skin fold thickness may be carried out in such cases);

viii. results of rechallenge treatments if performed;

ix. checking of strain sensitivity at regular intervals by using an appropriate control substance (as specified in OECD guidelines and EU Test Methods). Currently (2007), the recommended interval is 6 months.

The investigation of doubtful reactions in guinea pig tests, particularly those associated with evidence of skin irritation following first challenge, may benefit from rechallenge of the test animals. In cases where reactions may have been masked by staining of the skin, other reliable procedures may be used to assist with interpretation; where such methods are used, the submitting laboratory should provide evidence of their value.

Non-guideline compliant tests and refinements to the standard assays

The submitted dossier should include scientific justification for conducting any new test that is a modification or deviation from guideline methods. In such cases, it would be advisable to seek appropriate expert advice on the suitability of the assay before testing is begun.

For hazard identification, it may be possible to use a reduced LLNA (rLLNA) (Kimber et al., 2006) which reduces the use of animals by requiring only a single (high) dose group (≥10%) and a concurrent negative control group. A preliminary study or evaluation of existing acute toxicity/dermal irritation data is normally conducted to determine the highest concentration of test substance that is soluble in the vehicle, but that does not cause unacceptable local or systemic toxicity. As with the full LLNA, although a concurrent positive control group is not required, registrants would be required to submit historical positive control data supportive of their competence. The rLLNA should be used only in appropriate circumstances:

i. where hazard identification is the primary objective and

ii. where potency data are not required

As in the standard (OECD guideline-compliant) LLNA, group sizes should comprise four or five animals. A positive result in a rLLNA will suffice in circumstances where risk assessment and/or risk management is NOT required. Registrants should be aware that the rLLNA is less scientifically rigorous than the standard LLNA, with an associated increased level of uncertainty.

Historically, guinea pig studies that are not fully equivalent to OECD test protocols have been conducted and can provide useful hazard information. These studies include, but are not limited to, the following: Draize test, optimisation test, split adjuvant test, open epicutaneous test and the cumulative contact enhancement test. In the case of positive results the substance may be considered as a potential skin sensitizer. If, taking into account the above quality criteria, especially the positive and negative control data, there is a clear negative result, i.e. no animals displaying any signs of sensitisation reactions, then no further animal testing is required. Where there is a low level of response, the quality of the study is questionable, or where unacceptably low concentrations of the test material have been used for induction and/or challenge, further testing may be required.

R.7.3.4.2 Human data on skin sensitisation

When reliable and relevant human data are available, they can be useful for hazard identification and even preferable over animal data. However, lack of positive findings in humans does not necessarily overrule positive and good quality animal data.

Well conducted human studies can provide very valuable information on skin sensitisation. However, in some instances (due to lack of information on exposure, a small number of subjects, concomitant exposure to other substances, local or regional differences in patient referral etc) there may be a significant level of uncertainty associated with human data. Moreover, diagnostic tests are carried out to see if an individual is sensitised to a specific agent, and not to determine whether the agent can cause sensitisation.

For evaluation purposes, existing human experience data for skin sensitisation should contain sufficient information about:

- the test protocol used (study design, controls)
- the substance or preparation studied (should be the main, and ideally, the only substance or preparation present which may possess the hazard under investigation)
- the extent of exposure (magnitude, frequency and duration)
- the frequency of effects (versus number of persons exposed)
- the persistence or absence of health effects (objective description and evaluation)
- the presence of confounding factors (e.g. pre-existing dermal health effects, medication; presence of other skin sensitizers)
- the relevance with respect to the group size, statistics, documentation
- the *healthy worker* effect

Evidence of skin sensitising activity derived from diagnostic testing may reflect the induction of skin sensitisation to that substance or cross-reaction with a chemically very similar substance. In both situations, the normal conclusion would be that this provides positive evidence of the skin sensitising activity of the chemical used in the diagnostic test.

Human experimental studies on skin sensitisation are not normally conducted and are generally discouraged. Where human data are available, then quality criteria and ethical considerations are presented in ECETOC monograph no 32.

Ultimately, where a very large number of individuals (e.g.10^5) have frequent (daily) skin exposure for at least two years and there is an active system in place to pick up complaints and adverse reaction reports (including *via* dermatology clinics), and where no or only a very few isolated cases of allergic contact dermatitis are observed then the substance is unlikely to be a significant skin sensitizer. However, information from other sources should also be considered in making a judgement on the substance's ability to induce skin sensitisation.

It is emphasised that testing with human volunteers is strongly discouraged, but when there are good quality data already available they should be used as appropriate in well justified cases.

R.7.3.5 Information and its sources on respiratory sensitisation

R.7.3.5.1 Non-human data on respiratory sensitisation

Non-testing data on respiratory sensitisation

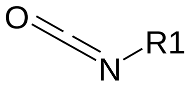
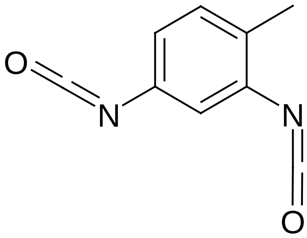
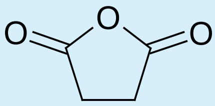
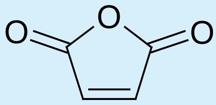
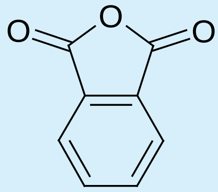
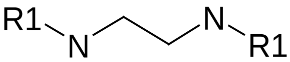
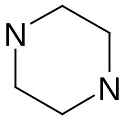
Attempts to model respiratory sensitisation have been hampered by a lack of a predictive test protocol for assessing chemical respiratory sensitisation. (Q)SAR models are available but these have largely been based on data for chemicals reported to cause respiratory hypersensitivity in humans. Examples of some structural alerts are shown in Table R.7.3–1.

Agius et al (1991) made qualitative observations concerning the chemical structure of chemicals causing occupational asthma. This work drew attention to the large proportion of chemical asthmagens with at least two reactive groups, e.g., ethylene diamine and toluene diisocyanate. The earlier work was followed up by a simple statistical analysis of the occurrence of structural fragments associated with activity, with similar conclusions (Agius et al, 1994 and 2000).

The MCASE group has developed three models for respiratory hypersensitivity (Karol et al., 1996, Graham et al., 1997, Cunningham et al., 2005). The Danish (Q)SAR Database has an in-house model for respiratory hypersensitivity for which estimates can be extracted from the on-line database (available at http://ecb.jrc.it/QSAR). Derek for Windows contains several alerts derived from a set of respiratory sensitisers/asthmogens (Payne et al., 1995).

Whilst the available structural alerts (SAR) are transparent and easily to apply (Aigus et al., 1991, 1994 and 2000, Payne et al., 1995), it should be stressed that these are derived on the basis of chemical asthmagens not specifically chemical respiratory allergens. A need therefore remains to develop new (Q)SARs as and when a robust predictive test method becomes available.

Table R.7.3–1 Examples of structural alerts for respiratory sensitisation

Structural Alert Description	Examples of structures
 isocyanate	 Toluene-2,4-diisocyanate
 cyclic anhydride	  maleic anhydride trimellitic anhydride
 diamine	 piperazine

Testing data for respiratory sensitisation

In vitro data

No *in vitro* tests specific for respiratory sensitisation are available yet, owing to the complexity of the mechanisms of the sensitisation process.

Efforts are still needed to identify the most relevant endpoints in the optimisation of existing tests. However, a combination of several *in vitro* tests, covering the relevant mechanistic steps of respiratory sensitisation, into a test battery could eventually lead to replacement of the *in vivo* tests.

Animal data

At present, although a number of test protocols has been published to detect respiratory allergenicity of low molecular weight compounds, none of these are validated nor are these widely accepted. One approach that might be of some value in characterising the likely respiratory sensitising activity of chemicals is application of the LLNA, or of other tests for measuring skin sensitisation potential. Although the LLNA was developed and validated for the identification of contact allergens, there is evidence that chemical respiratory allergens will also elicit positive responses in this assay (Kimber, 1995). That is, chemicals known to cause respiratory allergy and occupational asthma have been shown to test positive in the LLNA. Among such chemicals are acid anhydrides (such as trimellitic anhydride and phthalic anhydride), diisocyanates (including diphenylmethane diisocyanate and hexamethylene diisocyanate) and certain reactive dyes. In fact, the view currently is that most, if not all, chemical respiratory allergens are able to elicit positive responses in the LLNA, or in other tests for skin sensitisation, such as the M&K (guinea pig maximisation) test. This is true even of those chemical respiratory allergens, such as phthalic anhydride, for instance, that are implicated virtually exclusively with the induction of chemical respiratory allergy and have rarely, if ever, been shown to cause allergic contact dermatitis. Against this background and in combination with other data it might be possible to conclude in a WoE assessment that chemicals that (at an appropriate test concentration and test conditions, i.e. skin penetration should have occurred) are negative in the LLNA, as well as being considered as not being skin sensitizers, can also be regarded as lacking the potential to cause allergic sensitisation of the respiratory tract.

One approach that has been proposed for the identification of chemicals that have the potential to cause allergic sensitisation of the respiratory tract is one in which activity is measured as a function of the profiles of cytokines produced by draining lymph node cells in mice exposed more chronically (over a 2 week period) to the test chemical (Dearman et al., 2002). This method is predicated on an understanding that allergic sensitisation of the respiratory tract is favoured by selective Th_2-type immune responses and that in many instances chemical respiratory allergy and occupational asthma are associated with IgE antibody. Using this approach chemical respiratory allergens are identified as a function of their ability to stimulate in mice the selective development of preferential Th2-type immune responses associated with a predominance of type 2 cytokine secretion by draining lymph node cells (Dearman et al., 2002 and 2003). Specifically, chemical contact allergens promote Th1 responses characterised by an enhanced production of IFN-gamma, whereas chemical respiratory allergens promote Th_2 responses characterised by enhanced production of IL-4, IL-5 and IL-13. Many variables other than the compound itself, such as concentration used to induce sensitisation, duration of the sensitisation period, and presence or absence of mitogens to reveal differences in cytokine expression have all been noted to have impact on the outcome (Van Och et al., 2002). There are general guidelines now available for the conduct of the method (Dearman et al., 2003), however, this method has not yet been formally validated nor is it widely accepted.

Another, relatively simple approach may serve the purpose to specifically predict sensitisation of the respiratory tract: i.e. increases in total serum IgE antibodies after induction. This method is based on statistically significant increases in total serum IgE (see review by Arts and Kuper, 2007).

Methods that use both an induction and an inhalation elicitation or challenge phase and which include different parameters such as total and/or specific IgE antibody determinations, lung function testing, tests for a specific hyperreactivity (e.g. methacholine challenges), bronchoalveolar lavage measurements, and histopathological examination of the entire respiratory tract, may provide (additional) information on the potential of chemicals to cause respiratory sensitisation. These methods usually use high IgE-responding animal strains; to test for Th1-mediated responses low IgE-responding strains should typically be used. Several of these models have been reviewed recently (Arts and Kuper, 2007).

There are currently no predictive methods to identify chemicals that induce asthma through non-immunological mechanisms, however, when performing challenge tests including non-sensitised but challenged controls information can be obtained on non-immunological effects of these chemicals.

R.7.3.5.2 Human data on respiratory sensitisation

Human data on respiratory reactions (asthma, rhinitis, alveolitis) may come from a variety of sources:

- consumer experience and comments, preferably followed up by professionals (e.g. bronchial provocation tests, skin prick tests and measurements of specific IgE serum levels)

- records of workers' experience, accidents, and exposure studies including medical surveillance
- case reports in the general scientific and medical literature
- consumer tests (monitoring by questionnaire and/or medical surveillance)
- epidemiological studies

R.7.3.6 Evaluation of available information for respiratory sensitisation

R.7.3.6.1 Non-human data for respiratory sensitisation

Non-testing data for respiratory sensitisation

Given the lack of available (Q)SARs for respiratory sensitisation, it is not possible to provide any additional guidance.

Testing data for respiratory sensitisation

In vitro data

Presently (March 2007) there are no *in vitro* tests available to assess respiratory sensitisation. If such a method were to become available then it would need to be assessed for its relevance and reliability (Hartung et al., 2004).

Animal data

Although the LLNA does not represent a method for the specific identification of chemical respiratory allergens, there is evidence that chemical respiratory allergens will also elicit positive responses in this assay (Kimber, 1995). The interpretation is therefore that a chemical which fails to induce a positive response in the LLNA (at an appropriate test concentration) most probably lacks the potential for respiratory allergy. Conversely, it cannot be wholly excluded that a chemical that induces a positive response in the LLNA, might sensitise the respiratory tract upon inhalation or *via* dermal exposure. Any potential hazard for respiratory sensitisation could only be positively identified by further testing, although such testing is neither validated nor widely accepted.

One further approach to the identification of chemicals that have the potential to induce allergic sensitisation of the respiratory tract is *cytokine fingerprinting* (Dearman et al., 2002; see Section R.7.3.5.1). This method is predicated on an understanding that allergic sensitisation of the respiratory tract is favoured by selective Th_2-type immune responses and that in many instances chemical respiratory allergy and occupational asthma are associated with IgE antibody.

In addition, there are other approaches that have been proposed and these have been reviewed recently (Arts and Kuper, 2007) - although again it is important to emphasise that there are currently available no fully evaluated or validated animal models for the predictive identification of chemical respiratory allergens.

As indicated previously, some chemicals may have the potential to induce pulmonary reactions *via* Th1-type immune responses. Studies with typical skin allergens such as DNCB, DNFB and picryl chloride (trinitrochlorobenzene) in BALB/c mice, guinea pigs or Wistar rats have shown the potential of these chemicals to induce allergic reactions in the lungs that are independent of IgE (Garssen et al., 1991; Garcia et al., 1992; Buckley et al., 1994; Zwart et al., 1994;

Satoh et al., 1995; and see for a review Arts and Kuper, 2007). Sensitisation and challenge with DNCB resulted in laryngitis in low IgE-responding Wistar rats (Arts et al., 1998). [In addition, cellular immune responses to these sensitizers were shown to be associated with hyperreactivity of the airways to non-specific stimuli (Garssen et al., 1991).] For these reasons, it might be the case that people who are sensitised *via* the skin might suffer adverse pulmonary reactions if they were to inhale sufficient amounts of the contact allergen to which they were sensitised. As indicated previously, very few precedents for the elicitation of pulmonary reactions by skin sensitising chemicals in humans have been observed. In practice it appears not to represent a health issue.

R.7.3.6.2 Human data for respiratory sensitisation

Although human studies may provide some information on respiratory hypersensitivity, the data are frequently limited and subject to the same constraints as human skin sensitisation data.

For evaluation purposes, existing human experience data for respiratory sensitisation should contain sufficient information about:

- the test protocol used (study design, controls)
- the substance or preparation studied (should be the main, and ideally, the only substance or preparation present which may possess the hazard under investigation)
- the extent of exposure (magnitude, frequency and duration)
- the frequency of effects (versus number of persons exposed)
- the persistence or absence of health effects (objective description and evaluation)
- the presence of confounding factors (e.g. pre-existing respiratory health effects, medication; presence of other respiratory sensitizers)
- the relevance with respect to the group size, statistics, documentation
- the healthy worker effect

Evidence of respiratory sensitising activity derived from diagnostic testing may reflect the induction of respiratory sensitisation to that substance or cross-reaction with a chemically very similar substance. In both situations, the normal conclusion would be that this provides positive evidence for the respiratory sensitising activity of the chemical used in the diagnostic test.

For respiratory sensitisation, no clinical test protocols for experimental studies exist but tests may have been conducted for diagnostic purposes, e.g. bronchial provocation test. The test should meet the above general criteria, e.g. be conducted according to a relevant design including appropriate controls, address confounding factors such as medication, smoking or exposure to other substances, etc. Furthermore, the differentiation between the symptoms of respiratory irritancy and allergy can be very difficult. Thus, expert judgement is required to determine the usefulness of such data for the evaluation on a case-by-case basis.

Although predictive models are under validation, there is as yet no internationally recognized animal method for identification of respiratory sensitisation. Thus human data are usually evidence for hazard identification.

Where there is evidence that significant occupational inhalation exposure to a chemical has not resulted in the development of respiratory allergy, or related symptoms, then it may be possible to draw the conclusion that the chemical lacks the potential for sensitisation of the respiratory tract. Thus, for instance, where there is evidence that a large cohort of subjects have had opportunity for regular inhalation exposure to a chemical for a sustained period of time in the absence of respiratory symptoms, or related health complaints, then this will provide reassurance regarding the absence of a respiratory sensitisation hazard.

R.7.3.7 Conclusions on skin and respiratory sensitisation

The preceding paragraphs on skin and respiratory sensitisation are summarised in the separately provided summary tables. However, it is emphasised that the complete guidance text should be read in order to gain a correct and complete view of the described area.

R.7.3.7.1 Remaining uncertainty on sensitisation

Reliable data can be generated on skin sensitisation from well designed and well conducted studies in animals. The use of adjuvant in the GPMT may lower the threshold for irritation and so lead to false positive reactions, which can therefore complicate interpretation (running a pre-test with FCA treated animals can provide helpful information). In international trials, the LLNA has been shown to be reliable, but like the guinea pig tests is dependent on the vehicle used, and it can occasionally give false positive results with irritants. Careful consideration should be given to circumstances where exposure may be sub-optimal due to difficulties in achieving a good solution and/or a solution of sufficient concentration. In some circumstances inconsistent results from guinea pig studies, or between guinea pig and LLNA studies, might increase the uncertainty of making a correct interpretation. Finally, for existing human data consideration must be given to whether inter-individual variability is such that it is not scientifically sound to generalize from a limited test panel.

When considering whether or not a substance is a respiratory sensitizer, observations of idiosyncratic reactions in only a few individuals with hyper-reactive airways are not sufficient to indicate the need for classification.

Major uncertainties remain in our understanding of the factors that determine whether or not a substance is an allergen, and if so, what makes it a skin or a respiratory sensitizer.

R.7.3.7.2 Concluding on suitability for Classification and Labelling

REACH demands that all available information for a chemical is gathered and any lack of information is reported.

Skin sensitizers

Standard information required for skin sensitisation is described in Annex VII of REACH, i.e. for any substance manufactured or imported in quantity of 1 ton or more.

A substance can be classified as *skin sensitizer* following the flow chart for integrated testing strategy (ITS) reported in Table R.7.3–1 in Section R.7.3.8.3.

According to Directive 67/548/EEC[61], labelling for skin sensitisation is with symbol Xi, the indication of danger *irritant* and the risk phrase R43 (R43: May cause sensitisation by skin contact).

Respiratory sensitizers

In REACH, respiratory sensitizers are indicated for harmonised classification and labelling and regulated in Annex I of Directive 67/548/EEC. Annex XV in REACH lays down general principles for preparing dossiers to propose and justify harmonised classification and labelling of CMRs (carcinogenic, mutagenic, toxic for reproduction) and respiratory sensitizers.

Potential hazard for respiratory sensitisation cannot be easily addressed, as validated testing methods are currently not available. A probable hazard for respiratory sensitisation should be mentioned in the Safety Data Sheet.

Although no testing strategy is available, a substance could be classified as *respiratory sensitizer* by following the flow chart for integrated evaluation strategy (IES) reported in Section R.7.3.8.3 which is based on existing evidence.

According to Directive 67/548/EEC, labelling for *respiratory sensitizers* is with symbol Xn, the indication of danger *harmful* and the risk phrase R42 (R42: May cause sensitisation by inhalation). Concluding on suitability for chemical safety assessment: dose response assessment and potency

There is evidence that for both skin sensitisation and respiratory hypersensitivity dose-response relationships exist (although these are frequently less well defined in the case of respiratory hypersensitivity). The dose of agent required to induce sensitisation in a previously naïve subject or animal is usually greater than that required to elicit a reaction in a previously sensitised subject or animal; therefore the dose-response relationship for the two phases will differ. Little or nothing is known about dose-response relationships in the development of respiratory hypersensitivity by non-immunological mechanisms.

It is frequently difficult to obtain dose-response information from either existing human or guinea pig data where only a single concentration of the test material has been examined. With human data, exposure measurements may not have been taken at the same time as the disease was evaluated, adding to the difficulty of determining a dose response.

Dose-response data however, can be generated from local lymph node assays or, in exceptional cases, using specially designed guinea pig test methods. Such types of data can give data on induction and elicitation thresholds in these models, but it must be remembered these cannot be translated directly to human thresholds.

Measurement of potency

Appropriate dose-response data can provide important information on the potency of the material being tested. This can facilitate the development of more accurate risk assessments. This section refers to potency in the induction phase of sensitisation.

Neither the standard LLNA nor the GPMT/Buehler is specifically designed to evaluate the skin sensitising potency of test compounds, instead they are used to identify sensitisation potential for classification purposes. However, all could be used for some estimate of potency. The

---

[61] Directive 67/548/EEC will be repealed and replaced with the EU Regulation on classification, labelling and packaging of substances and mixtures, implementing the Globally Harmonized System (GHS). See section R.7 Introduction.

relative potency of compounds may be indicated by the percentage of positive animals in the guinea pig studies in relation to the concentrations tested. Likewise, in the LLNA, the EC3 value (the dose estimated to cause a 3-fold increase in local lymph node proliferative activity) can be used as a measure of relative potency (ECETOC, 2000). Often linear interpolation of a critical effect dose from the EC3 is proposed (ECETOC), but more advanced statistical approaches basing conclusions on the characteristic of the dose response curve and variability of the results is also used (Basketter et al., 1999, van Och et al., 2000). The dose-response data generated by the LLNA makes this test more informative than guinea pig assays for the assessment of skin sensitising potency. EC3 data correlate well with human skin sensitisation induction thresholds derived from historical predictive testing (Schneider et al., 2004; Griem, 2003; Basketter et al., 2005b). Accordingly, there are proposals for how this information may be used in a regulatory sense (Basketter et al., 2005b) and for risk assessment.

Derivation of a DNEL

Potency information, such as the LLNA EC3 value, can be utilised for the derivation of no-effect levels, that is – in this instance - the threshold required for the induction of skin sensitisation. It should be noted that thresholds for skin sensitisation should be expressed in terms of dose per unit area. As mentioned above, the EC3 value correlates well with thresholds observed in previously published human predictive test data and with clinical experience (reviewed in Basketter et al., 2007a). The EC3 value can then be extrapolated by the application of assessment factors (reflecting e.g. intra and inter-individual variability and vehicle matrix effects) to derive no-effect levels (expressed in µg/$cm^2$of skin) for use of specific skin sensitizers in defined exposure situations (Gerberick et al., 2001; Felter et al., 2002 and 2003; Basketter et al., 2006). The approach is commonly referred to as quantitative risk assessment (QRA) and has been deployed, with considerable effect, to identify safe exposure levels for a range of skin sensitising chemicals (Zachariae et al., 2003; Basketter et al., 2003). Most recently, this has been reported extensively for fragrance and preservative sensitizers (Api et al., 2007; Basketter et al., 2007b).

Guidance on how to use the potency information for qualitative assessment (see also Section E.3.4.2) and how to derive a DNEL as a second step in the safety assessment of sensitizers is given in Appendix R.8-10.

R.7.3.7.3 Additional considerations

Chemical allergy is commonly designated as being associated with skin sensitisation (allergic contact dermatitis), or with sensitisation of the respiratory tract (asthma and rhinitis). In view of this it is sometimes assumed that allergic sensitisation of the respiratory tract will result only from inhalation exposure to the causative chemical, and that skin sensitisation necessarily results only from dermal exposure. This is misleading, and it is important for the purposes of risk management to acknowledge that sensitisation may be acquired by other routes of exposure. Since adaptive immune responses are essentially systemic in nature, sensitisation of skin surfaces may theoretically develop from encounter with contact allergens *via* routes of exposure other than dermal contact (although in practice this appears to be uncommon). Similarly, there is evidence from both experimental and human studies which indicate that effective sensitisation of the respiratory tract can result from dermal contact with a chemical respiratory allergen. Thus, in this case, it appears that the quality of immune response necessary for acquisition of sensitisation of the respiratory tract can be skin contact with chemical respiratory allergens (Kimber et al., 2002). Such considerations have important implications for risk management. Thus, for instance, there is a growing view that effective prevention of respiratory sensitisation requires protection of both skin and respiratory tracts. This includes the cautious use of known contact allergens in products to which consumers are (or may be) exposed *via* inhalation, such as sprays. The generic advice is that appropriate strategies to minimise the risk of sensitisation to chemical allergens will require consideration of providing protection of all relevant routes of exposure.

R.7.3.7.4 Information not adequate

A WoE approach, comparing available adequate information with the tonnage-triggered information requirements by REACH, may result in the conclusion that the requirements are not fulfilled. In order to proceed in further information gathering the testing strategy given in the next Section R.7.3.8 can be adopted.

R.7.3.8 Integrated testing strategy (ITS) for sensitisation

R.7.3.8.1 Objective / General principles

Ensure that the objective of this testing strategy is to give guidance on a stepwise approach to hazard identification with regard to the endpoint; a key principle of the strategy is that the results of one study are evaluated before another is initiated. The strategy should seek to ensure that the data requirements are met in the most efficient and humane manner so that animal usage and costs are minimised.

R.7.3.8.2 Preliminary considerations

The guidance given in Sections R.7.3.2 to R.7.3.4 above will have enabled the identification of the data gaps that need to be filled in to meet the requirements of REACH as defined in Annexes VI to XI. Careful consideration of existing toxicological data, exposure characteristics and current risk management procedures is recommended to ascertain whether the fundamental objectives of the ITS (see above) have already been met. Give guidance on other factors that might mitigate data requirements for the endpoint of interest e.g. possession of other toxic properties, characteristics that make testing technically not possible.

R.7.3.8.3 Testing strategies for sensitisation

Develop a testing strategy for the endpoint that takes account of existing data on toxicity, exposure characteristics as well as the specific rules for adaptation from standard information requirements, as described in column 2 of Annexes VII-X, together with some general rules for adaptation from standard information requirements in Annex XI.

Figure R.7.3–1 Integrated testing strategy for skin sensitisation

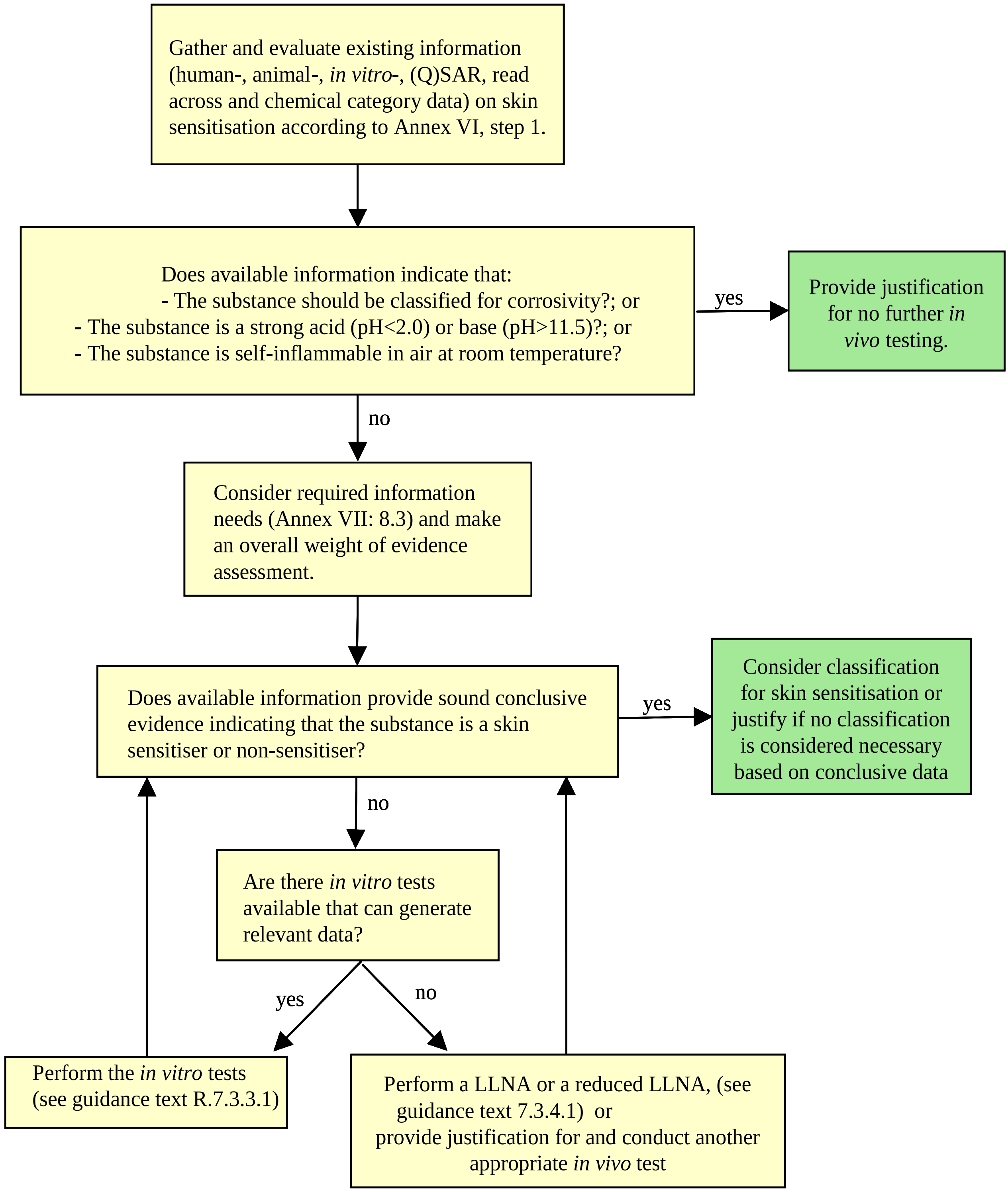
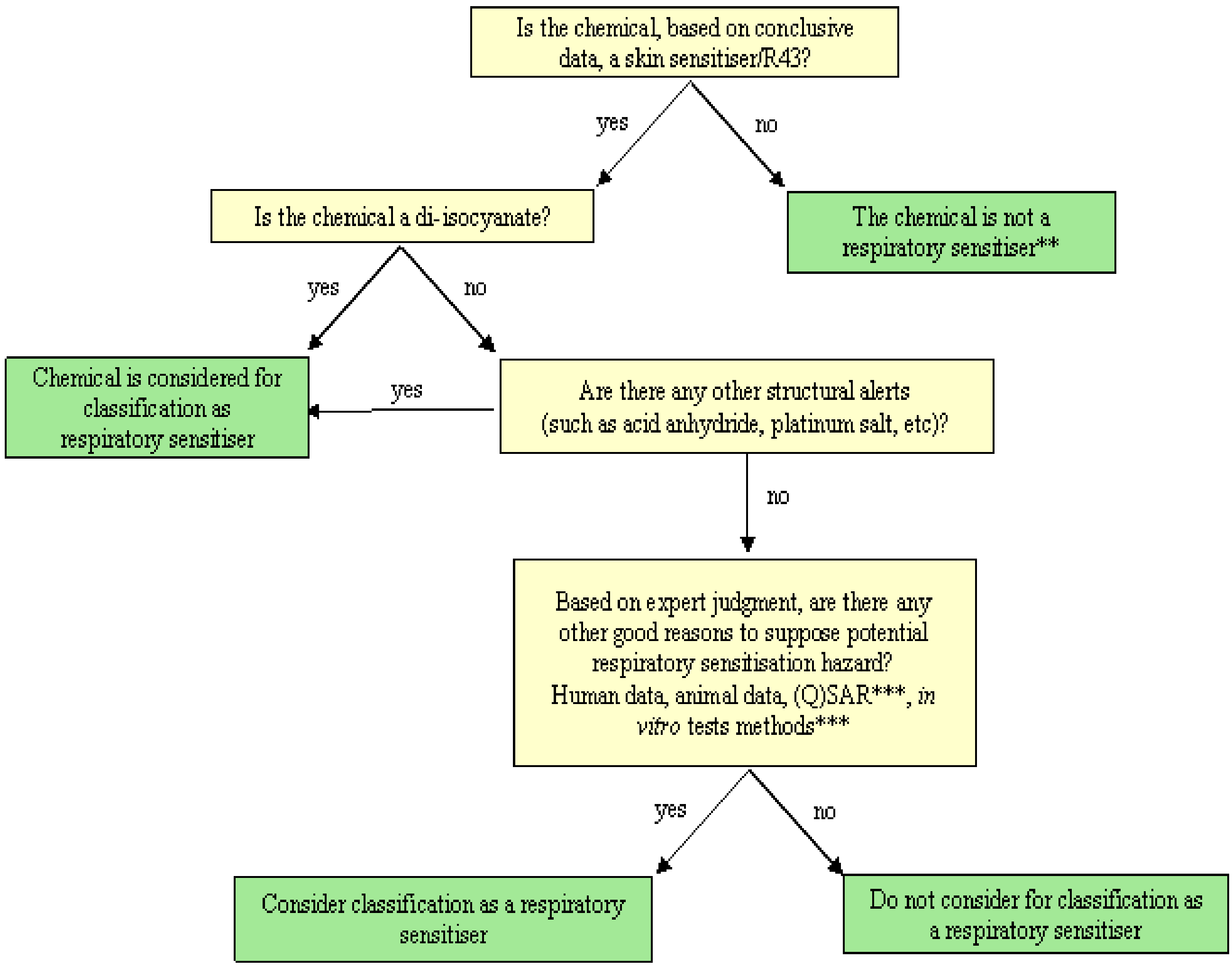


Figure R.7.3–2 Integrated evaluating strategy for respiratory sensitisation data*



* In contrast to tests for skin sensitisation, the performance of tests for respiratory sensitisation is currently not required under REACH. Therefore the present IES scheme depicts a strategy for evaluating existing data.
** This does not discount the possibility that the chemical may induce respiratory hypersensitivity through non-immunological mechanisms. Chemicals that act through such mechanisms are usually identified on the basis of evidence from human exposure.
*** not yet available

R.7.3.9 References on skin and respiratory sensitisation

Accelrys Inc., TopKat User Guide Version 6.02, San Diego: Accelrys Inc., 2004.

Aeby P, Wyss C, Beck H, Griem P, Scheffler, H and Goebel, C. 2004. Characterisation of the sensitizing potential of chemicals by in vitro analysis of dendritic cell activation and skin penetration. Journal of Investigative Dermatolog. 122, 1154-1164.

Agius RM, Nee J, McGovern B, Robertson A. 1991. Structure activity hypotheses in occupational asthma caused by low molecular weight substances. Ann Occup Hyg 35:129–137.

Agius RM, Elton RA, Sawyer L, Taylor P. 1994. Occupational asthma and the chemical properties of low molecular weight organic substances. Occup Med-State of the Art Rev 44:34–36.

Agius RM. 2000. Why are some low molecular weight agents asthmagenic? Occup Med-State of the Art Rev 50:369–384.

Aiba S, Manome H, Yoshino Y, Tagami H. 2000. In vitro treatment of human transforming growth factor-beta1-treated monocyte-derived dendritic cells with haptens can induce the phenotypic and functional changes similar to epidermal Langerhans cells in the initiation phase of allergic contact sensitivity reaction. Immunology 101, 68-75

Api AM, Basketter DA, Cadby PA, Cano M-F, Ellis G, Gerberick G F, Griem P, McNamee PM, Ryan CA, Safford B. Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients. Regulatory Toxicology and Pharmacology, submitted.

Aptula AO, Patlewicz G, Roberts DW. 2005. Skin Sensitisation: Reaction Mechanistic Applicability Domains for Structure-Activity Relationships. Chemical Research in Toxicology. 18: 1420-1426.

Aptula AO, Roberts DW. 2006a. Mechanistic applicability domains for nonanimal-based prediction of toxicological endpoints: general principles and application to reactive toxicity. Chemical Research in Toxicology 19(8):1097-105.

Aptula AO, Patlewicz G, Roberts DW, Schultz TW. 2006b. Non-enzymatic glutathione reactivity and in vitro toxicity: A non-animal approach to skin sensitisation, Tox in Vitro 20, 239-247

Arts JHE., Kuper CF, Spoor SM, Bloksma N. 1998. Airway morphology and function of rats following dermal sensitization and respiratory challenge with low molecular weight chemicals. Toxicol.Appl. Pharmacol. 152, 66-76

Arts JHE, Bloksma N, Leusink-Muis A, Kuper CF. 2003. Respiratory allergy and pulmonary irritation to trimellitic anhydride in Brown Norway rats. Toxicol. Appl. Pharmacol, 187, 38-49

Arts JHE, Kuper CF. 2007. Animal models to test respiratory allergy of low molecular weight chemicals: a guidance. Methods 41, 61-71

Ashby J, Basketter DA, Paton D, Kimber I. 1995. Structure activity relationships in skin sensitisation using the murine local lymph node assay. Toxicology 103: 177-194.

Azam P, Peiffer JL, Chamousset D, Tissier MH, Bonnet PA, Vian L, Fabre I and Ourlin JC. 2006. The cytokine-dependent MUTZ-3 cell line as an in vitro model for the screening of contact sensitizers. Toxicology and Applied Pharmacology 212, 14-23.

Balls M. et al. 1995. Practical aspects of the validation of Toxicity Test Procedures. ATLA 23:129-147

Barratt MD, Basketter DA, Chamberlain M, Admans GD, Langowski JJ. 1994. An expert system rulebase for identifying contact allergens. Toxicol in Vitro. 8: 1053-1060.

Barratt MD, Basketter DA, Roberts DW. 1997. Quantitative structure activity relationships. In Lepoittevin J-P, Basketter DA, Dooms-Goossens A, Karlberg A-T, eds, The Molecular Basis of Allergic Contact Dermatitis. Springer-Verlag Heidelberg. pp. 129-154.

Barratt MD, Langowski JJ. 1999. Validation and subsequent development of the Derek skin sensitisation rulebase by analysis of the BgVV list of contact allergens. J Chem Inf Comput Sci. 39: 294-298.

Barratt MD, Castell JV, Miranda MA, Langowski JJ. 2000. Development of an expert system rulebase for the prospective identification of photoallergens. J Photochem Photobiol B. 58: 54-61.

Basketter DA, Roberts DW, Cronin M Scholes EW. 1992. The value of the local lymph node assay in quantitative structure-activity investigations. Contact Dermatitis. 27: 137-142

Basketter DA, Lea LJ, Dickens A, Briggs D, Pate I, Dearman RJ, Kimber I. 1999. A comparison of statistical approaches to the derivation of EC3 values from local lymph node assay dose responses. J Appl Tox 19, 261-266.

Basketter DA, Angelini G, Ingber A, Kern P and Menné T. 2003. Nickel, chromium and cobalt in consumer products: revisiting safe levels in the new millennium. Contact Dermatitis, 49, 1 - 7.

Basketter DA, Clapp C, Jefferies D, Safford RJ, Ryan CA, Gerberick GF, Dearman RJ and Kimber I. 2005a. Predictive identification of human skin sensitisation thresholds. Contact Dermatitis, 53, 260 - 267.

Basketter DA, Andersen KE, Lidén C, van Loveren H, Boman A, Kimber I, Alanko K and Berggren E. 2005b. Evaluation of the skin sensitising potency of chemicals using existing methods and considerations of relevance for elicitation. Contact Dermatitis, 52: 39 - 43.

Basketter DA and Kimber I. 2006. Predictive test for irritants and allergens and their use in quantitative risk assessment. In “Contact Dermatitis”, 4th Edition. Eds, Frosch PJ, Menné T and Lepoittevin J-P, Springer Verlag, Heidelberg, pp 179 – 188.

Basketter DA, Gerberick GF and Kimber I. 2007a. The local lymph node assay EC3 value: status of validation. Contact Dermatitis, submitted.

Basketter DA, Clapp CJ, Safford BJ, Jowsey IR, McNamee PM, Ryan CA, Gerberick GF. 2007b. Preservatives and skin sensitisation quantitative risk assessment: risk benefit considerations. Dermatitis, subitted.

Baur X, Dewair M, and Fruhmann G. 1984. Detection of immunologically sensitized isocyanate workers by RAST and intercutaneous skin tests. J. Allergy Clin. Immunol., 73, 610-618.

Bentley AM, Maestrelli P, Saetta M, Fabbri LM, Robinson DS, Bradley BL, Jeffery PK, Durham SR, Kay AB. 1992. Activated T-lymphocytes and eosinophils in the bronchial mucosa in isocyanate-induced asthma. J. Allergy Clin. Immunol. 89, 821-829.

Bernstein IL, Bernstein DI, Chan-Yeung M, Malo J-L. 1993. Definition and classification of asthma. In: Bernstein, I.L., Chan-Yeung, M., Malo, J-L. and Bernstein D.I. (eds). Asthma in the Workplace. Marcel Dekker, New York, 1-4.

Buckley TL, Nijkamp FP. 1994. Mucosal exudation associated with a pulmonary delayed-type hypersensitivity reaction in the mouse. Role for the tachykinins. J Immunol. 153, 4169-4178

Casati et al. 2005. Dendritic cells as a tool for the predictive identification of skin sensitisation hazard. ATLA 33, 47-62

Chan-Yeung M and Malo J-L. 1993. Compendium 1: table of major inducers of occupational asthma. In: Bernstein, I.L., Chan-Yeung, M., Malo, J-L. and Bernstein, D.I., editors. Asthma in the Workplace. Marcel Dekker, New York, pp 595-623.

Corsini E, Primavera A, Marinovich M, Galli CL. 1998. Selective Induction of cell associated interleukin-1 in murine keratinocytes by chemical allergens, Toxicology 129, 193-200

Coquette A., Berna N, Vandenbosch A, Rosdy M, De Wever B, Poumay Y. 2003. Analysis of interleukin-1α (IL-1α) and interleukin-8 (IL-8) expression and release in in vitro reconstructed human epidermis for the prediction of in vivo skin irritation and/or sensitisation. Toxicol. In Vitro 17, 311-321.

Cronin MT, Basketter DA.1994. Multivariate QSAR analysis of a skin sensitisation database. SAR QSAR Environ Res. 2(3):159-79.

Cronin MTD, Jaworska JS, Walker JD, Comber MHI, Watts CD, Worth AP. 2003. Use of QSARs in International Decision-Making Frameworks to Predict Health Effects in Chemical Substances. Environmental Health Perspectives. 111(10): 1391-1401.

Cunningham AR, Cunningham SL, Consoer DM, Moss ST, Karol MH. 2005. Development of an information-intensive structure-activity relationship model and its application to human respiratory chemical sensitizers. SAR QSAR Environ Res. 16(3):273-85.

Curren, Southee, Spielmann, Liebsch, Fentem, Balls. 1995. The Role of Prevalidation in the Development, Validation and Acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217

Dearman RJ, Warbrick EV, Skinner R, Kimber I. 2002. Cytokine fingerprinting of chemical allergens: species comparisons and statistical analyses. Food Chem. Toxic. 40, 1881-1892

Dearman RJ, Betts CJ, Humphreys N, Flanagan BF, Gilmour NJ, Basketter DA, Kimber I. 2003. Chemical allergy: considerations for the practical application of cytokine profiling. Toxicol. Sci. 71, 137-145

Devillers J. 2000. A Neural Network SAR Model for Allergic Contact Dermatitis Toxicology Methods 10(3): 181-193.

Dimitrov SD, Low LK, Patlewicz GY, Kern PS, Dimitrova GD, Comber MH, Phillips RD, Niemela J, Bailey PT, Mekenyan OG. 2005. Skin sensitisation: modeling based on skin metabolism simulation and formation of protein conjugates. Int J Toxicol. 24(4):189-204.

Ebino K, Ueda H, Kawakatsu H, Shutoh Y, Kosaka T, Nagayoshi H, Lemus R, and Karol MH. 2001. Isolated airway exposure to toluene diisocyanate results in skin sensitization. Toxicol. Lett., 121, 79-85.

ECETOC Technical Report 89, 2003: (Q)SARs: Evaluation of the commercially available software for human health and environmental endpoints with respect to chemical management applications

Enslein K, Gombar VK, Blake BW, Maibach HI, Hostynek JJ, Sigman CC, Bagheri D. 1997. A quantitative structure-toxicity relationships model for the dermal sensitisation guinea pig maximization assay. Food Chem.Toxicol. 35: 1091-1098.

Eskes C, Zuang V eds. 2005. Alternative (Non Animal) Methods for Cosmetic Testing: Current Status and future Prospects. ATLA Vol 33 Supplement 1, 83-103.

Estrada E, Patlewicz G, Chamberlain M, Basketter D, Larbey S. 2003. Computer-aided knowledge generation for understanding skin sensitisation mechanisms: The TOPS-MODE approach. Chem. Res. Toxicol. 16: 1226-1235.

Farraj AK, Harkema JR and Kaminski NE. 2004. Allergic rhinitis induced by intranasal sensitisation and challenge with trimellitic anhydride but not with dinitrochlorobenzene or oxazolone in A/J mice. Toxicol. Sci. 79: 315-325.

Fedorowicz A, Zheng L, Singh H, Demchuk E. 2004. Structure-Activity Models for Contact Sensitisation. Int. J. Mol. Sci. 5: 56-66.

Fedorowicz A, Singh H, Soderholm S, Demchuk E. 2005. Structure-activity models for contact sensitisation. Chem Res Toxicol 18, 954-969

Felter SP, Robinson MK, Basketter DA and Gerberick GF. 2002. A review of the scientific basis for default uncertainty factors for use in quantitative risk assessment of the induction of allergic contact dermatitis. Contact Dermatitis, 47, 257-266.

Felter SP, Ryan CA, Basketter DA and Gerberick GF. 2003. Application of the risk assessment paradigm to the induction of allergic contact dermatitis. Regulatory Toxicol Pharmacol, 37, 1 - 10.

Garcia H, Salter-Cid L, Stein-Streilein J. 1992. Persistent interleukin-2 activity and molecular evidence for expression of lymphotoxin in the hapten-immune model for pulmonary interstitial fibrosis. Am. J. Respir. Cell. Mol. Biol. 6, 22-28.

Garssen J, Nijkamp FP, Vliet van der H, Loveren van H. 1991. T-cell mediated induction of airway hyperreactivity in mice. Am. Rev. Resp. Dis. 144, 931-938

Gealy R, Graham C, Sussman NB, Macina OT, Rosenkranz HS, Karol MH. 1996. Evaluating clinical case report data for SAR modeling of allergic contact dermatitis. Hum Exp Toxicol. 15(6): 489-93.

Gerberick GF, Robinson MK, Felter S, White I and Basketter DA. 2001. Understanding fragrance allergy using an exposure-based risk assessment approach. Contact Dermatitis, 45, 333-340.

Gerberick et al., Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP. 2004. Development of a peptide reactivity assay for screening contact allergens. Toxicological Sciences 81, 332-343.

Gerberick GF, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, Patlewicz GY, Basketter DA. 2005. Compilation of historical local lymph node data for evaluation of skin sensitisation alternative methods. Dermatitis 16(4):157-202.

Gerner I, Barratt MD, Zinke S, Schlegel K, Schlede E. 2004. Development and prevalidation of a list of structure-activity relationship rules to be used in expert systems for prediction of the skin-sensitising properties of chemicals. Altern Lab Anim. Nov;32(5):487-509.

Graham C, Gealy R, Macina OT, Karol MH, Rosenkranz HS. 1996. QSAR for allergic contact dermatitis. Quant Struct-Act Relat. 15: 224-229.

Graham C, Rosenkranz HS, Karol MH. 1997. Structure-activity model of chemicals that cause human respiratory sensitisation. Regul Toxicol Pharmacol 26:296–306.

Griem P, Goebel C, and Scheffler H. 2003. Proposal for a risk assessment methodology for skin sensitisation potency data. Reg. Tox. Pharmacol., 38: 269-290.

Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Hoffmann S, Janusch Roi A, Prieto P, Sabbioni E, Scott L, Worth A, Zuang V. 2004. A modular approach to the ECVAM Principles on Test Validity ATLA 32: 467-472

Health and Safety Executive. 1997. Critical assessments of the evidence for agents implicated in occupational asthma. HSE Books, Sudbury, Suffolk, UK

Herouet C, Cottin M, LeClaire J, Enk A. and Rousset F. 2000. Contact sensitizers specifically increase MHC class II expression on murine immature dendritic cells. In Vitro Molecular Toxicology 13, 113-123

Herrick CA, Das J, Xu L, Wisnewski AV, Redlich CA, Bottomly K. 2003. Differential roles for CD4 and CD8 T cells after diisocyanate sensitisation: genetic control of TH2-induced lung inflammation. J. Allergy Clin. Immunol. 111, 1087-1094.

Howes et al 1996. Methods for assessing percutaneous absorption ATLA 24: 81-106

Jowsey IR, Basketter DA, Westmoreland C, Kimber I. 2006. A future approach to measuring relative skin sensitising potency: a proposal. Journal of Applied Toxicology 26(4):341-50.

Johnson R, Macina OT, Graham C, Rosenkranz HS, Cass GR, Karol MH. 1997. Prioritizing testing or organic compounds detected as gas phase air pollutants: Structure-activity study for human contact allergens. Environmental Health Perspectives 105(9) 986-992.

Karol MH, Graham C, Gealy R, Macina OT, Sussman N, Rosenkranz HS. 1996. Structure-activity relationships and computer-assisted analysis of respiratory sensitisation potential. Toxicol Lett. 86:187-91

Kimber I. 1995. Contact and respiratory sensitization by chemical allergens: uneasy relationhips. Am. J. Contact Derm. 6, 34-39

Kimber I, Pichowski JS, Betts CJ, Cumberbatch M, Basketter D, Dearman R. 2001. Alternative Approaches to the Identification and Characterisaton of Chemical Allergens. Toxicology in vitro 15, 307-312

Kimber I and Dearman RJ. 2002. Chemical respiratory allergy: role of IgE antibody and relevance of route of exposure. Toxicology 181-182, 311-315.

Kimber I, Dearman RJ, Gerberick GF, Roggeband R and Basketter DA. 2003. Designation of substances as skin sensitising chemicals: a commentary. Human and Experimental Toxicology 22: 439-443.

Kimber I. and Dearman RJ. 2005. What makes a chemical a respiratory sensitizer. Current Opinion Allergy Clin. Immunol. 5, 119-124.

Kimber I, Dearman RJ, Betts CJ, Gerberick GF, Ryan CA, Kern PS, Patlewicz GY, Basketter DA. 2006. The local lymph node assay and skin sensitisation: a cut-down screen to reduce animal requirements? Contact Dermatitis 54: 181-5.

Kligman AM 1966. The identification of contact allergens by human assay. III. The maximization test: a procedure for screening and rating contact sensitizers. J. Invest. Derm. 47: 393-409.

Kubilus J, Cannon C, Neal P, Sennot H, Klausner M. 1986. Response of the EpiDerm Skin model to topically applied irritants and allergens. In vitro Toxicology 9, 157-166).

Li S, Fedorowicz A, Singh H, Soderholm SC. 2005. Application of the Random Forest method in Studies of Local Lymph Node Assay Based skin sensitisation data. J Chem Inf Model 45 952-964.

Li Y, Tseng YJ, Pan D, Liu J, Kern PS, Gerberick GF, Hopfinger AJ. 2007. 4D-Figerprint Categorical QSAR Models for skin sensitisation based on the classification of Local Lymph Node assay measures. Chem Res Tox 20(1):114-128.

Maestrelli P, Occari P, Turato G, Papiris SA, Di Stefano A, Mapp CE, Milani GF, Fabbri LM, Saetta M. 1997. Expression of interleukin (IL)-4 and IL-5 proteins in asthma induced by toluene isocyanate. Clin. Exp Allergy 27, 1292-1298.

Magnusson B. and Kligman AM. 1969. The identification of contact allergens by animal assay. The guinea pig maximisation test. J Invest Dermatol 52: 268-276.

Miller MD, Yourtee DM, Glaros AG, Chappelow CC, Eick JD, Holder AJ. 2005. Quantum Mechanical Structure-Activity Relationship Analyses for Skin Sensitisation. J. Chem. Inf. Model. 45: 924-929.

Van Och FM, Slob W, de Jong WH, Vandebriel RJ, van Loveren H. 2000. A quantitative method for assessing the sensitizing potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of the uncertainty margins. Toxicology 20;146(1):49-59.

Van Och FMM, Loveren van H, Jong de WH, Vandebriel RJ. 2002. Cytokine production induced by low-molecular-weight chemicals as a function of the stimulation index in a modified local lymph node assay: an approach to discriminate contact sensitizers from respiratory sensitizers. Toxicol. Appl.Pharmacol. 184, 46-56

Van Och FMM, Van Loveren H, Van Wolfswinkel JC, Machielsen AJC, Vandebriel RJ. 2005. Assessement of allergenic activity of low molecular weight compounds base don IL-1a and IL-18 production by a murine keratinocyte cell line. Toxicol 210, 95-109.

O'Neill R. 1995 Asthma at work. Causes, effects and what to do about them. Trades Union. Congres Sheffield Occupational Health Project Co-op Ltd., Sheffield, UK, p. 133.

Patlewicz G, Basketter DA, Smith CK,Hotchkiss SA, Roberts DW. 2001. Skin-sensitisation structure-activity relationships for aldehydes. Contact Dermatitis 44: 6 331-336.

Patlewicz G, Wright ZM, Basketter DA, Pease CK, Lepoittevin J-P, Arnau EG. 2002. Structure-activity relationships for selected fragrance allergens. Contact Dermatitis 47(4): 219-226.

Patlewicz G, Roberts DW, Walker JD. 2003. QSARs for the skin sensitisation potential of aldehydes and related compounds. QSAR Comb Sci. 22: 196-203.

Patlewicz G, Basketter DA, Pease CK, Wilson K, Wright ZM, Roberts DW, Bernard G, Gimenez Arnau E, Lepoittevin J-P. 2004. Further evaluation of quantitative structure-activity relationship models for the prediction of the skin sensitisation potency of selected fragrance allergens. Contact Dermatitis. 50: 91-97.

Payne MP, Walsh PT. Structure-activity relationships for skin sensitisation potential: development of structural alerts for use in knowledge-based toxicity prediction systems. J Chem Inf Comput Sci, 1994, 34, 154-161

Payne MP, Walsh PT. 1995. Structure-activity relationships for respiratory sensitisation. Proceedings, British Toxicology Society Meeting York, Yorkshire, UK, March/April 1995.

Potts RO, Guy RH. 1992. Predicting skin permeability. Pharmaceutical Research. 9(5): 663-669

Python, F., Goebel, C., Aeby, P. Assessment of the U937 cell line for the detection of contact allergens Toxicology and Applied Pharmacology (2007), doi: 10.1016/j.taap.2006.12.026 (in press).

Ren Y; Liu H; Xue C; Yao X; Liu M; Fan, B. 2006. Classification study of skin sensitizers based on support vector machine and linear discriminant analysis. Analytica Chimica Acta 572(2), 272-282.

Roberts, DW, Williams DL. 1982. The derivation of quantitative correlations between skin sensitisation and physico-chemical parameters for alkylating agents and their application to experimental data for sultones. J. Theor. Biol. 99: 807-825

Roberts DW. 1987. Structure-activity relationships for skin sensitisation potential of diacrylates and dimethacrylates. Contact Dermatitis. 17: 281-289.

Roberts DW, York M, Basketter DA. 1999. Structure-activity relationships in the murine local lymph node assay for skin sensitisation: alpha, beta-diketones. Contact Dermatitis 41(1):14-7.

Roberts, DW, Basketter DA. 2000. Quantitative structure-activity relationships: sulfonate esters in the local lymph node assay. Contact Dermatitis. 42(3):154-61.

Roberts DW, Patlewicz G. 2002. Mechanism based structure-activity relationships for skin sensitisation – the carbonyl group domain. SAR QSAR Env Res. 13(1): 145-152.

Roberts DW, Aptula AO, Patlewicz GY. 2006. Mechanistic Applicability Domains for Non-Animal Based Toxicological Endpoints. QSAR Analysis of the Schiff Base Applicability Domain for Skin Sensitisation. Chem Res Toxicol, 19(9):1228-33.

Roberts DW, Aptula AO, Cronin MTD, Hulzebos E, Patlewicz G. Global (Q)SARs for skin sensitization - assessment against OECD principles, SAR QSAR Environ. Res. 2007: in press.

Rosenkranz HS, Klopman G, Zhang YP, Graham C, Karol MH. 1999. Relationship between allergic contact dermatitis and electrophilicity. Environmental Health Perspectives 107: (2) 129-132

Ryan CA, Gerberick GF, Gildea LA, Hulette BC, Betts CJ, Cumberbatch M, Dearman RJ, Kimber I. 2005. Interactions of contact allergens with dendritic cells: opportunities and challenges for the development of novel approaches to hazard assessment. Toxicological Sciences 88, 4-11

Sakaguchi H, Ashikaga T, Miyazawa M. Yoshida Y, Yoneyama K, Hirota M, Itagaki H, Toyoda H, Suzuki H. 2006. Development of an in vitro skin sensitisation test using human cell lines; human Cell Activation Test (h-CLAT), I and II. Tox in Vitro 20, 767-773, 774-784.

Satoh T, Kramarik JA, Tollerud DJ, Karol MH. 1995. A murine model for assessing the respiratory hypersensitivity potential of chemical allergens. Toxicol. Lett. 78, 57-66.

Schultz TW, Carlson RE, Cronin MTD, Hermens JLM, Johnson R, O'Brien PJ, Roberts DW, Siraki A, Wallace KB, Veith GD. 2006. SAR and QSAR in Environmental Research 17(4) 413-428. A conceptual framework for predicting the toxicity of reactive chemicals" modelling soft electrophilicity.

Schlede E. and Eppler R. 1995. Testing for skin sensitisation according to the notification procedure for new chemicals: the Magnusson and Kligman test. Contact Dermatitis 32: 1-4.

Schneider K, and Akkan, Z. 2004. Quantitative relationship between the local lymph node assay and human skin sensitisation assays. Reg. Tox. Pharmacol. 39: 245-255.

Stadler J and Karol MH. 1984. Experimental delayed hypersensitivity following inhalation of dicyclohexylmethane-4,4'-diisocyanate: A concentration-response relationship. Toxicol. Appl. Pharmacol., 74, 244-249.

Steiling W, Basketter DA, Berthold K, Butler M, Garrigue J-L, Kimber I, Lea L, Newsome C, Roggeband R, Stropp G, Waterman S, Wieman C. 2001. Skin sensitisation testing - new perspectives and recommendations. Food and Chemical Toxicology 39: 293-301

Stotts J. 1980. Planning, conduct and interpretation of human predictive sensitisation patch tests. In: Current concepts in cutaneous toxicity. Academic Press, p. 41-53.

Tuschl H, Kovac R, Weber E. 2000. The Expression of Surface Markers on Dendritic Cells as Indicators for the Sensitising Potential of Chemicals. Toxicology in vitro, 14, 541-549

Vandebriel RJ, Van Och FMM, Van Loveren H. 2005. In vitro assessment of sensitizing activity of low molecular weight compounds. Toxicol Appl Pharmacol 207, S142-1

Website:
http://www.oecd.org/document/23/0,2340,en_2649_34365_33957015_1_1_1_1,00.htm

Wolfreys A, Basketter DA. 2004. Mutagens and Sensitizers – an unequal relationship? Journal of Toxicology Cutaneous and Ocular Toxicology 23(3): 197-205.

Worth A.P. and Balls M. 2001. The importance of the prediction model in the development and validation of alternative tests. ATLA 29, 135-143

A. Worth, M Balls eds 2002. Alternative (Non-Animal) Methods for Chemical Testing: Current Status and Future Prospects. ATLA Vol. 30 Supplement 1, 49-53

Worth AP, Bassan A, Gallegos A, Netzeva TI, Patlewicz G, Pavan M, Tsakovska I, Vracko M. 2005. The Characterisation of (Quantitative) Structure-Activity Relationships: Preliminary Guidance. ECB Report EUR 21866 EN Ispra, Italy: European Commission, Joint Research Centre.

Zachariae C, Rastogi S, Devantier C, Menne T and Johansen JD. 2003. Methyldibromo glutaronitrile: clinical experience and exposure-based risk assessment. Contact Dermatitis 48: 150-154.

Zinke S, Gerner I, Schlede E. Evaluation of a rule base for identifying contact allergens by using a regulatory database: Comparison of data on chemicals notified in the European Union with "structural alerts" used in the DEREK expert system. Altern Lab Anim. 2002 May-Jun;30(3):285-98.

Zwart A., Arts JHE, Kuper CF. 1994. Wave propagation: a new parameter in the description of mechanical airway impedance. Eur.Respir. Rev. 4, 203-209.

R.7.4 Acute toxicity

R.7.4.1 Introduction

Assessment of the acute toxic potential of a chemical is necessary to determine the adverse health effects that might occur following accidental or deliberate short-term exposure. The nature and severity of the acute toxic effects are dependent upon various factors, such as the mechanism of toxicity and bioavailability of the chemical, the route and duration of exposure and the total amount of chemical to which the person or animal is exposed.

R.7.4.1.1 Definition of acute toxicity

The term *acute toxicity* is used to describe the adverse effects, which may result from a single exposure (i.e. a single exposure or multiple exposures within 24 hours) to a substance. In the context of this guidance, exposure relates to the oral, dermal or inhalation routes. The adverse effects can be seen as clinical signs of toxicity (for animals, refer to OECD Guideline Document 19, 2000), abnormal body weight changes, and/or pathological changes in organs and tissues, which in some cases may result in death. In addition to acute systemic effects, some substances may have the potential to cause local irritation or corrosion of the gastro-intestinal tract, skin or respiratory tract following a single exposure. Acute irritant or corrosive effects due to the direct action of the chemical on the exposed tissue are not specifically covered by this document, although their occurrence may contribute to the acute toxicity of the chemical and must be reported. The endpoints of skin and eye irritation/corrosion and respiratory irritation are addressed in Section R.7.2.

At the cellular level acute toxicity can be related to three main types of toxic effect, (i) general basal cytotoxicity (ii) selective cytotoxicity and (iii) cell-specific function toxicity. Acute toxicity may also result from chemicals interfering with extracellular processes (ECVAM workshop report 16, 1996). Toxicity to the whole organism also depends on the degree of dependence of the whole organism on the specific function affected.

R.7.4.1.2 Objective of the guidance on acute toxicity

A chemical substance may induce systemic and/or local effects. This document is concerned with assessment of systemic effects following acute exposure.

Generally the objectives are to establish:

- whether a single exposure (or multiple exposures within 24 hours) to the substance of interest could be associated with adverse effects on human health; and/or
- what types of toxic effects are induced, their time of onset, duration and severity (all to be related to dose); and/or
- the dose-response relationships to determine the LD_{50}, the LC_{50}, the discriminating dose, or the acute toxic class; and/or
- when possible, the slope of the dose-response curve; and/or
- when possible, whether there are marked sex differences in response to the substance; and
- what information enables the classification and labelling of the substance for acute toxicity

The indices of LD_{50} and LC_{50} are statistically-derived values relating to the dose that is expected to cause death in 50% of treated animals in a given period; these values do not provide information on all aspects of acute toxicity. Indeed, information on lethality is not an essential requirement for the classification decision or risk assessment. Other parameters and observations and their type of dose response may yield valuable information. The potential to avoid acute toxicity testing should be carefully exploited by application of read-across or other non-testing means. Furthermore, there is an overriding obligation to minimize the use of animals in any assessment of acute toxicity.

For risk assessment, further considerations on the nature and reversibility of the toxic effects are necessary.

R.7.4.2 Information requirements for acute toxicity

The standard information requirements for acute toxicity under the REACH Regulations are as follows:

Annex VII (≥1 t/y): acute toxicity *via* the oral route of exposure is required;

Column 2 of Annex VII details specific rules for adaptation of these information requirements, notably allowing for the waiving of acute oral toxicity testing if the substance is corrosive to the skin or if a study on acute toxicity by the inhalation route is available.

Annex VIII -X (≥ 10 t/y): acute toxicity *via* the oral and dermal or inhalation route of exposure.

Column 2 of Annex VIII details specific rules for adaptation, notably requiring information on at least one other route of exposure depending on the nature of the substance and the likely route of human exposure (for details see Annex VIII Section 8.5); as for Annex VII, allowance is made for the waiving of acute oral toxicity testing if the substance is corrosive to the skin.

If there is any reason (alert from existing data) for a concern of acute toxicity at non-corrosive levels, one could point out needs to address this.

R.7.4.3 Information and its sources on acute toxicity

Information on acute toxicity, as detailed below, can be obtained from a variety of sources including unpublished studies, data bases and publications such as books, scientific journals, criteria documents, monographs and other publications (see Chapter R.3 of the *Guidance on IR&CSA* for further general guidance).

R.7.4.3.1 Non-human data on acute toxicity

Non-testing data on acute toxicity

Non-testing data can be provided by the following approaches: a) structure-activity relationships (SARs) and quantitative structure-activity relationships (QSARs), collectively called (Q)SARs; b) expert systems incorporating (Q)SARs and/or expert rules; and c) grouping methods (read-across and categories. These approaches can be used to assess acute toxicity if they provide relevant and reliable (adequate) data for the chemical of interest. Guidance on how to assess the relevance and reliability of non-testing data is provided in the general guidance on (Q)SARs in Section R.6.1 and on grouping approaches in Section R.6.2 of the *Guidance on IR&CSA*. Non-testing methods should be documented according to the appropriate reporting formats (see Sections R.6.1.9 and R.6.2.6). In the case of (Q)SARs and expert systems, a detailed description of available models is provided in the JRC QSAR Model Database (http://qsardb.jrc.it/).

Compared with some endpoints, there are relatively few (Q)SAR models and expert systems capable of predicting acute toxicity. Available approaches have been reviewed in the literature (Cronin *et al*., 1995,2003; Lessigiarska *et al*., 2005; Tsakovska et al., 2006). On the basis of these reviews, the following conclusions can be made: a) the relatively small number of models for *in vivo* toxicity is related to the nature of the endpoint – acute toxicity measurements are usually related to whole body phenomena and are therefore very complex. The complexity of the mechanisms involved leads to difficulties in the QSAR modelling process; b) most QSAR models identify hydrophobicity as a parameter of high importance for the modelled toxicity. In addition, many models indicate the role of the electronic and steric effects; c) most literature-based models are restricted to single classes of chemicals, such as phenols, alcohols, anilines. Models based on more heterogeneous data sets are those incorporated in the expert systems.

In the sections below some examples are given in order to illustrate the potential possibility for applying the (Q)SAR approaches for the acute toxicity endpoint for predictive purposes or to investigate the mechanisms of toxicity.

(Q)SAR models

QSARs on inhalation toxicity

Some simple regression models have been developed for predicting the inhalational toxicity of volatile substances, and these can be used reliably within their domains of applicability. Typically, parameters such as vapour pressure (VP) and boiling point (BP) have been found to be useful predictors of the acute toxic effect (e.g. LC_{50} value). These models are based on the assumption that toxicity occurs by the non-specific mechanism of narcosis, and that the LC_{50} data are based on tests in which a steady-state concentration has been reached in the blood. These models are suitable only for systemic acting volatile compounds.

For example, acute (non-lethal) neurotoxicity data for the neurotropic effects of some common solvents on both rats (whole-body exposures for 4h) and mice (whole-body exposures for 2h), taken from Frantik et al (1996), were subjected to QSAR analysis by Cronin (1996). Stepwise regression analysis of the 4-hr toxicity data causing the 30% depression in response ($\log 1/ECR_{30}$) in rats gave the following equation:

$$\log 1/ECR_{30} = 0.361\ ClogP – 0.117\ {}^{0}\chi - 1.76$$

$$n = 37\quad R^2 = 0.817\quad s = 0.280\quad F = 35.2$$

This relationship demonstrates a partial dependence of neurotoxicity with the octanol-water partition coefficient, logP. The negative correlation with the zero-order molecular connectivity ${}^{0}\chi$ is thought to be an indication that the membrane permeability of blood-brain barrier is reduced for large molecules.

Stepwise regression for mouse neurotoxicity gave the following equation:

$$\log 1/ECM_{30} = 0.212\ ClogP + 0.00767\ BP – 0.176\ {}^{0}\chi - 2.03$$

$$n = 39\quad R^2 = 0.811\quad s = 0.271\quad F = 22.4$$

in which BP is the boiling point of the substance (BP is inversely related to vapour pressure).

The application of principal components analysis (PCA), to separate compounds of high neurotoxicity from those of low neurotoxicity, suggested that in addition to partitioning through a membrane (determined by logP and molecular size), aqueous solubility and volatility are also important factors governing neurotoxicity (Cronin, 1996). Metabolism to more toxic compounds is suggested as a possible cause of compounds appearing as outliers in the QSARs.

QSARs for predicting LD_{50}

There are references in the literature to a few models for predicting LD_{50}, generally for small sets of compounds. For example, Hansch & Kurup (2003) developed the following QSAR to predict the toxicity of barbiturates (LD_{50}) in for female white mice, using toxicity data from Cope and Hancock (1939):

$$\log 1/LD_{50} = -1.44 \log P + 0.16\ NVE - 8.70$$

$$n = 11 \quad R^2 = 0.924 \quad s = 0.077 \quad R^2_{cv} = 0.879$$

where NVE is the number of valence electrons (used as a measure of polarisability).

QSARs for predicting human toxicity

The same descriptors were used to predict the LD_{100} of miscellaneous drugs to humans, using toxicity data from King (1985):

$$\log 1/C = 0.61 \log P + 0.017\ NVE + 1.44$$

$$n = 36 \quad R^2 = 0.850 \quad s = 0.438 \quad R^2cv = 0.817$$

QSARs for predicting *in vitro* effects

A number of QSAR models for predicting *in vitro* effects are cited in the literature (reviewed in Tsakovskaet al., 2006), but these are not directly relevant to the assessment of acute toxicity for regulatory purposes. In general, these models have been developed to investigate the mechanisms of cytotoxic action, and they outline the role of hydrophobicity as well electronic descriptors, including electrotopological state descriptors (Lessigiarska *et al*., 2006), bond dissociation energies (Selassie *et al*., 1999), and dissociation constants (Moridani *et al*., 2003). While these models are not directly relevant to the assessment of acute toxicity, the fact that reliable QSARs can be developed for the *in vitro* cytotoxicity of defined groups of chemicals indicates that the approach of modelling *in vitro* data should be further explored with a view to integrating such QSARs into the ITS for acute toxicity. For example, a battery of QSARs could be developed for predicting the *in vitro* data of a validated *in vitro* test, and then used to supplement or replace *in vivo* testing.

Expert systems

For heterogeneous groups of compounds, expert systems are available in which rule bases express generalised relationships between chemical structure and toxicity. In knowledge-based experts systems (see also Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*), such as HazardExpert, such rules are derived from human expert opinion. In statistically based expert systems, such as TOPKAT and MultiCASE, statistical methods were used to derive (Q)SAR models (see also Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*).

HazardExpert

HazardExpert is a module of Pallas software developed by CompuDrug Limited (http://www.compudrug.com). The program works by searching the query structure for known toxicophores, which are stored in the "Toxic Fragments Knowledge Base" and which include substructures exerting both positive and negative modulator effects. Once a toxicophore has been identified, this triggers estimates for a number of toxicity endpoints, including neurotoxicity. The default knowledge base of the system is based on a US-EPA report (Brink and Walker, 1987) and scientific information collected by CompuDrug Limited. This program can be linked to MetabolExpert, another module of the Pallas software, to predict the toxicity of the parent compound and its metabolites. Information on the validity of the model is not available. Investigations on the validity and applicability of HazardExpert are needed before recommendations can be made about its regulatory use.

TOPKAT

The TOPKAT software package employs cross-validated quantitative structure-toxicity relationship (QSTR) models for assessing various measures of toxicity (http://accelrys.com/products/discovery-studio/toxicology/). The Rat Oral LD_{50} module of TOPKAT includes 19 QSAR regression models for different chemical classes. The models are based on a number of structural, topological and electrophysiological indices, and they make predictions of the oral acute median lethal dose in the rat (LD_{50}).

The TOPKAT rat oral LD_{50} models are based on experimental data from the RTECS. Since RTECS lists the most toxic value when multiple values exist, the TOPKAT model tends to overestimate the toxicity of query structures.

The Rat Inhalation LC_{50} module of TOPKAT contains five submodels related to different chemical classes.

TOPKAT models, including the models for acute oral toxicity, have been used by Danish EPA to evaluate the dangerous properties of around 47 000 organic substances on the EINECS list [17]. An external evaluation of this model using 1840 chemicals not contained in the TOPKAT database gave poor results (R^2 = 0.31). However, 86% of estimations fall within a factor of 10 from test results (DK EPA study).

The Danish EPA concluded that the TOPKAT model is sufficient to give an indication of the least strict classification for acute toxicity, Xn; R22. An Internet version of the Danish QSAR database is accessible from the ECB website (http://qsardb.jrc.it).

MultiCASE

The MultiCASE software (http://www.multicase.com) contains an acute toxicity module, which consists of a rat LD_{50} model based on 7920 compounds from compilations by FDA, NTP and WHO data. Information on the validity of the model is not available. Investigations on the validity and applicability of MultiCASE are needed before recommendations can be made about its regulatory use.

Testing data on acute toxicity

In vitro data

There are currently no *in vitro* tests that have been officially adopted by the EU or OECD for assessment of acute toxicity.

However, a number of *in vitro* tests for acute toxicity are undergoing a validation process:

Two In vitro basal cytotoxicity assays for predicting starting doses for in vivo oral toxicity tests and lethal concentrations in man have undergone peer review by ICCVAM, namely the BALB/c 3T3 NRU & normal human keratinocyte (NHK) NRU assays (http://iccvam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm).

Two in vitro tests pre-validated: TER and PCP in 2 renal cell lines (test battery). The loss of monolayer integrity is often an early indicator of nephrotoxicity in intact renal epithelia in vitro and reflects loss of renal function in vivo. Trans-epithelial resistance (TER), coupled with enhanced paracellular permeability (PCP), is a good measure of this integrity. (Duff et al., 2002). These tests should be used in a WoE approach as alerts or correctors in respect to the basal cytotoxicity assays. Their contribution is under evaluation in A-Cute-Tox (see below).

A ECVAM validated test, the CFU-GM, to predict anticancer agents induced myelotoxicity in humans, is now under evaluation to widen its applicability domain to chemicals' induced toxicity (http://ecvam-dbalm.jrc.ec.europa.eu/). If sufficiently validated and suited to the purpose of assessment of acute toxicity, this could be included in a WoE.

The integrated project A-Cute-Tox (A 5-year 6th FP project initiated in 2005) is addressing the possible replacement of the acute oral systemic toxicity tests (http://www.acutetox.org/). Particular attention should be given in the future to results of the project.

Animal data

Data may be available, particularly for phase-in substances, from a wide variety of animal studies, which give different amounts of direct or indirect information on the acute toxicity of a substance; e.g.:

- OECD TG 420 (EU B.1 bis) Acute oral toxicity – Fixed dose procedure
- OECD TG 423 (EU B.1tris) Acute oral toxicity – Acute toxic class method
- OECD TG 425 Acute oral toxicity – Up-and-down procedure
- OECD 401 (EU B.1) Acute Oral Toxicity (method deleted from the OECD Guidelines for testing of chemicals and from Annex V to Directive 67/548/EEC; see below)
- OECD TG 402 (EU B.3) Acute dermal toxicity
- OECD TG 403 (EU B.2) Acute inhalation toxicity
- Draft OECD TG 433 "Acute Inhalation Toxicity, Fixed Dose Procedure";
- Draft OECD TG 436 "Acute Inhalation Toxicity, Acute Toxic Class Method";
- Draft OECD TG 434 "Acute Dermal Toxicity, Fixed Dose Procedure";
- ICH compliant studies;
- Mechanistic and toxicokinetic studies;
- Studies in non-rodent species.

Traditionally, acute toxicity tests on animals have used mortality as the main observational endpoint, usually in order to determine LD_{50} or LC_{50} values. These values were regarded as key information for hazard assessment and supportive information for risk assessment. However, derivation of a precise LD_{50} or LC_{50} value is no longer considered essential. Indeed, some of the current standard acute toxicity test guidelines, such as the fixed dose procedures (OECD 420, EU B.1 bis and draft OECD 433), use signs of non-lethal toxicity and have animal welfare advantages over the other guidelines.

Existing OECD TG 401 (EU B.1) data would normally be acceptable but testing using this deleted method must no longer be performed.

In addition to current regulatory methods, acute toxicity data on animals may be obtained by conducting a literature search and reviewing all available published and unpublished toxicological or general data, and the official/existing acute toxicological reference values. For more extensive general guidance see Section R.3.1 in Chapter R.3 of the *Guidance on IR&CSA*.

Utilising all the available information from sources such as those above, a *Weight of Evidence* approach should be taken to maximise use of existing data and minimise the commissioning of new testing.

When several data are available, a hierarchal strategy should be used to focus on the most relevant.

R.7.4.3.2 Human data on acute toxicity

Acute toxicity data on humans may be available from:

- Epidemiological data identifying hazardous properties and dose-response relationships;
- Routine data collection, poisons data, adverse event notification schemes, coroner's report;
- Biological monitoring/personal sampling;
- Human kinetic studies – observational clinical studies;
- Published and unpublished industry studies;
- National poisoning centres.

The main obstacles to the use of human data are their limited availability and often limited information on levels of exposure (ECETOC, 2004).

R.7.4.3.3 Exposure considerations for acute toxicity

With regard to acute toxicity, exposure considerations are detailed in column 2 in Annex VIII, but not in Annex XI. If there is only one demonstrated route of exposure, this route must be addressed. Where the potential for human exposure exists, the most likely route, or routes, of exposure should be determined so that the potential for acute toxicity by these routes can be assessed. Determination of the most likely route of exposure will have to take into account not only how the substance is manufactured and handled, including engineering controls that are in place to limit exposure, but also the physico-chemical properties of the substance, for instance, whether the substance is a solid or liquid, the particle size and proportion of respirable and inhalable particles, vapour pressure and log P.

R.7.4.4 Evaluation of available information on acute toxicity

The detailed generic guidance provided in Chapter R.4 of the *Guidance on IR&CSA* on the process of judging and ranking the available data for its adequacy (reliability and relevance), completeness and remaining uncertainty is relevant to information on acute toxicity.

R.7.4.4.1 Non-human data on acute toxicity

Non-testing data on acute toxicity

Physico-chemical properties[62]

It may be possible to infer from the physico-chemical characteristics of a substance whether it is likely to be corrosive or absorbed following exposure by a particular route and, produce acute toxic effects. Physico-chemical properties may be important in the case of the inhalation route (vapour pressure, MMAD, log K_{ow}), determining the technical feasibility of the testing and acting upon the distribution in the airways in particular for *local-acting substances*. Indeed, some physico-chemical properties of the substance or mixture could be the basis for waiving testing. In particular, it should be considered for low volatility substances, which are defined as having vapour pressures **<**1 x 10^{-5} kPa (7.5 x 10^{-5} mmHg) for indoor uses, and **<**1 x 10^{-4} kPa (7.5 x 10^{-4} mmHg) for outdoor uses. Furthermore, inhalable particles are capable of entering the respiratory tract *via* the nose and/or mouth, and are generally smaller than 100 µm in

62 Refer also to Tables R.12-1 to R.12-6 in Section R.7.12 of Chapter R.7c of the *Guidance on IR&CSA*.

diameter. Particles larger than 100 µm are less likely to be inhalable. In that way, particular attention should be driven on results of aerosol particle size determination.

In particular, for substances in powder form, particle size of the material decisively influences the deposition behaviour in the respiratory tract and potential toxic effects. Particle size considerations (determined by e.g. granulometry testing, OECD 110) can be useful for:

- selecting a representative sample for acute inhalation toxicity testing;
- assessing the respirable and inhalable fractions, preferably based on aerodynamic particle size;
- justifying derogations from testing, for instance, when read-cross (or chemical grouping approach) data can be associated with results from particle size distribution analyses (see Section R.6.2 in Chapter R.6 of the *Guidance on IR&CSA*).

Physico-chemical properties are also important for determination of the potential of exposure through the skin, for example, log K_{ow}, molecular weight and volume, molar refraction, degree of hydrogen bonding, melting point (Hostýnek, 1998).

Read-across to structurally or mechanistically similar substances (SAR)

Generic guidance on the application of grouping approaches is provided in Section R.6.2 in Chapter R.6 of the *Guidance on IR&CSA*.

(Q)SAR

Several (Q)SAR systems are available that can be used to make predictions about, for example, dermal penetration or metabolic pathways (see cross-cutting QSAR guidance for list of models). However, these systems have not been extensively validated against appropriate experimental data and it has not been yet verified if the results genuinely reflect the situation *in vivo*. That is why the modelled data can be used for hazard identification and risk assessment purposes only as part of a WoE approach.

The complexity of the acute toxicity endpoint (possibility of multiple mechanisms) is one of the reasons for limited availability and predictivity of QSAR models. In the absence of complete validation information, available models could be used as a part of the WoE approach for hazard identification and risk assessment purposes after precise evaluation of the information derived from the model.

Evaluation of the validity of the method

An evaluation of model validity according to OECD principles should be available, as described in Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*.

Evaluation of the reliability of the individual prediction

The reliability of individual (Q)SAR predictions should be evaluated, as described in Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*.
The evaluations of model validity and estimate reliability should be documented according to the appropriate reporting formats, as described in Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*.
In the case of grouping approaches, adequacy should be assessed and documented according to guidance described in Section R.6.2 in Chapter R.6 of the *Guidance on IR&CSA*.

Testing data on acute toxicity

In vitro data

The *in vitro* tests that are currently available provide supplementary information, which may be used to determine starting doses for *in vivo* studies, assist evaluation of data from animal studies, especially in identification of species differences, or to increase understanding of the toxicological mechanism of action of the substance. They cannot be used to replace testing in animals completely, although this may be possible in the future.

The outcome of the EU-US (ECVAM-ICCVAM) validation study on the Use of In Vitro Basal Cytotoxicity Test Methods For Estimating Starting Doses For Acute Oral Systemic Toxicity (http://iccvam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm) was that the Peer Review Panel agreed that the applicable validation criteria have been adequately addressed for using these *in vitro* test methods in a WoE approach to determine the starting dose for acute oral *in vivo* toxicity protocols. Moreover, on the basis of a preliminary analysis of data, there is the indication that the cytotoxicity tests might be useful in predicting low toxicity substances ($LD_{50} \geq$ 2g/kg body weight) and that they might therefore be used to filter these out in the future. This application needs to be validated with a wider range of compounds.

In vitro data may be useful for predicting acute toxicity in humans providing that the domain of applicability for the test method is appropriate for the class of chemical under evaluation and a range of test concentrations have been investigated that permit calculation of an IC_{50} (inhibitory concentration 50%) value. Indeed, on the basis of a preliminary comparison of data, there is the indication that the results of *in vitro* cytotoxicity tests may be more predictive of acute oral toxicity in humans than rat or mouse data. This aspect needs to be further investigated.

Generic guidance is given in Chapter R.4 of the *Guidance on IR&CSA* for judging the applicability and validity of the outcome of various study methods, assessing the quality of the conduct of a study (including how to establish whether the substance falls within the applicability domain of the method and the validation status for the given domain) and aspects such as vehicle, number of duplicates, exposure/ incubation time, GLP-compliance or comparable quality description.

Animal data

Acute toxicity tests on animals have primarily used mortality as the main observational endpoint, usually in order to determine LD_{50} or LC_{50} values, although some of the current standard protocols, such as the fixed dose procedure (OECD TG 420, EU B.1 bis), use evident signs of toxicity in place of mortality. In many cases, there will be little information on the cause of death or mechanism underlying the toxicity, and only limited information on pathological changes in specific tissues or clinical signs, such as behavioural or activity changes.

Many acute toxicity studies on chemicals of low toxicity are performed as limit tests. For more harmful chemicals choice of optimum starting dose will minimize use of animals. When multiple dose levels are assessed, characterisation of the dose-response relationship may be possible and signs of toxicity identified at lower dose levels may be useful in estimating LOAELs or NOAELs for acute toxicity. For local acting substances, mortality after inhalation may occur due to tissue damage in the respiratory tract. In these cases, the severity of local effects may be related to the dose or concentration level and therefore, it might be possible to identify a LOAEL or NOAEL. For systemic toxicity, there could be some evidence of target organ toxicity (pathological findings have to be documented) or signs of toxicity based on clinical observations.

Whichever approach is used in determining acute toxicity critical information needs to be derived from the data to be used in risk assessment. It is important to identify those dose levels which produce signs of toxicity, the relationship of the severity of these with dose and the level at which toxicity is not observed (i.e. the acute NOAEL).

In addition to current available OECD or EU test methods (see Section R.7.4.3), alternative *in vivo* test methods for assessment of acute dermal and inhalation toxicity are in the process for adoption and use for regulatory purposes. Whichever test is used to evaluate acute toxicity on animals, the evaluation of studies takes into account the reliability based on the approach of Klimisch *et al*. (1997) (standardised methods, GLP, detailed description of the publication), the relevance, and the adequacy of the data for the purposes of evaluating the given hazard from acute exposure (for more guidance see Section R.4.2 in Chapter R.4 of the *Guidance on IR&CSA*). The best studies are those that give a precise description of the nature and reversibility of the toxic effect, the number of subjects, gender, the number of animals affected by the observed effects and the exposure conditions (atmosphere generation for inhalation, duration and concentration or dose). The relevance of the data should be determined in describing the lethal or non-lethal endpoint being measured or estimated.

In addition, when several studies results are available for one substance, the most relevant one should be selected; data from others studies that have been evaluated should be considered as supportive data for the full evaluation of the substance.

The classification criteria for acute inhalation toxicity relate to a 4-hour experimental exposure period. If data for a 4-hour period are not available then extrapolation of the results to 4 hours are often achieved using Haber's Law (C.t = k). However, there are limits to the validity of such extrapolations, and it is recommended that the Haber's Law approach should not be applied to experimental exposure durations of less than 30 minutes or greater than 8 hours in order to determine the 4-hour LC_{50} for C&L purposes.

Nowadays a modification of Haber's Law is used (C^n.t = k) as for many substances it has been shown that n is not equal to 1 (Haber's Law). In case extrapolation of exposure duration is required, the *n* value should be considered. If this *n* value is not available from literature, a default value may be used. It is recommended to set n = 3 for extrapolation to shorter duration than the duration for which the LC_{50} or EC_{50} was observed and to set n = 1 for extrapolation to longer duration (ACUTEX TGD, 2006), also taking the range of approximately 30 minutes to 8 hours into account.

Experimentally, when concentration-response data are needed for specific purposes, OECD TG 403 (EU B.2) or the CxT approach could be taken into consideration. The OECD TG 403/(EU B.2 will result in a concentration-response curve at a single exposure duration, the CxT approach will result in a concentration-time-response curve, taking different exposure durations into account. The CxT approach (under consideration for the revision of OECD TG 403) uses two animals per CxT combination and exposure durations may vary from about 15 minutes up to approximately 6 hours. This approach may provide detailed information on the concentration-time-response relationship in particular useful for risk assessment and determination of NOAEL/LOAEL.

R.7.4.4.2 Human data on acute toxicity

When available, epidemiological studies, case reports, information from medical surveillance or volunteer studies may be crucial for acute toxicity and can provide evidence of effects that are undetectable in animal studies (e.g. symptoms like nausea or headache). Nevertheless, the conduct of human studies is not recommended.

Such data could also be useful to identify particular sensitive sub-populations like new born, children, patients with diseases (in particular with chronic respiratory, e. g. asthma, BPOC).

Additional guidance should be provided on the reliability and the relevance of human studies because there are no standardised guidelines for such studies (except for odour threshold determination) and these are not usually conducted according to GLP. Such guidance is provided in Section R.4.3.3 in Chapter R.4 of the *Guidance on IR&CSA*.

R.7.4.4.3 Exposure considerations on acute toxicity

Particular attention should be addressed to the potential routes of exposure in humans to select the appropriate testing strategy.

Generic aspects of data waivers based on exposure considerations are presented in Section R.5.1 in Chapter R.5 of the *Guidance on IR&CSA*. Information on the role of exposure information in the testing strategies for acute toxicity is presented in Section R.7.4.6.

R.7.4.4.4 Remaining uncertainty on acute toxicity

In most cases, remaining uncertainties will exist due to the absence of valid human acute toxicity data, and so appropriate assessment factors should be applied. Toxicokinetic data could help in deriving chemical-specific interspecies assessment factors. As acute toxicity testing does not usually include clinical chemistry, haematology and detailed histopathology and functional observations, an additional assessment factor may need to be applied when a NOAEL or LOAEL from these studies is used to derive DNELs (for more guidance on the setting of DNELs for acute toxicity, see Chapter R.8 of the *Guidance on IR&CSA*, Appendix R.8-8).

R.7.4.5 Conclusions on acute toxicity

R.7.4.5.1 Concluding on suitability for Classification and Labelling

In order to achieve classification and labelling, Annex VI of the Dangerous Substances Directive 67/548/EEC[63] must be applied. The criteria for classification are based on specific 'cut offs' based on the LD_{50} or LC_{50}, although determination of a precise LD_{50} or LC_{50} value is not essential for classification purposes. This is because the LD_{50}/LC_{50} is not an absolute value (Schütz, 1969) since many factors influence its reproducibility (Zbinden and Flury-Roversi, 1981).

Ideally, classification and labelling should be achieved using data generated from studies conducted in accordance with officially adopted OECD test guidelines, or test methods incorporated for the time being into Annex V of Directive 67/548/EEC[64]. Such studies will permit identification of the LD_{50}, LC_{50}, the discriminating dose (fixed dose procedures), or a range of exposure where lethality and/or severe toxicity is expected (acute toxic class methods). For materials of low toxicity (no mortalities expected at the upper dose limit) testing is restricted to this dose level (the limit test) and if absence of mortalities is confirmed, classification of the substance with respect to acute toxicity is not required.

In the Up-and-Down Procedure (OECD TG 425), where individual animals are dosed sequentially, estimation of the LD_{50} with a confidence interval is possible and this can be used for classification purposes. Data generated in the fixed dose/concentration procedures (OECD TG 420, draft 433 and 434 and EU B.1 bis) and the acute toxic class methods (OECD TG 423,

---

[63] Directive 67/548/EEC will be repealed and replaced with the EU Regulation on classification, labelling and packaging of substances and mixtures, implementing the Globally Harmonized System (GHS).

[64] The new Test Methods Regulation is currently (February 2008) under adoption and contains all the test methods previously included in Annex V to Directive 67/548/EEC.

draft TG 436 and EU B.1 tris) are equally sufficient for classification purposes. In the fixed dose/concentration procedures, the discriminating dose is identified as the dose causing evident toxicity but not mortality, and must be one of the four dose levels specified in the test method. Evident toxicity is a general term describing clear signs of toxicity such that at the next highest dose level, either severe pain and enduring signs of severe distress, moribund status or probable mortality can be expected in most animals. In the acute toxic class methods, the range of exposure where death is expected is determined by testing at one or more of the four fixed doses. The OECD and EU guidelines for fixed dose procedure and acute toxic class methods include flow charts that allow conclusions to be drawn with respect to GHS classification. In addition the flow charts in the acute toxic class methods allow identification of LD_{50} or LC_{50} cut offs. In the absence of GLP compliant data generated in accordance with OECD or EU methods, all other available information should be considered. Each individual set of data (e.g. a non-GLP study) must be assessed for reliability and relevance as stated in Section R.7.4.4 and any unsuitable data (i.e. that considered unreliable or not relevant) should be disregarded. When experimental data for acute toxicity are available in several animal species, scientific judgement should be used in selecting the most relevant data from among the valid, well-performed tests. When equally reliable data from several species are available, priority should be given to the data relating to the most sensitive species, unless there are reasons to believe that this species is not an appropriate model for humans. If definitive classification and labelling cannot be achieved from any individual source, but multiple sets of data all lead to the same conclusion, then, the WoE approach might be sufficient to classify and a robust proposal detailing this should be put forward.

Where evidence is available from both humans and animals and there is a conflict between the findings, the quality and reliability of the evidence from both sources shall be evaluated in order to resolve the question of classification. Generally, data of good quality and reliability in humans shall have precedence over other data. However, well designed and conducted epidemiological studies may lack the sufficient number of subjects to detect relatively rare, but nevertheless important, effects. Also, the interpretation of many studies is hampered by difficulties in identifying and taking account of confounding factors. Positive results from well-conducted animal studies are not necessarily negated by the lack of positive human experience but require an assessment of the robustness and quality of both the human and animal data.

If the existing data are contradictory, not concordant or insufficient to reliably determine the appropriate classification and labelling of the substance, additional *in vitro* studies, QSARs, read-across should be considered before conducting any OECD or EU compliant *in vivo* study. In that way *in vitro* data could have a supporting role in a read-across or chemical grouping approach. Study data, which permit an assessment of dose response relationship, should be considered for risk assessment and classification and labelling.

Of particular importance in classifying for inhalation toxicity is the use of well-articulated values in the high toxicity categories for dusts and mists. Inhaled particles between 1 and 4 microns mean mass aerodynamic diameter (MMAD) will deposit in all regions of the rat respiratory tract. This particle size range corresponds to a maximum dose of about 2 mg/L (draft OECD GD 39). In order to achieve applicability of animal experiments to human exposure, dusts and mists would ideally be tested in this range in rats. The cut off values in the table for dusts and mists allow clear distinctions to be made for materials with a wide range of toxicities measured under varying test conditions.

Currently, non-animal test data (e.g. *in vitro*, QSARs and read-across data) cannot be used as stand-alone for classification and labelling purposes, but can be used for classification to support a read-across argument. In future they might be used in different purposes when such methods have been formally validated and incorporated into official test guidelines, and when classification systems have been adapted to take account of such data.

R.7.4.5.2 Concluding on suitability for Chemical Safety Assessment

For chemical safety assessment, both standard OECD/EU test guideline data and all applicable data considered both reliable and relevant should be used. A quantitative rather than qualitative assessment is preferred to conclude on the risk posed by a substance with regards to acute toxicity dependent on the data available and the potential exposure to the substance during the use pattern/lifecycle of the substance. If quantitative data are not available, the nature and the severity of the specific acute toxic effects can be used to make specific recommendations with respect to handling and use of the substance.

Information on acute toxicity is not normally limited to availability of a LD_{50} or LC_{50} value. Additional information which is important for the chemical safety assessment will be both qualitative and quantitative and will include parameters such as the nature and severity of the clinical signs of toxicity, local irritant effects, the time of onset and reversibility of the toxic effects, the occurrence of delayed signs of toxicity, body weight effects dose response relationships (the slope of the dose response curve), sex-related effects, specific organs and tissues affected, the highest non-toxic and lowest lethal dose (adapted from ECETOC Monograph No. 6, 1985).

If a NOAEL can be identified this can be used in determination of a DNEL. However, depending upon the nature of the acute toxicity information available, this may not always be possible. For instance, data from an OECD/EU test method may permit calculation of an LD_{50}/LC_{50} value, or identification of the range of exposure where lethality is expected, or the dose at which evident toxicity is observed, but may not provide information on the dose level at which no adverse effects on health are observed. If the data permits construction of a dose-response curve, then derivation of the NOAEL may be possible. When a limit test has been conducted, and no adverse effects on health have been observed, then the limit dose can be regarded as the NOAEL. If adverse effects on health are seen at the limit dose then it is unlikely that lower dose levels will have been investigated and in this case identification of the NOAEL will not be possible. If data is available for several species, then the most sensitive species should be chosen for the purposes of the Chemical Safety Assessment, provided it is the most relevant to humans.

If human data on acute toxicity is available, it is unlikely that this will be derived from carefully controlled studies or from a significant number of individuals. In this situation, it may not be appropriate to determine a DNEL from this data alone, but the information should certainly be considered in the WoE and may be used to confirm the validity of animal data. In addition, human data should be used in the risk assessment process to be able to determine DNEL for particular sensitive sub-populations like new-born, children or those in poor health (patients).

More extensive guidance on the setting of DNELs for acute toxicity, see Chapter R.8 of the *Guidance on IR&CSA*, Appendix R.8-8.

The anticipated effects from physico-chemical properties and bioavailability data on the acute toxicity profile of the substance must also be considered in the Chemical Safety Assessment.

R.7.4.5.3 Information not adequate

A WoE approach, comparing available adequate information with the tonnage-triggered information requirements by REACH, may result in the conclusion that the requirements are not fulfilled.

In absence of data from test guidelines or equivalent methods, data from other endpoints could be helpful for the determination of acute toxicity potential. For example, data could be provided by subchronic toxicity or neurotoxicity studies, as in general the design of these studies includes a pilot study to determine dose of departure for the main test. In order to proceed with further information gathering the following testing strategy can be adopted.

R.7.4.6 Integrated Testing Strategy (ITS) for acute toxicity

R.7.4.6.1 Objective / General principles

The main objective of this Integrated Testing Strategy (ITS) is to provide advice on how the REACH Annex VII and VIII information requirements for acute toxicity can be met using the most humane methods. If the ITS is followed, the information generated will be sufficient to make a classification decision with respect to acute toxicity hazard and may provide data for the risk assessment and DNEL derivation. In addition, assessment of acute toxicity may provide information that is valuable for the conduct of repeated dose toxicity studies, such as identification of target organ toxicity and dose selection.

By adhering to the criteria outlined in the previous chapters, informed decisions may be made on whether sufficient data already exist to cover the objectives, or whether further testing is required.

If further testing is deemed necessary, the use of the most appropriate study in accordance with the REACH proposal is considered rather than a *one study fits all* approach. An overarching principle is that all data requirements are met in the most efficient and humane manner so that animal usage and costs are minimized.

R.7.4.6.2 Preliminary considerations

The standard information requirements for acute toxicity under the REACH regulations are given in Section R.7.4.2.

According to REACH, acute toxicity studies should not be conducted if a substance is known to be corrosive. However, if there are health concerns regarding exposure to non-corrosive concentrations, then acute toxicity assessment may be considered appropriate. In such cases, a specific protocol should be developed as standard LC_{50} or any other *in vivo* acute toxicity testing cannot be performed. For example, *in vitro* data on basal cytotoxicity could be used to establish the most appropriate range of concentrations to be tested.

Regardless of tonnage level, before any testing is triggered, careful consideration of existing toxicological data, exposure characteristics and current risk management procedures is recommended to ascertain whether the fundamental objectives of the ITS have already been met. This consideration should take account of discussions that have taken place under other regulatory schemes, such as ESR, DPD, BPD and the EU hazard classification scheme. If it is concluded that further testing is required, then a series of decision points are defined to help shape the scope of an appropriate testing program.

The following four-stage process has been developed for clear decision-making:

Stage 1. gather existing information according to Annex VI

Stage 2. consider information needs according to the relevant Annex VII to X

Stage 3. identify data gaps (and adequacy of all available data for classification and labelling and/or risk assessment, or to fulfil the criteria for waiving)

Stage 4. generate new data / propose testing strategy

R.7.4.6.3 Testing strategy for acute toxicity (see Figure R.7.4–1)

Stage 1. Gathering of existing information

The starting point of the ITS is the review of existing data (e.g. human or animal data, physico-chemical properties, (Q)SARs, *in vitro* test data). For non-corrosive substances, the results of skin and eye irritation and skin sensitisation studies (Annex VII) may provide useful information on the potential for systemic toxicity.

In the ITS, all existing human and test data (e.g. from clinical reports, poisoning cases, animal studies, corrosivity, physico-chemical properties) should be considered. Some information from the existing data e.g. *in vitro* studies (*de novo in vitro* basal cytotoxicity and dermal penetration studies), systemic effects observed in other studies, route of human exposure, physico-chemical properties, dermal or respiratory toxicity of structurally-related substances, might primarily be used for the selection of either an acute *in vivo* inhalation test or an acute *in vivo* dermal test. No specific reference is made to valid (Q)SAR models/approaches or to valid *in vitro* methods, but such data should be assessed when available or generated.

Section R.7.4.3 presents a detailed discussion of the sources that may provide relevant information for the assessment of acute toxicity.

Stage 2. Considerations on information needs

A detailed evaluation of the existing information collated in Stage 1 is conducted to allow an informed decision on the testing needs to fulfil the REACH requirements. It is important to ensure that the available data are relevant and reliable to fulfil these requirements.

It should be noted that if a substance is predicted to be corrosive then further consideration should be given as to whether or not an acute oral test can be justified (in particular in relation with animal welfare considerations). Justifications for conducting a study must be provided in order to minimise the animal use. If the substance is considered likely to be corrosive, no acute toxicity testing should normally be conducted (see above). Where information on corrosivity is not available then *in vitro* corrosivity tests should be conducted.

The standard information requirements for acute toxicity under the REACH regulations are given in Section R.7.4.2.

When acute toxicity *via* a second route is required, the choice of the second route (dermal or inhalation) depends on the nature of the substance and the likely route of human exposure. However, information on only one route of exposure may be sufficient and justified (based on physico-chemical, toxicokinetic or human data and review of all possible exposure scenarios; for example with gases only inhalation route could be evaluated as no relevant human exposure may occur by oral or dermal route; for liquid with high viscosity, no testing by inhalation route should be conducted).

If human exposure is possible *via* inhalation, or if physico-chemical properties indicate that such exposure may occur, then testing *via* this route for acute toxicity should be conducted. Data from skin/eye irritation, skin sensitisation and acute oral toxicity should be used as indicators to help testing *via* inhalation (for example, substance with only potential local toxicity; choice of exposure concentrations). If no systemic effects are shown during acute oral testing, then the requirement to conduct inhalation testing should be considered on a case-by-case basis.

Consideration of the need for assessment of acute dermal toxicity should be given if the inhalation route is not considered appropriate. In some cases, it may be possible to draw conclusions about the potential for acute dermal toxicity without further testing, on the basis of the data available from acute oral toxicity and/or dermal absorption studies. Evidence for the potential of high dermal absorption should be considered on a case-by-case basis taking into account physico-chemical properties e.g. Log Kow, water solubility, molecular weight and melting point of the substance. Testing for acute dermal toxicity is indicated if:

Systemic toxicity is observed in skin/eye irritation and/or skin sensitisation studies;

Death is observed in an acute oral toxicity test and there is potential for dermal absorption;

Systemic toxicity is observed in an acute oral toxicity test and there is potential for high dermal absorption (determined following e.g. OECD TG 428, EU B.45)

There is the potential for high dermal exposure (case-by-case basis)

Stage 3. Identification of data gaps / adequacy of data

The purpose of this step is to identify what additional information is required in order to classify the substance and to perform a risk assessment.

The available information may include data generated using study protocols that differ from the standard regulatory tests. The evaluation should include whether the available information meets or exceeds the data requirements from standard regulatory study protocols. Therefore it may be possible that the tonnage-driven minimum needs can be met through combined data obtained from several sources.

At this stage, it is also necessary to verify if the available information is adequate for hazard characterisation. For this process, all relevant information should be taken account of in a weight of evidence assessment. Quantitative data on the dose response relationship for the critical toxicological effects and/or estimations of the either the LC_{50}/LD_{50} values or the Discriminating Dose will be important for assessing the hazard classification and can be used in the risk assessment. Information from testing for other toxicological endpoints (e.g. repeated dose toxicity) may also be useful for the risk assessment (see also Chapter R.8 of the *Guidance on IR&CSA*, Appendix R.8-8). Mathematical modelling should be considered for estimating a threshold exposure level (e.g. benchmark dose), as an alternative to generating additional *in vivo* data.

For the inhalation route, standard protocols involve a 4-hour exposure. If data for other time periods are available (e.g. for 0.5 to 8 hours), extrapolation to a 4-hour exposure period can be achieved using a modification of Haber's Law ($C^n.t = k$). If this «n» value is not available from the literature, a default value may be used; it is recommended to set n = 3 for extrapolation to shorter duration than the duration for which the LC_{50} or EC_{50} was observed and to set n = 1 for extrapolation to longer duration (ACUTEX TGD, 2006). Experimentally, the value of n can be determined using the CxT approach (draft revision OECD TG 403).

If the data and subsequent decisions are deemed consistent with an adequate hazard characterisation and are sufficient to classify the substance or to conduct a risk assessment, then no further testing for acute toxicity is recommended.

In some cases, the substance may be excluded from acute toxicity testing if it does not appear as scientifically necessary (Annex XI). This might be the case for example if

A WoE analysis demonstrates that the available information is sufficient for an adequate hazard characterisation and the exposure to the substance is adequately controlled;

The substance is not bio-available *via* a specific route and possible local effects are adequately characterised (example, no dermal absorption for dermal route)

For inhalation route, no testing is required if it is not technically possible to generate a testing atmosphere, the vapour pressure is very low (<0.1 Pa at 20°C) or the particle size is > 100 μm

Finally, the conclusion that no further testing is required may be reached when the data meet the requirements for classification for toxic effects or if the substance has already been classified for acute toxic effects.

Where evidence is available from both humans and animals and there is a conflict between the findings, the evidence should be evaluated towards understanding the toxicological basis for these divergent findings. Issues relating to the quality and reliability of the data should also be taken into account. Generally, data of good quality and reliability in humans shall take precedence over other data. However, well-designed and conducted epidemiological studies may lack a sufficient number of subjects to detect relatively rare but still significant effects, to assess potentially confounding factors. Positive results from well-conducted animal studies are not necessarily negated by the lack of positive human experience but require an assessment of the robustness and quality of both the human and animal data.

If the remaining data are contradictory, not concordant or insufficient to determine reliably the appropriate classification and labelling of the substance, additional *in vitro* studies, QSARs, read-across should be considered before conducting any OECD compliant *in vivo* study. Study data, which permit an assessment of dose response relationship, should be considered particularly valuable for risk assessment purposes.

Stage 4. Generation of new data / proposal for testing strategy

If sufficient data for risk assessment and classification purposes are already available, no further testing will be required. If data gaps need to be filled, new data shall be generated (Annexes VII & VIII). Due to animal welfare considerations, new tests on animals should only be performed as a last resort when all other sources of information have been exhausted.

The standard OECD guidelines should normally be used as these provide the necessary information on acute toxicity hazard in a way that balances the need to protect human health with animal welfare concerns (see Section R.7.4.3 and the above guidance for Stage 3).

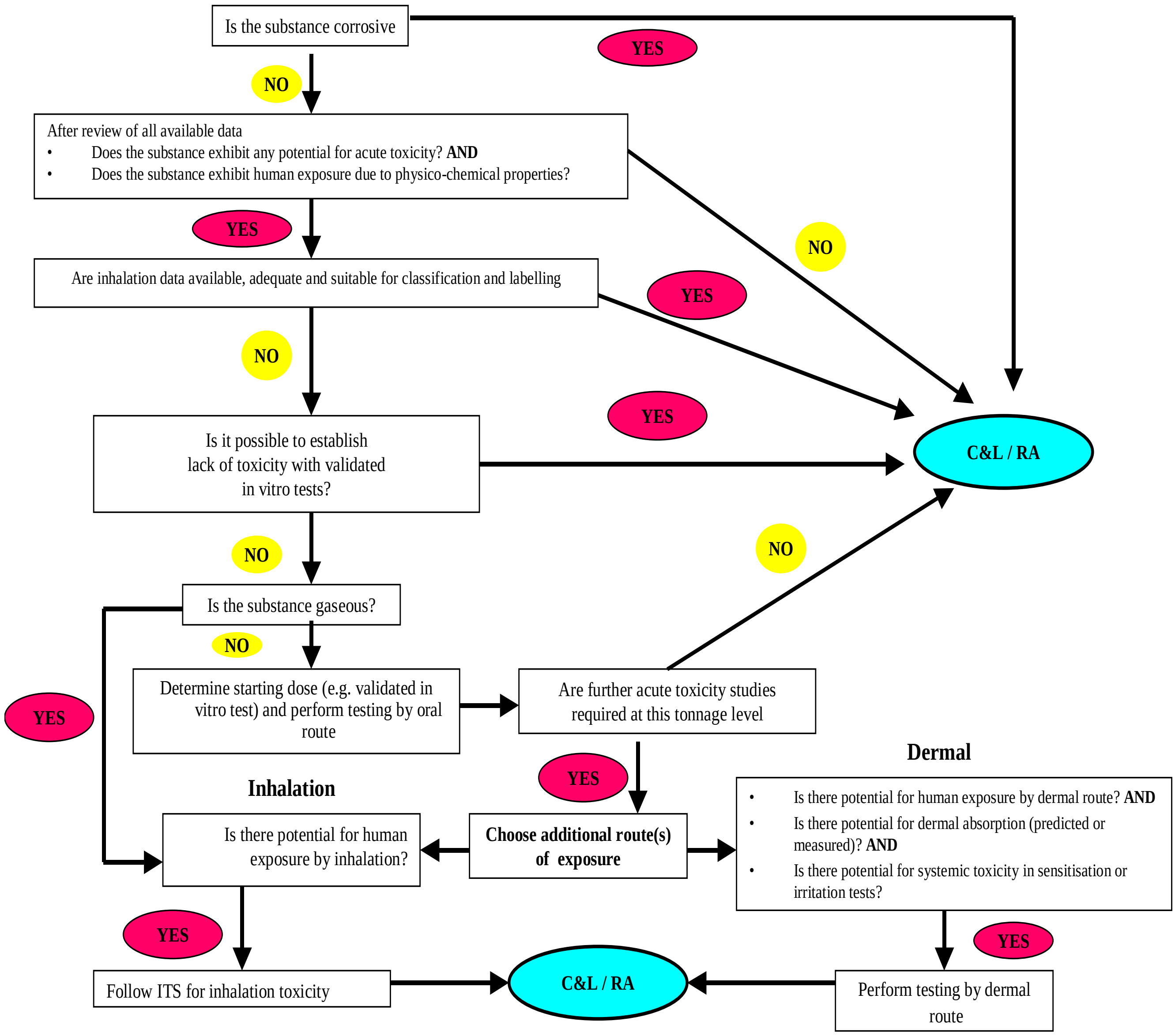
The route of exposure to be used for acute toxicity evaluation depends on the nature of the substance (e.g. gas or not, molecular weight, log K_{ow}) and should reflect the most likely route of human exposure. If any specific human exposure may be identified, further testing for risk assessment should be considered as proposed in Annex VIII. If any human exposure by inhalation is identified, then the testing strategy by inhalation should be proposed (Figure R.7.4–2).

First considerations should be based on defining the potential of the substance for acute toxicity. For such a question, information may be provided by existing data from SARs, QSARs, chemical categories approaches and available *in vitro* and *in vivo* data. If no potential for toxicity is shown, then no further testing is required and a decision on classification can be taken. Such information may also provide relevant information in risk assessment considerations.

Following the general testing strategy, dose selection appears to be an important aspect in order to select the most appropriate starting point. When validated *in vitro* tests are available, as shown by the joint ECVAM-ICCVAM study, these may provide relevant results, and help the dose selection for oral route testing (see Section R.7.4.4.1).

For substances in the ≥10 t/y tonnage band, testing by the dermal route should be considered if a human exposure is identified, or if results from physico-chemical properties and in particular skin irritation/sensitisation tests show any dermal absorption or any systemic toxicity. Depending on such information, dermal testing should be conducted or not following standard protocols (see Section R.7.4.3).

Figure R.7.4–1 ITS for acute toxicity endpoint

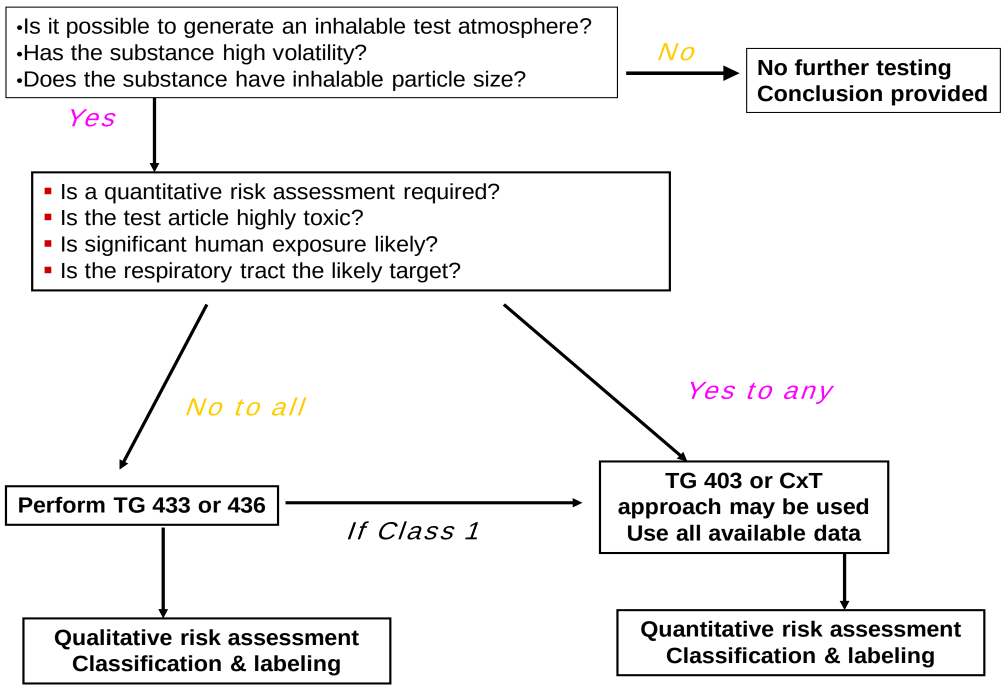


(*) if the substance is corrosive but there are health concerns regarding exposure to non-corrosive concentrations, then acute toxicity assessment may be considered appropriate
(**) Testing by inhalation may be required if the substance is a gas, a liquid or a solid with a high vapour pressure, or a solid with inhalable particle size (particular substances in powder form nanoparticles, fibres…)

A specific testing strategy (Figure R.7.4–2) is proposed for the inhalation route. Primary considerations should be based on the in(ability) to generate a suitable atmosphere depending on the physico-chemical properties (for example, low volatility, solid, particle size >100 µm (see also Section R.7.4.4.1). In this situation, no human exposure may be identified and no further testing is required.

Wherever possible, assessment of acute inhalation toxicity should be conducted in accordance with OECD TG's 433 and 436 (official adoption in process) since they have been designed to use less animals than OECD TG 403 and EU B.2. In addition, OECD TG 433 does not require mortality as endpoint. However, in some circumstances, i.e. if a dose response curve is needed for risk assessment purposes, testing according to OECD TG 403, EU B.2 or the CxT approach may be considered appropriate (see also draft OECD Guidance Document 39).

Figure R.7.4–2 ITS for acute inhalation toxicity endpoint (see also draft OECD GD 39)



R.7.4.7 References on acute toxicity

ACUTEX project. 2006. Technical guidance document (TGD) methodology to develop AETLs. (www.acutex.info)

Brink RH, Walker JD (1987). EPA TSCA ITC Interim Report, Dynamic Corporation, Rockville, 6/4/1987

Cope AC, Hancock EM (1939). Substituted Vinyl Barbituric Acids III. Derivatives Containing a Danish EPA (2001). Report on the Advisory list for self-classification of dangerous substances, Environmental Project no. 636, 2001. http://www.mst.dk/udgiv/publications/2001/87-7944-694-9/html/default_eng.htm

Dialkylvinyl Group having Five or More Carbon Atoms. J Am Chem Soc 61, 776-779

Cronin MTD, Dearden JC (1995). QSAR in toxicology. 2. Prediction of acute mammalian toxicity and interspecies correlations. Quant Struct-Act Relat 14, 117-120.

Cronin MTD (1996). Quantitative Structure-Activity Relationship (QSAR) Analysis of the Acute Sublethal Neurotoxicity of Solvents. Toxicology in vitro 10, 103-110

Cronin MTD, Dearden JC, Walker JD, Worth, AP (2003). Quantitative structure-activity relationships for human health effects: commonalities with other endpoints. Environ Toxicol Chem 22, 1829-1843

Duff T, Carter S, Feldman G, McEwan G, Pfaller W, Rhodes P, Ryan M, Hawksworth G. Transepithelial resistance and inulin permeability as endpoints in vitro nephrotoxicity testing. Altern Lab Anim. 2002 Dec;30 Suppl 2:53-9

ECETOC. 1985. Acute toxicity Tests, LD_{50} (LC_{50}) Determinations and Alternatives, Monograph No.6

ECETOC. 2004. Workshop on the use of human data in risk assessment; workshop report n°3; 23-24 February 2004, Cardiff.

ECVAM workshop report 16, 1996

European Commission. 2005. EU Technical Guidance Document in Support of Commission Directive 93/67/EEC on Risk Assessment for new notified substances and commission regulation (EC) No 1488/94 on Risk Assessment for existing substances. http://ecb.jrc.it

European Commission. 2005. European Technical Guidance Document on Risk Assessment, Human Risk Characterisation (revised chapter, final draft November 2005).

Frantik, E, Hornychova M, Horvath M (1994). Relative acute neurotoxicity of solvents: isoeffective air concentrations of 48 compounds evaluated in rats and mice. Environ Res 66, 173-185.8.

Hansch C, Kurup (2003). A. QSAR of chemical polarisability and nerve toxicity.hit3 2. J Chem Inf Comp Sci, 43, 1647-1651

Hostýnek,JJ (1998). Structure-Activity Relationships in Percutaneous Absorption. In: Dermatotoxicology Methods The Laboratory Worker's *Vade Mecum.* Taylor & Francis.

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), National Toxicology Program (NTP), Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Public Health Service, Department of Health and Human Services (2006). Peer Review Panel Report: The Use of In Vitro Basal Cytotoxicity Test Methods for Estimating Starting Doses for Acute Oral Systemic Toxicity Testing, http://iccvam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm

King LA (1985). Fergusons principle and the prediction of fatal drug levels in blood. Human Toxicol 4, 273-2

Klimisch HJ, Andreae M, Tillmann U. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regulatory Toxicology & Pharmacology* 25:1-5.

Lessigiarska I, Worth AP, Netzeva TI. (2005). Comparative review of QSARs for acute toxicity. EUR 21559 EN

Lessigiarska I, Worth AP, Netzeva TI, Dearden JC, Cronin MTD (2006). Quantitative structure-activity-activity and quantitative structure-activity investigations of human and rodent toxicity. Chemosphere, in press

Moridani MY, Siraki A, O'Brien PJ (2003). Quantitative structure toxicity relationships for phenols in isolated rat hepatocytes. Chem Biol Interact 145, 213-223

OECD. Organisation for Economic Co-operation and Development. Guidelines for the Testing of Chemicals 402, 403, 420,423,425, 433,434, 436.

OECD. Organization for Economic Co-operation and Development. Guidance Document n°19 on the recognition, assessment, and use of clinical signs as humane health endpoints for experimental animals used in safety evaluation. November 2000.

OECD. Organization for Economic Co-operation and Development. Draft Guidance Document n°39 on acute inhalation toxicity testing.

Schütz, E. (1969). On acute oral toxicity tests. Amer. Perf. Cosmet., 84,41

Selassie CD, Shusterman AJ, Kapur S, Verma RP, Zhang L, Hansch C (1999). On the Toxicity of Phenols to Fast Growing Cells. A QSAR Model for a Radical-Based Toxicity. Perkin Transactions 2, 2729-2733

Tsakovska I, Lessigiarska I, Netzeva T, Worth A. (2006). Review of QSARs for mammalian toxicity, 22486 EN

UNECE. 2005. Globally Harmonised System of Classification and Labelling of Chemicals (GHS). http://www.unece.org/trans/danger/publi/ghs/ghs_rev01/01files_e.html

Zbinden, G. and Flury-Roversi, M. (1981). Significance of the LD50 test for the toxicological evaluation of chemical substances. Arch. Toxicol., 47, 77.

R.7.5 Repeated dose toxicity

R.7.5.1 Introduction

Repeated dose toxicity studies provide information on possible adverse general toxicological effects likely to arise from repeated exposure to a substance. Furthermore, these studies may provide information on e.g. reproductive toxicity and carcinogenicity, even though they are not specifically designed to investigate these endpoints.

Organs and tissues investigated in repeated dose toxicity studies include vital organs such as heart, brain, liver, kidneys, pancreas, spleen, immune system, lungs etc. Effects examined may include changes in morphology, physiology, growth or life span, behaviour which result in impairment of functional capacity or impairment of capacity to compensate for additional stress or increase in the susceptibility to the harmful effects of other environmental influences. Therefore, it is important that the possible adverse general toxicological effects are assessed for chemical substances that may be present in the environment.

R.7.5.1.1 Definition of repeated dose toxicity

The term *repeated dose toxicity* comprises the general toxicological effects occurring as a result of repeated daily dosing with, or exposure to, a substance for a part of the expected lifespan (sub-acute or sub-chronic exposure) or for the major part of the lifespan, in case of chronic exposure.

The term *general toxicological effects* (in this report often referred to as *general toxicity*) includes effects on, e.g. body weight and/or body weight gain, absolute and/or relative organ and tissue weights, alterations in clinical chemistry, urinalysis and/or haematological parameters, functional disturbances in the nervous system as well as in organs and tissues in general, and pathological alterations in organs and tissues as examined macroscopically and microscopically. Repeated dose toxicity studies may also examine parameters, which have the potential to identify specific manifestations of toxicity such as e.g., neurotoxicity, immunotoxicity, endocrine-mediated effects, reproductive toxicity and carcinogenicity.

An *adverse effect* is a change in the morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub) population that results in an impairment of functional capacity, or an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences (OECD, 2003).

A chemical substance may induce systemic and/or local effects.

- A *local effect* is an effect that is observed at the site of first contact, caused irrespective of whether a substance is systemically available.
- A *systemic effect* is defined as an effect that is normally observed distant from the site of first contact, i.e., after having passed through a physiological barrier (mucous membrane of the gastro-intestinal tract or of the respiratory tract, or the skin) and becomes systemically available.
- It should be noted, however, that toxic effects on surface epithelia may reflect indirect effects as a consequence of systemic toxicity or secondary to systemic distribution of the substance or its active metabolite(s).

R.7.5.1.2 Objective of the guidance on repeated dose toxicity

The objectives of assessing repeated dose toxicity are to evaluate:

- whether exposure of humans to a substance has been associated with adverse toxicological effects occurring as a result of repeated daily exposure for a part of the expected lifetime or for the major part of the lifetime; these human studies potentially may also identify populations that have higher susceptibility;

- whether administration of a substance to experimental animals causes adverse toxicological effects as a result of repeated daily exposure for a part of the expected lifespan or for the major part of the lifespan; effects that are predictive of possible adverse human health effects;

- the target organs, potential cumulative effects and the reversibility of the adverse toxicological effects;

- the dose-response relationship and threshold for any of the adverse toxicological effects observed in the repeated dose toxicity studies;

- the basis for risk characterisation and classification and labelling of substances for repeated dose toxicity.

R.7.5.2 Information requirements for repeated dose toxicity

Section R.2.1 in Chapter R.2 of the *Guidance on IR&CSA* provides general guidance on the information requirements of REACH. For repeated dose toxicity, all available information relevant for the endpoint needs to be evaluated and classification considered at each tonnage level. The following standard information requirements on repeated dose toxicity are specified in REACH Annexes VII-X:

In **Annex VII** (≥ 1 t/y), no test requirements on repeated dose toxicity are specified additional to the available information relevant for repeated dose toxicity.

In **Annex VIII** (≥ 10 t/y), a short-term repeated dose toxicity study (28 days) is usually required, in one species, male and female, using the most appropriate route of administration, having regard to the likely route of human exposure.

In **Annex IX** (≥ 100 t/y), a sub-chronic repeated dose toxicity study (90-days) is usually required, in one species (90-day study: rodent), male and female, and a short-term repeated dose toxicity study (28 days) is the minimum requirement, using the most appropriate route of administration, having regard to the likely route of human exposure. It should be noted that the 28-day test is not required at this tonnage level if already provided as part of Annex VIII requirements or if the 90-day study is proposed at this tonnage level.

In **Annex X** (≥ 1000 t/y), no specific test requirements additional to those required in Annexes VIII-IX for repeated dose toxicity is required at this tonnage level.

Column 1 of the REACH Annexes VII to X establishes the standard information required for all chemical substances and Column 2 lists specific rules according to which the required standard information requirements for individual endpoints may be modified (adapted) by waiving requirement for certain information, or in certain cases, defining the need for additional or different information. (for further details see Section R.2.1 in Chapter R.2 of the *Guidance on IR&CSA*).

In addition to the specific rules for adaptation listed in column 2 of the Annexes VII to X, the required standard information may also be adapted according to Annex XI, which specifies general rules for adaptation of the standard testing requirements set out in Annexes VII-X in cases where 1) testing does not appear scientifically necessary, 2) testing is technically not possible, and 3) testing may be omitted based on the exposure scenarios developed in the CSA (substance-tailored exposure-driven testing) (see Section R.5.1 (Exposure based waiving) in Chapter R.5 of the *Guidance on IR&CSA*).

It should also be noted that the introductory sections to Annexes VII-X point at a specific adaptation to the standard information requirements as *in vivo* testing shall be avoided with corrosive substances at concentration/dose levels causing corrosivity.

Factors that can influence the standard information requirements include the results of other toxicity studies, immediate disintegration of the substance, accumulation of the substance or its metabolites in certain tissues and organs, failure to identify a NOAEL in the required test at a given tonnage level, toxicity of particular concern, exposure route, structural relationships with a known toxic substance, physico-chemical properties of the substance, and use and human exposure patterns. These adaptations are detailed in the stepwise ITS presented in Section R.7.5.6.

R.7.5.3 Information and its sources on repeated dose toxicity

Toxicological information, including repeated dose toxicity, can be obtained from unpublished studies, data bases and publications such as books, scientific journals, criteria documents, monographs and other publications (see Chapter R.3 of the *Guidance on IR&CSA* for further general guidance). Information relevant for repeated dose toxicity can also be obtained from data on other endpoints, structural analogues and physico-chemical properties.

Before new tests are carried out to determine the hazardous properties of a chemical substance, all available information, shall be assessed, according to REACH Annex VI, step 1. (See Chapter R.4 of the *Guidance on IR&CSA* for general guidance on evaluation of information).

R.7.5.3.1 Non-human data on repeated dose toxicity

Non-testing data on repeated dose toxicity

Physico-chemical data

The physico-chemical properties of a chemical substance are essential elements in deciding on the appropriate administration route to be applied in experimental *in vivo* repeated dose toxicity studies as well as to decide on exemption from testing in cases where testing is technically not possible.

(Q)SAR models

The OECD has recently prepared a report on the use of (Q)SAR in the various member countries (OECD, 2006), which provides clear insight in how these tools are being used in the various OECD member countries. A review conducted by ECETOC on the use of (Q)SARs within current regulatory decision-making frameworks in EU, North America, and Japan, and within industry concluded that applicability of currently available (Q)SARs for chronic mammalian toxicity, certainly as a stand-alone approach, was very limited at that time (ECETOC 2003).

The ECB has started building a freely accessible inventory of evaluated (Q)SAR models which help to identify valid (Q)SARs for regulatory purposes (see also cross cutting guidance on

(Q)SARs). If there are any models relevant for the underlying endpoint these will be included in the ECB inventory.

More extensive guidance on the availability and application of (Q)SARs is available in Section R.6.1 in Chapter R.6 of the *Guidance on IR&CSA*.

Structurally or mechanistically related substance(s) (read-across/chemical category)

The concept of grouping, including both read-across and the related chemical category concept has been developed under the OECD HPV program (OECD 2007a. This is an approach which might be used to fill data gaps without the need for conducting tests when specific conditions, as specified in REACH Annex XI Section 1.3, are met.

Extensive guidance on the application of chemical categories/read across is available in Section R.6.2 in Chapter R.6 of the *Guidance on IR&CSA*.

Testing data on repeated dose toxicity

In vitro data

Currently, no available alternatives to animal testing are accepted for regulatory purposes for detecting toxicity after repeated exposure. Numerous *in vitro* systems have been developed over the last decades and have been discussed and summarized in recent ECVAM reports on repeated dose toxicity testing (Worth & Balls 2002, Prieto et al., 2005, and Prieto et al., 2006). At present, the *in vitro* models listed in these reports are at research and development level and cannot be used for repeated dose toxicity predictive purposes, although they are very useful to study individual types of organ toxicity or in assessing mechanistic aspects of target organ toxicity, on the tissue, cellular and molecular level. Some of the drawbacks are for instance the limited possibilities of current cell culture systems to account for kinetics and biotransformation, and the difficulty to derive from *in vitro* systems values such as NOAELs. Further development and optimisation of current *in vitro* systems as well as the selection of endpoints relevant to general as well as cell-type-specific mechanisms of toxicity or expression of toxic effects *in vivo* is ongoing. New technologies such as genomics, transcriptomics, proteomics and metabolomics could help in the identification of specific markers of toxicity that occur early in the process of long-term toxic responses and that are mechanistically linked to the underlying pathology. A recent ECVAM workshop report (Prieto et al., 2006) includes a proposed approach to assess repeated dose toxicity *in vitro* by integrating physiologically-based kinetic (PBK) modelling, the use of biomarkers, and omics technologies. However, this integrated approach is still under development and evaluation and is not ready for regulatory purposes.

The latest information on the status of alternative methods that are under development can be obtained from the ECVAM website (current address: https://eurl-ecvam.jrc.ec.europa.eu/) and other international centres for validation of alternative methods.

Human *in vitro* data, particularly on kinetics and metabolism, may assist in study interpretation thereby avoiding the need for unnecessary animal experimentation.

At present, available *in vitro* test data from well-characterised target organ and target system models on, e.g. mode of action(s) / mechanism(s) of toxicity may be useful in the interpretation of observed repeated dose toxicity.

Animal data

The most appropriate data on repeated dose toxicity for use in hazard characterisation and risk assessment are primarily obtained from studies in experimental animals conforming to internationally agreed test guidelines. In some circumstances repeated dose toxicity studies not conforming to conventional test guidelines may also provide relevant information for this endpoint.

The information that can be obtained from the available EU/OECD test guideline studies for repeated dose toxicity is briefly summarised below. Table R.7.5–2 summarises the parameters examined in these OECD test guideline studies in more detail to facilitate overview of the similarities and differences between the various studies. It should be noted that the test guidelines given in Annex V to Directive 67/548/EEC[65] (http://ecb.jrc.it/testing-methods/) are generally comparable to the OECD test guidelines (http://www.oecd.org/env/ehs/testing/oecdguidelinesforthetestingofchemicals.htm). Further details of the study protocols are described in the respective test guidelines.

- Repeated dose 28-day toxicity studies:

Separate guidelines are available for studies using oral administration (EU B.7 / OECD TG 407), dermal application (EU B.9 / OECD TG 410), or inhalation (EU B.8 / OECD TG 412). The principle of these study protocols is identical although the OECD TG 407 protocol includes additional parameters compared to those for dermal and inhalation administration, enabling the identification of a neurotoxic potential, immunological effects or reproductive organ toxicity.

The 28-day studies provide information on the toxicological effects arising from exposure to the substance during a relatively limited period of the animal's life span.

- Repeated dose 90-day toxicity studies:

Separate guidelines are available for studies using oral administration (OECD TG 408/409 / EU B.26/B.27 in rodent/non-rodent species, respectively), dermal application (OECD TG 411/EU B.28), or inhalation (OECD TG 413/EU B.29). The principle of these study protocols is identical although the revised OECD TG 408 protocol includes additional parameters compared to those for dermal and inhalation administration, enabling the identification of a neurotoxic potential, immunological effects or reproductive organ toxicity.

The 90-day studies provide information on the general toxicological effects arising from sub-chronic exposure (a prolonged period of the animal's life span) covering post-weaning maturation and growth well into adulthood, on target organs and on potential accumulation of the substance.

- Chronic toxicity studies:

The chronic toxicity studies (OECD TG 452/EU B.30) provide information on the toxicological effects arising from repeated exposure over a prolonged period of time covering the major part of the animal's life span. The duration of the chronic toxicity studies should be at least 12 months.

The combined chronic toxicity / carcinogenicity studies (OECD TG 453/EU B.33) include an additional high-dose satellite group for evaluation of pathology other than neoplasia. The satellite group should be exposed for at least 12 months and the animals in the carcinogenicity

---

[65] All the test methods previously included in Annex V to Directive 67/548/EEC will be incorporated in a new Test Methods (TM) Regulation that is currently (February 2008) under adoption. The TM Regulation will be adapted to technical progress whenever a new test method has been developed, scientifically validated and accepted for regulatory use by the National Coordinators of the Member states

part of the study should be retained in the study for the majority of the normal life span of the animals.

Ideally, the chronic studies should allow for the detection of general toxicity effects (physiological, biochemical and haematological effects etc.) but could also inform on neurotoxic, immunotoxic, reproductive and carcinogenic effects of the substance. However, in 12-month studies, non-specific life shortening effects, which require a long latent period or are cumulative, may possibly not be detected in this study type. In addition, the combined study will allow for detection of neoplastic effects and a determination of a carcinogenic potential and the life-shortening effects.

- The combined repeated dose toxicity study with the reproduction/ developmental toxicity screening test:

The combined repeated dose toxicity / reproductive screening study (OECD TG 422[66]) provides information on the toxicological effects arising from repeated exposure (generally oral exposure) over a period of about 6 weeks for males and approximately 54 days for females (a relatively limited period of the animal's life span) as well as on reproductive toxicity. For the repeated dose toxicity part, the OECD TG 422 is in concordance with the OECD TG 407/EU B.7 except for use of pregnant females and longer exposure duration in the OECD TG 422 compared to the OECD TG 407/EU B.7.

- Neurotoxicity studies:

The neurotoxicity study in rodents (OECD TG 424/EU B.43) has been designed to further characterise potential neurotoxicity observed in repeated dose systemic toxicity studies. The neurotoxicity study in rodents will provide detailed information on major neuro-behavioural and neuro-pathological effects in adult rodents.

- Delayed neurotoxicity studies of organophosphorus substances:

The delayed neurotoxicity study (OECD TG 419/ EU Annex B.38) is specifically designed to be used in the assessment and evaluation of the neurotoxic effects of organophosphorus substances. This study provides information on the delayed neurotoxicity arising from repeated exposure over a relatively limited period of the animal's life span.

- Other studies providing information on repeated dose toxicity:

Although not aiming at investigating repeated dose toxicity per se, other available OECD/EU test guideline studies involving repeated exposure of experimental animals may provide useful information on repeated dose toxicity. These studies are summarised in Table R.7.5–1.

It should be noted that the repeated dose toxicity studies, if carefully evaluated, may provide information on potential reproductive toxicity and on carcinogenicity (e.g., pre-neoplastic lesions).

The one- and two-generation studies (OECD TG 415/416/EU B.34/B.35) may provide information on the general toxicological effects arising from repeated exposure over a prolonged period of time (about 90 days for parental animals) as clinical signs of toxicity, body weight, selected organ weights, and gross and microscopic changes of selected organs are recorded.

---

66 To date there is no corresponding EU testing method available.

The prenatal developmental toxicity study (OECD TG 414/EU B.31), the reproduction/developmental toxicity screening study (OECD TG 421[67]) and the developmental neurotoxicity study (draft OECD TG 426[67]) may give some indications of general toxicological effects arising from repeated exposure over a relatively limited period of the animals life span as clinical signs of toxicity and body weight are recorded.

The carcinogenicity study (OECD TG 451/EU B.32) will, in addition to information on neoplastic lesions, also provide information on the general toxicological effects arising from repeated exposure over a major portion of the animal's life span as clinical signs of toxicity, body weight, and gross and microscopic changes of organs and tissues are recorded.

Table R.7.5–1 Overview of other *in vivo* test guideline studies giving information on repeated dose toxicity

Test	Design	Endpoints (general toxicity)
OECD TG 416 (EU B.35) Two-generation reproduction toxicity study	Exposure before mating for at least one spermatogenic cycle until weaning of 2nd generation At least 3 dose levels plus control At least 20 parental males and females per group	Clinical observations Body weight and food/water consumption Gross necropsy (all parental animals) Organ weights (reproductive organs, brain, liver, kidneys, spleen, pituitary, thyroid, adrenal glands, and known target organs) Histopathology (reproductive organs, previously identified target organ(s) - at least control and high-dose groups
OECD TG 415 (EU B.34) One-generation reproduction toxicity Study	Exposure before mating for at least one spermatogenic cycle until weaning of 1st generation At least 3 dose levels plus control At least 20 parental males and females per group	As in TG 416

67 To date there is no corresponding EU testing method available.

OECD TG 414 (EU B.31) Prenatal developmental toxicity study	Exposure at least from implantation to one or two days before expected birth At least 3 dose levels plus control At least 20 pregnant females per group	Clinical observations Body weight and food/water consumption Macroscopical examination all dams for any structural abnormalities or pathological changes, which may have influenced the pregnancy
OECD TG 421[68] Reproduction/ developmental toxicity screening test	Exposure from 2 weeks prior to mating until at least post-natal day 4 At least 3 dose levels plus control At least 8-10 parental males and females per group	Clinical observations Body weight and food/water consumption Gross necropsy (adult animals, special attention to reproductive organs) Organ weights (all adult males: testes, epididymides) Histopathology (reproductive organs in at least control and high-dose groups)
OECD TG 426[68] Developmental neurotoxicity study (draft)	Exposure at least from implantation throughout lactation (PND 20) At least 3 dose levels plus control At least 20 pregnant females per group	Clinical observations Body weight and food/water consumption
OECD TG 451 (EU B.32) Carcinogenicity studies	Exposure for majority of normal life span At least 3 dose levels plus control At least 50 males and females per group	Clinical observations (special attention to tumour development) Body weight and food consumption Gross necropsy Histopathology (all groups - all grossly visible tumours or lesions suspected of being tumours; at least control and high-dose groups - brain, pituitary, thyroid, parathyroid, thymus, lungs, heart, salivary glands, liver, spleen, kidneys, adrenals, oesophagus, stomach, duodenum, jejunum, ileum, cecum, colon, rectum, uterus, urinary bladder, lymph nodes, pancreas, gonads, accessory sex organs, female mammary gland, skin, musculature, peripheral nerve, spinal cord, sternum with bone marrow and femur, eyes)

R.7.5.3.2 Human data on repeated dose toxicity

Human data adequate to serve as the sole basis for the hazard and dose-response assessment are rare. When available, reliable and relevant human data are preferable over animal data and can contribute to the overall *Weight of Evidence*. However, human volunteer studies are not recommended due to practical and ethical considerations involved in deliberate exposure of individuals to chemicals.

68 To date there is no corresponding EU testing method available.

The following types of human data may already be available, however:

- Analytical epidemiology studies on exposed populations. These data may be useful for identifying a relationship between human exposure and effects such as biological effect markers, early signs of chronic effects, disease occurrence, or long-term specific mortality risks. Study designs include case control studies, cohort studies and cross-sectional studies.

- Descriptive or correlation epidemiology studies. They examine differences in disease rates among human populations in relation to age, gender, race, and differences in temporal or environmental conditions. These studies may be useful for identifying priority areas for further research but not for dose-response information.

- Case reports describe a particular effect in an individual or a group of individuals exposed to a substance. Generally case reports are of limited value for hazard identification, especially if the exposure represents single exposures, abuse or misuse of certain substances.

- Controlled studies in human volunteers. These studies, including low exposure toxicokinetic studies, might also be of use in risk assessment.

Meta-analysis. In this type of study data from multiple studies are combined and analysed in one overall assessment of the relative risk or dose-response curve.

R.7.5.3.3 Exposure considerations on repeated dose toxicity

Information on exposure, use and risk management measures should be collected in accordance with Article 10 and Annex VI (Section 3) of REACH.

Such information may lead to adaptation of the extent and nature of information needed on repeated dose toxicity under REACH; three types of *adaptations* are possible due to exposure considerations: exposure-based waiving of a study, exposure-based triggering of further studies, or definition of appropriate exposure route.

More detailed guidance of exposure-based adaptations of the repeat dose toxicity information requirements is given in Sections R.7.5.4 (evaluation of available information) and R.7.5.6 (Integrated testing strategy).

R.7.5.4 Evaluation of available information on repeated dose toxicity

General guidance on how to evaluate the available information is given in Chapter R.4 of the *Guidance on IR&CSA*.

R.7.5.4.1 Non-human data on repeated dose toxicity

Non-testing data on repeated dose toxicity

Physico-chemical properties

The physico-chemical properties of a chemical substance under registration should always be considered before any new experimental *in vivo* repeated dose toxicity studies are undertaken.

The physico-chemical properties of a substance can indicate whether it is likely that the substance can be absorbed following exposure to a particular route and whether it (or an active metabolite) is likely to reach the target organ(s) and tissue(s). The physico-chemical properties are thus essential elements in deciding on the appropriate administration route to be applied in experimental *in vivo* repeated dose toxicity studies (see Section R.7.5.4.3).

The physico-chemical properties are also important in order to judge whether testing is technically possible. Testing for repeated dose toxicity may, as specified in Annex XI Section 2 of REACH, be omitted if it is technically not possible to conduct the study as a consequence of the properties of the substance, e.g. very volatile, highly reactive or unstable substances cannot be used, or mixing of the substance with water may cause danger of fire or explosion. The Annex further emphasises that the guidance given in the test methods referred to in REACH Article 13 (3), more specifically on the technical limitations of a specific method, shall always be respected.

Additional generic guidance on the use of physico-chemical properties is provided e.g. in Section R.7.12 on toxicokinetics, in Chapter R.7c of the *Guidance on IR&CSA*.

Read-across to structurally or mechanistically similar substances (SAR)

The potential toxicity of a substance, for which no data are available on a specific endpoint can, in some cases, be evaluated by read-across from structurally or mechanistically related substances for which experimental data exists. The read-across approach is based on the principle that structurally and/or mechanistically related substances may have similar toxicological properties. Note that there are no formal criteria to identify structural alerts for repeated dose toxicity or for read-across to closely related substances.

Based on structural similarities between different substances, the repeated dose toxicity potential of one substance or a group of substances can be extended (read-across) to a substance, for which there are no or limited data on this endpoint.

A mechanism of toxicity or mode of action identified for a substance and/or group of substances and causally related to adverse effects in a target organ can be extended (read-across) to a substance for which a similar mechanism or mode of action has been identified, but where no or limited data on repeated dose toxicity are available. In such cases, the substance under evaluation may reasonably be expected to exhibit the same pattern of toxicity in the target organ(s) and tissue(s).

The chemical category concept has been developed under the OECD HPV programme (OECD 2004) as an approach to fill data gaps without the need for conduction of tests. A chemical category is a group of chemicals whose physico-chemical and toxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity. In the category approach, not every substance needs to be tested for every endpoint. However, the information finally compiled for the category must prove adequate to support a hazard assessment, a risk assessment and a classification for the category and its members. That is, the final data set must allow one to assess the untested endpoints, ideally by interpolation between and among the category members.

When analogue data are used to fill the data gaps for repeated dose toxicity, the data for the analogues must be compared and discussed in relation to the substance under evaluation in order to shed light on the similarities and differences in the toxicological profile of the substance under evaluation and its analogue(s).

Specific guidance regarding use of analogues is available in Section R.6.2 (in Chapter R.6 of the *Guidance on IR&CSA*) in order to decide on when further *in vivo* repeated dose toxicity studies shall be proposed (Annex VIII) or may be proposed (Annex X) as well as to decide on when analogue data can replace *in vivo* testing (Annex XI Section 1.3).

(Q)SAR

A (Q)SAR analysis for a substance may give indications for a specific mechanism to occur and identify possible organ or systemic toxicity upon repeated exposure. The reliability, applicability and overall scope of (Q)SAR science to identify chemical hazard and assist in risk assessment have been evaluated by various groups and organizations. Guidance on this issue is presented in Section R.6.1 (in Chapter R.6 of the *Guidance on IR&CSA*) and in OECD Monograph No. 69. (OECD 2007b).

Overall, (Q)SAR approaches are currently not well validated for repeated dose toxicity and consequently no firm recommendations can be made concerning their routine use in a testing strategy in this area. There are a large number of potential targets/mechanisms associated with repeated dose toxicity that today cannot be adequately covered by a battery of (Q)SAR models. Therefore, a negative result from current (Q)SAR models without other supporting evidence cannot be interpreted as demonstrating a lack of a toxicological hazard or a need for hazard classification. Another limitation of QSAR modelling is that dose-response information, including the N(L)OAEL, is not provided. Similarly, a validated QSAR model might identify a potential toxicological hazard, but because of limited confidence in this approach, such a result would not be adequate to support hazard classification.

In some cases, QSAR models could be used as part of a *Weight of Evidence* approach, when considered alongside other data, provided the applicability domain is appropriate. Also, QSAR's can be used as supporting evidence when assessing the toxicological properties by read-across within a substance grouping approach, providing the applicability domain is appropriate. Positive and negative QSAR modelling results can be of value in a read-across assessment and for classification purposes.

Testing data on repeated dose toxicity

In vitro data

As mentioned earlier in Section R.7.5.3.1 available *in vitro* data, at present, is not useful on its own for regulatory decisions such as risk assessment and C&L. However, such data may be helpful in the assessment of repeated dose toxicity, for instance to detect local target organ effects and/or to clarify the mechanisms of action. Since, at present, there are not validated and regulatory accepted *in vitro* methods, the quality of each of these studies and the adequacy of the data provided should be carefully evaluated.

Generic guidance is given in Chapters R.4 and R.5 for judging the applicability and validity of the outcome of various study methods, assessing the quality of the conduct of a study, reproducibility of data and aspects such as vehicle, number of replicates, exposure/incubation time, GLP-compliance or comparable quality description.

Animal data

The basic concept of repeated dose toxicity studies to generate data on target organ toxicity following sub-acute to chronic exposure is to treat experimental animals for 4 weeks, 13 weeks or longer. These studies are mentioned in Section R.7.5.3.1 and summarised in Table R.7.5–2. In addition, other studies performed in experimental animals may provide useful information on repeated dose toxicity. While at this time most alternative methods remain in the research and development stage and are not ready as surrogates for sub-chronic/chronic animal studies there are opportunities to improve data collection for risk assessment providing greater efficiency and use of fewer animals and better use of resources. Although not required by REACH, other opportunities include early development of kinetic data, in conjunction with early repeat dose toxicity testing thus ensuring that the maximum amount of information is drawn from the animal studies and for use in the risk assessment process.

The number of repeated dose toxicity studies available for a substance under registration is likely to be variable, ranging from none, a dose-range finding study, a 28-day repeated dose toxicity guideline study, to a series of guideline studies for some substances, including sub-chronic and/or chronic studies. There may also be studies employing different species and routes of exposure. In addition, special toxicity studies investigating further the nature, mechanism and/or dose-relationship of a critical effect in a target organ or tissue may also have been performed for some substances.

The following general guidance is provided for the evaluation of repeated dose toxicity data and the development of the *Weight of Evidence*:

- Studies on the most sensitive animal species should be selected as the significant ones, unless toxicokinetic and toxicodynamic data show that this species is less relevant for human risk assessment.

- Studies using an appropriate route, duration and frequency of exposure in relation to the expected route(s), frequency and duration of human exposure have greater weight.

- Studies enabling the identification of a NOAEL, and a robust hazard identification have a greater weight.

- Studies of a longer duration should be given greater weight than a repeated dose toxicity study of a shorter duration in the determination of the most relevant NOAEL.

- If sufficient evidence is available to identify the critical effect(s) (with regard to the dose-response relationship(s) and to the relevance for humans), and the target organ(s) and/or tissue(s), greater weight should be given to specific studies investigating this effect in the identification of the NOAEL. The critical effect can be a local as well as a systemic effect.

While data available from repeated dose toxicity studies not performed according to conventional guidelines and/or GLP may still provide information of relevance for risk assessment and classification and labelling such data require extra careful evaluation. REACH Annex XI specifically identifies circumstances where use of existing studies not carried out according to GLP or test methods referred to in Article 13(3) (guideline studies) can replace *in vivo* testing performed in accordance with Article 13(3). Data from non-guideline studies shall be considered to be equivalent to data generated by corresponding test methods referred to in Article 13(3) if the following conditions are met:

- adequate for the purpose of classification and labelling and/or risk assessment,

- adequate and reliable coverage of the key parameters foreseen to be investigated in the corresponding test methods referred to in Article 13(3),

- exposure duration comparable to or longer than the corresponding test methods referred to in REACH Article 13(3) if exposure duration is a relevant parameter, and

- adequate and reliable documentation of the study is provided.

In all other situations, non-guideline studies may contribute to the overall weight of the evidence but cannot stand alone for a hazard and risk assessment of a substance and thus, cannot serve as the sole basis for an assessment of repeated dose toxicity as well as for exempting from the standard information requirements for repeated dose toxicity at a given tonnage level, i.e. cannot be used to identify a substance as being adequately controlled in relation to repeated dose toxicity.

If sufficient information from existing studies is available on the repeated dose toxicity potential of a substance in order to perform a risk assessment as well as to conclude on classification and labelling for repeated dose toxicity (R48), no further *in vivo* testing is needed. The existing information is considered sufficient when, based on a *Weight of Evidence* analysis, the critical effect(s) and target organ(s) and tissue(s) can be identified, the dose-response relationship(s) and NOAEL(s) and/or LOAEL(s) for the critical effect(s) can be established, and the relevance for human beings can be assessed.

It should be noted that potential effects in certain target organs (e.g., the thyroid) following repeated exposure may not be observed within the span of the 28-day study. Attention is also drawn to the fact that the protocols for the oral 28-day and 90-day studies include additional parameters compared to those for the 28-day and 90-day dermal and inhalation protocols.

Where it is considered that the existing data as a whole is inadequate to provide a clear assessment of this endpoint, the need for further testing should be considered in view of all available relevant information on the substance, including use pattern, the potential for human exposure, physico-chemical properties, and structural alerts. The testing strategy is presented in Section R.7.5.6.3.

Specific investigations such as studies for neurotoxicity or immunotoxicity are also elements in the testing strategy presented in REACH.

Regarding neurotoxicity and immunotoxicity, standard oral 28-day and 90-day toxicity studies include endpoints capable of detecting such effects. Indicators of neurotoxicity include clinical observations, a functional observational battery, motor activity assessment and histopathological examination of spinal cord and sciatic nerve. Indicators of immunotoxicity include changes in haematological parameters, serum globulin levels, alterations in immune system organ weights such as spleen and thymus, and histopathological changes in immune organs such as spleen, thymus, lymph nodes and bone marrow. Where data from standard oral 28-day and 90-day studies identify evidence of neurotoxicity or immunotoxicity other studies may be necessary to further investigate the effects. It should be noted that endpoints capable of detecting neurotoxicity and immunotoxicity are not examined in the standard 28-day and 90-day dermal or inhalation repeated dose toxicity studies.

More focus has also been put on endocrine disrupters during the latest decade. In relation to hazard and risk assessment, there are currently no test strategies or methods available, which specifically detect all effects, which have been linked to the endocrine disruption mechanism. It should be noted that work is on-going with the purpose of updating the present oral 28-days study (OECD TG 407/EU B.7) with more emphasis to be placed on detection of endocrine effects.

If data are not available from an oral standard 28-day repeated dose toxicity guideline study (OECD TG 407/EU B.7), the minimum repeated dose toxicity data requirement (28-day study) at tonnage levels from 10 t/y may in certain circumstances be met by results obtained from *the combined repeated dose toxicity study with the reproduction / developmental toxicity screening test* (OECD TG 422[69]). An advantage of this approach is obtaining information on repeated dose toxicity and reproductive toxicity in a single study providing an overall saving in the number of animals used for testing. In addition, the number of animals is higher (10 per sex compared to 5 per sex in the standard oral 28-day study) and the dosing period is longer

[69] To date there is no corresponding EU testing method available.

in the combined study than in the standard oral 28-day study. Therefore, more information on repeated dose toxicity could be expected from the combined study. Potential complications in using the combined study include selecting adequate dose levels to examine adequately both repeated dose toxicity and reproductive toxicity. In addition, interpretation of the results may be complicated due to differences in sensitivity between pregnant and non-pregnant animals, and an assessment of the general toxicity may be more difficult especially when serum and histopathological parameters are not evaluated at the same time in the study. Consequently, where the combined study is used for the assessment of repeated dose toxicity, the use of data obtained from such a study should be clearly indicated. Despite such complications, the use of the combined study is recommended for the initial hazard assessment of the repeated dose toxicity potential of a substance when this study is relevant also for reproductive toxicity assessment.

In general, results from toxicological studies requiring repeated administration of a test substance (see also Section R.7.5.3.1) such as *reproduction and developmental toxicity studies* as well as *carcinogenicity studies* can contribute to the assessment of repeated dose toxicity. However, such toxicological studies rarely provide the information obtained from a standard repeated dose toxicity study and therefore, cannot stand alone as the sole basis for the assessment of repeated dose toxicity or for exempting from the standard information requirements for repeated dose toxicity at a given tonnage level.

Studies such as *acute toxicity and irritation studies* as well as *in vivo genotoxicity studies* contribute limited information to the overall assessment of the repeated dose toxicity. However, such studies may be useful in deciding on the dose levels for use in repeated dose toxicity.

Guidance on the dose selection for repeated dose toxicity testing (see also Table R.7.5–2) is provided in detail in the EU and OECD test guidelines. Unless limited by the physical-chemical nature or biological effects of the test substance, the highest dose level should be chosen with the aim to induce toxicity but not death or severe suffering.

Although not required by REACH, toxicokinetic studies may be helpful in the evaluation and interpretation of repeated dose toxicity data, for example in relation to accumulation of a substance or its metabolites in certain tissues or organs as well as in relation to mechanistic aspects of repeated dose toxicity and species differences. Toxicokinetic information can also assist in the selection of the dose levels. When conducting repeated dose toxicity studies it is necessary to ensure that the observed treatment-related toxicity is not associated with the administration of excessive high doses causing saturation of absorption and detoxification mechanisms. The results obtained from studies using excessive doses causing saturation of metabolism are often of limited value in defining the risk posed at more relevant and realistic exposures where a substance can be readily metabolised and cleared from the body. It is suggested that a key resource in designing better repeated dose toxicity studies is to select appropriate dose levels based on results from useful metabolic and toxicokinetic investigations. Further details on the application of toxicokinetic information in the design and evaluation of repeated dose toxicity studies is available in Section R.7.12 on toxicokinetics, in Chapter R.7c of the *Guidance on IR&CSA*.

Table R.7.5–2 Overview of *in vivo* repeated dose toxicity test guideline studies

Test	Design	Endpoints
OECD TG 407 (EU B.7) Repeated dose 28-day oral toxicity study in rodents	Exposure for 28 days At least 3 dose levels plus control At least 5 males and females per group Preferred rodent species: rat	Clinical observations Functional observations (4th exposure week – sensory reactivity to stimuli of different types, grip strength, motor activity) Body weight and food/water consumption Haematology (haematocrit, haemoglobin, erythrocyte count, total and differential leucocyte count, platelet count, blood clotting time/potential) Clinical biochemistry Urinalysis (optional) Gross necropsy (full, detailed, all animals) Organ weights (all animals - liver, kidneys, adrenals, testes, epididymides, thymus, spleen, brain, heart) Histopathology (full, at least control and high-dose groups - all gross lesions, brain, spinal cord, stomach, small and large intestines, liver, kidneys, adrenals, spleen, heart, thymus, thyroid, trachea and lungs, gonads, accessory sex organs, urinary bladder, lymph nodes, peripheral nerve, a section of bone marrow)
OECD TG 410 (EU B.9) Repeated dose dermal toxicity: 21/28-day study	Exposure for 21/28 days At least 3 dose levels plus control At least 5 males and females per group Rat, rabbit or guinea pig	Clinical observations Body weight and food/water consumption Haematology (haematocrit, haemoglobin, erythrocyte count, total and differential leucocyte count, clotting potential) Clinical biochemistry Urinalysis (optional) Gross necropsy (full, detailed, all animals) Organ weights (all animals - liver, kidneys, adrenals, testes) Histopathology (full, at least control and high-dose groups - all gross lesions, normal and treated skin, liver, kidney)
OECD TG 412 (EU B.8) Repeated dose inhalation toxicity: 28-day or 14-day study	Exposure for 28 or 14 days At least 3 concentrations plus control At least 5 males and females per group Rodents: preferred species - rat	Clinical observations Body weight and food/water consumption Haematology (haematocrit, haemoglobin, erythrocyte count, total and differential leucocyte count, clotting potential) Clinical biochemistry Urinalysis (optional) Gross necropsy (full, detailed, all animals) Organ weights (all animals - liver, kidneys,

Test	Design	Endpoints
		adrenals, testes) Histopathology (full, at least control and high-dose groups - all gross lesions, lungs, liver, kidney, spleen, adrenals, heart)
OECD TG 408 (EU B.26) Repeated dose 90-day oral toxicity study in rodents	Exposure for 90 days At least 3 dose levels plus control At least 10 males and females per group Preferred rodent species: rat	Clinical observations Ophthalmological examination Functional observations (towards end of exposure period – sensory reactivity to stimuli of different types, grip strength, motor activity) Body weight and food/water consumption Haematology (haematocrit, haemoglobin, erythrocyte count, total and differential leucocyte count, platelet count, blood clotting time/potential) Clinical biochemistry Urinalysis Gross necropsy (full, detailed, all animals) Organ weights (all animals - liver, kidneys, adrenals, testes, epididymides, uterus, ovaries, thymus, spleen, brain, heart) Histopathology (full, at least control and high-dose groups - all gross lesions, brain, spinal cord, pituitary, thyroid, parathyroid, thymus, oesophagus, salivary glands, stomach, small and large intestines, liver, pancreas, kidneys, adrenals, spleen, heart, trachea and lungs, aorta, gonads, uterus, accessory sex organs, female mammary gland, prostate, urinary bladder, gall bladder (mouse), lymph nodes, peripheral nerve, a section of bone marrow, and skin/eyes on indication)
OECD TG 409 (EU B.27) Repeated dose 90-day oral toxicity study in non-rodents	Exposure for 90 days At least 3 dose levels plus control At least 4 males and females per group Preferred species: dog	Clinical observations Ophthalmological examination Body weight and food/water consumption Haematology (as in TG 408) Clinical biochemistry Urinalysis Gross necropsy (full, detailed, all animals) Organ weights (as in TG 408 - additional: gall bladder, thyroid, parathyroid) Histopathology (as in TG 408 – additional: gall bladder, eyes)
OECD TG 411 (EU B.28) Subchronic dermal toxicity: 90-day	Exposure for 90 days At least 3 dose levels plus control At least 10 males and	Clinical observations Ophthalmological examination Body weight and food/water consumption

Test	Design	Endpoints
study	females per group Rat, rabbit or guinea pig	Haematology (haematocrit, haemoglobin, erythrocyte count, total and differential leucocyte count, clotting potential) Clinical biochemistry Urinalysis Gross necropsy (full, detailed, all animals) Organ weights (all animals - liver, kidneys, adrenals, testes) Histopathology (full, at least control and high-dose groups - all gross lesions, normal and treated skin, and essentially the same organs and tissues as in TG 408)
OECD TG 413 (EU B.29) Subchronic inhalation toxicity: 90-day study	Exposure for 90 days At least 3 concentrations plus control At least 10 males and females per group Rodents: preferred species - rat	Clinical observations Ophthalmological examination Body weight and food/water consumption Haematology (haematocrit, haemoglobin, erythrocyte count, total and differential leucocyte count, clotting potential) Clinical biochemistry Urinalysis Gross necropsy (full, detailed, all animals) Organ weights (all animals - liver, kidneys, adrenals, testes) Histopathology (full, at least control and high-dose groups - all gross lesions, respiratory tract, and essentially the same organs and tissues as in TG 408)
OECD TG 452 (EU B.30) Chronic toxicity studies	Exposure for at least 12 months At least 3 dose levels plus control Rodents : At least 20 males and females per group Non-rodents: At least 4 males and females per group Preferred rodent species: rat Preferred non-rodent species: dog	Clinical observations, including neurological changes Ophthalmological examination Body weight and food/water consumption Haematology (haematocrit, haemoglobin, erythrocyte count, total leucocyte count, platelet count, clotting potential) Clinical biochemistry Urinalysis Gross necropsy (full, detailed, all animals) Organ weights (all animals - brain, liver, kidneys, adrenals, gonads, thyroid/parathyroid (non-rodents only)) Histopathology (full, at least control and high-dose groups - all grossly visible tumours and other lesions, as well as essentially the same organs and tissues as in the 90-day studies (TG 408/409))

Test	Design	Endpoints
OECD TG 453 (EU B.33) Combined chronic toxicity / carcinogenicity studies	Exposure for at least 12 months (satellite groups) or majority of normal life span (carcinogenicity part) At least 3 dose levels plus control At least 50 males and females per group Satellite group: At least 20 males and females per group Preferred species: rat	Essentially as in TG 452
OECD TG 422[70] Combined repeated dose toxicity study with the reproduction/developmental toxicity screening test	Exposure for a minimum of 4 weeks (males) or from 2 weeks prior to mating until at least post-natal day 4 (females – at least 6 weeks of exposure) At least 3 dose levels plus control At least 10 males and females per group	Clinical observations as in TG 407 Functional observations as in TG 407 Body weight and food/water consumption Haematology as in TG 407 Clinical biochemistry Urinalysis (optional) Gross necropsy (full, detailed, all adult animals) Organ weights (testes and epididymides - all males; liver, kidneys, adrenals, thymus, spleen, brain, heart - in 5 animals of each sex per group, i.e. as in TG 407) Histopathology (ovaries, testes, epididymides, accessory sex organs, all gross lesions - all animals in at least control and high-dose groups; brain, spinal cord, stomach, small and large intestines, liver, kidneys, adrenals, spleen, heart, thymus, thyroid, trachea and lungs, urinary bladder, lymph nodes, peripheral nerve, a section of bone marrow - in 5 animals of each sex in at least control and high-dose groups, i.e. as in TG 407)
OECD TG 424 (EU B.43) Neurotoxicity study in rodents	Exposure for at least 28 days Dose levels: not specified At least 10 males and females per group Preferred rodent species: rat Generally oral route of administration	Detailed clinical observations Functional observations (sensory reactivity to stimuli of different types, grip strength, motor activity, more specialized tests on indication) Ophthalmological examination Body weight and food/water consumption Haematology (haematocrit, haemoglobin, erythrocyte count, total and differential leucocyte count, platelet count, blood clotting

[70] To date there is no corresponding EU testing method available

Test	Design	Endpoints
		time/potential) Clinical biochemistry Histopathology: at least 5 animals/sex/ group) for neuropathological examinations (brain, spinal cord, and peripheral nerves); remaining animals to be used either for specific neurobehavioural, neuropathological, neurochemical or electrophysiological procedures that may supplement the histopathology or alternatively, for routine pathological evaluations according to the guidelines for standard repeated dose toxicity studies
OECD TG 419 (EU B.38) Delayed neurotoxicity of organophosphorus substances: 28-day repeated dose study	Exposure for 28 days At least 3 dose levels plus control At least 12 birds per group Species: domestic laying hen	Detailed clinical observations Body weight and food/water consumption Clinical biochemistry (NTE activity, acetylcholinesterase activity Gross necropsy (all animals) Histopathology (neural tissue)

R.7.5.4.2 Human data on repeated dose toxicity

Human data in the form of epidemiological studies or case reports can contribute to the hazard identification process as well as to the risk assessment process itself. Criteria for assessing the adequacy of epidemiology studies include an adequate research design, the proper selection and characterisation of the exposed and control groups, adequate characterisation of exposure, sufficient length of follow-up for the disease as an effect of the exposure to develop, valid ascertainment of effect, proper consideration of bias and confounding factors, proper statistical analysis and a reasonable statistical power to detect an effect. These types of criteria have been described in more detail (Swaen, 2006 and can be derived from Epidemiology Textbooks (Checkoway *et al*, 1989; Hernberg, 1991; Rothman, 1998).

The results from human experimental studies are often limited by a number of factors, such as a relatively small number of subjects, short duration of exposure, and low dose levels resulting in poor sensitivity in detecting effects.

In relation to hazard identification, the relative lack of sensitivity of human data may cause particular difficulty. Therefore, negative human data cannot be used to override the positive findings in animals, unless it has been demonstrated that the mode of action of a certain toxic response observed in animals is not relevant for humans. In such a case a full justification is required. It is emphasised that testing with human volunteers is strongly discouraged, but when there are good quality data already available they can be used in the overall *Weight of Evidence*.

R.7.5.4.3 Exposure considerations for repeated dose toxicity

Three types of *adaptations* from testing are possible due to exposure considerations: exposure-based waiving of a study, exposure-based triggering of further studies, or selection of appropriate exposure route. More information on exposure-based waiving is available in Section R.5.1 in Chapter R.5 of the *Guidance on IR&CSA*. More detailed guidance of exposure-based adaptations of the testing for repeated dose toxicity is given below and in Section R.7.5.6 (Integrated Testing Strategy).

Comparison of exposure and effect data should consider the existing (or most likely expected) *exposure patterns* for humans (e.g. daily exposure during life-time or repeated short or medium periods of exposures) and the most adequate DNEL (Derived No Effect Level) that reflects the specific exposure route and time pattern for each human population group at exposure. For instance, short-term exposure estimates should be compared to a descriptor of short-term toxicity whereas repeated daily exposure estimates should be compared to a corresponding descriptor of chronic toxicity. In all cases actually experienced daily human exposures are to be used in this comparison instead of daily exposures obtained by averaging over exposed and non-exposed days.

Concerning repeated dose toxicity testing the oral route is the preferred one. However, dependent on the physico-chemical properties of a substance as well as on the most relevant route of human exposure, the dermal or the inhalation route could also be appropriate as specified in REACH Annex VIII and IX.

The dermal route is appropriate if the physico-chemical properties suggest potential for a significant rate of absorption through the skin. The inhalation route is appropriate if exposure of humans *via* inhalation is the most relevant route of human exposure taking into account the vapour pressure of the substance and/or the possibility of exposure to aerosols, particles or droplets of an inhalable size.

According to Annex VIII-X further studies shall be proposed by the registrant or may be required by the Agency for example if there is particular concern regarding exposure, e.g. use in consumer products leading to exposure levels which are:

- close to the dose levels at which toxicity to humans may be expected (Annex VIII) i.e. a dose lower than, but in the vicinity of, the dose levels at which toxicity to humans may be expected

- high relative to the dose levels at which toxicity to humans may be expected (Annex IX), i.e. exposure levels higher than the dose levels at which toxicity to humans may be expected

- close to the dose levels at which toxicity is observed (Annex X); i.e. a dose lower than, but in the vicinity of, the dose levels at which toxicity is observed from animal studies.

Any of the exposure-triggered studies proposed by the registrant or required by the Agency should be considered on a case-by-case basis.

Various types of exposure considerations are possible for *waiving* of repeated dose toxicity studies. For instance, it is stated in REACH Article 13 and Annex XI:3 that testing in accordance with Annex VIII, Sections 8.6 and 8.7 (i.e. repeated dose toxicity and reproductive toxicity), Annex IX and X may be omitted based on the exposure scenario(s) developed in the Chemical Safety Report. Adequate justification and documentation shall in all cases be provided (see Section R.5.1 in Chapter R.5 of the *Guidance on IR&CSA*).

Further, the sub-chronic toxicity study (90-days study) does not need to be conducted according to Annex IX of REACH if: "the substance is unreactive, insoluble and not inhalable and there is no evidence of absorption and no evidence of toxicity in a 28-days *limit test*, particularly if such a pattern is *coupled with limited human exposure*. In order to omit the study the prerequisites interpreted above have to be considered jointly since the word "and" is used in between them. In addition, limited human exposure would strengthen the possibility for waiving.

The interpretation of *un-reactive* can be that it relates to the inherent chemical reactivity and as such, is an indicator of lack of local effects and mutagenicity, *insoluble and not inhalable* can be interpreted as indicators of low exposure potential and should be further defined, and *no evidence of absorption* that there has to be evidence for lack of absorption in order to omit the study. Further *no evidence of toxicity in a 28-days limit test* can be interpreted as it has to be at least a 28-days limit test available in order to waive the 90-days study, and this 28-days study should not show any sign of toxicity at 1000 mg/kg.

Limited exposure should consider the level of exposure, the frequency and/or the duration of exposure. Therefore, limited exposure must be considered on a case-by-case basis.

Finally, according to REACH Annex VIII testing of repeated dose toxicity (28-days study) does not need to be conducted if: *relevant human exposure can be excluded*.

Relevant human exposure depends on the inherent properties of the substance, if the population comes into contact with the substance or not, and how the substance is used. Thus, waiving might be considered on a case-by-case basis.

The concept of the Threshold of Toxicological Concern (TTC) might be applied to reduce the use of animals and other evaluation resources (Kroes *et al*., 2004); Use of the TTC concept may also be seen as a driving force for deriving exposure information of adequate quality. However, there are a number of limitations or drawbacks that should be taken into consideration in deciding if the concept is to be applied for industrial chemicals and further discussions on the cut-off values are needed before integration into into the guidance (see Appendix R.7-1 to Chapter R.7, in Chapter R.7c of the *Guidance on IR&CSA*; TemaNord, 2005*)*.

R.7.5.4.4 Remaining uncertainty on repeated dose toxicity

The key requirement for a CSA is the DNELs per exposure scenario (box 5 of Figure R.7.5–1). The DNEL for repeated dose toxicity is the threshold of the critical effect derived in a *Weight of Evidence* assessment of the available repeated dose toxicity data and an overall assessment factor (AF) that takes into account any uncertainty. The following elements contribute to the uncertainty in the determination of a threshold for the critical effects and the selection of the AF (further guidance on deriving a DNEL and application of AFs is provided in Chapter R.8 of the *Guidance on IR&CSA*).

Threshold of the critical effect

In the determination of the overall threshold for repeated dose toxicity all relevant information is evaluated to determine the lowest dose that induces an adverse effect (i.e. LOAEL or LOAEC) and the highest level with no biologically or statically significant adverse effects (i.e. NOAEL or NOAEC). In this assessment all toxicological responses are taken into account and the critical effect is identified. The uncertainty in the threshold depends on the strength of the data and is largely determined by the design of the underlying experimental data. Parameters such as group size, study type/duration or the methodology need to be taken into account in the assessment of the uncertainty in the threshold of the critical effect(s).

The NOAEL is typically used as the starting point for the derivation of the DNEL. In case a NOAEL has not been achieved, a LOAEL may be used, provided the available information is sufficient for a robust hazard assessment and for Classification and Labelling. The Bench Mark Dose (BMD) may also be used as the starting point for the derivation of the DNEL (Chapter R.8 of the *Guidance on IR&CSA*).

The selection of NOAEL or LOAEL is usually based on the dose levels used in the most relevant toxicity study, without considering the shape of the dose response curve. Therefore, the NOAEL/LOAEL may not reflect the true threshold for the adverse effect. On the other hand, the

BMD is a statistical approach for the determination of the threshold and relies on the dose response curve. Alternatively, mathematical curve fitting techniques or statistical approaches exist to determine the threshold for an adverse effect. The use of such approaches (e.g. Benchmark Dose) to estimate the threshold should be considered on a case-by-case basis. For further guidance see Chapter R.8 of the *Guidance on IR&CSA*.

Overall AF

Variability in sensitivity across and within species is another source of uncertainty for repeated dose toxicity. These inter- and intraspecies differences, respectively, are linked with variations in the toxicokinetics and dynamics of a substance. Information derived from non-testing, *in vitro* or *in vivo* methods may lead to an improvement of the understanding of the relevance of animal data for human risk assessment and may lead to a replacement of adopted standard default AF for these differences.

The quality of the whole database should be assessed for reliability and consistency across different studies and endpoints and taking into account the quality of the testing method, size and power of the study design, biological plausibility, dose-response relationships and statistical association. Missing test data might be substituted by non-testing data obtained from physico-chemical properties, read-across to structurally or mechanistically related substances (SAR/chemical category) or by quantitative structure-activity relationships (QSARs). Also *in vitro* data might be used to fill in data gaps as well as *in vivo* non-standard animal experimental tests. Such data in combination with toxicity tests according to standard OECD/EU guidelines may in some cases lead to an improved understanding to the toxicological effect resulting in a reduction in the overall uncertainty. On the other hand information solely based on *in-vitro* and non-testing data are at present insufficient to act as a surrogate for repeated dose toxicity data and the uncertainty is sufficiently large that such information is unsuitable for use in a CSA and for classification and labelling. In the case of chemical categories information from non-testing methods or *in vitro* data may used to fulfil the data requirements on repeated dose toxicity and lead to improvement in the overall reliability and consistency for the read-across within a category of substances.

Since the adequacy and/or completeness of different data may vary, lack of quality and completeness of the overall database should be compensated for with an assessment factor for remaining uncertainty.

Besides AF addressing these differences (inter- and intraspecies, quality of the whole database), other uncertainties relating to differences between human and animal exposure conditions (e.g. route, and duration), and dose response characteristics are taken into account in the more extensive guidance on deriving a DNEL (see Section R.8.4.3 in Chapter R.8 of the *Guidance on IR&CSA*).

Other considerations

Another situation may arise when testing is not technically possible, a waiving option indicated in Annex XI(2) (see also Chapter R.5 of the *Guidance on IR&CSA*). In such cases approaches such as QSAR, category formation and read-across may be helpful in the hazard characterisation; they should also be considered for information that might be suitable as a surrogate for a dose descriptor. Alternatively, generic threshold approaches, e.g. the Threshold of Toxicological Concern, TTC might be considered for the starting point of a risk characterisation (see Appendix R.7-1 to Chapter R.7, in Chapter R.7c of the *Guidance on IR&CSA*).

R.7.5.5 Conclusions on repeated dose toxicity

The evaluation of all available toxicological information for repeated dose toxicity (step 3 in Figure R.7.5–1 should include an assessment whether the available information as a whole (i.e. testing and non-testing, and relevant information from studies addressing other endpoints) meets the tonnage driven data requirements necessary to fulfil the REACH requirements. A *Weight of Evidence* approach should be used in assessing the database for a substance. This approach requires a critical evaluation of the entire body of available data for consistency and biological plausibility. Potentially relevant studies should be judged for quality and studies of high quality given more weight than those of lower quality. When both epidemiological and experimental data are available, similarity of effects between humans and animals is given more weight. If the mechanism or mode of action is well characterised, this information is used in the interpretation of observed effects in either human or animal studies. *Weight of Evidence* is not to be interpreted as simply tallying the number of positive and negative studies, nor does it imply an averaging of the doses or exposures identified in individual studies that may be suitable as starting points for risk assessment. The study or studies used for the starting point are identified by an informed and expert evaluation of all the available evidence.

The available repeated dose toxicity data should be evaluated in detail for a characterisation of the health hazards upon repeated exposure. In this process an assessment of all toxicological effect(s), their dose-response relationships and possible thresholds are taken into account. The evaluation should include an assessment of the severity of the effect, whether the observed effect(s) are adverse or adaptive, if the effect is irreversible or not or if it is a precursor to a more significant effect or secondary to general toxicity. Correlations between changes in several parameters, e.g. between clinical or biochemical measurements, organ weights and (histo)pathological effects, will be helpful in the evaluation of the nature of effects. Further guidance to this issue can be found in publications of the International Programme on Chemical Safety (IPCS 1994, 1999) and ECETOC (2002).

The effects data are also analysed for indications of potential serious toxicity of target organs or specific organ systems (e.g. neurotoxicity or immunotoxicity), delayed effects or cumulative toxicity. Furthermore, the evaluation should take into account the study details and determine if the exposure conditions and duration and the parameters studied are appropriate for an adequate characterisation of the toxicological effect(s).

If an evaluation allows the conclusion that the information of the repeated dose toxicity is adequate for a robust characterisation of the toxicological hazards, including an estimate of a dose descriptor (NOAEL/LOAEL/BMD), and the data are adequate for risk assessment and classification and labelling, no further testing will be necessary unless there are indications for further risk, according to column 2 of Annexes VIII-X of REACH.

Another consideration to be taken into account is whether the study duration has been appropriate for an adequate expression of the toxicological effects. If the critical effect involves serious specific system or target organ toxicity (e.g. haemolytic anaemia, neurotoxicity or immunotoxicity), delayed effects or cumulative toxicity and a threshold has **not** been established dose extrapolation may not be appropriate and further studies are required. In this case a specialised study is likely to be more appropriate for an improved hazard characterisation and should be considered instead of a standard short-term rodent or sub-chronic toxicity test at this stage.

In the identification of the NOAEL, other factors need to be considered such as the severity of the effect, the presence or absence of a dose- and time-effect relationship and/or a dose- and time-response relationship, the biological relevance, the reversibility, and the normal biological variation of an effect that may be shown by representative historical control values (IPCS, 1990).

R.7.5.5.1 Concluding on suitability for Classification and Labelling

In order to conclude on the suitability for classification and labelling (C&L), the data requirements in Annex VI of the dangerous substances Directive 67/548/EEC[71] have to be considered (box 4 in Figure R.7.5–1).

A decision on classification and labelling will affect downstream events/Directives under REACH. Therefore, it is important that the data are adequate for checking against the classification criteria in order to ensure safe use under REACH.

Basically the following conclusions can be obtained from the assessment of adequacy for C&L for repeated dose toxicity:

- Data are considered adequate for the purpose of C&L and can be checked against the criteria (boxes 6 and 11 in Figure R.7.5–1)[72].

- Data are considered as inadequate for the purpose of C&L and cannot be checked against the criteria (inconclusive or lacking data). In this case testing should be considered in relation to the risk management of the substance.

R.7.5.5.2 Concluding on suitability for Chemical Safety Assessment

In order to be suitable for CSA (box 5 of Figure R.7.5–1) appropriate DNELs have to be established for each exposure scenario. Typically, the derivation of the DNEL takes into account a dose descriptor, modification of the starting point and application of assessment factors (see Chapter R.8 of the *Guidance on IR&CSA*).

Identification of the so-called dose descriptor: i.e. an appropriate threshold dose for the critical effect as the starting point for DNEL derivation, i.e. a NOAEL or BMD. If a NOAEL can not be identified, the LOAEL may be used instead provided the data are adequate for a robust hazard assessment.

It is to be noted that the dose descriptor should be route-specific. Thus, in case only animal data with oral exposure are available and humans are exposed mainly *via* skin and/or inhalation, a DNEL for dermal route and/or DNEL for inhalation route are needed: i.e. route-to-route extrapolation is needed, if allowed. Guidance for this route-to-route extrapolation is provided in Section R.8.4.2 in Chapter R.8 of the *Guidance on IR&CSA*.

If this route-to-route extrapolation is not allowed, route-specific information is needed, possibly including testing, as a last resort (see Section R.7.5.6.3).

---

[71] Directive 67/548/EEC will be repealed and replaced with the EU Regulation on classification, labelling and packaging of substances and mixtures, implementing the Globally Harmonized System (GHS).

[72] It should be noted that although the exposure assessment and risk characterisation need not to be performed, when a substance is not classified (see Part A, section A.1.2), for potency-based endpoints like repeated dose toxicity, there could still potentially be a risk. Therefore one might consider performing an exposure assessment and risk characterisation on voluntary basis, to ensure safe handling and use.

Derivation of a DNEL from this dose descriptor by applying AFs (to address uncertainty in the available data) is described elsewhere (see Section R.8.4.3 in Chapter R.8 of the *Guidance on IR&CSA*; see also Section R.7.5.4.4).

R.7.5.5.3 Information not adequate

A *Weight of Evidence* approach comparing available adequate information with the tonnage-triggered information requirements by REACH may result in the conclusion that the requirements are not fulfilled. In order to proceed in further information gathering the testing strategy described in Section R.7.5.6.3 can be adopted.

R.7.5.6 Integrated Testing Strategy (ITS) for repeated dose toxicity

R.7.5.6.1 Objective / General principles

The objective in this testing strategy is to give guidance on a stepwise approach to hazard identification with regard repeated dose toxicity. A principle of the strategy is that the results of one study are evaluated before another study is initiated. The strategy seeks to ensure that the data requirements are met in the most efficient and humane manner so that animal usage and costs are minimised.

The core objectives of the Integrated Testing Strategy (ITS) for repeated dose toxicity are to generate sufficient information to allow:

- Characterisation of the hazard profile and the dose-response of a substance upon repeated exposure.
- Performance of a chemical safety assessment for repeated dose toxicity.

Information generated in this strategy should be suitable for Classification and Labelling according to the criteria given in Annex VI to Directive 67/548/EEC[73].

In addition, information from repeated dose toxicity studies can give valuable information to other endpoints based on repeated exposure (e.g. reproductive and developmental toxicity), and are valuable for other *in vivo* studies.

R.7.5.6.2 Preliminary considerations

On the basis of the objectives outlined above, a framework has been developed so that informed decisions can be made on the need for further testing. If generation of further data is deemed necessary, the information needs should be met efficiently in terms of resources and animal use. This means the use of the most appropriate study type in accordance with the tonnage-driven requirements stipulated by the REACH information requirements and taking into account modifications due to considerations of exposure, grouping and category formation. The data requirements may be increased or decreased taking into account exposure considerations or the level of concern noted during any of the stages in the testing strategy.

Testing for repeated dose toxicity is not required for chemicals produced at tonnage levels less than 10 tonne per annum (t/y). At higher production volumes, standard data requirements

[73] Directive 67/548/EEC will be repealed and replaced with the EU Regulation on classification, labelling and packaging of substances and mixtures, implementing the Globally Harmonized System (GHS).

are, in general, increased with each tonnage band (see Section R.7.5.2); maintaining flexibility to adopt the most appropriate testing regime for any single chemical is a key component of the ITS. However, regardless of whether testing for repeated dose toxicity is required or not at a specific tonnage level, all existing test data, and all other available and relevant information on the substance should be collected.

R.7.5.6.3 Testing strategy for repeated dose toxicity

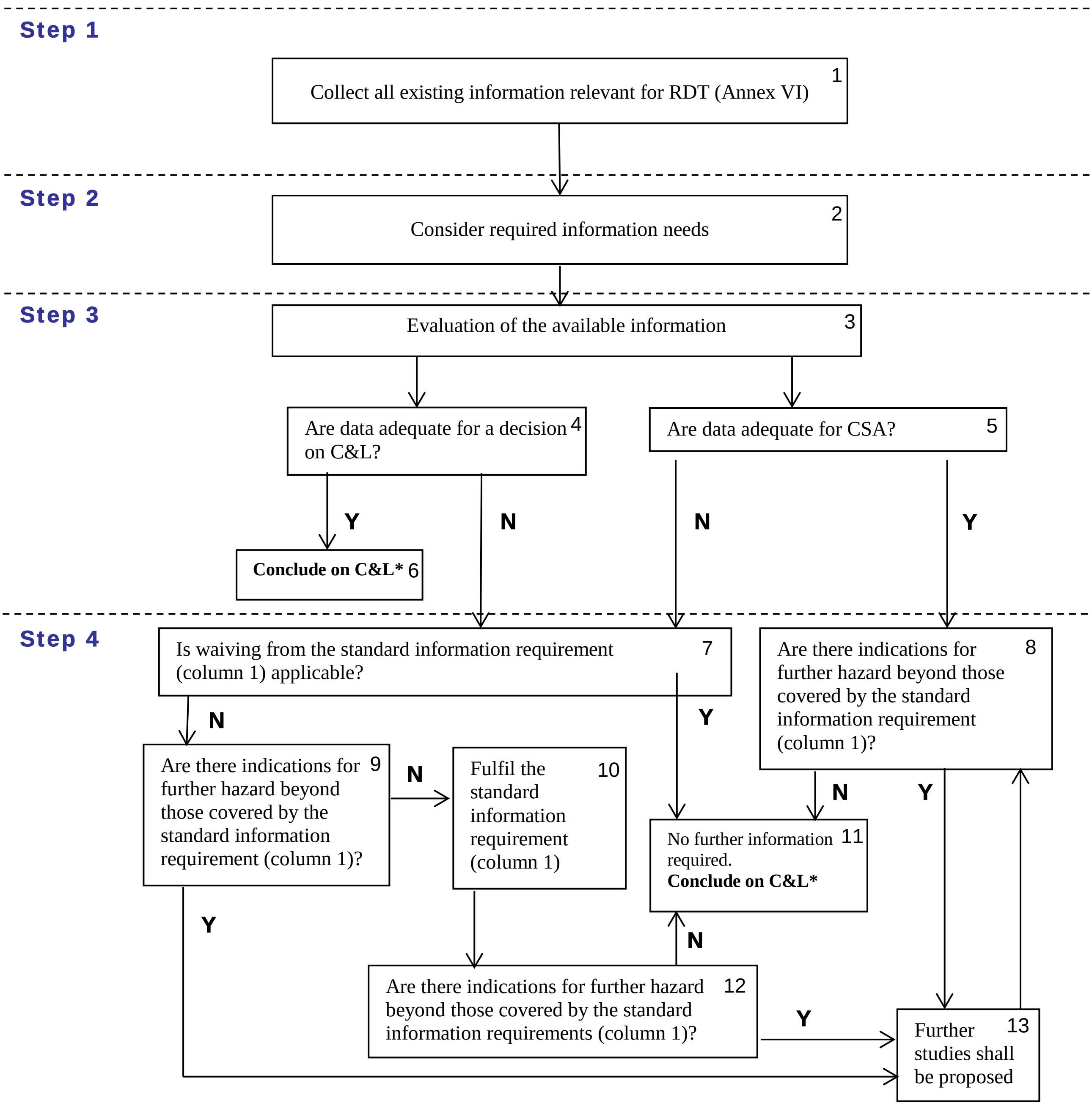
In order to proceed in further information gathering the following testing strategy is out-lined (step 4 in Figure R.7.5–1).

Before testing is initiated the available information should be scrutinised for evidence that may indicate severe effects, serious specific system or target organ toxicity (e.g. neurotoxicity or immunotoxicity), delayed effects or cumulative toxicity (boxes 8, 9 and 12 in Figure R.7.5–1). These indications may provide a trigger for specialised study protocols instead of the standard protocols for the short-term and/or (sub)chronic toxicity (box 13 in Figure R.7.5–1). These specific protocols should be designed on a case-by-case basis, such that they enable an adequate characterisation of these hazards, including the dose-response, threshold for the toxic effect and an understanding of the nature of the toxic effects. An example of such an approach is given in Appendix R.7.5–1.

Annexes VII-X of the REACH regulation provide the standard information requirements in Column 1 (box 10 of Figure R.7.5–1) and specify triggering and waiving possibilities for the specific endpoints in Column 2. Different descriptors used for repeated dose toxicity in these annexes varying from *limited* (Annex IX) to *no relevant exposure* (Annex VIII). In addition, Annex XI of the REACH regulation contains basic approaches, or rules for adaptation of the standard testing regime, set out in Annexes VII-IX (see Chapter R.5 of the *Guidance on IR&CSA*; for waiving see box 7 in Figure R.7.5–1).

Exposure considerations at this stage may trigger a need for additional data if the applications include wide dispersive uses to a large population (e.g. consumer products) and if a particular concern exists for a low margin of exposure (box 13 in Figure R.7.5–1). The data to be generated at this stage should aim to improve the risk quotient and could therefore be a trigger for an improved exposure characterisation or an improved hazard characterisation. In the latter case the required information might include a special study leading to an improved characterisation of the critical toxic endpoint thereby decreasing the uncertainty in the NOAEL for repeated dose toxicity. An example of such a testing approach applied to neurotoxicity is given in Appendix R.7.5–1.

Figure R.7.5–1 Integrated Testing Strategy for repeated dose toxicity



Utilisation of the different tests at each of the different tonnage levels is summarised below:

10 t/y or more (Annex VIII)

At this tonnage level a short-term (28-day) toxicity test (OECD TG 407/EU B.7) is usually required. The use of a combined repeated dose toxicity study with the reproduction/developmental toxicity screening test (OECD TG 422[74]) is recommended if an initial assessment of repeated dose toxicity and reproductive toxicity is required. The route of

[74] To date there is no corresponding EU testing method available.

exposure in these tests is oral unless the predominant route of human exposure or the physico-chemical properties indicate that the dermal or inhalational route may be a more appropriate route of exposure to assess the repeated dose toxicity test (requiring OECD TG 410 or 412/EU B.9 or B.8).

If the results of a short-term rodent toxicity study (OECD TGs 407; 410, 412, 422) are adequate for a dose response characterisation and C&L and risk assessment, and if there are no indications for further risks, no further testing is required (see Section R.7.5.5.2 for a detailed discussion of the criteria for a robust hazard characterisation).

At this tonnage level the short-term toxicity study (28 days) does not need to be conducted if:

- a reliable sub-chronic (90 days) or chronic toxicity study is available, provided that an appropriate species, dosage, and route of administration were used; or
- where a substance undergoes immediate disintegration and there are sufficient data on the cleavage products; or
- relevant human exposure can be excluded in accordance with Annex XI Section 3.

It should be noted that any of the rules for adaptation according to Annex XI also apply (see Chapter R.5 of the *Guidance on IR&CSA*). For further details see this section under Annex XI (below).

According to REACH (Annex IX, 8.6.2), the sub-chronic toxicity study (90 days) shall be proposed by the registrant if:

- the frequency and duration of human exposure indicates that a longer term study is appropriate;

and one of the following conditions is met:

- other available data indicate that the substance may have a dangerous property that cannot be detected in a short-term toxicity study; or
- appropriately designed toxicokinetic studies reveal accumulation of the substance or its metabolites in certain tissues or organs which would possibly remain undetected in a short-term toxicity study but which are liable to result in adverse effects after prolonged exposure.

REACH also specifies that further studies shall be proposed by the registrant or may be required by the Agency in accordance with Article 40 or 41 in case of:

- failure to identify a NOAEL in the 28 or the 90 days study, unless the reason for the failure to identify a NOAEL is absence of adverse toxic effects; or
- toxicity of particular concern (e.g., serious/severe effects); or
- indications of an effect for which the available evidence is inadequate for toxicological and/or risk characterisation. In such cases it may also be more appropriate to perform specific toxicological studies that are designed to investigate these effects (e.g., immunotoxicity, neurotoxicity); or
- the route of exposure used in the initial repeated dose study was inappropriate in relation to the expected route of human exposure and route-to-route extrapolation cannot be made; or

- particular concern regarding exposure (e.g. use in consumer products leading to exposure levels which are close to the dose levels at which toxicity to humans may be expected); or

- effects shown in substances with a clear relationship in molecular structure with the substance being studied, were not detected in the 28 or the 90 days study.

It should be pointed out that a failure to identify a NOAEL does not lead to a data gap in every case and should not trigger additional studies by default. If the data are sufficient for a robust hazard assessment and for Classification and Labelling, the LOAEL may be used as the starting point for the CSA (see also Sections R.7.5.4.4 and R.7.5.5 and Chapter R.8 of the *Guidance on IR&CSA*).

A specialised study is likely to be more appropriate for an improved hazard characterisation and should be considered instead of a standard short-term rodent or sub-chronic toxicity test at this stage.

100 t/y or more (Annex IX)

At this tonnage level, the following information is required (REACH Annex IX, Sections 8.6.1 and 8.6.2):

- a short-term study (28 day) in a single rodent species is the minimum requirement. The default route of exposure in these tests is oral (OECD TG 407/EU B.7; TG 422[75]) unless the predominant route of human exposure or the physico-chemical properties indicates that the dermal or inhalational route (OECD TG 410, 412/EU B.9, B.8) is a more appropriate route of exposure in the repeated dose toxicity tests.

- a sub-chronic toxicity study (90-day) in a single rodent species is usually required. The default route of exposure in these tests is oral (OECD TG 408/EU B.26) unless the predominant route of human exposure or the physico-chemical properties indicates that the dermal or inhalational route (OECD TG 411, 413/EU B.28, B.29) is a more appropriate route of exposure in the repeated dose toxicity tests.

According to REACH, at this tonnage level the sub-chronic toxicity study (90 days) does not need to be conducted if:

- a reliable short-term toxicity study (28 days) is available showing severe toxicity effects according to the criteria for classifying the substance as R48, for which the observed NOAEL-28 days, with the application of an appropriate assessment factor, allows the extrapolation towards the NOAEL-90 days for the same route of exposure; or

- a reliable chronic toxicity study is available, provided that an appropriate species and route of administration were used; or

- a substance undergoes immediate disintegration and there are sufficient data on the cleavage products (both for systemic effects and effects at the site of uptake); or

- the substance is unreactive, insoluble and not inhalable and there is no evidence of absorption and no evidence of toxicity in a 28-day limit test, particularly if such a pattern is coupled with limited human exposure;

[75] To date there is no corresponding EU testing method available.

It should be noted that any of the rules for adaptation according to Annex XI also apply. For further details see the section on Annex XI below.

In case human exposure is limited or different in frequency and duration from that used in the test protocol for repeated dose toxicity, the sub-chronic toxicity study may not be necessary if the data for the short-term toxicity study are adequate for a robust hazard characterisation, a risk assessment and classification and labelling. This adaptation requires full justification by the registrant.

In case the weight of the evidence indicates that the available information is adequate to characterise the short-term toxicity and sufficiently robust for proper dose-selection of the 90-day study, a dedicated 28-day study is not necessary at this stage.

No further testing is required if the available data, which may include a sub-chronic rodent toxicity study (OECD TG 408, 411, 413/EU B.26, B.28, B.29) are adequate for a dose response characterisation and C&L and risk assessment.

In case data are inadequate for hazard characterisation and risk assessment further studies shall be proposed by the registrant or may be required by the Agency in accordance with REACH Articles 40 or 41: According to REACH Annex IX Section 6.6.2 such a situation may arise if there is:

- failure to identify a NOAEL in the 90 days study unless the reason for the failure to identify a NOAEL is absence of adverse toxic effects; or
- toxicity of particular concern (e.g. serious/severe effects); or
- indications of an effect for which the available evidence is inadequate for toxicological and/or risk characterisation; In such cases it may also be more appropriate to perform specific toxicological studies that are designed to investigate these effects (e.g. immunotoxicity, neurotoxicity); or
- particular concern regarding exposure (e.g. use in consumer products leading to exposure levels which are high relative to the dose levels at which toxicity to humans occurs)

A specialised study is likely to be more appropriate for an improved hazard characterisation and should be considered instead of a standard short-term rodent or sub-chronic toxicity test. An example of such an approach given in Appendix R.7.5–1.

It should be pointed out that a failure to identify a NOAEL does not lead to a data gap in every case and should not be a default trigger for additional studies. If the data are sufficient for a robust hazard assessment or for Classification and Labelling, the LOAEL may be used as the starting point for the CSA (see also Sections R.7.5.4.4 and R.7.5.5 and Chapter R.8 of the *Guidance on IR&CSA*).

1000 t/y or more (Annex X)

There is no default testing requirement for repeated dose toxicity at this tonnage level beyond those recommended for the level 100 t/y or more (see above). However, in accordance with REACH Articles 40 and 41, if the frequency and duration of human exposure indicates that a long-term study is appropriate and one of the following conditions is met a long-term repeated toxicity test (≥12 months) may be proposed:

- serious or severe toxicity effects of particular concern were observed in the 28-days or 90-days study for which available evidence is inadequate for toxicological evaluation or risk characterisation; or

- effects shown in substances with clear relationship in molecular structure with the substance being studied were not detected in the 28-days or 90-days study; or
- the substance may have a dangerous property that cannot be detected in a 90-days study.

In addition, further studies shall be proposed by the registrant or may be required by the Agency in accordance with REACH Articles 40 or 41, in case of:

- toxicity of particular concern (e.g. serious/severe effects); or
- indications of an effect for which the available evidence is inadequate for toxicological evaluation and/or risk characterisation; In such cases it may also be more appropriate to perform specific toxicological studies that are designed to investigate these effects (e.g. immunotoxicity, neurotoxicity); or
- particular concern regarding exposure (e.g. use in consumer products leading to exposure levels which are close to the dose levels at which toxicity is observed).

In some cases a specialised study might the most appropriate study in case an improved hazard characterisation is necessary and should be considered instead of a standard sub-chronic or chronic toxicity test. An example of such an approach given in Appendix R.7.5–1.

No further testing is required if the results of a sub-chronic rodent toxicity study (OECD TG 408, 410, 411, 412, 413 or EU B.26, B.9, B.28, B.8, B.29) are adequate for a robust hazard characterisation and suitable for risk assessment and classification and labelling (see step 3 Identify data gaps for a detailed discussion of the criteria for a robust hazard characterisation).

Also, the testing requirements can be adapted if any of the rules according to REACH Annex XI apply: For further details see this Section under *REACH Annex XI* (below).

As there is no standard test requirement at this tonnage level, column 2 also had no waiving options.

REACH Annex XI adaptations of the standard testing regime for repeated dose toxicity

General guidance on the application of the Annex XI adaptations to information requirements is given in Chapter R.5 of the *Guidance on IR&CSA*. For repeated dose toxicity the following additional guidance applies.

Testing does not appear scientifically necessary

Some substances may be excluded from testing for repeated dose toxicity if it does not appear scientifically necessary (Annex XI Section 1). This might be the case for example if:

- a *Weight of Evidence* analysis demonstrates that the available information is sufficient for an adequate hazard characterisation, and a CSA where the exposure to the substance is adequately controlled;
- a substance is not bio-available *via* a specific route and possible local effects have been adequately characterised;
- the vapour pressure is sufficiently low that inhalational exposures are unlikely to be of significance, or if human exposure is limited to dusts or aerosols unlikely to be inhalable
- for substances belonging to a group or a category of substances that have a common functionality and/or breakdown products or sufficient information for a qualitative and

quantitative understanding of the toxicological properties, testing of all individual category members may not be necessary (Annex XI Section 1.5). The criteria for application of read-across for a category of substances and detailed guidance can be found in Sections R.4.3.2 and R.6.2 of the *Guidance on IR&CSA*.

Testing is technically not possible

There may also be cases where it is technically not possible to conduct a repeated dose toxicity test (Annex XI Section 2). This might be the case if

- The substance ignites in air at ambient conditions.
- The substance undergoes immediate disintegration. In such a case the information requirements for the cleavage products should be assessed following an approach similar to that outlined in this document.
- The substance is corrosive in the dose range of interest for the study. Also, for reasons of animal welfare such studies should be avoided.

Substance-tailored exposure-driven testing

Exposure considerations may also lead to adaptation of the testing requirements (Annex XI Section 3). This might be the case if:

Testing requirements may be adapted based on a substance-specific exposure-assessment according to Annex XI Section 3. In this case testing for short-term repeated dose toxicity (Annex VIII, 8.6.1) may be waived at the 10-100 tonnage level if relevant human exposure can be excluded (see Section R.7.5.4.3).

Human exposure is limited at the tonnage level of 100 t/y or more (Annexes IX and X). The need for a sub-chronic study should be considered if the substance is only handled in industrial or commercial installations using closed systems and/or handled only as preparations at low concentrations.

Appendix R.7.5-1 to Section R.7.5

Appendix R.7.5–1 Testing strategy for specific system / organ toxicity.

Content of Appendix R.7.5-1

1. General aspects
2. Structure-activity considerations
3. Assessment of available information or results from initial testing
4. Recommendations from the WHO/FAO Joint Meeting of Experts on Pesticide Residues (JMPR)
5. Further neurotoxicity testing

Mechanisms of respiratory irritation

1. General aspects

For some specific system/organ effects the testing methods of EU Annex V or the OECD may not provide for adequate characterisation of the toxicity. There may be indications of such effects in the standard studies for systemic toxicity, or from SAR. For adequate characterisation of the toxicity and, hence, the risk to human health, it may be necessary to conduct studies using other published test methods, *in-house* methods or specially designed tests. Some references are given in Error! Reference source not found.. Before initiating a study to investigate specific organ/system toxicity, it is important that the study design is presented to the Agency, in order that the need for (and scope/size of) studies using live animals should be particularly carefully considered.

Specific investigation of organ/systemic toxicity is to some extent undertaken as part of the repeated dose toxicity tests conducted according to test guidelines of the OECD and Annex V to Directive 67/548/EEC[76]. Specific investigation (or further investigation) of any organ/system toxicity (e.g. immune, endocrine or nervous system) may sometimes be necessary and should be addressed on a case-by-case basis. As an example of a testing strategy the approach for neurotoxicity is given below.

Definition of neurotoxicity

Neurotoxicity is the induction by a chemical of adverse effects in the central or peripheral nervous system, or in sense organs. It is useful for the purpose of hazard and risk assessment to differentiate sense organ-specific effects from other effects which lie within the nervous system. A substance is considered *neurotoxic* if it induces a reproducible lesion in the nervous system or a reproducible pattern of neural dysfunction.

The starting point for the testing strategy are the REACH requirements specified in Annex VIII, IX and X and detailed in Section R.7.5.6.3 Depending on the tonnage level, these requirements may trigger a 28-day and/or a 90-day test (e.g. OECD TG 407, 408/EU B.7, B.26). These

[76] All the test methods previously included in Annex V to Directive 67/548/EEC will be incorporated in a new Test Methods (TM) Regulation that is currently (February 2008) under adoption. The TM Regulation will be adapted to technical progress whenever a new test method has been developed, scientifically validated and accepted for regulatory use by the National Coordinators of the Member states

protocols include a number of nervous system endpoints (e.g. clinical observations of motor and autonomous nervous system activity, histopathology of nerve tissue), which should be regarded as the starting point for evaluation of a substance potential to cause neurotoxicity. It should be recognised that the standard 28-/90-day tests only measure some aspects of nervous system structure and function e.g. Functional Observational Battery, while other aspects, e.g. learning and memory and sensory function is not or only superficially tested. SAR considerations may prompt the introduction of additional parameters to be tested in standard toxicity tests or the immediate request of studies such as delayed neurotoxicity (OECD TG 418 or 419/EU B.37 or B.38,; see below).

If there are no indications of neurotoxicity from available information i.e. adequately performed repeated dose toxicity tests, other testing systems (e.g. *in vitro*), non-testing systems ((Q)SAR and read-across) or human data, it will not be necessary to conduct any special tests for neurotoxicity.

The approach presented below is a hierarchical, step-wise strategy to investigate the potential neurotoxicity of a substance. It should be pointed out that the requirements outlined in steps 1 and 2 are met by the tonnage-based information requirements in Annex VIII, IX and X of REACH.

2. Structure-activity considerations

Structural alerts are only used as a positive indication of neurotoxic potential. Substance classes with an alert for neurotoxicity may include organic solvents (for chronic toxic encephalopathy); organophosphorus compounds (for delayed neurotoxicity), and carbamates (for cholinergic effects). Several estimation techniques are available, one of which is the rule-based DEREK (Deductive Estimation of Risk from Existing Knowledge) system. The rulebase comprises the following hazards and structural alerts: Organophosphate (for direct and indirect anticholinesterase activity); N-methyl or N,N-dimethyl carbamate (for direct anticholinesterase activity); gamma-diketones (for neurotoxicity).

3. Assessment of available information or results from initial testing

Signs of neurotoxicity in standard acute or repeated dose toxicity tests may be secondary to other systemic toxicity or to discomfort from physical effects such as a distended or blocked gastrointestinal tract. Nervous system effects seen at dose levels near or above those causing lethality should not be considered, in isolation, to be evidence of neurotoxicity. In acute toxicity studies where high doses are administered, clinical signs are often observed which are suggestive of effects on the nervous system (e.g. observations of lethargy, postural or behavioural changes), and a distinction should be made between specific and non-specific signs of neurotoxicity.

Neurotoxicity may be indicated by the following signs: morphological (structural) changes in the central or peripheral nervous system or in special sense organs; neurophysiological changes (e.g. electroencephalographic changes); behavioural (functional) changes; neurochemical changes (e.g. neurotransmitter levels).

A *Weight of Evidence* approach should be taken into account for the assessment of the neurotoxicity and the type, severity, number and reversibility of the effect should be considered. A consistent pattern of neurotoxic findings rather than a single or a few unrelated effects should be taken as persuasive evidence of neurotoxicity.

It is important to ascertain whether the nervous system is the primary target organ. The reversibility of neurotoxic effects should also be considered. The potential for such effects to occur in exposed humans (i.e. the exposure pattern and estimated level of exposure are *acute*) should be considered in the risk characterisation. Reversible effects may be of high concern depending on the severity and nature of effect. In this context it should be kept in mind that effects observed in experimental animals that appear harmless might be of high concern in humans depending on the setting in which they occur (e.g. sleepiness in itself may not be harmful, but in relation to operation of machinery it is an effect of high concern). Furthermore the possibility that a permanent lesion has occurred cannot be excluded, even if the overt effect is transient. The nervous system possesses reserve capacity, which may compensate for the damage, but the resulting reduction in the reserve capacity should be regarded as an adverse effect. Irreversible neurotoxic effects are of high concern and usually involve structural changes, though, at least in humans, lasting functional effects (e.g. depression, involuntary motor tremor) are suspected to occur as a result of neurotoxicant exposure, apparently without morphological abnormalities.

For the evaluation of organophosphate pesticides, the WHO/FAO Joint Meeting of Experts on Pesticide Residues (JMPR) has published recommendations on "Interpretation of Cholinesterase Inhibition" (FAO, 1998; 1999). The applicability of these recommendations, outlined below, could also be extended to other substances that inhibit cholinesterase. It should be pointed out that for substances that may have a structural alert for cholinesterase inhibition, the measurement of acetylcholinesterase activity as recommended by JMPR can be included in the list of parameters for the standard 28- or 90 day testing protocols required by REACH, irrespective of the route of exposure.

4. Recommendations from the WHO/FAO Joint Meeting of Experts on Pesticide Residues (JMPR)

The inhibition of brain acetylcholinesterase activity and clinical signs are considered to be the primary endpoints of concern in toxicological studies on compounds that inhibit acetylcholinesterases. Inhibition of erythrocyte acetylcholinesterase is also considered to be an adverse effect, insofar as it is used as a surrogate for brain and peripheral nerve acetylcholinesterase inhibition, when data on the brain enzyme are not available. The use of erythrocyte acetylcholinesterase inhibition as a surrogate for peripheral effects is justified for acute exposures resulting in greater acetylcholinesterase inhibition in erythrocytes than in the brain. However, reliance on inhibition of erythrocytic enzyme in studies of repeated doses might result in an overestimate of inhibition on peripheral tissues, because of the lower rate of resynthesis of the enzyme in erythrocytes than in the nervous system. Plasma acetylcholinesterase inhibition is considered not relevant. Regarding brain and erythrocyte acetylcholinesterase inhibition, the experts defined that statistically significant inhibition by 20% or more represents a clear toxicological effect and any decision to dismiss such findings should be justified. JMPR also agreed on the convention that statistically significant inhibition of less than 20% or statistically insignificant inhibition above 20% indicate that a more detailed analysis of the data should be undertaken. The toxicological significance of these findings should be determined on a case-by-case basis. One of the aspects to consider is the dose-response characteristic.

5. Further neurotoxicity testing

If the data acquired from the standard systemic toxicity tests required by REACH provide indications of neurotoxicity which are not adequate for a hazard assessment, risk characterisation or classification and labelling, the nature of further investigation will need to be considered. If a 90-day study is triggered to meet the requirements of Annex IX following a standard 28-day study, a number of endpoints assessing the nervous system endpoints should be included,. irrespective of the administration route. In some cases, it may be necessary to conduct a specific study such as a neurotoxicity test using the OECD TG 424 with possible inclusion of a satellite group for assessment of reversibility of effects. The OECD TG 424 is intended for confirmation or further characterisation of potential neurotoxicity identified in previous studies. The OECD guideline allows for a flexible approach, in which the number of simple endpoints which duplicate those already examined during standard testing may be minimised, and where more effort is put into in-depth investigation of more specific endpoints by inclusion of more specialised tests. Adjustment of dose levels to avoid confounding by general toxicity should be considered.

If data from standard toxicity studies are clearly indicative of specific neurotoxicity, e.g. neurotoxicity occurring at lower dose levels than systemic toxicity, further specific neurotoxicity testing is required to confirm and extend the findings from the general toxicity studies and to establish an NOAEL for neurotoxicity. Again, the neurotoxicity test according to OECD TG 424 is considered appropriate for this situation.

Certain substances and/or certain effects are best investigated in particular species. Pyridine derivatives are neurotoxic to humans and primates but not to rats. Among other neurotoxic compounds, organophosphorus compounds are a group with known delayed neurotoxic properties, which need to be assessed in a specified test for delayed neurotoxicity, to be performed preferentially in the adult laying hen according to EU B.37 or OECD TG 418 (Delayed neurotoxicity of organophosphorus substances following acute exposure) and B.38 or OECD TG 419 (Delayed neurotoxicity of organophosphorus substances: 28-day repeated dose study). Such studies are specifically required for biocidal substances of similar or related structures to those capable of inducing delayed neurotoxicity. If anticholineesterase activity is detected, a test for response to reactivating agent may be required.

Standard exposure conditions may not always be adequate for neurotoxicity studies. The duration of exposure needed to induce specific neurotoxic effects in an animal experiment will depend on the underlying mechanism of action. Short-term peak exposures can be important for certain types of substance/effect. When the test compound is administered as a bolus *via* the intravenous, subcutaneous or oral route it is essential to determine the time-effect course, and to perform measurements of neurotoxicity parameters preferentially at the time of peak effect.

For example, the neurotoxicity associated with short-term exposure to some volatile organic solvents has largely been identified following human exposure - particularly occupational exposure. Acute inhalation studies, using protocols designed to detect the expected effects, are ideal for such substances/effects. For some neurotoxic substances a long exposure period is necessary to elicit neurotoxicity.

The most appropriate methods for further investigation of neurotoxicity should be determined on a case-by-case basis, guided by the effects seen in the standard systemic toxicity tests and/or from SAR-based predictions. Extensive coverage of methods which may be used is given in OECD (2004a), IPCS (1986) and ECETOC (1992), and some are summarised in Table R.7.5–3.

Table R.7.5–3 Methods for investigation of neurotoxicity

Effect	Methods available	References*
Morphological changes	Neuropathology. Gross anatomical techniques. Immunocytochemistry. Special Stains	Krinke, 1989; Odonoghue, 1989; Mattson et al., 1990
Physiological changes	Electrophysiology (e.g. nerve conduction velocity (NCV), Electroencephalogram (EEG), evoked potentials	Fox et al., 1982; Rebert, 1983; Mattson and Albee, 1988
Behavioural changes	Functional observations. Sensory function tests. Motor function tests (e.g. locomotor activity). Cognitive function tests	Robbins, 1997; Tilson et al., 1980; Cabe and Eckerman, 1982; Pryor et al., 1983 Moser and McPhail, 1990; Moser 1995
Biochemical changes	Neurotoransmitter analysis. Enzyme/protein activity. Measures of cell integrity.	Dewar and Moffet, 1977; Damstra and Bondy, 1982; Cooper et al., 1986; Costa, 1998.

*Given in full in ECETOC (1982), IPCS (1986) or Mitchell (1982)

R.7.5.7 References on repeated dose toxicity

Checkoway H, Pearce NE, Crawford-Brown DJ. Research methods in Occupational Epidemiology Oxford University Press 1989.

ECETOC (1992). Evaluation of the neurotoxic potential of chemicals. European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC), Monograph No. 18, Brussels.

ECETOC (2002). Recognition of, and Differentiation between, Adverse and Non-adverse Effects in Toxicology Studies. Technical Report No. 85, The European Centre for Ecotoxicology and Toxicology of Chemicals, December 2002.

ECETOC (2003). (Q)SARs: Evaluation of the commercially available software for human health and environmental endpoints with respect to chemical management applications. Technical Report No. 89, The European Centre for Ecotoxicology and Toxicology of Chemicals, September 2003.

FAO (1998). Pesticide Residues in Food. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide. Residues in Food and the Environment and the WHO Core Assessment Group.

FAO (1999). Plant Production and Protection Paper, No. 148, 17-19.

Hernberg S. Introduction to Occupational Epidemiology. Lewis Publishers 1991.

IPCS (1986). Principles and Test Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals. World Health Organisation (WHO), International Programme on Chemical Safety (IPCS), Environmental Health Criteria 60, Geneva.

Kroes, R., Renwick, AG, Cheeseman, M. et al., 2004. Structure-based thresholds of toxicological concerns (TTC): Guidance for application to substances present at low levels in the diet. Food Chem. Toxicol. 42, 65-83.

Mitchell CL (ed) (1982). Nervous System Toxicology. Raven Press, New York, NY.

OECD (2003) Descriptions of Selected Key Generic Terms used in Chemical Hazard/Risk

Assessment; OECD Series on Testing and Assessment: No 44. (ENV/JM/MONO(2003)15). 30 Oct 2003.

OECD (2004a). Guidance Document for Neurotoxicity Testing. Organization for Economic Cooperation and Development (OECD), Environment Directorate, OECD Environmental Health and Safety Publications, Series on Testing and Assessment No 20. Paris.

OECD (2006). Report on the regulatory uses and applications in OECD member countries of (Quantitative) Structure activity relationship (Q)SAR models in the assessment of new and existing chemicals. (ENV/JM/MOM(2006)25). 11 August 2006

OECD (2007a) Manual for Investigation of HPV Chemicals. http://www.oecd.org/document/7/0,2340,en_2649_34379_1947463_1_1_1_1,00.html

OECD (2007b) Guidance Document on the Validation of (Quantitative)Structure-Activity

Relationships [(Q)Sar] Models, Monograph No. 69, (ENV/JM/MONO(2007)2). 30 Mar 2007

Prieto P, Clemedson C, Meneguz A, Pfaller W, Sauer UG and Westmoreland C. (2005). 3.6 Subacute and subchronic toxicity. ATLA **33**, Suppl. 1, 109-116.

Prieto P, Baird AW, Blaauboer BJ, Castell Ripoll JV, Corvi R, Dekant W, Dietl P., Gennari A, Gribaldo A, Griffin JL, Hartung T, Heindel JJ, Hoet P, Jennings P, Marocchio L, Noraberg J, Pazos P, Westmoreland C, Wolf A, Wright J, Pfaller W. (2006). The Assessment of Repeated Dose Toxicity In Vitro: A Proposed Approach. The Report and Recommendations of ECVAM Workshop 56. ATLA. ATLA 34, 315-341.

Rennen MAJ, De Heer C and Houben GF (1999). Prediction of Local Effects upon Dermal and Respiratory Repeated Exposure to Chemicals. Organisation for Applied Scientific Research (TNO), TNO Nutrition and Food. Research Institute, TNO Report V98.1267, Zeist, The Netherlands.

Rothman KJ, Greenland S. Modern Epidemiology Lippincott-Raven 1998.

Swaen GMH. A framework for using epidemiological data for risk assessment. Human and Experimental Toxicology 2006; 25:147-155,

TemaNord 2005: 559Threshold of Toxicological Concern (TTC), Literature review and applicability ISBN 92-893-1196-7

WHO, 1990. IPCS Environmental Health Criteria 104; Pesticide Residues in food, principles for the toxicological assessment,World Health Organization, Geneva 1990

WHO, 1994. IPCS Environmental Health Criteria 170; Guidance values for human exposure limits, World Health Organization, Geneva 1994

WHO, 1999. IPCS Environmental Health Criteria 210, Principles for the assessment of Risks to human health from exposure to chemicals. World Health Organization, Geneva 1999

WHO 2004. IPCS Risk assessment terminology. World Health Organization, Geneva, 2004

Worth, A.P. & Balls, M., eds (2002). Alternative (non-animal) methods for chemicals testing: current status and future prospects. A report prepared by ECVAM and the ECVAM working group on chemicals. ALTA 30, Suppl. 1, 71.

R.7.6 Reproductive toxicity

R.7.6.1 Introduction

Reproductive hazards of chemicals are of obvious concern for the general population. Similarly, to the individual, an impairment of the ability to reproduce and the occurrence of developmental disorders are self-evidently serious health constraints. Therefore it is important that the potential hazardous properties and risks with respect to reproduction are established for substances. The REACH information requirements have two core objectives:

- to have adequate information in order to decide whether classification and labelling, including categorisation, as a reproductive toxicant is warranted;
- to have sufficient information for the purpose of risk assessment.

REACH information requirements for reproductive toxicity were amended in 2015[77] and the recitals of that amendment describe the motivation of the legislator. Recitals are considered a complementary part of the guidance aiming to allow a comprehensive understanding of the objectives of the legislation. Some of them are referred to in this guidance as necessary.

The terminology used in various legislation and in context related to reproductive toxicity differs. In this guidance document the term "reproductive toxicity" is used to cover both the effects on fertility and development. Fertility is seen as a broad concept covering all the effects on the reproductive cycle except for developmental toxicity. Development, referred to as "developmental toxicity" is defined in the text below.

In REACH, the Chemical Safety Report (CSR) format includes the terms "effects on fertility" and "developmental toxicity" under the main heading of "toxicity to reproduction". Also in other texts in REACH, such as in the REACH Annexes, reproductive toxicity is divided into fertility and developmental toxicity[78]. It is worth noting that in IUCLID the main heading for reproductive toxicity (7.8) is "Toxicity to reproduction", the subheading for fertility (7.8.1) is "Toxicity to reproduction" and the subheading for developmental toxicity (7.8.2) is "Developmental toxicity / teratogenicity".

In Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures (CLP Regulation), the term "*reproductive toxicity*" as defined in CLP Annex I, is used to describe the adverse effects induced (by a substance) on sexual function and fertility in adult males and females, the development of the offspring and adverse effects on or mediated *via* lactation. Thus, in the CLP Regulation, the differentiation within reproductive toxicity differs from the one stipulated in REACH, namely that lactation effects are considered separately. Hence, for the purpose of classification, reproductive toxicity is divided into three main differentiations, which relate to (i) impairment of male and female reproductive functions or capacity (fertility), (ii) the induction of non-heritable harmful effects on the progeny (developmental toxicity), and (iii) effects on or *via* lactation.

It is necessary to distinguish as far as possible effects on fertility and developmental toxicity for a substance and information on both types of effects is required by REACH above certain tonnage levels. The term "*fertility*" is used in the present guidance document instead of "*sexual function and fertility*" as explained above in order to follow the terminology used in REACH. The term "*sexual function and fertility"* is not used in REACH, however, in specific

---

[77] Commission Regulation (EU) 2015/282 of 20 February 2015 amending Annexes VIII, IX and X to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards to Extended one-generation reproductive toxicity study

[78] in Column 2 (see REACH Annexes VIII, IX and X, 8.7.1, Column 2).

places, where classification and labelling is discussed, "*sexual function and fertility*" is used as a hazard class in the same context as "*fertility*" used alone. It is to be noted that fertility (as a REACH endpoint) covers functional fertility, morphological and histological changes related to reproductive organs in males and females as well as the ability to produce offspring and to nurse them.

In the following text, endpoints for fertility and developmental toxicity are explained based on the description provided in the CLP Regulation. In practical terms, reproductive toxicity is characterised by multiple diverse endpoints, which relate to impairment of male and female reproductive functions or capacity (fertility), the induction of non-heritable harmful effects on the progeny (developmental toxicity), and effects on or *via* lactation.

Adverse effects on sexual function and fertility include any effect of a substance that has the potential to interfere with sexual function and fertility. This includes, but is not limited to, alterations to the female and male reproductive system, adverse effects on onset of puberty, gamete production and transport, reproductive (oestrus) cycle normality, sexual behaviour, fertility, gestation length, parturition, pregnancy outcomes, premature reproductive senescence, or modifications in other functions that are dependent on the integrity of the reproductive system.

Developmental toxicity includes, in its widest sense, any effect interfering with normal development of the organism, before or after birth and resulting from exposure of either parent prior to conception, or exposure of the developing organism during prenatal development, or postnatal development, to the time of sexual maturation – thus generally speaking, these effects can be manifested at any point in the life span of the organism. However, it is considered that classification under the heading of developmental toxicity is primarily intended to provide a hazard warning for pregnant women, and for men and women of reproductive capacity. The major manifestations of developmental toxicity include (1) death of the developing organism, (2) structural abnormality, (3) altered growth, and (4) functional deficiency.[79]

This guidance provides advice on how the registrant can address the reproductive toxicity of the substance and how the information requirements of REACH can be met, thereby providing data on the hazardous properties that can be used for classification purposes and in the risk assessment.

R.7.6.2 Information requirements and testing approaches for reproductive toxicity

Article 10 of REACH specifies the information that is to be submitted for general registration purposes. This information includes minimum information requirements on physicochemical, toxicological and ecotoxicological properties, which are dependent on the tonnage of the registration (Article 10(a) (vi) and (vii) read with Article 12(1) of REACH).

The standard information requirements for the lowest tonnage level are given in Annex VII of REACH. Whenever a higher tonnage level is reached, the minimum requirements of the corresponding REACH Annex (i.e. the REACH Annex for the higher tonnage level) have to be fulfilled in addition to those in all preceding REACH Annexes (see Annex VI of REACH).

For reproductive toxicity, as for any endpoint, all available information must be collected, including data from literature searches. This should then be evaluated with regard to its

[79] As written in 3.7.1.3 and 3.7.1.4 in Annex I to CLP (the definition for developmental toxicity is shortened here).

reliability and relevance, and whether it fulfils the information requirements and their adaptations (triggers and waivers), as well as its use for the purpose of classification, risk assessment and risk management measures.

R.7.6.2.1 REACH information requirements

To examine effects on reproduction, REACH requires information on fertility and developmental toxicity via the "standard information requirements" which are specified in Column 1 of the respective REACH Annexes.

These standard information requirements are minimum information requirements. If there are concerns ("triggers" or "conditions") further testing might be needed to assure availability of appropriate information for chemical safety assessment (including risk characterisation, classification and labelling and other risk management measures).

The term "triggers" is used here as a general term instead of various other possible terms (such as alert, condition, indication, indication of concern, serious concern, or a particular concern) which are used in the REACH Regulation and some of which are used in this Guidance document as described below. A discussion on the evaluation of triggers is given in Appendix R.7.6–5 of this Guidance. For clarification purposes when reading this Guidance document, the terms are used as follows:

- triggers: general term covering all other terms describing findings/conditions which raise concerns;
- alerts: previous term used in this guidance; means the same as triggers but may also include aspects relating to waiving;
- conditions: a specific term used e.g. in REACH Annex IX/X for triggering the extension of Cohort 1B, and which includes aspects which are not findings.

Certain specific adaptation rules described in Column 2 for reproductive toxicity specify when further testing is needed or may be needed at that tonnage level.

REACH information requirements can also be fulfilled by adaptations that reduce the requirement for testing. Adaptation possibilities are either specified in Column 2 of the information requirement or in REACH Annex XI.

An approach on how to fulfil the information requirements is presented in Section R.7.6.2.3 "*Adaptation and testing approaches*" of this Guidance.

The information requirements specified in Column 1 (standard information requirements) are generally cumulative with increasing tonnage levels. Column 2 adaptations are linked with the corresponding Column 1 requirement in the respective REACH Annex and should be considered together with the Column 1 requirement. For reproductive toxicity the standard information requirements (Column 1) combined with specific Column 2 adaptations that require different or further testing are as follows:

REACH Annex VIII (applicable for any registration of 10 tonnes or more per year)

- *Screening for reproductive/developmental toxicity*[80]*, one species (OECD TGs 421 or 422*[81]*) if there is no evidence from available information on structurally related substances, from (Q)SAR estimates or from in vitro methods that the substance may be a developmental toxicant;*

If there are serious concerns about the potential for adverse effects on fertility or development, the registrant may propose:

an extended one-generation reproductive toxicity study (B.56 of the Commission Regulation on test methods as specified in Article 13(3) or OECD TG 443) if there are serious concerns about the potential for adverse effects on fertility or peri-postnatal development;

or

a prenatal developmental toxicity study (B.31 of the Commission Regulation on test methods as specified in Article 13(3) or OECD TG 414) if there are serious concerns about the potential for adverse effects on prenatal development[82];

REACH Annex IX (applicable for any registration of 100 tonnes or more per year)

- *Prenatal developmental toxicity study, one species, most appropriate route of administration, having regard to the likely route of human exposure*[82] *(B.31 of the Commission Regulation on test methods as specified in Article 13(3) or OECD TG 414);*

and if Column 2 of REACH Annex IX, Section 8.7.2 applies for a second species:

- *Prenatal developmental toxicity study, second species (B.31 of the Commission Regulation on test methods as specified in Article 13(3) or OECD TG 414);*
- *Extended one-generation reproductive toxicity study (B.56 of the Commission Regulation on test methods as specified in Article 13(3) or OECD 443), basic test design (cohorts 1A and 1B without extension to include a F2 generation), one species, most appropriate route of administration, having regard to the likely route of human exposure*[12]*, if the available repeated dose toxicity studies (e.g. 28-day or 90-day studies, OECD TGs 421 or 422 screening studies) indicate adverse effects on reproductive organs or tissues or reveal other concerns in relation with reproductive toxicity.*

see REACH Annex IX, Section 8.7.3, Column 2 for the triggers (conditions) when to extend the Cohort 1B to mate the F1 animals and produce the F2 generation, and the triggers (conditions) when to include the Cohorts 2A/2B and/or Cohort 3. For further information on the study design see Appendix R.7.6–2 of this Guidance.

and if Column 2 of REACH Annex IX, Section 8.7.3 applies for a second species/strain:

- *Extended one-generation reproductive toxicity study on a second strain or a second species (exceptional cases only).*

---

[80] Later referred also as a screening study.

[81] To date there are no corresponding EU test methods available.

[82] It is strongly recommended that the registrant considers conducting a screening study in addition to the prenatal developmental toxicity study to cover the fertility and early peri/post natal development if an extended one-generation reproductive toxicity study is not conducted.

It should be noted that regarding the requirement of a second species, the EU B.56, OECD TG 443 prefers the rat and notes that if another species is to be used, justification should be given and appropriate modifications to the protocol will be necessary. There is currently (at the time of publication July 2015), still very limited experience of the protocol and only in rats. This will of course change in the future and registrants should check for new protocols and updates. It is stated in the OECD TG 443 paragraph 9 that "*When a sufficient number of studies are available to ascertain the impact of this new study design, the Test Guideline will be reviewed and if necessary revised in light of experience gained*."

REACH Annex X (applicable for any registration of 1000 tonnes or more per year)

- *Developmental toxicity study, one* [additional] *species, most appropriate route of administration, having regard to the likely route of human exposure (OECD TG 414);*
- *Extended one-generation reproductive toxicity study (B.56 of the Commission Regulation on test methods as specified in Article 13(3) or OECD 443), basic test design (cohorts 1A and 1B without extension to include a F2 generation), one species, most appropriate route of administration, having regard to the likely route of human exposure, unless already provided as part of REACH Annex IX requirements.*

see REACH Annex X, Section 8.7.3, Column 2 for the triggers (conditions) when to extend the Cohort 1B to mate the F1 animals and produce the F2 generation, and the conditions when to include the Cohorts 2A and 2B and/or Cohort 3. For further information on the study design see Appendix R.7.6–2 of this Guidance.

and if Column 2 of REACH Annex IX, Section 8.7.3. applies for a second species/strain:

- *Extended one-generation reproductive toxicity study on a second strain or a second species, in exceptional cases if not already provided as part of REACH Annex IX requirements. (for further explanation see REACH Annex IX above).*

A simplified summary of the information requirements for reproductive toxicity is presented in the following

Table R.7.6–1. The standard information requirements of REACH Annexes VIII to X, Section 8.7 Column 1 are indicated, combined with specific Column 2 adaptations that require different or further testing.

Table R.7.6–1 Summary of information requirements for reproductive toxicity in REACH (Annexes VII to X).

Study	Annex VII (<10 t/yr)	Annex VIII (≥10 t/yr)	Annex IX (≥100 t/yr)	Annex X (≥1000 t/yr)
Screening test for reproductive /developmental toxicity (OECD TGs 421 or 422)		Required. If a prenatal developmental toxicity study is available or proposed, it is strongly recommended to consider conducting a screening study in addition to the prenatal developmental toxicity[1] study. If an extended one-generation reproductive toxicity study is available or is proposed, a screening study may not need to be conducted.	Strongly recommended if no higher tier study (such as OECD TG 443) is/will be available to address fertility and peri/post natal development	(a higher tier study is required)
Prenatal developmental toxicity study (EU B.31, OECD TG 414)		May be proposed in cases of serious concern[2] for prenatal developmental toxicity instead of the screening study.	Required in one species; second species may be triggered[3]	Required in two species
Extended one-generation reproductive toxicity study (EU B.56, OECD TG 443)[4]		May be proposed in cases of serious concern for fertility instead of the screening study[2]	Required in one species if triggered[5]; second species/strain may be triggered in exceptional cases	Required in one species unless already conducted at previous Annex level; second species/strain may be triggered in exceptional cases

NOTES for Table R.7.6-1

[1] See discussion at Stage 4.3 (i) Reproduction/developmental toxicity screening test under Section R.7.6.2.3.2 of this Guidance.

[2] Column 1 and Column 2 provisions at REACH Annex VIII, 8.7.1 need to be considered together. Serious concern reflects a high likelihood for adverse effects on reproductive health.

[3] For discussion on triggers see Stage 4.4 (ii), Prenatal developmental toxicity study under Section R.7.6.2.3.2 of this Guidance.

[4] Basic study design addressing fertility, and developmental toxicity effects manifested after birth, with Cohort 1A and Cohort 1B without extension of Cohort 1B, see Stage 4.4 (iii) and Stage 4.5 (ii) Extended one-generation reproductive toxicity study of this Guidance under Section R.7.6.2.3.2 for an overview and Appendix R.7.6–2 and Appendix R.7.6–3 for details and when the study needs to be expanded.

[5] For description of triggers see Stage 4.4 (iii), Extended one-generation reproductive toxicity study under Section R.7.6.2.3.2 of this Guidance.

R.7.6.2.2 Key objectives and information produced by the test methods referred to in REACH

Key objectives and information produced by the test methods referred to in the REACH Regulation for reproductive toxicity are explained in short below in the text and in Table R.7.6–2. More information on how these studies are to be used in a REACH context and important aspects to consider during planning and evaluation are described in Section R.7.6.4.2 of this Guidance.

REACH Annex IX and REACH Annex X level studies and other studies considered not to be screening level studies, require a testing proposal.

R.7.6.2.2.1 Reproduction / Developmental Toxicity Screening Test

The purpose of the reproduction/developmental toxicity screening tests (OECD TGs 421 and 422) is to provide initial information of the effects on male and female reproductive performance such as gonadal function, mating behaviour, conception and parturition and histopathological information on reproductive organs. Initial information on the offspring is limited to mortality, abnormal behaviour and body weight of pups after birth, a macroscopic examination and additional parameters for endocrine disrupting modes of action as given in the revised TGs (2015)[83]. These screening tests are not meant to provide complete information on all aspects of reproduction and development.

R.7.6.2.2.2 Prenatal developmental toxicity study

The prenatal developmental toxicity study (EU B.31, OECD TG 414) provides a focused evaluation of potential effects following prenatal exposure, although only effects that are manifested before birth can be detected. More specifically, this study is designed to provide information on substance-induced effects on growth and survival of the foetuses, and increased incidences in external, skeletal and soft tissue malformations and variations in foetuses.

R.7.6.2.2.3 Extended one-generation reproductive toxicity study

The extended one-generation reproductive toxicity study (EOGRTS, EU B.56, OECD TG 443) allows evaluation of effects of the test substance on the integrity and performance of the adult male and female reproductive system, prenatal effects manifested postnatally and postnatal effects of substances on development as well as a thorough evaluation of systemic toxicity in pregnant and lactating females and young and adult offspring. The study also includes certain parameters for endocrine disrupting modes of action. The extended one-generation reproductive toxicity study is a modular study design with various investigational options.

[83] OECD TGs 421 and 422 are in the process of being revised: adoption and publication is expected by the end of 2015.

The basic study design, which is the standard information requirement at REACH Annexes IX and X[84], focuses on evaluation of the fertility of parental animals (F0 animals) and of defined parameters on postnatal development of F1 animals until adulthood (see the test method, EU B.56, OECD TG 443). The basic study design does not include mating of F1 animals (extension of Cohort 1B) or cohorts for developmental neurotoxicity (Cohorts 2A and 2B) or developmental immunotoxicity (Cohort 3). Conditions for triggering extension of Cohort 1B and Cohorts 2 and 3 are adaptations to the standard information requirement, and must be proposed by the registrant if the triggers (conditions) described in Column 2 are met. A check list for information that should be presented in the dossier in order to establish the existence or the nonexistence of the conditions and triggers specifying the study design for an extended one-generation reproductive toxicity study regarding the extension of Cohort 1B, inclusion of Cohort 2 and/or Cohort 3 is provided in Appendix R.7.6–1 of this Guidance. More detailed information and examples of triggers and conditions for extension of Cohort 1B and the need to include Cohort 2 and/or Cohort 3, are presented in Appendix R.7.6–2 of this Guidance.

The focus of the study in the REACH Annexes is on fertility[85], which should be considered in the study design of the extended one-generation reproductive toxicity study. Thus, as a starting point, a ten-week premating exposure duration and a highest dose level with the aim to induce some toxicity for all variant study designs of an extended one-generation reproductive toxicity study should be proposed. However based on substance specific justifications the premating exposure duration may be shorter than ten weeks but should not be shorter than two weeks (see Appendix R.7.6–3 of this Guidance). Regarding the highest dose level, it is important to ensure that toxicity in both female and male animals is considered to ensure that reproductive toxicity in either gender is not overlooked.

The extension of the Cohort 1B (mating of the Cohort 1B animals to produce the F2 generation) provides information on the fertility of the offspring, (i.e. the F1 generation), which has been exposed already during primordial germ cell and germ line formation, pre-implantation, *in utero* and postnatal periods. The fertility of Cohort 1B animals, if mated, is evaluated after exposure of full spermatogenesis.

Cohorts 2A and 2B provide information on developmental neurotoxicity and Cohort 3 on developmental immunotoxicity; this information is not covered by any other study within REACH requirements, but might be useful for further hazard and risk assessment.

---

[84] Recital (6) of Commission Regulation (EU) 2015/282 of 20 February 2015 amending Annexes VIII, IX and X to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards the Extended one-generation reproductive toxicity study: “*The standard information requirement in Annexes IX and X to Regulation (EC) No 1907/2006 should be limited to the basic configuration of EOGRTS. Nevertheless, in certain specific cases, where justified, the registrant should be able to propose and the European Chemicals Agency (ECHA) should be able to request the performance of the F2 generation, as well as the DNT and DIT cohorts.”.*

[85] Recital (7) of Commission Regulation (EU) 2015/282 of 20 February 2015 amending Annexes VIII, IX and X to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards the Extended one-generation reproductive toxicity study: “*It should be ensured that the reproductive toxicity study carried-out under point 8.7.3 of Annexes IX and X to Regulation (EC) No 1907/2006 will allow adequate assessment of possible effects on fertility. The premating exposure duration and dose selection should be appropriate to meet risk assessment and classification and labelling purposes as required by Regulation (EC) No 1907/2006 and Regulation (EC) No 1272/2008 of the European Parliament and of the Council.*”

Table R.7.6–2 Overview of *in vivo* EU test methods and OECD test guidelines for reproductive toxicity referred to in REACH

Test	Design	Focus of examination
Reproduction/ developmental toxicity screening test (OECD TGs 421 and 422)	Exposure from 2 weeks prior to mating (P) until a specified post-natal day (F1) 3 dose levels plus control Preferred species rat Preferred route oral[1] N = 10 mating pairs per dose group	*Parental (P) generation:* Growth, survival, fertility (limited) Pregnancy length and litter size Histopathology and weight of reproductive organs Histopathology and weight of major non-reproductive organs (OECD TG 422 only) *Offspring (F1):* Growth and survival until a specified post-natal day Certain parameters for endocrine modes of action.[86]
Prenatal developmental toxicity study (EU B.31, OECD TG 414)	Maternal exposure at least from implantation to one or two days before expected delivery 3 dose levels plus control Preferred species rat and rabbit Preferred route oral[1] N = 20 pregnant females per dose group	*Maternal animals:* Growth, survival, (effects on implantation only if dosing is started before implantation), maintenance of pregnancy *Offspring:* Resorptions, foetal deaths foetal growth Morphological variations and malformations (external, skeletal and visceral)
Extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) REACH requires a "basic study design" with a focus on fertility and defines specific conditions for the extension of Cohort 1B and/or inclusion of Cohorts 2A and 2B and/or Cohort 3 (see Section R.7.6.4.2.3 and Appendix R.7.6–2 of this Guidance)	Exposure of 10 weeks prior to mating[2] (P) until post-natal day 90-120 (Cohorts 1A and 1B). If the extension of Cohort 1B is triggered, then until post-natal day 4 or 21 (F2)[3]. 3 dose levels plus control; highest dose level must be chosen with the aim to induce some toxicity. Preferred species rat Preferred route oral[1] N = sufficient mating pairs to produce 20 pregnant animals per dose group (P generation)	*Parental (P) generation:* Growth, survival, fertility Oestrus cyclicity and sperm quality Pregnancy length and litter size Histopathology and weight of reproductive and non-reproductive organs Haematology and clinical chemistry *Offspring (F1):* Growth, survival and sexual maturation Histopathology and weight of reproductive and non-reproductive organs (Cohort 1A) Weight of reproductive organs and optional histopathology (Cohort 1B) Haematology and clinical chemistry

[86] OECD TGs 421 and 422 are in the process of being revised: adoption and publication is expected by the end of 2015.

Test	Design	Focus of examination
	N = 20 mating pairs (extension of Cohort 1B, if triggered) N = 10 males and 10 females per dose group (Cohorts 2A, 2B and 3, if triggered)	Fertility of F1 animals to produce F2 generation (extension of Cohort 1B) under certain conditions Developmental neurotoxicity (Cohorts 2A and 2B or a separate study) in cases of a particular concern Developmental immunotoxicity (Cohort 3 or a separate study) in cases of a particular concern Certain parameters for endocrine modes of action

NOTES for Table R.7.6-2

[1] See Stage 4.1 (iv) for discussion on route of administration (Section R.7.6.2.3.2 of this Guidance).

[2] Unless data to support a shorter pre-mating period (see discussion in Appendix R.7.6–3 of this Guidance).

[3] According to the test method EU B.56 (OECD TG 443) the F2 generation may be terminated on postnatal day 4 or 21. For further details see Section R.7.6.4.2.3 of this Guidance, under “Further aspects”.

R.7.6.2.3 Adaptation and testing approaches

R.7.6.2.3.1 Overview

This section describes how to use testing approaches and adaptations to achieve the core objectives of REACH (to fulfil information requirements for adequate risk assessment and classification and labelling purposes) with effective use of the gathered information and for designing potential actions needed to fulfil information requirements and to ensure the safe use of substances.

While Column 1 describes the standard information requirements, Column 2 sets certain rules if further or different information is triggered or if information may be omitted thus, Column 2 specific adaptation rules should be considered together with Column 1 standard information requirements. Adaptation may mean further or less information needs than specified in Column 1. If, where the specific adaptation rules in Column 2 or general adaptation rules in REACH Annex XI are not met, the standard information requirements must be fulfilled.

The Registrant is guided in a step-by-step tiered manner on how to meet the information requirements within the production tonnage and influenced by triggers (or conditions). These may increase the need for information or conditions which may allow adaptation of standard information requirements by means of replacing, omitting or adapting in another way. Adaptations of information requirements always need to be clearly stated and supported by adequate justification demonstrating the fulfilment of applicable conditions established by REACH.

As an initial step, all available information relevant to reproductive toxicity must be collected for substances manufactured or imported at tonnage levels ≥1 t/y (REACH Annexes VII-X)(see REACH Annex VI, Step 1). Information from literature may assist identifying the presence or absence of hazardous properties of the substance. In addition, information on exposure, uses and risk management measures should be collected. This information needs to be evaluated with regard to relevance and reliability and to decide if it is adequate for the purposes of risk assessment and classification for reproductive toxicity, including a comparison with the criteria for classification (Annex I of CLP); (see also the *Guidance on the Application of the CLP criteria* and *Guidance on IR&CSA* Chapter R.3 on Information gathering and Chapter R.4 on Evaluation of available information). Considering all the information together, the registrant will be able to determine the need to generate further information in order to fulfil the information requirements.

Consistent with the information requirements defined within REACH Annexes VII to X, testing for reproductive toxicity is not required as a standard approach for registrations of chemicals for the manufacture or import at tonnage levels below 10 tonnes per year (REACH Annex VII). At higher production volumes (i.e. ≥10 t/y, ≥100 t/y or ≥1000 t/y), standard information requirements are staggered according to tonnage levels of the registrations. Flexibility to adopt the most appropriate testing regime for any single substance is maintained by using adaptation rules provided by Column 2 and REACH Annex XI. The adaptation rules are the key components of the testing approaches.

However, regardless of tonnage level, before any testing is carried out careful consideration by the registrants of the following is required: all the available toxicological data, the classification for reproductive toxicity, carcinogenicity and germ cell mutagenicity (EU harmonised or self-classification), human exposure characteristics and current risk management procedures; these are necessary to ascertain whether the information requirements can already be met (see the Guidance on IR&CSA Chapter R.5 on *Adaptation of information requirements*). If it is concluded that testing is required in order to fulfil the information requirements, for reasons such as triggers, data gaps which cannot be adapted (for the purpose of classification and/or risk assessment), or increases in production volumes resulting in an REACH Annex upgrade. A series of decision points are defined and described below to help shape the scope of an appropriate testing programme. The REACH approach provides a four-stage process for clear decision-making, relevant for all tonnage levels.

Stage 1: Consider hazardous CMR properties meeting the classification criteria to Category 1A or 1B to decide on the need for further reproductive toxicity testing. Based on Column 2 adaptation of Section 8.7 in REACH Annexes further information on reproductive toxicity may be omitted in certain conditions described in Column 2. Therefore, dependent on the outcome of this analysis, it is possible that some chemicals may not progress beyond Stage 1.

Stage 2: Clarify the standard information requirements relevant for manufactured/imported tonnage level of a single registrant or a SIEF[87].

Stage 3: Evaluate the available toxicology database and consider reproductive toxicity findings and conditions that may serve as triggers or allow omitting further studies. This evaluation should also consider information from substances with a similar structure or causing toxicity via similar mechanisms/modes of action. The aim of this stage is to ensure that the applicable REACH information requirements are identified and to determine the scope of the reproductive toxicity testing necessary to adequately clarify the reproductive toxicity properties. Following this review in conjunction with the analysis in Stage 1 or if sufficient data for risk assessment/risk management and classification purposes are available allowing adaption based on Column 2 or REACH Annex XI adaptation rules, it is possible that no further testing may be necessary.

If the specific adaptation rules in Column 2 or general adaptation rules in REACH Annex XI are not met, the standard information requirements must be fulfilled. Thus, any scientific or other substance-specific justifications for adaptation must follow Column 2 or REACH Annex XI adaptation rules.

Stage 4: Plan and conduct a screening study or plan and propose a prenatal developmental toxicity study or an extended one-generation reproductive toxicity study or specific other

---

87 SIEF is a substance information exchange forum.

studies in exceptional cases. In accordance with Article 12.1d-e/Article 22.1h of REACH, a testing proposal must be submitted to ECHA.

R.7.6.2.3.2 Procedure for adaptations and testing approaches

Collection of data

At all REACH Annex levels, the available information from human, animal and non-animal studies and testing approaches need to be collected, including data from literature searches which needs to be evaluated and documented (see REACH Annex I, Step 1 of REACH).

Stage 1: Genotoxic carcinogenicity, germ cell mutagenicity and reproductive toxicity (CMR- properties) to be considered before deciding whether any testing for reproductive toxicity potential is required (relevant for all tonnage levels)

If the answer at the Stage 1.1 and/or Stage 1.2 is yes, i.e. the substance has been already classified to Category 1 for any of the CMR property (as described below), no further testing for reproductive toxicity may be needed if the conditions are fulfilled and appropriate risk management measures are in place.

Stage 1.1

Has the substance already been classified[88] for effects on sexual function and fertility *and* developmental toxicity (Reproductive toxicity Category 1A or 1B (H360FD))?

If the answer is no, proceed to Stage 1.2: if the answer is yes and the available data are adequate to support a robust risk assessment, then no further testing may be necessary. However, if the substance is classified for fertility only, further testing for developmental toxicity must be considered and if the substance is classified for developmental toxicity only, further testing for fertility must be considered; then proceed to Stage 2 via Stage 1.2. If the available data are not adequate to support a robust risk assessment then proceed to Stage 2.

Stage 1.2

Is the substance known to be[89] a genotoxic carcinogen (Carcinogenicity Category 1A and at least Germ cell mutagenicity Category 2; or Carcinogenicity Category 1B and at least Germ cell mutagenicity Category 2) or as a germ cell mutagen (Germ cell mutagenicity Category 1A or 1B) and appropriate risk management measures are implemented?

If the answer is no, proceed to Stage 2. If the answer is yes, it is important to establish that appropriate risk management measures addressing potential carcinogenicity, genotoxicity and reproductive toxicity have been implemented and therefore further specific testing for reproductive and/or developmental toxicity will not be necessary.

Stage 2: Clarify the standard information requirements

At this stage it is necessary to understand what the standard information requirements are at the tonnage level relevant to the registrant. The registrant must fulfil the standard information requirements unless the Column 2 or REACH Annex XI adaptions rules are met to omit the study. In addition to standard information requirements presented in Column 1, Column 2 adaptation rules may indicate triggers (or conditions) for further studies or if certain study design must be proposed.

---

[88] Harmonised classification or self-classification meeting the classification criteria.

[89] Harmonised classification or self-classification meeting the classification criteria.

Stage 3: Conduct a detailed review of the available relevant toxicological data to identify conditions to adapt standard information requirements for reproductive toxicity

At Stage 3, the available relevant data is examined to verify if any of the adaptations rules beyond "CMR classification adaptations" explained at Stage 1 are met. Adaptation rules may allow omitting the study or indicate when further information may be needed or must be proposed.

Before any testing is conducted, a thorough data review should be conducted.

Following the adaptation based on CMR classification considered in Stage 1, further general adaptation possibilities of REACH Annex XI and specific adaptation possibilities for omitting the testing provided in Column 2 of the REACH Annexes should be explored. These adaptation rules are described in Stage 3.1 in Appendix R.7.6–4 of this Guidance. These adaptation rules apply to substances for which standard information requirements apply because they passed the Stage 1.

It is important to consider both Column 2 and REACH Annex XI adaptation possibilities because new tests on vertebrates must only be conducted or proposed as a last resort when all other data sources have been exhausted (REACH Annex VI, Step 4).

If sufficient data are available to permit an adaptation according to Column 2 and/or REACH Annex XI rules, then no further testing is required. If the rules for adaptation according to Column 2 or REACH Annex XI are not met and there is a data gap, then the testing strategy for reproductive and/or developmental toxicity in Stage 4 should be followed.

Standard information requirements are described in Column 1 at each REACH Annex. At REACH Annex IX, if there are triggers for reproductive toxicity (fertility and postnatal development) an extended one-generation reproductive toxicity study must be proposed. For definition of triggers and how to evaluate them, see Appendix R.7.6–5 of this Guidance. The examples for triggers for an extended one-generation reproductive toxicity study at REACH Annex IX are described in this Section, under Stage 4.4 (iv), extended one-generation reproductive toxicity study.

If the data are insufficient, which study (or studies) is most appropriate? This decision must take account of both the tonnage-related standard information requirements, the nature of the trigger(s) and total assessment of data.

REACH standard information requirements are minimum information requirements and triggers for reproductive toxicity may indicate a need for further information. Where there is an information gap that needs to be filled, new data must be generated (REACH Annexes VII and VIII) or a testing approach must be proposed (REACH Annexes IX and X). Note that other data sources need to be explored and new tests on vertebrates must only be conducted or proposed as a last resort when all other data sources have been exhausted (REACH Annex VI, Step 4). Whether the registrant must or should or may propose/conduct further information beyond the standard information requirements depends on the REACH Annex level and the provisions in Column 2 and any further concerns. These are further explained at Stage 3.2 and Appendix R.7.6–5 of this Guidance.

Stage 3.1 Substances for which the standard information requirements apply after Stage 1 – options for adaptation rules which may apply instead of conducting new studies

These are substances which are not classified as Category 1 for CMR properties as described in Stage 1 (i.e. are not genotoxic Category 1 carcinogens, germ cell Category 1 mutagens or Category 1 reproductive toxicants (fertility and development)). See Appendix R.7.6–4 of this Guidance for details of adaptation possibilities for these substances. In Appendix R.7.6–4, Stages 3.1.1-3.1.7 describe REACH Annex XI adaptations based on:

1) existing information from non-GLP or test methods not referred in the test method regulation;

2) existing historical human data;

3) existing information in a *Weight-of-Evidence* approach;

4) non-animal approaches such as QSAR approaches and *in vitro* methods;

5) grouping and read across;

6) technical reasons, and substance-tailored exposure driven testing.

Stage 3.1.8 describes adaptations based on Column 2 rules others than based on CMR classification described at Stage 1.

Stage 3.2 Substances for which there are triggers for further information needs beyond the standard information requirements (Column 1)

Whereas Column 1 describes the standard information requirements (and triggers for those), Column 2 includes triggers for further information needs (in addition to provision to omit studies which are described at Stage 3.1.8 in Appendix R.7.6–4 of this Guidance).

Column 2 triggers may have various levels of requirements/consequences:

1) the registrant must act;

2) the registrant should act;

3) the registrant may act.

The consequence level depends on the wording in Column 2. If there is further concern on reproductive toxicity beyond the information requirements (Column 1 and 2 provisions), it is the responsibility of the registrant to consider how to address the concern to ensure the safe use of that substance. The various triggers related to reproductive toxicity and how to evaluate them are described in Appendix R.7.6–5 of this Guidance, Evaluation of Triggers, giving further information needs beyond the standard information requirements.

Stage 4. Reproductive toxicity tests triggered by tonnage level or by findings/conditions which raise concerns for further studies identified in Stages 1-3

Stage 4.1 Preliminary considerations

(i) Introduction

It has to be noted that if studies listed in REACH Annexes IX and X like the prenatal developmental toxicity study or the extended one-generation reproductive toxicity study are intended to be performed, a testing proposal must be submitted to ECHA. Furthermore, before the result from a study for which a testing proposal is submitted to ECHA will be available, risk management measures have to be put in place, recorded in the chemical safety report and recommended to downstream users according to REACH Annex I, 0.5.

A brief description of the protocols for the studies listed in REACH Annexes is presented at Stages 4.2, 4.3 and 4.4 according to registration tonnage levels. When planning any reproductive toxicity studies, considerations such as the properties of the substance, dose levels, vehicle, adequate study design, route and animal species, are needed. Some of these considerations which are especially relevant for reproductive toxicity testing are presented below.

(ii) Range-finding studies

It is recommended that the dose range-finding studies are reported together with the main studies (in IUCLID) to provide sufficient information and justification for the doses selected for testing. The findings from a range-finding study may also support the interpretation of the results from the main study.

(iii) Selection of vehicle

Most of the test methods guide on selection of vehicle if that is needed. For use of all other vehicles except for water a justification is needed and has to be documented. The vehicle should not cause any adverse effects itself as that may interfere with the interpretation of the results and may invalidate the study. Also, the vehicle must not react with the substance or interfere with toxicokinetics of the substance or affect significantly the nutritional status of the animals. The control group should receive the same vehicle and at the same dosing volume as the treated groups.

(iv) Route of administration for reproductive toxicity studies

REACH specifies that the reproductive toxicity studies should be conducted via the "*most appropriate route of administration, having regard to the likely route of human exposure*". "*Likely routes of human exposure*" within REACH are oral, inhalation and dermal. The selection of the "most appropriate route of administration" focuses on identification of hazards (see the Introduction to this Guidance, R7a and sub-section "Selection of the appropriate route of administration for toxicity testing", under R.7.2 Human health properties or hazards) and depends on the most appropriate route for identification of the intrinsic properties of the substance for reproductive hazard.

According to the test methods for reproductive toxicity which focus on the detection of reproductive hazards, the oral route (gavage, in diet, or in drinking water) is the "default" route, except for gases. For the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) dietary administration may be an appropriate route to model human exposure. If another route of administration other than oral is used, the registrant should provide justification and reasoning for its selection. In practice, testing via the oral route is usually performed with liquids and dusts and testing via inhalation route is usually performed with gases and liquids with very high vapour pressure. Testing via dermal route might be necessary under specific circumstances, for example for substances with high dermal penetration and indications for a specific toxicity following dermal absorption. Dermal application or inhalation route using nose-only administration may need specific considerations to ensure that the administration can be adequately conducted without causing confounding factors, for example, cause additional stress to the pregnant animals. Case-specific deviations from the default approach must be justified, such as in the case of available information on route-specific toxicity or toxicokinetics indicating that the use of oral administration of substance would not be relevant for assessing the human health hazards via inhalation, which would be the main route of exposure.

It is to be noted that corrosive or highly irritating substances should be tested preferentially via the oral route, however it must be noted that in vivo testing with corrosive substances at concentration/dose levels causing corrosivity must be avoided (see REACH Annex VII-X preamble). The vehicle should be chosen to minimise gastrointestinal irritation. For some substances dietary administration may allow adequate dosing without irritation compared with oral gavage dosing. In certain cases, testing of neutral salts of alkaline or acidic substances may be appropriate and allows investigation of intrinsic properties at adequate dose levels. If immediate hydrolysis of a substance occurs, it may be possible to provide information on all the cleavage products. For this read-across approach adequate justification and documentation is needed according to REACH Annex XI, 1.5. For corrosive or irritating vapours or gases for which oral testing is not possible, the highest concentration for inhalation should be chosen carefully to induce some toxicity (or mild irritation).

(v) Selection of species

The most common species used for reproductive toxicity testing is the rat. There is good historical background information for various rat strains which may be used to support the interpretation of the results. The strain selected should have an adequate fecundity and not too high an incidence of spontaneous malformations or any other specific feature that may reduce the adequacy of the strain to study reproductive toxicity of a substance in question. In order to make integrated data interpretation including information from other studies, it is recommended to use the same strain both in reproductive toxicity testing as well as repeated dose toxicity studies.

For prenatal developmental toxicity studies, testing in two species is a standard information requirement for registrations at 1000 or more tonnes per year (and might be triggered at lower tonnage levels). According to the test methods (EU B.31, OECD TG 414), the rat is the preferred rodent species and the rabbit the preferred non-rodent species. The extended one-generation reproductive toxicity study may need to be conducted using a second strain or species in certain exceptional cases. (For details see Stage 4.5 (ii) under Section R.7.6.2.3.2 of this Guidance). The most sensitive species and/or strain should be used as a first species taking into account the human relevancy, if known.

However, in choosing the appropriate species or strain of animal, consideration must be given to the suitability of the species and strain for the test protocol, and the availability of background information on the species and strain for the test protocol. The species/strain selection should be justified if the default species referred to in a test method is not used.

(vi) Dose level selection

Like in repeated dose toxicity studies the highest dose level should be chosen with the aim to induce some toxicity unless limited by physical or chemical properties of the substance (e.g. flammability and explosivity limits). Regarding the highest dose level, it is important to ensure that toxicity in both female and male animals is considered to ensure that reproductive toxicity in either gender is not overlooked. Generally at least three dose levels and a concurrent control must be used, except where a limit test (1000 mg/kg bw/day which is generally referred to as the oral limit dose level) is conducted. Expected human exposure may indicate the need for a higher dose level to be used than a 1000 mg/kg bw/day[90]. The conditions for applicability of a limit test are provided in the individual test methods for reproductive toxicity. For inhalation exposure, OECD Guidance document 39 may be used.

Dose level selection is assisted by the information from existing studies as well as from specific dose range-finding studies that may need to be conducted. Toxicokinetic information may provide reasons to adjust for example, the dosing route and regime. In addition, it should be considered that toxicity and toxicokinetics in pregnant animals may differ to that in non-pregnant animals. This may cause challenges in selecting the highest dose level for the study as at various phases of the study the sensitivity of the animals may differ.

For fertility as well as developmental toxicity it is important to investigate whether these reproductive toxicity effects are considered to be a secondary non-specific consequence of other toxic effects seen, such as, maternal toxicity, which may occur at the same dose level as the reproductive effects. However, in general, all findings on reproductive toxicity should be

[90] CLP, Annex I, Sections 3.7.2.5.7 – 3.7.2.5.9 state on the limit dose and very high dose levels the following: “There is general agreement about the concept of a limit dose, above which the production of an adverse effect is considered to be outside the criteria which lead to classification, but not regarding the inclusion within the criteria of a specific dose as a limit dose. However, some guidelines for test methods, specify a limit dose, others qualify the limit dose with a statement that higher doses may be necessary if anticipated human exposure is sufficiently high that an adequate margin of exposure is not achieved. Also due to species differences in toxicokinetics, establishing a specific limit dose may not be adequate for situations where humans are more sensitive than the animal model.” Section 3.7.2.5.8: “In principle, adverse effects on reproduction seen only at very high dose levels in animal studies (for example doses that induce prostration, severe inappetence, extensive mortality) would not normally lead to classification, unless other information is available, e.g. toxicokinetics information indicating that humans may be more susceptible than animals, to suggest that classification is appropriate. Please also refer to the section on maternal toxicity (3.7.2.4) for further guidance in this area.” And section 3.7.2.5.9 continues: “However, specification of an actual ‘limit dose’ will depend upon test method that has been employed to provide the test results, e.g. in the OECD Test Guideline for repeated dose toxicity studies by oral route, an upper dose of 1000 mg/kg has been recommended as a limit dose, unless expected human response indicates the need for a higher dose level.”

considered for classification purposes even if they are seen in the presence of parental toxicity. A comparison between the severity of the effects on fertility/development and the severity of other toxicological findings must then be performed[91]. Thus, it is important to get information about the reproductive toxicity profile of a substance including the spectrum of reproductive toxicity effects related to different dose levels as well as information to allow evaluation of the potency for reproductive toxicity of a substance. Therefore, the highest dose level should be intended to produce some toxicity to provide adequate information on reproductive toxicity for the purpose of both classification (including categorisation within the Reproductive toxicity hazard class) and risk assessment. For further information and clarification see the CLP criteria for classification (Section 3.7, Annex I of CLP) and Section 3.7 in the *Guidance on the Application of the CLP criteria*.

In reproductive toxicity studies local irritating effects at the site of administration may not allow investigating the reproductive toxicity in relation to systemic toxicity. In addition the irritation may affect the behaviour of the animals confounding the interpretation. Therefore, testing of corrosive or highly irritating substances at dose levels causing corrosivity or irritation must be avoided as far as possible (see REACH Annex VII-X preamble).

Dose level selection (and vehicle used) must be justified and documented to allow independent evaluation of the choice made.

Stage 4.2 Registrations of 1 to 10 tonnes per year (REACH Annex VII)

For substances manufactured or imported at tonnage levels ≥1-<10 t/y (REACH Annex VII) there are no specific standard information requirements for reproductive toxicity. However, the available relevant information needs to be evaluated and the classification for reproductive toxicity should be considered and applied if the classification criteria are met. If no information on reproductive toxicity is available, relevant non-animal approaches like validated *in vitro* tests, (Q)SAR predictions, or other available *in vivo* studies with the substance or with structurally related substances may be used to evaluate if there are triggers for reproductive toxicity. If the available information indicates a concern (trigger) for reproductive toxicity and relevant human exposure occurs, an animal study like the reproduction/developmental toxicity screening test (OECD TGs 421 or 422) should be considered to be performed to address the concern as an option. If an REACH Annex IX or X level study, such as prenatal development toxicity study (EU B.31, OECD TG 414) or extended-one-generation reproductive toxicity study (EU B.56, OECD TG 443) is considered necessary to address the concern, a testing proposal should be submitted to ECHA. A thorough scientific justification on how the concern has been addressed should be adequately documented.

Stage 4.3 Registrations of 10 to 100 tonnes per year (REACH Annexes VII and VIII)

At this tonnage level, progression beyond Stages 1-3, will trigger the reproduction/ developmental toxicity screening test (OECD TG 421) or a combined repeated dose toxicity study with the reproduction/developmental toxicity screening test (OECD TG 422).

(i) Reproduction/developmental toxicity screening test

If a 28-day study (EU B.7, OECD TG 407) is not already available, the conduct of a combined repeated dose toxicity study with the reproduction/developmental toxicity screening test (OECD TG 422) is preferred to the reproduction/developmental toxicity screening test (OECD

91 See the *Guidance on the Application of the CLP criteria*, i.e. the intro to section 3.7.2.2.1.1 "*Effects to be considered in the presence of marked systemic effects*"

TG 421). This approach offers the possibility to avoid carrying out a 28-day study, because the OECD TG 422 can at the same time fulfil the information requirement of REACH Annex VIII, 8.7.1 and that of REACH Annex VIII, 8.6.1.

If available information indicates serious concerns[92] (trigger(s)) about the potential of a substance for adverse effects on fertility or development, a screening test (OECD TG 421 or 422; REACH Annex VIII, Section 8.7.1) may not need to be performed. Instead, a testing proposal for either a prenatal developmental toxicity study (EU B.31, OECD TG 414; REACH Annex IX, Section 8.7.2) or an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443; REACH Annex IX, Section 8.7.3) may be submitted to ECHA depending on the type of trigger(s). Trigger(s) indicating serious concerns that the substance may be toxic to reproduction could stem from non-animal approaches[93] or *in vivo* information with the substance under consideration or from structurally related substances. Triggers for fertility (see Appendix R.7.6–5 of this Guidance for discussion on triggers), could also stem for example, from existing repeated dose toxicity studies showing histopathological changes in gonads, and/or effects in sperm parameters. The correct study to be proposed depends on the concern: if there is a concern for hazardous effects on fertility and/or development leading to developmental toxicity effects manifested after birth, an extended one-generation study should be proposed; if there is a concern for hazardous effects on embryonic or foetal development, a prenatal developmental toxicity study should be proposed. However, since the fertility and reproductive performance and developmental toxicity manifested shortly after birth are not assessed in a prenatal developmental toxicity study, it is strongly recommended to also conduct an OECD TGs 421 or 422 screening study as already discussed earlier (a testing proposal is not needed for a screening study). If an extended one-generation reproductive toxicity study is proposed (all various study designs) then it covers all the same parameters, exposure duration and statistical power compared to that of the screening study and thus, an additional screening study is not required.

If a reproduction/developmental toxicity screening test (OECD TGs 421 or 422) for an REACH Annex VIII substance provides no triggers for reproductive and developmental toxicity, then further testing for reproductive toxicity is not required at this tonnage level. Similarly, if a clear and unequivocal reproductive and/or developmental toxicity effect is observed in a screening test which is deemed sufficient to enable a scientifically robust decision on classification and categorisation to 1B for reproductive toxicity and risk assessment, then no further testing beyond the screening test is recommended at this tonnage level.

However, if a screening test (OECD TGs 421 or 422) shows effects which are deemed not sufficient to enable a scientifically robust decision on classification and risk assessment, further studies may be considered. Based on the type of trigger, a testing for either a prenatal developmental toxicity study (REACH Annex IX, Section 8.7.2) or an extended one-generation study (REACH Annex IX, Section 8.7.3) may be proposed. Specifically, if a clear and unequivocal reproductive and/or developmental toxicity effect is observed in a screening test which is deemed sufficient for classification in Category 2 for reproductive toxicity, then this is a serious concern and either a prenatal developmental toxicity study (REACH Annex IX, Section 8.7.2) or an extended one-generation study (REACH Annex IX, Section 8.7.3) may be proposed.

---

[92] Serious concern reflects a high likelihood for adverse effects on reproductive health.

[93] In order to be considered providing "*serious concern*", information from non-animal approaches should be reliable, relevant and from validated studies with appropriate applicability domain (for QSAR models a formal validation process is not required) . Based on case-by-case scientific justification results from non-validated studies and non-guideline tests, it may be acceptable. Generally several information sources may be needed.

Stage 4.4 Registrations of 100 to 1000 tonnes per year (REACH Annexes VII to IX)

At this tonnage level, progression beyond Stages 1-3 will trigger a prenatal developmental toxicity study in a first species (EU B.31, OECD TG 414) and, if the available repeated dose toxicity studies indicate adverse effects on reproductive organs or tissues or reveal other concerns in relation to reproductive toxicity, will also trigger an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443). For further information on triggers for an extended one-generation toxicity study at REACH Annex IX level, see point (iii) below.

If the results from existing studies (prenatal developmental toxicity test or repeated-dose studies) are sufficient to support classification to Category 1B for effects on developmental toxicity and/or sexual function and fertility and the risk assessment, the Column 2 adaptation rules for REACH Annex IX, point 8.7 should be followed. If the classification criteria for sexual function and fertility are met, then further testing for developmental toxicity must be considered and vice versa. For details, see Stage 1.

(i) Reproduction/developmental toxicity screening test

A reproduction/developmental toxicity screening test (OECD TGs 421 or 422) is a standard information requirement at REACH Annex VIII level. Since the Column 1 requirements in the REACH Annexes are cumulative, a screening test should also be available at REACH Annex IX and X level. However, if a prenatal developmental toxicity study, a two-generation reproductive toxicity study or an extended one-generation study is available, the screening study can be omitted based on REACH Annex VIII, Section 8.7.1., Column 2 adaptation rules (at REACH Annex VIII).

Where a screening test is omitted based on a prenatal developmental toxicity study and an extended one-generation reproduction toxicity study is not triggered at REACH Annex IX level, then information on fertility would be limited to evaluation of the reproductive organs after repeated dosing, if those studies are available. Where information from a reproductive toxicity study addressing a fertility endpoint is not available, it is strongly recommended that a screening study is considered to fulfil this endpoint.

(ii) Prenatal developmental toxicity study

A prenatal developmental toxicity study (EU B.31, OECD TG 414), conducted in one species, is a standard data requirement at REACH Annex IX level.

Consideration of existing information and the testing approach is required to select the appropriate species for the prenatal developmental toxicity study (see especially Stage 4.1(v) above). According to the test methods (EU B.31, OECD TG 414), the rat is the preferred rodent species and the rabbit the preferred non-rodent species. Since most of the toxicity studies (e.g. acute, repeated-dose, and toxicokinetic studies) are conducted in the rat, it may be considered that the first prenatal developmental toxicity study should also be conducted in this species. Findings from previous studies may be useful in dose selection, or the identification of additional endpoints for evaluation. In addition, the outcome of the prenatal developmental toxicity study may be helpful in the interpretation of other reproductive toxicity studies, for which the rat is generally the preferred species.

In certain cases the rabbit might be selected as the species for the first prenatal developmental toxicity study. This may be done for example, if the rabbit is considered to be a more sensitive species than the rat for that specific substance. The selection of the species for the prenatal developmental toxicity study should be made taking into account substance-specific aspects. If a species other than the rat and the rabbit is selected as the first or second species, the selection should be justified.

A decision on the need to perform a study on a second species at REACH Annex IX level should be based on the outcome of the first study and all other relevant available data. A study on a second species might be necessary if the available data contain triggers for prenatal developmental toxicity. For example, performance of a prenatal developmental toxicity study in a second species may be justified if developmental effects that are not sufficient to meet classification criteria to Category 1B reproductive toxicant (but maybe sufficient to Category 2 reproductive toxicant) were observed in the prenatal developmental toxicity study with the first species. Further triggers may stem from non-animal approaches, structurally similar substances, mechanisms/modes of action or results from a screening study. However, if there are no triggers and no indication of prenatal developmental toxicity in the first prenatal developmental toxicity study, no study on a second species is necessary at REACH Annex IX level.

If a study on a second species is found to be necessary by the registrant, a testing proposal needs to be submitted. Testing in a second species should be performed in a non-rodent species (rabbit) if the first species was a rodent species (rat) and vice versa. Further considerations on the species selection are provided in Section R.7.6.4.2.2 of this Guidance.

(iii) Extended one-generation reproductive toxicity study

An extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) is required at REACH Annex IX level if the available repeated dose toxicity studies (e.g. 28- or 90-days studies or OECD TGs 421 or 422 screening tests) indicate adverse effects on reproductive organs or tissues or reveal other concerns in relation with reproductive toxicity. Information from non-animal approaches are thus not listed as triggers for this study at REACH Annex IX level in the REACH Annex text. However, if there is a serious concern based on available information from non-animal approaches or structurally analogous substances, the study may be triggered.

Triggers for the study at REACH Annex IX level

A detailed review of the available data is required to identify any reproductive toxicity triggers (see Appendix R.7.6–5 of this Guidance for evaluation and determination of triggers), furthermore, examples of triggers for an extended one-generation reproductive toxicity study at REACH Annex IX level are provided below.

The legal text does not especially specify that the adverse effects should be seen in intact animals, however, it is considered that findings observed in non-intact animals should generally be used as triggers unless there is evidence that the findings would not be relevant for intact animals and/or humans. Experiments with non-intact animals may include animals with removal of an endocrine organ, such as ovary (ovariectomy). Another possibility is hormonal manipulation, for example causing decrease or increase of organ weight. These animal models may be very sensitive to detect a change in hormonal response, however, it should be considered whether the same applies in intact animals.

Examples (not an exhaustive list) of triggers to conduct an extended one-generation reproductive toxicity study at REACH Annex IX level (considered as adverse, and which are in line with other data and not considered secondary to systemic or maternal toxicity) are as follows:

From a screening study or equivalent:

- Changes in reproductive or other endocrine organ weight in intact animals;
- Effects in spermatogenesis or folliculogenesis *in vivo* and/or histopathological findings in reproductive organs and/or accessory sex organs;
- Effects in histopathology of the thyroid;
- Effects on sperm parameters analysis or oestrous cycle;
- Biologically relevant changes in hormone levels *in vivo* (related to reproductive toxicity);
- Reduced mating, fertility or litter size;
- Increased incidence of abortions compared to controls;
- Changes in gestation length;

- Reduced survival of offspring;
- Reduced body weight of offspring independent of litter size;
- Reduced maternal care;
- Changes in anogenital distance unrelated to body weight/size;
- Changes in nipple retention;
- Indication of other endocrine disrupting modes of action related to reproductive toxicity.

From a repeated dose toxicity study:

- Changes in reproductive or other endocrine organ weight in intact animals;
- Effects in spermatogenesis or folliculogenesis *in vivo* and/or histopathological findings in reproductive organs and/or accessory sex organs;
- Effects on sperm parameters analysis or oestrous cycle
- Biologically relevant changes in hormone levels (related to reproductive toxicity)*;*
- Indication of other endocrine disrupting modes of action related to reproductive toxicity.

From in vivo studies from non-intact animals (if the findings are considered relevant for intact animals/humans):

- Changes in reproductive or other endocrine organ weight.
- Indication of other endocrine disrupting modes of action related to reproductive toxicity

Study design for the extended one-generation reproductive toxicity study

If triggers are identified that require performance of an extended one-generation reproductive toxicity study, the appropriate study design as described in Column 1 and 2 and in Recital (7) of Commission Regulation (EU) 2015/282 amending REACH, needs to be defined, justified and documented. Specification is required for 1) length of the premating exposure duration and dose level selection, 2) the need to extend Cohort 1B and termination time for F2 generation, 3) the need to include Cohorts 2A and 2B, and 4) the need to include Cohort 3.

The study design of the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) specified in REACH in Column 1 as a standard information requirement, is the so called "basic" study design and a one-generation study including Cohorts 1A and 1B. Recital (7) of Commission Regulation (EC) 2015/282 amending REACH, states that the extended one-generation reproductive toxicity study should allow adequate assessment of fertility and that premating exposure duration and dose levels should be appropriate to meet the risk assessment and classification and labelling purposes (including categorisation)[94]. The focus of the study in the REACH Annexes is on fertility, which should be considered in the study design of the extended one-generation reproductive toxicity study, thus, as a starting point, a ten-week premating exposure duration and a highest dose level with the aim to induce some toxicity for all variant study designs of an extended one-generation reproductive toxicity study should be proposed. Regarding the highest dose level, it is important to ensure that toxicity in

---

[94] Recital (7) of Commission Regulation (EU) 2015/282 of 20 February 2015 amending Annexes VIII, IX and X to Regulation (EC) No 1907/2006 of the European parliament and of the Council on the Registration, evaluation, Authorisation and Restriction of Chemicals (REACH) as regards the Extended one-generation reproductive toxicity study: "*It should be ensured that the reproductive toxicity study carried-out under point 8.7.3 of Annexes IX and X to Regulation (EC) No 1907/2007 will allow adequate assessment of possible effects on fertility. The premating exposure duration and dose selection should be appropriate to meet risk assessment and classification and labelling purposes as required by Regulation (EC) No 1907/2006 and Regulation (EC) No 1272/2008 of the European parliament and of the Council.*"

both female and male animals is considered to ensure that reproductive toxicity in either gender is not overlooked. The basic study design, including the premating exposure duration according to Appendix R.7.6–3 of this Guidance, should be proposed by registrants unless the conditions specified in Column 2 are met.

The extension of the Cohort 1B (mating of the Cohort 1B animals to produce the F2 generation) must be proposed by the registrant if the conditions specified in Column 2 are met. Based on specific triggers for neurotoxicity defined in Column 2, developmental neurotoxicity cohorts (Cohorts 2A and 2B) must be proposed by the registrant. Respectively, based on specific triggers for immunotoxicity defined in Column 2, developmental immunotoxicity cohort (Cohort 3) must be proposed by the registrant.

The registrant may also propose a separate developmental neurotoxicity and/or developmental immunotoxicity study instead of the cohorts for developmental neurotoxicity and/or developmental immunotoxicity.

The conditions specifying the study design are listed in REACH Annex IX, 8.7.3, Column 2 and explained in more detail in Appendix R.7.6–2 of this Guidance and discussed in Section R.7.6.4.2.3 "Extended one-generation reproductive toxicity study" of this Guidance. Appendix R.7.6–1 of this Guidance lists the information that should be considered and, where available, presented in the dossier in order to establish the existence or the nonexistence of the conditions (triggers) specifying the study design. It is the registrant's responsibility to evaluate all the available information and to propose an adaptation of the standard information requirement following conditions described in Column 2 of REACH Annex IX/X, 8.7.3.

The justification of the study design that is most appropriate for evaluation of the reproductive toxicity of a substance must be adequately documented. This documentation must include justifications why the registrant holds the conditions of deviations from the basic study design not to be fulfilled taking into account all the available information.

A study on a second species or strain

REACH Annex IX specific rules for adaptation states that the need to perform an extended one-generation reproductive toxicity study (EU B.56; OECD TG 443) in a second strain or a second species, either at this tonnage level or the next, may be considered, and a decision should be based on the outcome of the first test and any other relevant available data.

It is recognised that the extended one-generation reproductive toxicity study is designed to be conducted in rats and it may be challenging to use other species. Thus, it has been made possible to conduct a second study using another rat strain instead of a second species. The need to conduct the study using a second species or strain will be in exceptional cases only.

A study on a second strain or species might be necessary if the available data contain triggers which have not been addressed in the study on the first species. For example, performance of a study in a second strain or species may be justified if effects were observed in the study with the first species cause further serious concern but are not sufficient to meet classification criteria to Category 1B reproductive toxicant. Further triggers may stem from validated[95] non-animal approaches, reliable and relevant QSAR models with adequate applicability domain, structurally similar substances, modes of action or results from a screening study. However, if there are no triggers and no indication of adverse effects on reproductive toxicity in the first study and other available data, no study on a second species or strain is necessary at REACH Annex IX level.

If a study on a second species or strain is found to be necessary by the registrant, a testing proposal should be submitted.

---

[95] Case-by-case scientific justification must be provided when non-validated or non-guideline methods are used.

Stage 4.5 Registrations of 1000 tonnes or more per year (REACH Annexes VII to X)

Progression beyond Stage 1-3 will trigger a prenatal developmental toxicity study (EU B.31, OECD TG 414) on a second species, if not conducted at the previous tonnage level, and an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443), if not already conducted at the previous tonnage level.

(i) Prenatal developmental toxicity study

At REACH Annex X level, a prenatal developmental toxicity study (EU B.31, OECD TG 414) conducted on a second species is a standard information requirement in addition to a prenatal developmental toxicity study in a first species that is required at REACH Annex IX level. Availability of information on two species allows a more comprehensive evaluation of prenatal developmental toxicity. The prenatal developmental toxicity study in a second species can be omitted, if, taking into account the outcome of the first test and all other relevant available data, an adaptation pursuant to REACH Annex X, Section 8.7, Column 2 or pursuant to REACH Annex XI can be justified.

According to the test methods (EU B.31, OECD TG 414), the rat is the preferred rodent species and the rabbit the preferred non-rodent species. Depending on whether the rat or the rabbit is selected as a first species, and/or is already available, the other should be the preferred second species. In certain cases the rabbit might be selected as the species for the first prenatal developmental toxicity study. This may be done for example if the rabbit is considered to be the more sensitive species than the rat for that specific substance. The selection of the species for the prenatal developmental toxicity study should be made taking into account substance-specific aspects. If a species other than the rat and the rabbit is selected as the first or second species, the selection must be justified.

(ii) Extended one-generation reproductive toxicity study

The extended one-generation reproductive toxicity study (EU B.56; OECD TG 443) is a standard information requirement at REACH Annex X level.

<u>Study design for the extended one-generation reproductive toxicity study</u>

The criteria for the study design for the extended one-generation reproductive toxicity study are the same at REACH Annex IX and X levels. Thus, the description of the study design here is identical to that at REACH Annex IX level (Stage 4.4 (iii)).

The appropriate study design as described in Column 1 and 2 and in Recital (7) of Commission Regulation (EU) 2015/282 amending REACH, needs to be defined, justified and documented. Specification is required for 1) length of the premating exposure duration and dose level selection, 2) the need to extend Cohort 1B and termination time for F2 generation, 3) the need to include Cohorts 2A and 2B, and 4) the need to include Cohort 3.

The study design of the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) specified in REACH in Column 1 as a standard information requirement is the so called “basic” study design and a one-generation study including Cohorts 1A and 1B. Recital (7) of Commission Regulation (EC) 2015/282 amending REACH, states that the extended one-generation reproductive toxicity study should allow adequate assessment of fertility and that premating exposure duration and dose levels should be appropriate to meet the risk assessment and classification and labelling purposes[96]. The focus of the study in the REACH Annexes is on fertility, which should be considered in the study design of the extended one-generation reproductive toxicity study. Thus, as a starting point, a ten -week premating exposure duration and a highest dose level with the aim to induce some toxicity for all variant study designs of an extended one-generation reproductive toxicity study should be proposed. Regarding the highest dose level, it is important to ensure that toxicity in both female and male animals is considered to ensure that reproductive toxicity in either gender is not overlooked. The basic study design, including the premating exposure duration according to Appendix R.7.6–3 of this Guidance, should be proposed by registrants unless the conditions specified in Column 2 are met.

The extension of the Cohort 1B (mating of the Cohort 1B animals to produce the F2 generation) must be proposed by the registrant if the conditions specified in Column 2 are met. Based on specific triggers for neurotoxicity defined in Column 2, developmental neurotoxicity cohorts (Cohorts 2A and 2B) must be proposed by the registrant. Respectively, based on specific triggers for immunotoxicity defined in Column 2, developmental immunotoxicity cohort (Cohort 3) must be proposed by the registrant.

The registrant may also propose a separate developmental neurotoxicity and/or developmental immunotoxicity study instead of the cohorts for developmental neurotoxicity and/or developmental immunotoxicity.

The conditions specifying the study design are listed in REACH Annex X, 8.7.3, Column 2 and are explained in more detail in Appendix R.7.6–2 of this Guidance and discussed in Section R.7.6.4.2.3 “Extended one-generation reproductive toxicity study” of this Guidance. Appendix R.7.6–1 of this Guidance lists the information that should be considered and, where available, presented in the dossier in order to establish the existence or the non-existence of the conditions (triggers) specifying the study design. It is the registrant’s responsibility to evaluate all the available information and to propose an adaptation of the standard information requirement following conditions described in Column 2 of REACH Annex IX/X, 8.7.3.

The justification of the study design that is most appropriate for evaluation of the reproductive toxicity of a substance must be adequately documented. This documentation must include justifications why the registrant holds the conditions of deviations from the basic study design not to be fulfilled taking into account all the existing information.

A study on a second species or strain

REACH Annex IX specific rules for adaptation states that the need to perform an extended one-generation reproductive toxicity study (EU B.56; OECD TG 443) in a second strain or a second species, either at REACH Annex IX tonnage level or at REACH Annex X tonnage level, may be considered and a decision should be based on the outcome of the first test and any other relevant available data. It is recognised that the extended one-generation reproductive toxicity study is designed to be conducted in rats and it may be challenging to use other species. Thus, it has been made possible to conduct a second study using another rat strain instead of a second species. The study on a second species or strain is needed in exceptional cases only.

---

[96] Recital (7) of Commission Regulation (EU) 2015/282 of 20 February 2015 amending Annexes VIII, IX and X to Regulation (EC) No 1907/2006 of the European parliament and of the Council on the Registration, evaluation, Authorisation and Restriction of Chemicals (REACH) as regards the Extended one-generation reproductive toxicity study: “*It should be ensured that the reproductive toxicity study carried-out under point 8.7.3 of Annexes IX and X to Regulation (EC) No 1907/2007 will allow adequate assessment of possible effects on fertility. The premating exposure duration and dose selection should be appropriate to meet risk assessment and classification and labelling purposes as required by Regulation (EC) No 1907/2006 and Regulation (EC) No 1272/2008 of the European parliament and of the Council.*”

A study on a second strain or species might be necessary if the available data contain triggers which have not been addressed in the study on first species. For example, performance of a study in a second strain or species may be justified if effects were observed in the study with the first species cause further serious concern but are not sufficient to meet classification criteria to Category 1B reproductive toxicant. Further triggers may stem from validated[97] non-animal approaches, reliable and relevant QSAR methods with adequate applicability domain, structurally similar substances, modes of action or results from a screening study. However, if there are no triggers and no indication of adverse effects on reproductive toxicity in the first study and other available data, no study on a second species or strain is necessary at REACH Annex X level.

If a study on a second species or strain is found to be necessary by the registrant, a testing proposal must be submitted.

R.7.6.3 Information sources on reproductive toxicity

Information on reproductive toxicity can be obtained from various source categories, which are indicated below. Examples from each source category are provided. Evaluation of this information is described in Section 0 of this Guidance. Where *in vivo* testing is required, registrants must follow the EU Directive 2010/63 in selecting the test(s) requiring fewest animals and the least suffering.

R.7.6.3.1 Information on reproductive toxicity from non-animal approaches

Limited information of supportive nature may be inferred from numerous non-animal approaches (tests not using whole animals including embryos and foetuses after a certain developmental stage). For evaluation of the quality of the information, see Section 0 of this Guidance where reference to ECHA guidance on evaluation of available information is given (*Guidance on IR&CSA, Chapter R.4 "Evaluation of available information"*):

- physico-chemical characteristics of a substance (distribution, accumulation);
- information on structurally analogue substances and (Q)SAR models;
- *in silico* and *in chemico* models (with adequate applicability domain);
- *in vitro* tests (with relevant concentrations) in reproductive toxicity or relevant modes on action; e.g.:
  - Performance-based test guideline for stably transfected transactivation *in vitro* assays to detect oestrogen receptor agonists (OECD TG 455, updated 2012);
  - BG1Luc Estrogen receptor transactivation test method for identifying oestrogen receptor agonists and antagonists (OECD TG 457);
  - H295R steroidogenesis assay (EU B.57, OECD TG 456);
  - *in vitro* embryotoxicity tests;
  - *in vitro* organ and cell cultures.

[97] Case-by-case scientific justification must be provided when non-validated or non-guideline methods are used.

- Where possible, well developed and justified reverse toxicokinetic models may be used to support results from *in vitro* tests to estimate exposures needed to achieve bioactive blood concentrations.

Approaches combining various methodologies, e.g. from adverse outcome pathway (AOP) concept (OECD GD 184).

R.7.6.3.2 Information on reproductive toxicity in humans

If human information is available, it must be presented and if possible in the form of a table as stated in REACH Annex I, 1.2.

Information may stem from epidemiological and/or occupational studies, medical records, case studies and accidents. For evaluation of the quality of the information, see Section 0 of this Guidance where reference to ECHA guidance on evaluation of available information is given (*Guidance on IR&CSA, Chapter R.4 "Evaluation of available information"*).

R.7.6.3.3 Information on reproductive toxicity from *in vivo* animal studies

Data may be available from a wide variety of animal studies, with standard or non-standard study design, which give different amounts of direct or indirect information on the potential reproductive toxicity of a substance. For evaluation of the quality of the information, see Section 0 of this Guidance where reference to ECHA guidance on evaluation of available information is given (*Guidance on IR&CSA*, *Chapter R.4 "Evaluation of available information"*).

In vivo studies referred to in REACH and providing information on reproductive toxicity:

- Extended one-generation reproductive toxicity study (EU B.56, OECD TG 443);
- Two-generation reproductive toxicity study (EU B.35, OECD TG 416);[98]
- Prenatal developmental toxicity study (EU B.31, OECD TG 414).

In vivo studies referred to in REACH and providing preliminary information on reproductive toxicity:

- A reproduction/developmental toxicity screening test (OECD TG 421);[99]
- Combined repeated dose toxicity study with the reproductive/developmental toxicity screening test (OECD TG 422)[100].

Other *in vivo* study on reproductive toxicity with EU and OECD test guidelines:

- One-generation reproductive toxicity study (EU B.34, OECD TG 415).

Repeated dose toxicity studies which may include parameters relevant for reproductive toxicity:

- 28- and 90-day repeated-dose toxicity studies (EU B.7; EU B.10), where relevant parameters are included, for example semen analysis, oestrous cyclicity, organ weights of reproductive organs and accessory sex organs, and/or reproductive organ histopathology.

---

[98] Existing two-generation reproductive toxicity studies (EU B.35, OECD TG 416) started before 15 March 2015 fulfil the standard information requirement for Annex IX/X, 8.7.3 but new studies for REACH must be proposed according to an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) as described in Annex IX/X, 8.7.3.

[99] To date there are no corresponding EU testing methods available.

[100] At the time of publication (July 2015), there are no corresponding EU testing methods available.

Short-term *in vivo* tests on endocrine disrupting modes of action in intact or non-intact animals, e.g.:

- Uterotrophic bioassay in rodents: a short-term screening test for oestrogenic properties (EU B.54, OECD TG 440; OECD GD 71 for anti-oestronicity);
- Hershberger bioassay in rats: a short-term screening assay for (anti)androgenic properties (EU B.55, OECD TG 441 and GD 115);
- Studies on juvenile/peripubertal animals.

Other studies which may provide relevant information, e.g.:

- Chernoff/Kavlock tests (see Hardin *et al.,* 1987);
- a modified one-generation study by NTP (National Toxicology Program, U.S. Department of Health and Human Services; http://ntp.niehs.nih.gov/testing/types/mog/index.html)
- Reproductive Assessment by Continuous Breeding (RACB) protocol (e.g. Chapin and Sloane 1997);
- peri-postnatal studies;
- male or female fertility studies of non-standard design;
- dominant lethal assay (EU B.22, OECD TG 478);
- mechanistic studies;
- toxicokinetic studies (EU B.36, OECD TG 417);
- studies in fish (e.g. Fish Sexual Development Test (OECD TG 234);
- studies in amphibians (e.g. Amphibian Metamorphoses Assay (OECD TG 231) or Larval Amphibian Growth and Development Assay (under development));
- studies in other non-mammalian species.

Studies with focus on developmental neurotoxicity and developmental immunotoxicity:

- developmental neurotoxicity studies (such as EU B.53, OECD TG 426);
- developmental immunotoxicity studies (see Section R.7.6.4.2.7 of this Guidance for references).

R.7.6.4 Evaluation of available information for reproductive toxicity

This section provides information on evaluation of the available data including aspects which influence the study designs. Both non-human (non-animal approaches and *in vivo* animal studies) and human data are considered. Under this section the studies required as standard information requirements are described as well as how to evaluate the conditions described in Column 2 to trigger a study or to adapt the study design. In addition, the evaluation of information from other internationally accepted *in vivo* studies are briefly described.

The generic guidance on the evaluation of available information gathered in the context of REACH Annexes VI-XI is provided in the *Guidance on IR&CSA, Chapter R.4:* "*Evaluation of available information*". The information should be evaluated for its completeness and quality for the purpose of REACH to assess whether (see the detailed wording in Chapter R.4):

- It fulfils the information requirements;
- It is appropriate for hazard classification and risk assessment.

The evaluation process of data quality by judging and ranking the available data for its relevance, reliability and adequacy is provided in Chapter R.4. Chapter R.4 applies to all kinds of information; human, animal and non-animal sources and it is also applicable to information for reproductive toxicity endpoint. OECD guidance document 43 may be consulted for aid in the interpretation of reproductive and neurotoxicity results (see OECD GD 106 for histologic evaluation, OECD GD 57 and 207 for thyroid hormone modulation assays, and OECD retrospective performance assay for developmental neurotoxicity, No 89 (OECD, 2008)).

In the present document some additional scientific aspects relevant for reproductive toxicity have been highlighted in context of the relevant information sources.

The main principles for evaluation of non-human information (information from animal studies and non-animal approaches) are presented in REACH Annex I, 1.1 and it must be comprised of:

- Hazard identification for the effect based on all available non-human information;
- Establishment of the quantitative dose (concentration) response (effect) relationship.

Robust study summaries are necessary for key data on reproductive toxicity. If possible the information should be provided in the form of table(s) (see further details in REACH Annex I, 1.1.3.).

R.7.6.4.1 Non-animal data

For reproductive toxicity, a grouping and category approach and weight of evidence adaptation are the best fit-for-purpose tools for non-animal approaches for the time being to adapt the (standard) information requirements for reproductive toxicity. However, appropriate justification and documentation must be provided. In addition, non-animal approaches may be used for prioritisation and screening chemical inventories.

Information on the current developments of *in vitro* tests and methodology can be found on the ECVAM website (http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam) and other international centres for validation of alternative methods. ECHA's website is also updated with new internationally accepted non-animal approaches (http://www.echa.europa.eu/support/oecd-eu-test-guidelines). However, the regulatory acceptance of these studies and approaches to replace the animal testing for reproductive toxicity has not been achieved as they do not provide equivalent information and thus, cannot be used alone for classification and labelling and/or risk assessment. In spite of this, they may serve as elements in categories/read across and weight of evidence adaptation. They may also provide important information on mechanisms and modes of action, or preliminary screening information which can be used in planning further testing.

R.7.6.4.1.1 Physico-chemical properties

It may be possible to infer from the physico-chemical characteristics of a substance whether it is likely to be absorbed following exposure by a particular route and, furthermore, whether it (or an active metabolite) is likely to cross the placental, blood-brain or blood-testes barriers, or be secreted in milk. Information on the physico-chemical properties may contribute to a Column 2 adaptation (e.g. indicate concern on prolonged phase before reaching a steady state which is part of the conditions triggering extension of Cohort 1B in the extended one-generation reproductive toxicity study) or a weight of evidence adaptation according to REACH Annex XI, 1.2.

R.7.6.4.1.2 (Q)SAR

There are a large number of potential targets/mechanisms associated with reproductive toxicity which, on the basis of current knowledge, cannot normally be adequately covered by a battery of QSAR models. In principle QSAR models are potential adaptation possibilities according to REACH Annex XI, 1.3, but they should adequately cover the endpoint in question – all the key aspects/parameters should be covered.

QSAR models are usually trained (developed) to give binary results; the substance is predicted to have or not have a particular property, e.g. developmental toxicity. If the substance is predicted to have that property, the result of a QSAR prediction is considered as positive. Similarly, if the substance is predicted not to have a particular property, the result of the QSAR prediction is considered negative. QSAR approaches are currently not well fitted-for-purpose for reproductive toxicity and consequently no firm recommendations can be made concerning their routine use in a testing strategy in this area. A particular challenge for this endpoint is the complexity and amount of information needed from various functions and parameters to evaluate the effects on reproduction. Not all necessary aspects can be covered by a QSAR prediction. Therefore, a negative result from current QSAR models predicting that the substance has not a particular property, cannot be interpreted as demonstrating the absence of a reproductive hazard unless there is other supporting evidence. Another limitation of QSAR modelling is that dose response information, for example the N(L)OAEL, required for risk assessment is not provided.

However, a positive result from a reliable and relevant QSAR model with an appropriate applicability domain predicting that the substance has a particular property could provide a trigger for further testing beyond the standard information requirement (e.g. one element to trigger the extension of Cohort 1B in an extended one-generation reproductive toxicity study). For evaluation of the triggers see Appendix R.7.6–5 of this Guidance. Due to the limited confidence in this approach such a result would not normally be adequate for making a decision on classification on its own. It may, although not normally used, provide supportive information that can be used when concluding on the appropriate classification (see 3.7.2.5.4, Annex I, CLP).

Provided the applicability domain is appropriate, the results from using QSAR models may be used in a weight of evidence analysis where such data are considered alongside other relevant data (for classification and labelling and as one element for weight of evidence adaptation approach according to REACH Annex XI, 1.2). Also, the results from using QSAR models can be used as supporting evidence when assessing the toxicological properties by read-across in a grouping approach, providing the applicability domain is appropriate. Both positive and negative QSAR modelling prediction results concerning the existence or non-existence of a particular property, respectively, may be of value in supporting a read-across assessment.

R.7.6.4.1.3 *In vitro* data and Adverse Outcome Pathways (AOPs)

The design of alternatives to *in vivo* testing for reproductive toxicity is especially challenging in view of the complexity of the reproductive process and large number of potential targets/mechanisms associated with this broad area of toxicity. In addition, many *in vitro* approaches do not include elements of biotransformation which, in addition, may differ depending on organ.

Currently there are only three officially adopted EU test methods or OECD test guidelines for *in vitro* tests of relevance to modes of action for reproductive toxicity: two measuring oestrogenicity (OECD TG 455 and OECD TG 457) and the other measuring steroidogenesis (EU B.57, OECD TG 456). Most assays under development and international validation are focusing on agonist/antagonistic properties measured by binding and activating or blocking a steroid (or a thyroid) hormone receptor.

Three *in vitro* embryotoxicity tests to predict developmental toxicity have been validated but have not been accepted for regulatory use (Genschow *et al.,* 2002, Piersma *et al.,* 2004, Spielmann *et al.*, 2004 and 2006). These three tests, the embryonic stem cell test, the limb bud micromass culture and the whole embryo culture, showed high predictivity for certain strongly embryotoxic chemicals. However, due to the nature of the methods and limitations in their predictivity, they may be used only as supporting information along with other more reliable data to predict the developmental toxicity. The value of these validated methods could be increased by incorporating molecular based markers through the application of proteomic and toxicogenomic approaches (Piersma, 2006; van Dartel *et al.*, 2010). The embryonic stem cell method may be combined with Physiologically Based Biokinetic modelling in order to derive quantitative points of departure *in vitro*, which are then extrapolated to *in vivo* points of departure for use in risk assessment (Worth *et al.*, 2014).

The combination of assays in a tiered and/or battery approach may improve predictivity, but the *in vivo* situation remains more than the sum of the areas modelled by a series of *in vitro* assays (see Piersma, 2006 for review). Therefore, a negative result predicting absence of a particular property for a substance with no supporting information cannot be interpreted as demonstrating the absence of a reproductive hazard with the same confidence as an animal study. Another limitation of *in vitro* tests is that an N(L)OAEL and other dose-response information required for a risk assessment is not provided.

However, a positive result predicting a particular reproductive hazard in a validated *in vitro* test could provide a justification for the need of further testing beyond the standard information requirement, dependent on the effective concentration and taking account of what is known about the toxicokinetic profile of the substance. However, because of limited confidence in this approach at this time, such a result in isolation would not be adequate to support hazard classification.

Additionally, validated and non-validated *in vitro* tests, provided the applicability domain is appropriate, could be used with other data in a weight of evidence adaptation according to REACH Annex XI, 1.2 to gather information on hazardous properties. *In vitro* techniques can be used in mechanistic investigations, which can also provide support for regulatory decisions. Also, *in vitro* tests can be used as supporting evidence when assessing the toxicological properties by read-across within a substance grouping approach, providing the applicability domain is appropriate. Positive and negative *in vitro* test results may be of value in a read-across assessment and in category approach as one element.

Current developments on adverse outcome pathways (AOPs) to build a combination of studies and investigations to cover key events from an initiating molecular event to an adverse outcome may provide information on certain pathways, especially in developmental toxicity for certain malformations. Approaches may combine various different methods (e.g. *in vitro* tests, QSARs, *in chemico* assays etc). As these pathways do not cover all potential mechanisms/modes of action, negative results predicting absence of a particular property from those approaches do not provide enough confidence for regulatory decision making to demonstrate absence of a reproductive hazard. In addition, currently they do not provide an N(L)OAEL value or other dose-response information for risk assessment. However, they may provide necessary support for read across justification and categories and contribute to a weight of evidence adaptation according to REACH Annex XI, 1.2.

R.7.6.4.2 Animal data

In general, all findings on reproductive toxicity should be considered for classification purposes irrespective of the level of concurrent parental toxicity, see the *Guidance on the Application of the CLP criteria* (Setion 3.7.2.2.1, classification in the presence of parental toxicity).

For evaluation of the results of a reproductive toxicity study, it is important, where possible, to distinguish between a specific effect on reproduction (fertility and/or pre- and postnatal development) as a consequence of an intrinsic property of the substance and an adverse reproductive effect which is a secondary non-specific consequence to the general toxicity. Inclusion of additional parameters for general toxicity may enhance this interpretation. According to the criteria for classification, reproductive toxic effects should be considered if they occur in the absence of other (systemic) toxic effects or if they occur together with other toxic effects, are considered not to be a secondary non-specific consequence of the other toxic effects (see 3.7.2, Annex I of CLP).

R.7.6.4.2.1 Reproduction/developmental toxicity screening test

The screening studies provide initial information of the effects on male and female reproductive performance as well as on developmental toxicity during and shortly after birth, as well as certain additional parameters for endocrine disrupting mode of action including anogenital distance, nipple/areola retention, thyroid hormone levels as given in the revised TGs [101] (2015). These screening tests are not meant to provide complete information on all aspects of reproduction and development. However, the screening test (OECD TGs 421 or 422) is a standard information requirement for reproductive toxicity at REACH Annex VIII level. Thus, a negative study result at REACH Annex VIII is considered adequate although the screening study does not provide similar confidence than more comprehensive studies on reproduction toxicity. An evaluation of the screening tests (OECD TGs 421 or 422) has confirmed that these tests are useful for initial hazard assessment and can contribute to decisions on further test requirements (Reuter *et al.*, 2003, Gelbke *et al.*, 2004, Beekhuiisen *et al.*, 2014).

With regard to male and female fertility, the number of parameters investigated are less than in the more comprehensive generation study designs such as the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) or the two-generation reproductive toxicity study (EU B.35, OECD 416), and the statistical power is much lower due to a lower number of animals per dose group. Furthermore, the pre-mating exposure duration in these screening studies may not be sufficient to detect all effects on the spermatogenic cycle or folliculogenesis. The two weeks premating exposure duration used in this study is equivalent to the time for epididymal transit of maturing spermatozoa and thus allows for the detection of post-testicular effects on sperm at mating (during the final stages of spermiation and epididymal sperm maturation). The two weeks premating exposure duration for females, covers 2-3 oestrous cycles and effects on cyclicity may be detected. Thus, the full spermatogenesis and folliculogenesis are not covered at the time of mating or together before and after the mating, as they take 70 and 62 days in rats, respectively.

Because exposure during the full spermatogenic period and folliculogenesis are not covered at the time of mating, effects at earlier stages of spermatogenesis and folliculogeneiss cannot be reflected in the functional fertility examination. For instance, earlier stages of the spermatogenesis (spermatogonia) and/or specific cell types (Sertoli cell and Leydig cells), are sensitive to many chemicals (see e.g review by Bonde, 2010). With a two-week premating exposure, the effects on functional fertility of exposure to these early stages of developing spermatozoa will not be covered. In addition, steady state may not be reached in all organs

[101] OECD TGs 421 and 422 are in the process of being revised: adoption and publication is expected by the end of 2015.

(see also discussion in Appendix R.7.6–3 of this Guidance). Histopathological data will be limited because the duration of the study itself does not cover the full spermatogenesis or folliculogenesis. Depending on the tonnage level, results from the 90-day study may be available with investigations of histopathology of gonads, however sperm parameters or oestrous cycles are usually not investigated. Histopathology of gonads may be among the most sensitive parameters to detect adverse effects on male fertility and the most sensitive parameter may be used to derive the NOAEL. However, the clarity of the effects rather than the sensitivity of the effects observed, are important for classification and labelling and will affect the category into which the substance is classified. Thus, to address the fertility also for the classification and labelling purposes, including the categorisation, it is necessary to consider how well all the available parameters address the fertility endpoint.

Due to its limitations, a screening study cannot be used to fulfil the information requirement of the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443). It should also be noted that these screening studies do not provide relevant information on post-natal developmental toxicity like a one- or two-generation reproductive toxicity studies (EU B.34/OECD 415 or EU B.56/OECD TG 443 or EU B.35/OECD 416) because the screening studies are already terminated at an earlier developmental stage than those more comprehensive studies.

With regard to developmental toxicity, these screening tests do not provide sufficient information on prenatal developmental toxicity because the pups are not examined for external, skeletal and visceral anomalies as in the prenatal developmental toxicity study (EU B.31, OECD TG 414). In addition, the pups in the screening studies are delivered naturally and the dams may cannibalise malformed pups. In the prenatal developmental toxicity study caesarean section is performed to avoid any cannibalism and to allow an appropriate evaluation of the foetuses. In addition, the statistical power of the screening study is lower than that of the prenatal developmental toxicity study. Therefore, a screening study cannot be used to fulfil the standard information requirement of a prenatal developmental toxicity study (EU B.31, OECD TG 414).

Depending on the tonnage level or based on adaptations, a screening study might be the only available reproductive toxicity study. However, the screening studies were not designed as an alternative or a replacement of the higher tier reproductive toxicity studies (EU B.31, OECD TG 414 and EU B.56, OECD TG 443). Therefore, the results of a screening study should be interpreted with caution and even statistically non-significant effects may be indicators for an impairment of reproduction. A result showing no effects in a OECD TGs 421 or 422 screening test does not provide reassurance of the absence of any hazardous property for reproductive toxicity. Further information on reproduction toxicity may be available to assist the interpretation of the results.

The observation of clear evidence of adverse effects on reproduction or on reproductive organs in these tests may be sufficient to meet the information needs for classification and labelling and risk assessment (using an appropriate assessment factor), and may provide an N(L)OAEL from which a DNEL can be identified (by adding an additional assessment factor due to higher uncertainty involved than in more comprehensive studies).

Effects observed in the screening study may serve as triggers, leading to more comprehensive reproductive toxicity studies or they may constitute conditions which specify the study design of an extended one-generation reproductive toxicity study. For instance EU B.56 (OECD TG 443) may be triggered based on evidence indicating concern on reproductive toxicity, (see Section R.7.6.2.3.2, Stage 4.4 REACH Annex IX, extended one-generation reproductive toxicity study, of this Guidance). A screening study may provide useful information when considering dose level selection for an extended one-generation reproductive toxicity study.

R.7.6.4.2.2 Prenatal developmental toxicity study

The prenatal developmental toxicity study (EU B.31, OECD TG 414) provides a focused evaluation of potential effects on prenatal development, although only effects that are manifested before birth can be detected. Detailed information on external, skeletal and visceral malformations and variations and other developmental effects are provided. Cesarean section allows precise evaluation of the number of foetuses affected.

For a comprehensive assessment of prenatal developmental toxicity, information from two species, one rodent (usually the rat) and one non-rodent (usually the rabbit) is assessed. However, depending on the REACH tonnage level, there might only be a standard information requirement for a prenatal developmental toxicity in one species (REACH Annex IX) or for none (REACH Annex VII and VIII). Under such circumstances, it needs to be evaluated if testing beyond the standard information requirement is triggered. If both or one of the default species (the rat or the rabbit) are not suitable species for prenatal developmental toxicity testing, a more suitable species considering the human relevancy should be selected for testing. An adequate justification must be provided for other species other than the rat and the rabbit. The results from prenatal developmental toxicity studies are considered relevant to humans unless there is substance-specific toxicokinetic or toxicodynamic evidence showing otherwise.

For evaluation, developmental effects should be considered in relation to adverse effects occurring in the parents, for further information see the *Guidance on the Application of the CLP criteria* (Section 3.7).

It should be noted that a prenatal developmental toxicity study (EU B.31, OECD TG 414) does not provide information on postnatal development or sufficient information on female fertility. However, some findings might raise concerns; if exposure started on gestation day 0, effects on preimplantation or implantation could indicate effects on female fertility. Also effects on maintenance of pregnancy and potentially on gestation length may be identified if significantly affected.

If a study is conducted according to an old test method and thus uses a shorter administration period than current test methods, it is important that there is no indication challenging the exposure period used. Thus, if there is a concern suggesting that a longer exposure period would have revealed developmental toxicity or more profound findings affecting also lower dose levels that were not observed using shorter exposure duration, this should be addressed; for example, by using an additional assessment factor which lowers the NOAEL to the next lower dose level or divides it by two if there is no lower dose level; or if a serious concern, a new study with longer exposure duration should be proposed. These indications challenging the exposure duration used may stem from fertility studies such as screening studies (OECD TGs 421 or 422) or from an extended one-generation reproductive toxicity study or from information on mechanisms/modes of action or structurally similar substances. It is to be noted that screening studies (OECD TGs 421 or 422) or the extended one-generation reproductive toxicity study do not provide equivalent information on prenatal developmental toxicity to that from the prenatal developmental toxicity study. Thus, if the indication of challenging the exposure duration rises from other available data, the results from these fertility studies may not always, depending on the case, provide sufficient confidence to conclude that there is no prenatal developmental toxicity.

Prenatal developmental toxicity studies may provide triggers for further reproductive toxicity studies, for example, in the form of foetotoxicity or foetal findings. In addition, some findings, such as increased foetal weight or placental weight, considered in light of litter size, may indicate an endocrine disrupting mode of action. Although there is no toxicological need to differentiate endocrine disrupting modes of action from other modes of action for developmental toxicity, in REACH the reproductive effects may trigger an extended one-generation reproductive toxicity study at REACH Annex IX and the indication of endocrine disrupting modes of action are one element in triggering the extension of Cohort 1B in an extended one-generation reproductive toxicity study.

R.7.6.4.2.3 Extended one-generation reproductive toxicity study

Introduction

The test method of the extended one-generation reproductive toxicity study (EOGRTS, EU B.56, OECD TG 443) describes a flexible modular study design with several investigational options allowing each jurisdiction to decide on the study design required for the respective regulatory context. The study design for REACH is described in detail in Appendix R.7.6–2 of this Guidance.

The extended one-generation reproductive toxicity study allows evaluation of the effects of the test substance on the integrity and performance of the adult male and female reproductive system and offspring viability, health and some aspects of physical and functional development until adulthood. The extension of the Cohort 1B (to mate the F1 animals to produce the F2 generation) also provides information on the fertility of the offspring (F1 generation), thus addressing the potential effects after exposure of the most sensitive life stages (i.e. *in utero* and early postnatal period). Therefore, mating of the Cohort 1B animals will cover information on the complete reproductive cycle.

In REACH the standard information requirement only includes Cohorts 1A and 1B for reproductive toxicity (without extension to produce the F2 generation). Thus, the basic study design is a one-generation study providing information on the fertility of the parental animals (P0 or F0 animals) and extended postnatal development of F1 animals. In addition, for REACH purposes it is necessary that the study design allows the adequate assessment of possible effects on fertility for risk assessment and classification and labelling purposes, including categorisation. To ensure that the study design adequately addresses the fertility endpoint, the duration of premating exposure period and the selection of the highest dose level are key aspects to be considered, see Appendix R.7.6–3 of this Guidance for further details. Regarding the highest dose level, it is important to ensure that toxicity in both female and male animals is considered to ensure that reproductive toxicity in either gender is not overlooked.

If the Column 2 conditions at REACH Annex IX/X are met, (for further information see Appendix R.7.6–2 of this Guidance) Cohort 1B must be extended, which means that the F2 generation is produced by mating the Cohort 1B animals. This extension also provides information on the mating, fertility and reproductive performance of the F1 animals. F1 animals are exposed *in utero* and during the early postnatal period allowing a comprehensive assessment of effects induced during these sensitive life stages. Similarly developmental neurotoxicity (Cohorts 2A and 2B) and/or developmental immunotoxicity (Cohort 3) cohorts need to be conducted if the triggers for such expansion of the basic study design, (which are provided in Column 2 of REACH Annex IX/X, 8.7.3) are fulfilled. These cohorts provide information on neurotoxic or immunotoxic potency of substances after exposure during sensitive life stages. When there are triggers for developmental neurotoxicity, both the Cohorts 2A and 2B are to be conducted as they provide complementary information. Considerations for evaluation of developmental neurotoxicity and developmental immunotoxicity are provided later in this section (see Sections R.7.6.4.2.6 and R.7.6.4.2.7 of this Guidance).

It is recommended that results from a range-finding study (or range-finding studies) for an extended one-generation reproductive toxicity study are reported with the main study. This will support the justifications of the dose level selections and interpretation of the study results.

If a range-finding study indicates adverse effects on fertility but the effects do not meet the criteria for Reproductive toxicity Category 1B, it is recommended that the main study should be designed to confirm the findings from the range-finding study. However, if the results from the range-finding study meet the criteria for Reproductive toxicity Category 1B reproductive toxicants, the adaptation of Column 2 may apply and further studies (including the main study) may not be needed.

General considerations related to investigation of (developmental) neurotoxicity and/or immunotoxicity

If triggers for neurotoxicity or immunotoxicity are identified at REACH Annex VIII or IX level but an extended one-generation reproductive toxicity study is not triggered, a separate neurotoxicity or immunotoxicity study in the developing organism or in adults must be proposed in line with the Column 2 adaptation to Section 8.6.1 of REACH Annex VIII or Section 8.6.2 of REACH Annex IX[102]. Depending on the cases, inclusion of additional parameters to the repeated dose toxicity study (including screening study), if not yet conducted may be considered to further characterise the effect.

Whether the neurotoxic and/or immunotoxic properties should be investigated in adults or in the developing organisms at REACH Annex VIII or REACH Annex IX level if an extended one-generation reproductive toxicity study is not triggered, should be considered on a case by case basis taking into account the various aspects affecting the decision, for example, the target population, toxicokinetics and mode of action. Generally, a study in developing organisms is recommended as a more conservative approach.

At REACH Annex X, the extended one-generation reproductive toxicity study is a standard information requirement, and if there are triggers for the (developmental) neurotoxicity and/or (developmental) immunotoxicity meeting the triggers described in Column 2, Section 8.7.3, the registrant must propose Cohorts 2A and 2B to address the concern for developmental neurotoxicity or Cohort 3 to address the concern for developmental immunotoxicity. The general evaluation of triggers is presented in Appendix R.7.6–5 of this Guidance. Instead of these cohorts, the registrant may propose separate developmental toxicity studies to address these concerns, as explained below in this section under "Proposals for developmental neurotoxicity or immunotoxicity studies". Likewise at REACH Annex IX, if an extended one-generation reproductive toxicity study is triggered, these cohorts, or separate studies, must be proposed by the registrant to address the concern in question.

It should be noted that neurotoxicity and/or immunotoxicity observed in adult animals may trigger developmental neurotoxicity and/or developmental immunotoxicity cohorts in an extended one-generation reproductive toxicity study or in separate studies unless substance specific information is provided why these effects or mode of action would not be relevant in a developing organism (for evaluation of triggers see Stage 3.2.1). In addition, if the classification criteria for STOT are met, based on studies in adults, this is not an adaptation rule allowing the omission of investigations on developmental neurotoxicity and/or developmental immunotoxicity. This is due to expected higher sensitivity of the developing organisms (see e.g. Dietert, 2014), which may lead to a lower DNEL. In addition, a classification to Repr. 1B or 2 may be necessary if the effects are considered to be of developmental origin, i.e. exposure during development. Sensitivity has been evaluated in animal studies for nine reviewed (immuno)toxicants and, according to the authors, the developing immune system was found to be at least as sensitive or more sensitive than the general (developmental) toxicity parameters (Hessel *et al*., 2015).

Proposals for developmental neurotoxicity or immunotoxicity studies

REACH specifies that "*Other studies on developmental neurotoxicity and/or developmental immunotoxicity instead of cohorts 2A/2B (developmental neurotoxicity) and/or cohort 3*

102 Column 2 at Annex VIII, 8.6.1 and Annex IX, 8.6.2: "Further studies shall be proposed by the registrant or may be required by the Agency in accordance with Article 40 or 41 in case of: …- indications of an effect for which the available evidence is inadequate for toxicological and/or risk characterisation. In such cases it may also be more appropriate to perform specific toxicological studies that are designed to investigate these effects (e.g. immunotoxicity, neurotoxicity), …")

(developmental immunotoxicity) of the Extended One-Generation Reproductive Toxicity Study may be proposed by the registrant in order to clarify the concern on developmental toxicity."

The cohorts for developmental neurotoxicity and developmental immunotoxicity included in the extended one-generation reproductive toxicity study provide information on these endpoints. Information on developmental neurotoxicity and developmental immunotoxicity are not standard information requirements in REACH but they must be proposed when particular concerns as specified in Column 2 are met. An advantage of this approach is that fewer animals are needed compared to running three separate studies (reproductive toxicity study, developmental neurotoxicity and developmental immunotoxicity study).

Other studies on developmental neurotoxicity

The registrant has a choice to propose a separate developmental neurotoxicity study instead of Cohorts 2A and 2B if the conditions for a particular concern for developmental neurotoxicity are met. The concern should be related to developmental neurotoxicity specifically. The study design for developmental neurotoxicity should follow the EU B.53 (OECD TG 426) protocol. The selection between the choices should be based on scientific and substance specific considerations taking into account which method adequately addresses the scientific concern with least amount of animals and investigations. However, practical limitations in testing laboratories can also be a reason to propose separate studies. Some examples of aspects of these considerations are presented below.

The developmental neurotoxicity cohort integrated into an extended one-generation reproductive toxicity study contains no endpoints for social or cognitive dysfunctions (e.g. autism, attention deficient hyperactivity disorders, attenuated learning and/or memory), thus, if there are signs of behavioural disturbances from adult animal studies, the design of the developmental neurotoxicity cohort in an extended one-generation reproductive toxicity study might have to be adjusted. Optionally EU B.53 (OECD TG 426) may be the preferred study design.

It should be borne in mind that, when it comes to developmental neurotoxicity, the outcome of a developmental neurotoxicity study (OECD 426) may differ from that of the developmental neurotoxicity Cohorts 2A and 2B in an extended one-generation reproductive study, considering the different exposure scenarios. For example, recent publications point at the importance of a healthy immune system of the mother during pregnancy for brain development of her offspring (Smith *et al.*, 2007); in other words, the maternal impact in the cohort study on nervous system development may be larger than that in the OECD 426 study (exposure from gestation day 6 to PND 21) due to a longer exposure period and the extent of effect often is unknown.

If an extended one-generation reproductive toxicity study is not triggered or a standard information requirement is not met but there are triggers for neurotoxicity, separate studies must be proposed according to REACH Annex VIII, 8.6.1, REACH Annex IX, 8.6.2, or REACH Annex X, 8.6.4.

Other studies on developmental immunotoxicity

The registrant has a choice to propose a separate developmental immunotoxicity study instead of Cohort 3 if the conditions for a particular concern for developmental immunotoxicity are met. The concern should be related to developmental immunotoxicity specifically. For developmental immunotoxicity there is currently no available internationally accepted protocol and thus the registrant must include the proposed protocol in his testing proposal until internationally accepted methods are available. For references to study designs for developmental immunotoxicity see Section R.7.6.4.2.7 of this guidance. The selection between the choices should be based on scientific and substance specific considerations taking into account which method adequately addresses the scientific concern with least amount of animals and investigations. Some examples of aspects of these considerations are presented below.

The nature and/or severity of the triggers may provide guidance to select between a separate study or a cohort. Other aspects to consider may include statistical power and the investigations included. It should be considered whether the cohorts or a separate study best address the particular concern identified (see also Appendix R.7.6–5 of this Guidance).

The outcome of a separate developmental immunotoxicity study may differ from that of the developmental immunotoxicity Cohort 3 in an extended one-generation reproductive study, if the exposure scenarios and set ups are different.

If an extended one-generation reproductive toxicity study is not triggered or a standard information requirement is not met but there are trigger(s) for immunotoxicity, separate studies must be proposed according to REACH Annex VIII, 8.6.1, REACH Annex IX, 8.6.2, or REACH Annex X, 8.6.4.

Common to both developmental neurotoxicity and immunotoxicity studies

Conflicts may arise to decide on the dose levels and premating exposure duration in an extended one-generation reproductive toxicity study. The adequacy of the study design to assess the effects on fertility should be ensured. Thus, the dose level selection should be based upon the fertility endpoint with the developmental neurotoxicity/immunotoxicity being tested at the same dose levels. The fertility endpoint is the only endpoint where *in vivo* data are typically available to make decisions on selecting dose levels for an extended one-generation reproductive toxicity study.

Even if there are trigger(s) for developmental neurotoxicity/immunotoxicity, the dose level setting must not compromise an appropriate investigation of the fertility endpoint. The challenge in deciding the dose levels and length for the premating exposure duration is that there may be a risk that in reducing fertility not enough pups will be produced for example, at the highest dose level for the evaluation of the potential developmental neurotoxicity/immunotoxicity at all dose levels. However, results from lower dose levels can still be used. Another possibility is to add an additional dose level or to address the developmental neurotoxicity/immunotoxicity in (a) separate stud(y)ies.

Evaluation of findings from developmental neurotoxicity and developmental immunotoxicity cohorts

Currently there is not much experience on the interpretation of the results of developmental neurotoxicity (see some considerations under R.7.6.4.2.6 of this Guidance) and developmental immunotoxicity cohorts included in extended one-generation reproductive toxicity studies. Guidance will be developed after gathering more experience. Until further experience on these cohorts, experiences from existing protocols on developmental neurotoxicity and developmental immunotoxicity can be used although all of them may not be standardised and internationally acceptable protocols yet. For evaluation of the results from separate studies, see Sections R.7.6.4.2.6 for developmental neurotoxicity and R.7.6.4.2.7 for developmental immunotoxicity of this Guidance.

Further aspects

The OECD GD 151 provides guidance for conducting the extended one-generation reproductive toxicity study as agreed at OECD level (OECD 2013) but does not for example, define the study design or criteria for the extension of Cohort 1B or the inclusion of cohorts. Thus, the study design should be defined to meet the REACH requirements. OECD GD 117 includes the internal triggers for extension of the Cohort 1B, however, these triggers are not used in REACH as such. The registrant may expand the study based on new information indicating a concern which needs to be addressed. The justification for the expansion must be documented.

For REACH purposes, the focus of the study should be on assessment of the effects on fertility and thus, a ten-week premating exposure duration and dose level setting based on toxicity are required as a starting point as explained above. In addition, for REACH the conditions which specify the extension of the Cohort 1B and the inclusion of Cohorts 2A, 2B and 3 are listed in Column 2 of REACH Annex IX/X, 8.7.3. EU B.56 (OECD TG 443) and OECD GD 151 should be followed only in conducting the study modules. It is recommended that results from a range-finding study (or range-finding studies) for an extended one-generation reproductive toxicity study are reported with the main study. This should support the justifications of the dose level selections, duration of the premating exposure and interpretation of the study results.

The study design of EU B.56 (OECD TG 443) selected must be adequately justified and documented in all cases[103].

In general, all findings on reproductive toxicity should be considered for classification purposes irrespective of the level of concurrent parental toxicity, see the *Guidance on the Application of the CLP criteria* (Chapter 3.7).

Most of the parameters investigated in the 90-day study are also included in the extended one-generation reproductive toxicity study. However, the results obtained may not be equivalent for several reasons and it may not be adequate to adapt the information requirement of a 90-day study by information from an extended one-generation reproductive toxicity study. This is because the 90-day study and the extended one-generation study have different aims. A 90-day study is meant to provide relevant information on systemic and organ-specific toxicity after a subchronic exposure and relevant route especially considering exposure conditions and non-pregnant animals are to be used. Usually the dose level selection for a 90-day study is higher when based on toxicity than the dose levels which can be used in an extended one-generation reproductive toxicity study. This is because the exposure is longer and pregnant animals (and offspring) may be more sensitive than non-pregnant animals. In addition, haematological, clinical chemistry, urinary and histological samples may be collected after a shorter exposure period in an extended one-generation reproductive toxicity study (8-10 weeks) than in a 90-day study (13 weeks) and if conducted in F1 animals, the exposure history and the developmental stages of the animals are different from that in a separate 90-day study. A very careful evaluation is needed when considering whether the information from an extended one-generation reproductive toxicity study can be used to adapt the information requirement of a 90-day study. In certain cases with adequate exposure levels and durations the results from an extended one-generation reproductive toxicity study may support for example, an older but with somewhat limited results, 90-day study.

Information from a 90-day study may be valuable in deciding the dose levels of an extended one-generation reproductive toxicity study.

The extended one-generation reproductive toxicity study provides information on peri-postnatal development but does not address the same parameters as those in the prenatal developmental toxicity study and thus does not provide equivalent information.

R.7.6.4.2.4 Two-generation reproductive toxicity study

Two-generation reproductive toxicity studies are no longer standard information requirements (EU B.35, OECD TG 416) in REACH but those studies initiated before 13 March 2015[104] are considered appropriate to address the standard information requirement for REACH Annex IX/X, 8.7.3. The two-generation reproductive toxicity study was the standard information

---

[103] REACH Art 3(28): "robust study summary: means a detailed summary of the objectives, methods, results and conclusions of a full study report providing sufficient information to make an independent assessment of the study minimising the need to consult the full study report;"

[104] Commission Regulation (EU) 2015/282, Recital (11) and Article 2.

requirement for REACH until the amendment of REACH Annexes IX and X[105]. Two-generation reproductive toxicity studies initiated before the date indicated above are considered appropriate to address the standard information requirement and therefore fulfil the Column 1 requirements, however they do not automatically meet the adaptation criteria described in Column 2. If the available information shows triggers for developmental neurotoxicity and/or developmental immunotoxicity according to Column 2, these particular concerns must be addressed by proposing a separate developmental neurotoxicity and/or a separate developmental immunotoxicity study, respectively (see Section R.7.6.4.2.3, under "Proposals for developmental neurotoxicity or immunotoxicity studies", and R.7.6.4.2.6, and R.7.6.4.2.7 of this Guidance).

Although the two-generation reproductive toxicity study may lack information on some parameters which are part of EU B.56 (OECD TG 443), it addresses the fertility endpoint in two-generations and is adequate for risk assessment and classification and labelling, including categorisation when conducted according to the EU B.35 (OECD TG 416).

From the legal text it is clear that two-generation reproductive toxicity studies initiated after the date indicated in the legislation are not considered appropriate to address the standard information requirement at REACH Annex IX/X, 8.7.3, including the study design adaptation described in Column 2. This means that testing proposals for two-generation reproductive toxicity studies to fulfil the (standard) information requirement at REACH Annex IX/X, 8.7.3 cannot be accepted. If the study already exists and was initiated after March 13, 2015, the registrant may explore the possibilities to adapt the information requirement by substance specific justifications according to REACH Annex XI adaptation rules.

When considering the relevance of old non-guideline compliant two(multi)-generation reproductive toxicity studies to address the fertility endpoint (REACH Annex IX/X, 8.7.3), these studies will be assessed in line with REACH Annex XI, 1.1.2 adaptation rules for existing information. Thus, old existing non-guideline studies may fulfil the Column 1 standard information requirement or may serve as elements in a weight of evidence adaptation according to REACH Annex XI, 1.2 to identify hazardous properties or support a category approach.

R.7.6.4.2.5 One-generation reproductive toxicity study

The one-generation reproductive toxicity study (EU B.34, OECD TG 415) is not an appropriate study to fulfil the information requirement for an extended one-generation reproductive toxicity study because of limited postnatal exposure duration and inadequate coverage of key aspects/parameters (REACH Annex XI, 1.1.2).

This study does not correspond to any REACH standard information requirement but could potentially be enhanced with certain parameters to fulfil the information requirement of the screening study. Compared to the screening study it has a higher statistical power, it addresses the functional fertility by covering the spermatogenesis and folliculogenesis before mating and reproductive performance until weaning. However, the test method lacks requirements of various important parameters as compared with the extended one-generation reproductive toxicity study. Existing studies may be used as one element in a weight of evidence approach according to REACH Annex XI, 1.2 to adapt the standard information requirement of REACH Annex IX/X, 8.7.3 together with other information or to support a category approach.

---

[105] Commission Regulation (EU) 2015/282 of 20 February 2015 amending Annexes VIII, IX and X to Regulation (EC) No 1907/2006 of the European parliament and of the Council on the Registration, evaluation, Authorisation and Restriction of Chemicals (REACH).

Existing studies according to a modified one-generation study protocol may also provide adaptation possibilities. A modified one-generation study is a flexible study design developed by NTP (National Toxicology Program, U.S. Department of Health and Human Services) which has no respective OECD Test guideline or EU Test method available. This study design provides information on reproductive toxicity after exposure from gestation day 6 of the parental animals up to mid gestation of the F1 generation.

R.7.6.4.2.6 Developmental neurotoxicity studies

Developmental neurotoxicity studies are not standard information requirements but may be triggered by REACH Annex VIII point 8.6.1 or REACH Annex XI point 8.6.2 or REACH Annex X point 8.6.4 based on Column 2 adaptation rules[106]. There, the Column 2 adaptation requires the registrant to propose further studies if there are indications of an effect for which the available evidence is inadequate for toxicological evaluation and/or risk characterisation. A separate developmental neurotoxicity study may also be proposed by the registrant instead of the developmental neurotoxicity cohorts (Cohorts 2A and 2B) in an extended one-generation reproductive toxicity study, if these cohorts are triggered.

Developmental neurotoxicity studies (e.g. EU B.53, OECD TG 426) are designed to provide information on the potential functional and morphological hazards of the nervous system arising in the offspring from exposure of the mother during pregnancy and lactation. These studies investigate changes in structure and function of the central nervous system (CNS) and the peripheral nervous system (PNS) using extensive neuropathology (structure) and behavioural (function) surveys. Advanced neuropathology may be assessed including quantitative structural measures as changes in cell structures related to for example, delayed development which may be of quantitative rather than qualitative nature. Such quantitative changes may be significant, but may still go unrecognised without quantification (De Groot *et al.*, 2005). To investigate behaviour a range of parameters, such as a behavioural test battery addressing different functions (domains) of the nervous system, motor activity and more advanced tests addressing cognitive behaviour, are performed. As behaviour may also be affected by the function of other organs such as liver, kidneys and the endocrine system, toxic effects on these organs in the offspring may also be reflected in general changes in behaviour. No single behaviour is able to reflect the entire complex and intricate function of behaviour and so integration of findings of different tests is deemed relevant to evaluate the relevance of the results on substance exposure. Likewise, it may be helpful for the interpretation to review behavioural (functional) changes in light of the neuropathology (structural) findings.

The severity and nature of the effect should be considered. Generally a pattern of effects (e.g. impaired learning during several consecutive trials) is more persuasive evidence of developmental neurotoxicity than one or a few unrelated changes. The reversibility of effects should be considered too. Important to mention in this context is that ‘development’ of an organism *a priori* goes with ‘normal’ structural and functional changes. Under toxic or pathologic circumstances a substance or disease may disturb ‘normal’ development and ‘toxic’ changes are built on top of ‘normal’ developmental changes. The nervous and immune systems are still under development up to and after birth. Moreover, different time-windows have been recognised for speed of developmental growth which in turn, may differ for different parts and structures of the developing nervous and immune systems. As a consequence, the vulnerability of these organ-systems differs during different time-windows of exposure. The nervous system

---

[106] Column 2 at Annex VIII, 8.6.1, Annex IX, 8.6.2, and Annex X, 8.6.4: “Further studies shall be proposed by the registrant or may be required by the Agency in accordance with Article 40 or 41 in case of: …- indications of an effect for which the available evidence is inadequate for toxicological and/or risk characterisation. In such cases it may also be more appropriate to perform specific toxicological studies that are designed to investigate these effects (e.g. immunotoxicity, neurotoxicity), …”)

possesses reserve capacity for repairing. We may for example, find the nervous system impaired during puberty, whereas the adult nervous system seems intact. In such a case, however, one should still realise that not only the trajectory from birth to puberty differed between control and substance-exposed individuals, but the trajectory from puberty to adulthood also differed. So even when a developmental neurotoxicant may not show adverse effects in the adult, the trajectories towards adulthood have been affected and the consequences of this are so far unknown. The nervous system may compensate for damage but the resulting reduction in reserve capacity is of concern and neurotoxicity occurring during development should be regarded as an adverse effect. If developmental neurotoxicity is only observed during part of the lifespan then compensation should be suspected. Also, effects observed for example during the beginning of a learning task but not at the end, should not be interpreted as reversible effects; rather the results may indicate that the speed of learning is decreased.

The experience of offspring especially during infancy may affect their later behaviour. For example, frequent handling of rats during infancy may alter the physiological response to stress and the behaviour in tests for emotionality and learning. In order to control environmental experiences, the conditions under which the offspring are reared should be standardised within experiments with respect to variables such as noise level, handling and cage cleaning. The performance of the animals during the behavioural testing may be influenced by for example, the time of day, and the stress level of the animals. Therefore, the most reliable data are obtained in studies where control and treated animals are tested alternatively and environmental conditions are standardised.

In interpreting the results, maternal toxicity should be taken into account as the development of pups may be affected by maternal toxicity. During early postnatal period pups are dependent of maternal care and maternal toxicity for example, in way of CNS depression, may compromise the survival and development of the pups. In addition, dams and pups should not be separated other than for very short periods of time during the first five postnatal days (e.g. for dose administration) and also later dams should not be moved from cages more than necessary (e.g. for inhalation exposure). In practise this would mean than for inhalation exposure, a whole-body exposure may be considered instead of nose-only exposure.

Adverse effects observed in a development neurotoxicity study will be relevant to hazard classification and the human health risk assessment, providing an N(L)OAEL, unless there is information to show that effects seen in these studies could not occur in humans. Due to a complexity of the endpoint, adversity should preferably be based on a holistic analysis of data by grouping similar parameters.

For more detailed reviews of how to interpret the developmental neurotoxicity results see OECD TG 426, OECD GD 43 and Tyl *et al.* (2008).

R.7.6.4.2.7 Developmental immunotoxicity studies

Developmental immunotoxicity studies are not standard information requirements but may be triggered by REACH Annex VIII point 8.6.1 or REACH Annex IX point 8.6.2 or REACH Annex X, point 8.6.4 based on Column 2 adaptation rules[107]. There, the Column 2 adaptation requires the registrant to propose further studies if there are indications of an effect for which the

[107] Column 2 at Annex VIII, 8.6.1, Annex IX, 8.6.2, and Annex X, 8.6.4: "Further studies shall be proposed by the registrant or may be required by the Agency in accordance with Article 40 or 41 in case of: …- indications of an effect for which the available evidence is inadequate for toxicological and/or risk characterisation. In such cases it may also be more appropriate to perform specific toxicological studies that are designed to investigate these effects (e.g. immunotoxicity, neurotoxicity), …").

available evidence is inadequate for toxicological evaluation and/or risk characterisation. A separate developmental immunotoxicity study may be proposed by the registrant instead of the developmental immunotoxicity cohort (Cohort 3) in an extended one-generation reproductive toxicity study, if these cohorts are triggered.

Developmental immunotoxicity studies are designed to provide information on the potential functional and morphological hazards to the immune system arising in the offspring from exposure of the mother during pregnancy and lactation. Currently there is no OECD test guideline for developmental immunotoxicity testing. Recent reviews provide information on the available approaches and considerations (Gupta (2011), page 219-225; WHO, 2012; De Jong and Van Loveren 2007; DeWitt *et al*., 2012a and 2012b; Dietert and DeWitt, 2010; Dietert and Holsapple, 2007; Holsapple *et al*., 2005; Rooney *et al*., 2009; Boverhof *et al*., 2014).

These studies investigate changes in immune response due to effects on the innate or acquired immune system. As immune response may also be affected by the function of other organs such as liver, kidneys and the endocrine system, toxic effects on these organs in offspring may also be reflected in changes in immune response. No single immune parameter is able to reflect the entire complex and intricate function of immune system and so, integration of findings of different tests is relevant to evaluate the relevance of the results on substance exposure.

Effects considered as adverse will be relevant to hazard classification and the human health risk assessment, providing an N(L)OAEL, unless there is information to show that effects seen in these studies could not occur in humans. Due to a complexity of the endpoint, adversity should preferably be based on a holistic analysis of data by grouping similar parameters.

R.7.6.4.2.8 Repeated-dose toxicity studies

Although not aimed directly at investigating reproductive toxicity, repeated-dose toxicity studies are standard information requirements (e.g. the 28-day study EU B.7, OECD TG 407 or the 90-day study EU B.26, OECD TG 408) and may reveal clear effects on reproductive organs in adult animals. In addition to histopathology of reproductive organs and changes in organ weights, parameters evaluated, such as sperm analysis and measurements of oestrous cycle, may provide relevant information for reproductive toxicity or indicate a concern (trigger(s)). However, no observed effects in measured parameters predicting fertility in repeated dose toxicity studies do not rule out the possibility that the substance may have the capacity to affect fertility. At REACH Annex IX level, triggers for reproductive toxicity from repeated dose toxicity studies trigger an extended one generation reproductive toxicity study (EU B.56, OECD TG 443). At REACH Annex VIII level the registrant may consider proposing an extended one-generation reproductive toxicity study instead of a screening study, based on triggers from a 28-day study.

The observation of effects on reproductive organs in repeated-dose toxicity studies may also be sufficient to be used for classification and labelling and for identifying an N(L)OAEL for use in the risk assessment. It should, however, be noted that the sensitivity of repeated-dose toxicity studies for detecting effects on reproductive organs may be less than reproductive toxicity studies because of the lower number of animals per group (lower statistical power). In addition, a number of cases have demonstrated that effects on the reproductive system may occur at lower doses when animals are exposed during the development or as young animals rather than as adults. Consequently, if there are adverse effects on the reproductive organs in adult animals in the absence of reproductive toxicity studies, an increased assessment factor may be considered in the risk assessment process at REACH Annex VII-VIII levels. An extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) may be triggered based on findings from a repeated dose toxicity study at lower tonnage REACH Annexes, and must be proposed at REACH Annex IX.

The adversity of some effects seen in repeated dose toxicity studies may be difficult to interpret, for example changes in sex hormone levels, and may need to be investigated further as part of studies that may be required to meet standard REACH information requirements (for example EU B.26 (OECD TG 408) or other repeated-dose toxicity studies), rather than serve as a trigger for the immediate conduct of an extended one-generation reproductive toxicity study. Whether or not a finding will serve as a trigger depends on the reliability of the finding and if it can be considered as adverse (see discussions in Appendix R.7.6–5 of this Guidance). It may be considered that statistically significant changes from relevant studies can be considered as triggers; however, sometimes a statistically non-significant change can be also considered as biologically relevant if not contradicting to other available information.

Repeated-dose toxicity studies may also provide indications of a particular concern to evaluate the need to investigate developmental neurotoxicity or developmental immunotoxicity endpoints. The potential triggers for these cohorts in an extended one-generation reproductive toxicity study or separate studies are described in the context of the extended one-generation reproductive toxicity study (Section R.7.6.4.2.3 of this Guidance).

R.7.6.4.2.9 *In vivo* assays for endocrine disruption mode of action

The endocrine system has a critical role in the control of all aspects of the reproductive cycle and therefore endocrine disruption is a potential mechanism for reproductive toxicity. None of the available *in vivo* assays only focusing on identification of endocrine disrupting potency, such as Uterotrophic assay (EU B.54, OECD TG 440) and Herschberger assay (EU B.55, OECD TG 441), correspond to standard REACH information requirements. These studies involve dosing of immature or ovarectomised/castrated animals, and the weighing of oestrogen/ androgen dependent tissues (e.g. uterus or prostate). The methods can be used to identify (anti)oestrogenic or (anti)androgenic modes of action and the results may serve as triggers for further studies in certain cases. These animal models are sensitive to detect the hormonal mode of action. However, only investigation in intact animals proves if the mode of action is relevant in non-manipulated conditions. A comprehensive collection of screening tests and tests for endocrine disrupting chemicals are presented in OECD GD 150 and are included within the "*OECD Conceptual Framework for the Screening and Testing of Endocrine Disrupting Chemicals*".

A result in the uterotrophic assay in a well conducted dose-response study showing no effect indicates that the test substance is not an oestrogen receptor (ER)-ligand in those *in vivo* conditions. Equally, a result in the Hershberger assay showing no effect indicates that the test substance is neither an androgen receptor (AR)-ligand nor a 5-alpha reductase inhibitor in those *in vivo* conditions. A test substance not causing effect in these assays may, however, still have endocrine disrupting properties as well as a potential for reproductive toxicity mediated through other mechanisms. The uterotrophic and Hershberger assays may be used to provide NOEL/LOELs for these endocrine disruption modes of action only if immature (intact) animals are used. The results may also support findings from other studies or serve as triggers for further studies and examinations.

A number of assays in experimental animals may provide information on the ability of a substance to act on the production of steroids and the pubertal assays and the intact male assay may provide information about the endocrine disruption potency of the substance *in vivo* (OECD GD 150). Effects on the various endpoints included in these assays may be considered adverse and/or as representing an effect on a mechanism relevant for humans and serve as triggers for further studies and examinations.

In summary, while these *in vivo* assays in intact animals may be considered predictive for adverse effects on reproduction, they do not provide adequate information on reproductive toxicity for risk assessment and classification and labelling. The repeated dose 28-day oral toxicity study (EU B.7, OECD TG 407) has been updated (2008) to include parameters aiming to identify substances acting through (anti)oestrogenic, (anti)androgenic and (anti)thyroid mechanisms. Validation studies indicate that enhanced design can reliably identify substances with strong potential to act through endocrine modes of action on the gonads and thyroid. A result suggesting no effects in such a study up to the highest dose tested provides some evidence of the absence of potent endocrine activity. However, effects induced by a lower endocrine disrupting potency cannot be ruled out and therefore a result showing no effects does not provide reassurance of the absence of the capability to cause reproductive toxicity via the mechanism of endocrine disruption. Notably in this context, prolongation of exposure from 28 days up to 90 days is unlikely to improve the detectability of endocrine effects (Gelbke *et al*., 2006). Evidence of effects on reproductive organs potentially via endocrine disrupting mode of action seen in a repeated-dose toxicity study provides a trigger for the conduct of a more comprehensive study, i.e. the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) at REACH Annex IX.

The potential triggers related to endocrine disrupting modes of action to be used to define the study design of an extended one-generation reproductive toxicity study are presented along with other triggers in Appendix R.7.6–2 of this Guidance.

The screening studies (OECD TGs 421 or 422) may be adopted [108] with additional parameters for endocrine disrupting modes of action, including measurements of anogential distance, nipple/areolae retention, and thyroid hormone) levels These parameters indicate endocrine disrupting mode of action and may be predictive for adverse effects on reproduction. A statistically significant change in anogenital distance that cannot be explained by the body weight/size of the animal indicates an antiandrogenic mode of action and should be used for setting the NOAEL. To support the adversity of this parameter an association with reduced human reproduction has been reported (Jain and Singal, 2013; Eisenberg *et al*., 2011 and 2012; Mendiola *et al*., 2011). A statistically significant change in nipple/areolae retention indicates also an antiandrogenic mode of action but likely via other spectrum of mechanisms than that of anogenital distance. Due to the difference in biology in controlling the final number of nipples between male rats and human, it is not possible to study the association between nipple/areolae retention findings in rats and adversity in humans as for anogenital distance. However, as the assumed mode of action (antiandrogenicity) and potential underlining mechanisms affecting nipple/areolae retention in rats are also relevant to humans, although not causing similar effects, this finding can be considered likely to predict an adverse effect and used to set the NOAEL. Nipple/areolae retention measures the same mode of action (antiandrogenicity) as anogenital distance but due to different tissue specific underlining mechanisms and possibly toxicokinetic differences nipple/areolae retention may be more or less sensitive than anogenital distance. It is recommended that these endpoints are evaluated together.

As the extended one-generation reproductive toxicity study is a more comprehensive reproductive toxicity study which includes certain parameters to detect endocrine disrupting modes of action, it may be possible a) to identify an endocrine disrupting mode of action, b) to identify an adverse effect on reproduction, and c) for both (a) and (b) not necessarily indicating a causal relationship. If an endocrine disrupting mode of action is identified without an adverse effect on reproduction (e.g. reduced thyroid hormone level in pups), further studies or actions may be considered. If the findings on reproduction meet the classification criteria to Category 1B reproductive toxicant, irrespective indications of an endocrine disrupting mode of action, the substance should be classified accordingly.

R.7.6.4.3 Human data on reproductive toxicity

Epidemiological data require a detailed critical appraisal that includes an assessment of the adequacy of controls, the quality of the health effects and exposure assessments, and of the influence of bias and confounding factors. Epidemiological studies can generally only provide

[108] OECD TGs 421 and 422 are in the process of being revised: adoption and publication is expected by the end of 2015.

associations, not causality because, although it may be possible to show the link and estimate the likelihood of the causality,it cannot give a final proof.

Epidemiological studies, case reports and clinical data may provide sufficient hazard and dose-response evidence for classification of chemicals as reproductive toxicants in Category 1A and for risk assessment, including the identification of a NAEL or LAEL. In such cases, there will not normally be a need to test the substance. However, convincing human evidence of reproductive toxicity for a specific substance is rarely available because it is often impossible to identify a population suitable to study that is/was exposed only to the substance of interest. Human data may provide limited evidence of reproductive toxicity that indicates a need for further studies of the substance; the test method selected should be based on the potential effect suspected.

When evidence of a reproductive hazard has been derived from animal studies it is unlikely that the absence of evidence of this hazard in an exposed human population will negate the concerns raised by the animal model. This is because there will usually be methodological and statistical limitations to the human data. For example, statistical power calculations indicate that a prospective study with well-defined exposure during the first trimester with 300 pregnancies could identify only those developmental toxins that caused at least a 10-fold increase in the overall frequency of malformations; a study with around 1000 pregnancies would have power to identify only those developmental toxins that caused at least a 2-fold increase (EMA/CHMP Guideline, 2006). Extensive, high quality and preferable prospective, data are necessary to support a conclusion that there is no risk from exposure to the chemical. Thus, the absence of effects in humans at a dose level below the dose levels inducing reproductive toxicity in animals, will not negate the concerns raised by the animal model.

R.7.6.4.4 Derivation of DNELs and DMELs

Identification of DNEL(s) are referred to in REACH Annex I, 1.4. Depending on the available information and the exposure scenario(s), it may be necessary to identify different DNELs for each relevant human population (consumers, professional, workers, humans exposed indirectly via the environment and certain vulnerable subpopulations (children, pregnant woman) and for different routes of exposure and all routes combined. In certain cases exposure from various sources may need to be considered. For reproductive toxicity endpoints it is especially relevant to consider deriving the different DNELs for vulnerable subpopulations.

Generally, effects on reproduction have been considered as effects having a threshold and thus allowing derivation of a DNEL. However, in certain cases, the possibility for a non-threshold mode of action may need to be considered (e.g. if a substance has (anti)hormonal activity similar to a hormone having a primary biological control role and there is a concern of lack of body's regulation capacity). For these cases derivation of DMEL may need to be considered.

In order to be suitable for CSA appropriate DNELs (a DNEL for fertility and a DNEL for development) have to be established for each exposure scenario and each population exposed. Typically, the derivation of the DNEL takes into account a dose descriptor, modification of the starting point and application of assessment factors – see the *Guidance on IR&CSA, Chapter R.8 Characterisation of dose [concentration]-response for human health* (Section R.8.4 and Appendix R.8-12) and Section R.7.6.4.3 of this Guidance. Appendix R.8-12 Reproductive toxicity provides specific advice for reproductive toxicity studies.

R.7.6.5 Classification and labelling

Guidance on classification and labelling is given in the *Guidance on the Application of the CLP criteria* (Section 3.7) and specifically for parental toxicity see Section 3.7.2.2.1 *Classification in the presence of parental toxicity*.

R.7.6.6 Conclusions on reproductive toxicity

Reproductive toxicity endpoints should be considered separately for establishing the relevant endpoint(s) and NOAEL(s) to be used in risk assessment (for fertility and developmental toxicity endpoints) and for classification (for sexual function and fertility; developmental toxicity; and lactation). The study or studies giving rise to the highest concern must normally be used to establish the DNEL(s) (see REACH Annex I, 1.2.4). If another study / other studies are used, an acceptable justification for this exception needs to be provided. Derivation of DMEL needs to be considered if adverse effects are likely to be induced via a non-threshold mode of action.

Risk assessment and determination of classification involves the consideration of all data that is available and may be relevant to reproductive toxicity (see Section 0 of this Guidance for different data sources). There can be no firm rules on how to conduct the risk assessment and determination of classification for hazards as this process involves expert judgment and also because the mix and reliability of information available for a particular substance will probably be unique. Also data resulting from studies on other hazards, for example, repeated dose toxicity, can be relevant for consideration in the risk assessment and determination of classification of reproductive toxicity.

In order to conclude on a hazard classification and category, all the available information needs to be taken into account, and compared with the criteria in Annex I of the CLP Regulation (see also the *Guidance on the Application of the CLP criteria*). If the information is not adequate to decide on classification and labelling, the registrant must indicate and justify the action or decision he has taken as a result of inadequate data (REACH Annex VI, 4.1 and REACH Annex VI, 1.3.2).

If the substance has an EU harmonised classification for Reproductive toxicity (included in Annex VI, CLP) or meets the classification criteria and is subject to self-classification, exposure scenarios should be established and the risk characterisation ratio (RCR) calculated to indicate the safe use of the substance.

R.7.6.7 Integrated Testing Strategy (ITS) for reproductive toxicity

Section R.7.6.2 of this Guidance, includes guidance on how to define and generate relevant information on substances in order to meet the information requirements and address the concerns related to intrinsic properties of substances related to reproductive health.

An integrated testing strategy (ITS) may be defined as an approach which combines one or more non-animal methods with animal studies to fulfil the information requirements or could only include non-animal methods which covered all key aspects of reproductive toxicity. Thus, REACH Annex XI adaptations (with the exception of Section 3.2.a – substance tailored exposure-driven testing) play an important role in ITSs for reproductive toxicity. An ITS must produce information usable for a robust risk assessment and/or for classification and labelling. A definition for ITS is given by Blaauboer *et al*., (1999)[109]. The ITS concept is similar to that of IATA, Integrated Approaches to Testing and Assessment. In principle, ITS and IATA are approaches where information is collected, evaluated and weighted with the aim to provide a sufficient amount of information by development of the weight of evidence. ITS and IATA could be used with a view to generate information in a step-wise approach, allowing for justifying an adaptation of one or more standard information requirements according to REACH REACH

[109] "An Integrated Testing Strategy is any approach to the evaluation of the hazard which serves to reduce, refine or replace an existing animal procedure, and which is based on the use of two or more of the following: physicochemical data, in vitro data, human data (for example, epidemiological, clinical case reports), animal data (where unavoidable), computational methods (such as quantitative structure activity relationships (QSARs) and biokinetic models" (Blaauboer *et al*., 1999).

Annex XI, 1.2. (weight of evidence) taking into account that REACH Annex XI, 1.2 is a hazard-based approach and exposure and risk-based consideration cannot be used.

A comprehensive use of ITS for reproductive toxicity endpoint requires knowledge on all different mechanistic steps and processes involved in the outcome of a possible adverse effect. Reproductive toxicity relates to a number of potential target tissues and comprises a very large number of interacting processes, which are not even known in their entirety and which at present are far from being fully understood in their complexity. Another particular challenge in the identification of reproductive toxicity effects relates to the potential impact of systemic toxicity on the fertility and maternal toxicity on the development of the offspring. The existence of windows of particular sensitivity during the development of the embryo is another characteristic feature of reproductive toxicity. However, currently adverse outcome pathways (AOPs) are under development each covering one specific effect for example, vasculogenesis and cleft palates. It is to be noted that also the specific effects like clefts can be formed via several different mechanisms and AOPs increasing the complexity. AOPs may form a basis for ITS/IATA in describing the key events in toxicity pathways that need to be addressed by and ITS/IATA.

Combined approaches including various methods may be used as preliminary steps only because they do not provide equivalent information on the standard information requirements. In addition they may be elements in a weight of evidence adaptation according to REACH Annex XI, 1.2 approach or supporting categories and read across according to REACH Annex XI, 1.5 approach. However, as these combined approaches include more uncertainty due to missing parts of information; this should be addressed when such approaches are proposed. As all the potential molecular mechanisms and regulatory mechanisms are not covered these approaches may not be appropriate to prove the absence of an effect. Currently derivation of a NOAEL is not possible with these methods.

R.7.6.8 References on reproductive toxicity

Adori M, Kiss E, Barad Z, Barabas K, Kiszely E, Schneider A, Sziks E, Abraham IM, Matko J and Saramay G (2010) Estrogen augments the T cell-dependent but not the T-independent immune response. Cell Mol Life Sci 67:1661-74.

Bailey SA, Zidell RH, Perry RW (2004) Relationships between organ weight and body/brain weight in the rat: what is the best analytical endpoint? Toxicol Pathol 32:448-66.

Beekhuijzen M, de Raaf MA, Zmarowski A, van Otterdijk F, Peter B and Emmen H (2014) The underestimated value of OECD 421 and 422 repro screening studies: putting it in the right perspective. Reprod Toxicol 48:81-7.

Blagojević J, Jovanović V, Stamenković G, Jojić V, Bugarski-Stanojević V, Adnađević T and Vujošević M (2012) Age Differences in Bioaccumulation of Heavy Metals in Populations of the Black-Striped Field Mouse, Apodemusagrarius (Rodentia, Mammalia) Int J Environ Res 6:1045-52.

Blaauboer BJ, Barratt MD, Houston JB (1999) The integrated use of alternative methods in toxicological risk evaluation . ECVAM Integrated Testing Strategies Task Force Report 1. Altern Lab Anim 27:229-37.

Boverhof DR, Ladics G, Luebke B, Botham J, Corsini E, Evans E, Germolec D, Holsapple M, Loveless S, Lu H, van der Laan JW, White Jr KI and Yang Y (2014) Approaches and considerations for the assessment of immunotoxicity for environmental chemicals: A workshop summary. Regul Toxicol Pharmacol 68:96-107.

Bonde JP (2010) Male reproductive organs are at risk from environmental hazards. Asian J Androl 12:152.

Chapin RE and Sloane RA (1997) Reproductive assessment by continuous breeding: evolving study design and summaries of ninety studies. Environ Health Perspect 105 Suppl 1:199-205.

De Groot DM, Bos-Kuipers MH, Kaufmann WS, lammers JH, O'Callahan JP, Pakkenberg B, Pelgrim MT, Waalkens-Berendsen ID, Waanders MM, Gundersen HJ (2005) Regulatory developmental neurotoxicity testing: a model study focussing on conventional neuropathology endpoints and other perspectives. Environ Toxicol Pharmacol 19:745-55.

De Groot MW, Westerink RH, Dingemans MM (2013) Don't judge a neuron only by its cover: neuronal function in in vitro developmenatl neurotoxicity testing. Toxicol Sci 132:1-7.

De Jong W, Van Loveren H, (Eds) (2007) Animal Models in Immunotoxicology. Methods Special Issue 41:1-142.

DeWitt JC, Peden-Adams MM, Keil DE and Dietert RR (2012a) Current status of developmental immunotoxicity: early-life patterns and testing. Toxicol pathol 40:230-36.

DeWitt JC, Peden-Adams MM, Keil DE and Dietert RR (2012b) Developmental immunotoxicity (DIT): assays for evaluating effects of exogenous agents on development on the immune system. Curr Protoc Toxicol Chapter 18:Unit 18.15.

Dietert RR (2014) Developmental immunotoxicity, perinatal programming, and noncommunicable disease: Focus on human studies. Advances in Medicine Vol 2014:18 pages.

Dietert RR and DeWitt J (2010) Developmental immunotoxicity (DIT): the why, when, and how of DIT testing. Methods Mol Biol 598:17-25.

Dietert RR and Holapple MP (2007) methodologies fro developmental immunotoxicity (DIT testing. Methods 41:123-131.

ECHA: Guidance on the Application of the CLP Criteria [http://echa.europa.eu/web/guest/guidance-documents/guidance-on-clp]

ECHA: Guidance on information requirements and chemical safety assessment: [http://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment]

EMA (European Agency for the Evaluation of Medicinal Products) (2006) Guideline on risk assessment of medicinal products on human reproduction and lactation: from data to labelling. EMEA/CHMP/203927. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003307.pdf

Eisenberg ML, Hsieh MH, Walters RC, Krasnow R and Lipshultz LI (2011) The relationship between anogenital distance, fatherhood, and fertility in adult men. PLoS One 6(5):e18973.

Eisenberg ML, Jensen TK, Walters RC, Skakkebaek NE, Lipshultz LI (2012) The relationship between anogenital distance and reproductive hormone levels in adult men. J Urol 187:594-8.

Evgeni E, Charalabopoulos K, Asimakopoulos B (2014) Human Sperm DNA Fragmentation and its Correlation with Conventional Semen Parameters. J Reprod Infertil 15:2–14.

Fryer CA, Bo E, Calamandrei G, Calzà L, Dessì-Fulgheri F, Fernández M, Fusani L, Kah O, Kajta M, Le Page Y, Patisaul HB, Venerosi A, Wojtowicz AK and Panzica GC (2012) Endocrine disrupters : a review of some sources, effects, and mechanisms of actions on behaviour and neuronedocrine systems. J Neuroendocrinol 24:144-59.

Genschow E, Spielmann H, Scholz G, Pohl I, Seiler A, Clemann N, Bremer S and Becker K (2004) Validation of the embryonic stem cell test in the international ECVAM validation study on three *in vitro* embryotoxicity tests. Altern Lab Anim 32:209-44.

Genschow E, Spielmann H, Scholz G, Seiler A, Brown N, Piersma A, Brady M, Clemann N, Huuskonen H, Paillard F, Bremer S and Becker K (2002) The ECVAM international validation study on *in vitro* embryotoxicity tests: results of the definitive phase and evaluation of prediction models. Altern Lab Anim 30:151-76.

Gelbke H-P, Fleig H, Meder M on behalf of German Chemical Industry Association (2004) SIDS Reprotoxicity screening test update: testing strategy and use. Reg Toxicol Pharm 39:81-6.

Gelbke HP, Hofmann A, Owens JW and Freyberger A (2006) The enhancement of the subacute repeat dose toxicity test OECD TG 407 for the detection of endocrine active chemicals: comparison with toxicity tests of longer duration. Arch Toxicol 81:227-50.

Gupta RC (Ed) (2011) Reproductive and Developmental Toxicology. Elsevier Inc Academic Press, Amsterdam, The Netherlands.

Hardin BD, Becker RA, Kavlock RJ, Seidenberg JM, Chernoff N (1987) Workshop on the Chernoff/Kavlock preliminary developmental toxicity test. Teratog Carcinog Mutagen 7:119-27.

Hessel EV, Tonk ECM, Bos PM, van Loveren H and Piersma AH (2015) Developmental immunotoxicity of chemicals in rodents and its possible regulatory impact. Crit Rev Toxicol 45:68-82.

Holsapple MP, Burns-Naas LA, Hastings KL, Ladics GS, Lavin AL, Makris SL, Yang Y and Luster MI (2005) A proposed testing framework for developmental immunotoxicity (DIT). Toxicol Sci 83:18-24.

Hood RD (ed) (2011): Developmental and Reproductive Toxicology: A practical Approach, Third Edition. Informa Healthcare, London, UK.

Jacobson-Kram D and Keller KA (Eds) (2006) Toxicological Testing Handbook: Principles, Applications and Data Interpretation, Second Edition. Informa Healthcare, New York, NY, USA. pp 220-221.

Jain VG and Singal AK (2013) Shorter anogenital distance correlates with undescended testis: a detailed genital anthropometric analysis in human newborns. Hum Reprod 9:2343–9.

Kerr JB, Loveland KL, O'Bryan MK, de Kretser DM (2006) Chapter 18 - Cytology of the Testis and Intrinsic Control Mechanisms. *In:* Neill JD, Plant TM, Pfaff DW, Challis JRG, de Kretser DM, Richards JS and Wassarman PM (eds) Knobil and Neill's Physiology of Reproduction, Third Edition, pp. 827-947. Elsevier Inc Academic Press, Amsterdam, The Netherlands.

Korach KS (Ed) (1998) Reproductive and Developmental Toxicology. Marcel Dekker Inc, New York, USA.

McGee EA and Hsueh AJ (2000) Initial and cyclic recruitment of ovarian follicles. Endocr Rev 21:200-14.

Mendiola J, Stahlhut RW, Jørgensen N, Liu F and Swan SH (2011) Shorter anogenital distance predicts poorer semen quality in young men in rochester, new york. Environ Health Perspect 119:958-63.

Nielsen E, Ostergaard G and Larsen JC (2008) Toxicological Risk Assessment of Chemicals: A Practical Guide, CRC Press, Boca Raton, FL, USA, p. 143.

NTP modified one-generation studies. National Toxicology Program U.S. Department of Health and Human Services. http://ntp.niehs.nih.gov/testing/types/mog/index.html

OECD (2006) OECD Guidance Document No 57. Detailed review paper on thyroid hormone disruption assays. ENV/JM/MONO(2006)24. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2007) Series on testing and assessment, No. 71. Guidance document on the uterotrophic bioassay – procedure to test for antioestrogenicity. ENV/JM/MONO(2007)15. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2008a) OECD Guidance Document No 43. Guidance document on mammalian reproductive toxicity testing and assessment. ENV/JM/MONO(2008)16. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2008) Series on testing and assessment, Number 89. Retrospective performance assessment of the test guideline 426 on developmental neurotoxicity. ENV/JM/MONO(2008)15. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2009a) OECD Guidance Document No 39. Guidance document on acute inhalation toxicity testing. ENV/JM/MONO(2009)28. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2009b) OECD Guidance Document No 106. Guidance document for histologic evaluation of endocrine and reproductive tests in rodents. ENV/JM/MONO(2009)11. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2009c) OECD Guidance document on the weanling Hersberger bioassays in rats: a short-term screening assays for (anti)androgenic properties. Series testing and assessment, number 115. ENV/JM/MONO(2009)41. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2011) OECD Guidance Document No 117. Guidance document 117 on the current implementation of internal triggers in the test guideline 443 for an extended one-generation reproductive toxicity study, in the United States and Canada. ENV/JM/MONO(2011)21. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2012) OECD Guidance Document No 150. Guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption. ENV/JM/MONO(2012)22. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2013a) OECD Guidance Document No 151. Guidance document supporting OECD test guideline 443 on the extended one-generation reproductive toxicity test. ENV/JM/MONO(2013)10. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2013b) OECD Guidance Document No 184. Guidance document on developing and assessing adverse outcome pathways. ENV/JM/MONO(2013)6. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD (2014) Series on testing and assessment, No. 207. New scoping document on *in vitro* and *ex vivo* assays for the identification of modulators of thyroid hormone signalling. ENV/JM/MONO(2014)23. Available at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

Pallarés ME, Adrover E, Imsen M, González D, Fabre B, Mesch V, Baier CJ and Antonelli MC (2014) Maternal administration of flutamide during late gestation affects brain and reproductive organs development in the male rat offspring. Neurosci 278:122-35.

Piersma AH, Genschow E, Verhoef A, Spanjersberg MQ, Brown NA, Brady M, Burns A, Clemann N, Seiler A and Spielmann H (2004) Validation of the postimplantation rat whole-embryo culture test in the International ECVAM validation study on three in vitro embryotoxicity tests. Altern Lab Anim 32:275-307.

Piersma AH (2006) Alternatives to animal testing in developmental toxicity testing. Basic Clin Toxicol Pharmacol 98:427-31.

Reuter U, Heinrich-Hirsch B, Hellwig J, Holzum B and Welsch F (2003) Evaluation of OECD screening tests 421 (reproduction/developmental toxicity screening test) and 422 (combined repeated-dose toxicity study with the reproduction/developmental toxicity screening test). Reg Toxicol Pharm 38:17-26.

Rooney AA, Yang Y and Makris SL (2009) Recent progress and diverge effects in developmental immunotoxicology: overview of a symposium at the 46th Annual SOT Meeting, Charlotte, NC. J Immunotoxicol 5:395-400. Erratum in J Immunotox 6:74.

Russell LD, Ettlin RA, Sinha Hikim AP and Clegg ED (Eds) Histological and Histopathological Evaluation of the Testis, Cache River Press, Clearwater, FL (1990).

Sellers RS, Morton, D, Michael B, Roome N, Johnson JK, Yano BL, Perry R and Schafer K (2007) Society of Toxicological Pathology position paper: organ weight recommendations for toxicology studies. Toxicol Pathol 35:751-5.

Sjöblom T, Parvinen M and Lähdetie J (1995) Stage-specific DNA synthesis of rat spermatogenesis as an indicator of genotoxic effects of vinblastine, mitomycin C and ionizing radiation on rat spermatogonia and spermatocytes. Mutat Res 1995 331:181-90.

Smith SEP, Li J, Garbett K, Mirnics K and Patterson PH (2007) maternal immune activation alters fetal brain development through Interleukin-6. J Neurosci 27:10695–702.

Spielmann H, Genschow E, Brown NA, Piersma AH, Verhoef A, Spanjersberg MQ, Huuskonen H, Paillard F and Seiler A (2004) Validation of the rat limb bud micromass test in the international ECVAM validation study on three in vitro embryotoxicity tests. Altern Lab Anim 32:245-74.

Spielmann H, Seiler A, Bremer S, Hareng L, Hartung T, Ahr H, Faustman E, Haas U, Moffat GJ, Nau H, Vanparys P, Piersma A, Riego Sintes J and Stuart J (2006) The practical application of three validated *in vitro* embryotoxicity tests. The report and recommendations of an ECVAM/ZEBET workshop (ECVAM workshop 57). Altern Lab Anim 34:527-38.

Trigunaite A, Dimo J, Jørgensen Tn (2015) Suppressive effets of androgens on the immune system. Cell Immunol 294:87-94.

Tyl RW, Crofton K, Moretto A, Moser V, Sheets LP and Sobotka TJ (2008) Identification and interpretation of developmental neurotoxicity effects: a report from the ILSI Research Foundation/Risk Science Institute Expert Working Group on Neurodevelopmental Endpoints. Neurotoxicol Teratol 30:349-81.

Van Dartel DAM, Pennings JLA, van Schooten FJ, Piesrma AH (2010) Transcriptomics-based indentification of developmentl toxicants through their interference with cardiomyocyte differentiation of embryonic stem cells. Toxicol Applied Phamacol 243:420-28.

Worth A, Barroso J, Bremer S, Burton J, Casati S, Coecke S, Corvi R, Desprez B, Dumont C, Gouliarmou V, Goumenou M, Gräpel R, Griesinger C, Halder M, Janusch Roi A, Kienzler A, Madia F, Munn S, Nepelska M, Paini A, Price A, Prieto P, Rolaki A, Schäffer M, Triebe J, Whelan M, Wittwehr C, Zuang V (2014) Alternative methods for regulatory toxicologu – a state-of-the-art review. European Commission, JRC Science and policy reports. Report EUR 26797 EN. https://ec.europa.eu/jrc/sites/default/files/echa_jrc_sla_report_public_05-09-14_withcover_ipo.pdf

WHO/IPCS (2012) Guidance for Immunotoxicity Risk Assessment for Chemicals, IPCS Harmonisation Project Document No 10. Available at: http://www.who.int/ipcs/methods/harmonization/areas/immunotoxicity/en/.

Appendices R.7.6-1 to 5 to Section R.7.6

NOTE to the reader: The references cited in the Appendices are given in Section R.7.6.8 References on Reproductive Toxicity

Appendix R.7.6–1 A check list for information that contributes to EOGRTS design

This is a "check list" for information (sources) that should be checked in order to establish the existence or the nonexistence of the triggers and conditions specifying the study design of an extended one-generation reproductive toxicity study for REACH. Please note that this is not advice on how to conduct an evaluation of the data.

The information is expected to be derived from the substance itself but if it is a surrogate, such as a component of a multiconstituent substance, the triggers from all the components and metabolites must be considered and justified.

More details and examples of triggers are provided in Appendix R.7.6–2 (EOGRTS study design) of this Guidance and length of the premating exposure duration is discussed in Appendix R.7.6–3 of this Guidance.

Condition / trigger	Where to find the information to decide on the existence or nonexistence of the triggers and conditions
E1: Uses leading to significant exposure of consumers or professional, taking into account inter alia consumer exposure from articles	Consumer and/or professional uses (one very wide uses or several limited uses) of a substance as neat(concentrate), in a mixture, in an article with intended release, or in an article with unintended migration from the matrix. The registrant must record and justify the existence or nonexistence of any of the conditions above. If any of these exist together with any of the other three conditions below (E2, E3 or E4), fulfilling the criteria detailed in Appendix R.7.6–2 of this Guidance, then the extension of the Cohort 1 B must be proposed.
E2: Genotoxicity potentially meeting classification criteria to Mutagen Category 2	Results from *in vivo* mutagenicity studies (if one of the *in vitro* tests is positive, then an *in vivo* somatic cell mutagenicity test must have been conducted). The registrant must record the findings and justify the existence or nonexistence of the condition.
E3: Extended exposure is needed to reach the steady state kinetics.	Indications on the exposure duration needed to reach the steady state can be obtained from various sources. • toxicokinetic studies in animals • human data, e.g. substance or metabolite(s) level(s) in blood or organs. • existing *in vivo* studies with long exposure duration showing unexpected severity or occurrence of findings compared with studies with short exposure. • Any other indication of potential to accumulate, such as prediction from log Pow, non-animal approaches (QSAR predictions), information from eco-toxicity (elevated levels in biota, high levels at the top of food chain, very slow depuration, bioaccumulation potency (B or vB, or similar concern), biomagnifications) All the components and metabolites of the multicomponent substance (multiconstituent or UVCB substances) must be considered and justified. The registrant must record the findings and justify the existence or nonexistence of the condition.
E4: Indications of modes of action related to endocrine disruption from *in vivo* or non-animal approaches	Repeated dose toxicity studies and reproductive toxicity studies may provide indication of endocrine disrupting modes of action. Check the parameters related to endocrine modes of action. Check *in vivo* assays for endocrine (disrupting) modes of action.

Condition/trigger	**Where to find the information to decide on the existence or nonexistence of the triggers and conditions**
	Check the non-animal approaches for prediction to endocrine (disrupting) modes of action. Check data from eco-toxicity testing for predicting endocrine (disrupting) modes of action The registrant must record the findings and justify the existence or nonexistence of the condition.
N1: Information on neurotoxicity from *in vivo* studies or non-animal approaches.	*In vivo* toxicity studies may provide information on neurotoxicity. Check all the parameters related to nervous system. Check the non-animal approaches for prediction of (developmental) neurotoxicity. The registrant must record the findings and justify the existence or nonexistence of the triggers and particular concern for developmental neurotoxicity.
N2: Specific mechanism/modes of action with association to (developmental) neurotoxicity.	Some studies may include measurements which reveal the mechanism, or there may be specific mechanistical studies (*in vivo* or *in vitro*) available. The registrant must record the findings and justify the existence or nonexistence of the triggers and particular concern for developmental neurotoxicity.
N3: Existing information on (developmental) neurotoxicity from structurally analogous substances	Structurally analogous substances should be identified and existing information on effects showing (developmental) neurotoxicity must be checked. The registrant must record the findings and justify the existence or nonexistence of the triggers and particular concern for developmental neurotoxicity.
I1: Information on immunotoxicity from *in vivo* studies or non-animal approaches.	*In vivo* toxicity studies may provide information on immunotoxicity. Check all the parameters related to immune system. Check the non-animal approaches for prediction of (developmental) immunotoxicity. The registrant must record the findings and justify the existence or nonexistence of the triggers and particular concern for developmental neurotoxicity.
I2: Specific mechanism/modes of action with association to (developmental) immunotoxicity.	Some studies may include measurements which reveal the mechanism or there may be specific mechanistical studies (*in vivo* or *in vitro*) available. The registrant must record the findings and justify the existence or nonexistence of the triggers and particular concern for developmental immunotoxicity.
I3: Existing information on (developmental) immunotoxicity from structurally analogous substances	Structurally analogous substances should be identified and existing information on effects showing (developmental) immunotoxicity must be checked. The registrant must record the findings and justify the existence or nonexistence of the triggers and particular concern for developmental immunotoxicity.

Appendix R.7.6–2 EOGRTS Study Design

The registrant must propose the study design for an extended one-generation reproduction toxicity study with the following specifications. Relevant justifications are needed for the study design, including the existence or nonexistence of the conditions for extension of the Cohort 1B and trigger(s) for the Cohorts 2A and 2B, and Cohort 3.

Specifications for study designs in REACH are needed for the following aspects:

1) Premating exposure duration and dose level selection;
2) The need to extend the reproduction toxicity Cohort 1B and to define the termination time for F2;
3) The need to include the developmental neurotoxicity Cohorts 2A and 2B;
4) The need to include the developmental immunotoxicity Cohort 3.

In the following text the specifications and triggers (conditions) are presented for each study design. The Table in Appendix R.7.6–1 of this Guidance provides a check list for the registrants in order to provide a short list of studies/tests which could provide information on triggers to specify the study design of an extended one-generation reproductive toxicity study. The existence or the nonexistence of triggers (conditions) must be recorded in order to allow an independent evaluation.

The study design should be decided before the study is started. For REACH the in-study triggers are not recommended. However, the registrant may expand the study based on new information (that arises after the ECHA Evaluation decision has been issued) indicating a concern which needs to be addressed. The justification for the expansion must be documented.

The OECD guidance document GD 151 provides guidance for the conduct of cohorts of the extended one-generation reproductive toxicity study (OECD 2013) but the study design applicable for REACH and CLP is outlined in REACH Annexes IX and X and Recital (7) of Commission Regulation (EC) 2015/282 amending REACH and described in more detail in this guidance.

Specifications needed in testing proposals:

1) Premating exposure duration and dose level selection

Recital (7) of Commission Regulation (EC) No 2015/282 of 20 February 2015 amending REACH states that an extended one-generation reproductive toxicity study should allow adequate assessment of fertility and that premating exposure duration and dose levels should be appropriate to meet the risk assessment and classification and labelling purposes[110].

Both the length of premating exposure duration and dose level setting are aspects which influence the possibility to adequately assess potential adverse effects on fertility. In order to adequately address the assessment of the fertility endpoint, the starting point for deciding on the length of premating exposure period should be ten weeks to cover the full

---

[110] Recital (7) of Commission Regulation (EU) No 2015/282 Of 20 February 2015 amending Annexes VIII, IX and X to Regulation (EC) No 1907/2006 of the European parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) as regards the Extended one-generation reproductive toxicity study: "*It should be ensured that the reproductive toxicity study carried-out under point 8.7.3 of Annexes IX and X to Regulation (EC) No 1907/2007 will allow adequate assessment of possible effects on fertility. The premating exposure duration and dose selection should be appropriate to meet risk assessment and classification and labelling purposes as required by Regulation (EC) No 1907/2006 and Regulation (EC) No 1272/2008 of the European parliament and of the Council.*"

spermatogenesis and folliculogenesis before the mating, allowing meaningful assessment of the effects on fertility. The exposure can be started when the animals are around 5 weeks old and mate them around 15 weeks of age. However, based on substance specific justifications a shorter premating exposure duration may be proposed, but it should not be shorter than two weeks. Further discussion on premating exposure duration is provided in Appendix R.7.6–3 of this Guidance. If the registrant prefers another length of premating exposure duration, an acceptable substance-specific scientific justification must be provided.

The highest dose for an extended one-generation reproductive toxicity study should be selected with the aim to induce some toxicity (or to use the limit dose of 1000 mg/kg bw/day if humans are not exposed to higher dose levels), in order to allow a conclusion on whether effects on reproduction are considered to be secondary, non-specific consequence of other toxic effects seen (see also the dose level selection under Section R.7.6.2.3.2, Stage 4.1(6) of this Guidance). Only in this way is it possible to assess if the substance is a reproductive toxicant and/or if the effects on reproduction are potentially associated with systemic toxicity and to what extent.

The possibility to select the highest dose level, based on the toxicokinetic data as mentioned in EU B.56 (OECD TG 443) and in the OECD GD 151, may not allow comparison of adverse effects on fertility with systemic toxicity and, thus, does not support production of data for classification and labelling purposes, including categorisation. Regarding the highest dose level, it is important to ensure that toxicity in both female and male animals is considered to ensure that reproductive toxicity in either gender is not overlooked.

Both the ten weeks premating exposure duration and the highest dose level meeting the requirement of inducing toxicity, should allow conclusion on classification and labelling, including categorisation, for the hazard endpoint for sexual function and for fertility according to CLP.

2) Extension of Cohort 1B and termination time for F2

REACH specifies that the extension of cohort 1B to include the F2 generation shall be proposed by the registrant or may be required by the Agency if:

a) *"the substance has uses leading to significant exposure of consumers or professionals, taking into account, inter alia, consumer exposure from articles, and*

b) any of the following conditions are met:

- *the substance displays genotoxic effects in somatic cell mutagenicity tests in vivo which could lead to classifying it as Mutagen Category 2, or*
- *there are indications that the internal dose for the substance and/or any of its metabolites will reach a steady state in the test animals only after an extended exposure, or*
- *there are indications of one or more relevant modes of action related to endocrine disruption from available in vivo studies or non-animal approaches."*

In the following lists, examples are provided for the criteria when the registrant must propose the extension of Cohort 1B to mate the Cohort 1B animals to produce a F2 generation:

<u>Guidance for uses leading to significant exposure:</u>

- If the substance is intended to be used[111] in the EU by consumers (i.e. members of the public) or professionals, either neat or in a chemical mixture and there is one very

[111] Registrant to provide data to support his registration.

wide use or several limited uses potentially affecting many consumers and/or professionals, then this is considered as meeting the criterion;

- If the substance is in an article used by consumers or professionals in the EU the criterion would be met if the substance is intended to be released from the article during use of the article by the consumers or professionals and there is one very wide use or several limited uses potentially affecting many consumers and/or professionals;
- Use of a substance in consumer articles exhibiting significant migration from the matrix and for which dermal absorption is relevant.

Guidance for substance specific toxicity conditions to be used together with criteria for uses leading to significant exposure:

(i) "*The substance displays genotoxic effects in somatic cell mutagenicity tests in vivo which could lead to classifying it as Mutagen Category* 2":

- Genotoxicity/mutagenicity observed *in vivo* potentially meeting the classification criteria to Mutagen Category 2:
  - Note: If the substance meets the criteria to Mutagen Category 1A/1B and the adequate risk management measures are in place then the reproductive toxicity studies need not to be conducted (according to adaptation possibilities in REACH Annex IX/X, point 8.7, Column 2);
  - *An in vivo* mutagenicity study should be available if one of the *in vitro* mutagenicity studies is positive (i.e. predicts mutagenicity). If one of the *in vitro* mutagenicity studies is positive, an *in vivo* mutagenicity study should be conducted before deciding on the study design of an extended one-generation reproductive toxicity study, if the other criteria for extending the Cohort 1B are not met.

(ii) "*There are indications that the internal dose for the substance and/or any of its metabolites will reach a steady state in the test animals only after an extended exposure*":

Extended time to reach the steady state may be indicated by available toxicokinetic information, physico-chemical properties and information from (eco)toxicological data. The effect of sex and life stages could be also considered[112]. Information can be obtained from:

- Assessment of toxicokinetic behaviour of the substance:
  - Generally, duration of longer than a week to reach the steady state may be considered as extended (in practise a steady state can be considered to be achieved after 4 to 6 half-lives)[113];
  - Attention needs to be also given to indications of very slow clearance (e.g. perfluorooctanoic acid (PFOA) which is a Category 1B reproductive toxicant).
- Physico-chemical properties of the substance:
  - An octanol-water partition coefficient (log Kow) value (e.g. above 4.5) indicates (bio)accumulative potential (determined experimentally or estimated by QSAR

[112] See e.g. Blagojević, J *et al.*, Age Differences in Bioaccumulation of Heavy Metals in Populations of the Black-Striped Field Mouse, *Apodemusagrarius* (Rodentia, Mammalia) *Int. J. Environ. Res., 6(4):1045-1052, Autumn 2012).*

[113] Steady state is achieved when the rate of elimination equals the rate of administration. Accumulation factor is 2 for a substance given once every half-live. Accumulation can be expected for a substance with slow elimination; e.g., with high octanol-water coefficient and no predicted hydrophilic metabolites. For lipophilic substances excretion may be impossible if there is no metabolism.

models) of the substance and/or its metabolites unless the substance is fully metabolised to hydrophilic metabolites.

- Indications on (bio)accumulation in animals or from human biomonitoring data:
  - High levels of substance/metabolites in human body fluids or tissues, such as blood, milk or fat are indicative of a concern on accumulation and persistence. Substances of purely endogenous origin and high levels only due to high exposure are excluded;
  - Bioaccumulation potency, for example if the substance properties meet the bioaccumulation screening criteria described in Table C.4-1 of the *Guidance on IR&CSA - Part C: PBT/vPvB assessment*. The assessment approach is described further in Section R.11.4.1.2 of the *Guidance on IR&CSA - Chapter R.11: PBT/vPvB assessment*;
  - If the substance fulfils the bioaccumulation criterion (B or vB) described in REACH Annex XIII;
  - Indications of biomagnifications (high levels of the substance in biota or terrestrial animals in the top of food chains, resulting from the effective accumulation of the substance in organisms and the slow elimination (not from high releases). This is further discussed under 'Field data and biomagnification', page 52, Section R.11.4.1.2 of the *Guidance on IR&CSA - Chapter R.11 PBT/vPvB assessment*.
- Indications from existing *in vivo* studies that after longer exposure duration the effects are more severe/occurring at lower dose than would be expected based on assessment factors generally used to extrapolate the dose descriptor between studies with different exposure duration:
  - e.g. if the NOAEL/LOAEL of a subchronic study (90-day) is more than 3 times lower than the NOAEL/LOAEL from a subacute study (28-day), taking the dose level selection and other differences into account;
  - Effects observed only at a later time point in chronic studies, thus indicating a need to have a longer exposure time to cause the toxicity likely, due to accumulation of a substance or its metabolites.

(iii) "*There are indications of one or more relevant modes of action related to endocrine disruption from available in vivo studies or non-animal approaches*".

Indications of endocrine disrupting mode(s) of action[114] such as (anti)oestrogenicity, (anti)androgenicity or influence on thyroid hormone activity or other modes of action related to endocrine disrupting properties relevant to reproductive toxicity. These modes of action have been associated with adverse effects on fertility, reproductive performance or development of offspring. See Appendix R.7.6–5 of this Guidance for evaluation of triggers:

- Endocrine disrupting modes of action may be indicated from *in vivo* studies by 1) changes in organ weight sensitive to endocrine disrupting activity (intact and/non-intact animals), 2) (increased) body weight, 3) measurements of hormone levels, or 4) effects on reproduction associated to endocrine (disrupting) modes of action;
- Repeated dose toxicity studies, especially the 28-day repeated dose toxicity study (EU B.7, OECD TG 407) updated in 2008, may provide indication of endocrine (disrupting) modes of action. Check the parameters related to endocrine modes of action; e.g.:

---

114 A comprehensive collection of screens and tests for endocrine disrupting chemicals are presented in OECD GD 150, covering the oestrogen receptor, androgen receptor and thyroid hormone mediated and steroidogenesis interference modalities. Both the test results for toxicity and ecotoxicity may be relevant.

- Changes in reproductive organs and other endocrine organs (e.g. ovaries, testes, uterus, cervix, epididymides, seminal vesicles, coagulating glands, prostate, vagina, pituitary, mammary gland, thyroid and adrenal gland);
- Changes in body weight (increase);
- Alterations in oestrus cycle;
- Changes in relevant hormone levels.

- Reproductive toxicity studies (e.g. a screening study) may provide indication of endocrine modes of action. Check the parameters related to endocrine modes of action; e.g.:
  - Changes in reproductive organs and other endocrine organs (see above);
  - Changes in anogenital distance, nipple retention, mammary gland histopathology or in any indicators of hormonal modes of action;
  - Changes is oestrus cycle;
  - Changes in gestation length;
  - Changes in body weight (increase);
  - Changes in pup body weight (increase not secondary to reduced litter size);
  - Other effects showing a likely endocrine disrupting mode of action.
- Endocrine effects from ecotoxicology studies and tests predicting endocrine (disrupting) modes of action (especially thyroid, see OECD GD 150);
- Non-animal approaches and specific animal studies may provide mechanistic data, information on receptor binding, epigenetics or other regulatory mechanism for endocrine (disrupting) modes of action, e.g.:
  - Uterotrophic assay (EU B.54, OECD TG 440) ;
  - Hershberger assay (EU B.55, OECD TG 441);
  - Performance-based test guideline for stably transfected transactivation *in vitro* assays to detect oestrogen receptor agonists (OECD TG 455);
  - H295R steroidogenesis assay (OECD 456);
  - BG1Luc Estrogen receptor transactivation test method for identifying oestrogen receptor agonists and antagonists;
  - Yeast Estrogen Screening (YES) and Yeast Androgen Screening (YAS) Tests;
  - Androgen receptor binding study;
  - Aromatase assay;
  - Endocrine organ cultures;
  - QSAR and computational predictions considered adequately reliable to serve as trigger(s).

The identified triggers should not be contradicted by other findings in the available data. The relevance and quality of triggers from the *in vivo* studies and non-animal approaches used must be adequately documented and justified. Case by case considerations are needed in evaluating trigger(s); evaluation is discussed in Appendix R.7.6–5 of this Guidance.

Further aspects to consider related to extension of the Cohort 1B and termination time for F2:

An extension of Cohort 1B to F2 is considered relevant in the context for classification and labelling and categorisation especially if the effect in P0 parental/F1 offspring is significant but not meeting classification criteria to Repr. 1B and more severe effects are seen in the F1 mating pairs/F2 offspring, thus affecting both P0 parental/F1 offspring and F1 mating pairs/F2 offspring but being more prominent or with a broader/different spectrum in F1 mating pairs/F2 offspring This could lead to a change in the classification from Repr. 2 to Repr. 1B.

Substances meeting the classification to Mutagen Category 2 are considered to have properties which increase the concern for reproductive toxicity and especially to the vitality and health of the second generation. The substance may have adverse effects on primordial germ cell development, proliferation and migration during *in utero* development, which may then be observed as reduced fertility in the F1 animals. Many genotoxic substances are also reproductive toxicants.

The test method for the extended one-generation reproductive toxicity study provides the possibility to terminate the F2 generation on postnatal day (PND) 4 based on a weight of evidence approach (integrated evaluation of the existing data). A weight of evidence adaptation approach according to REACH Annex XI, 1.2 could be usedfor example, if the results already meet the classification criteria to Repr 1B and it is highly likely that results from the rest of the lactation period (PND 5-21) would not lead to a lower NOAEL value. To cover the remaining uncertainty, an additional assessment factor may be applied.

The decision on whether or not to extend the Cohort 1B to F2 generation is/should be done before starting the study when the specified conditions are met. The testing proposal submitted by the registrant must include the study design proposed with justifications. During conduct of the experimental study the registrant is responsible for implementing the overall design of the study as requested, conduct of the study and interpretation of the results in order to meet the regulatory requirements and to insure the scientific integrity of the study in line with the test method.

So called internal triggers or in-study triggers for mating the Cohort 1B animals to produce the F2 generation (as those described in OECD TG 117) are not recommended to be used as such in REACH. However, the registrant may expand the study based on new information indicating a concern which needs to be addressed. The justification for the expansion must be documented.

3) Inclusion of Cohorts 2A and 2B

The main concepts of the triggers (conditions) for Cohort 2 (developmental neurotoxicity, DNT) are based on a particular concern for (developmental)[115] neurotoxicity[116]. A particular concern means that the concern should be specific to (developmental) neurotoxicity but also that the concern needs to reach a certain level of severity. Based on text inREACH Annex VIII, 8.6.1 for example, it can be understood that a particular concern may be indicated, such as by serious

---

[115] Both particular concerns for neurotoxicity as well as for developmental neurotoxicity may be addressed. See discussion in Section R.7.6.4.2.3 of this Guidance, under “General considerations related to investigation of (developmental) neurotoxicity and/or immunotoxicity” and “Proposals for developmental neurotoxicity or immunotoxicity studies”.

[116] (Nielsen *et al*., 2008) “Signs of neurotoxicity in standard acute or repeated dose toxicity tests may be secondary to other systemic toxicity or to discomfort from physical effects such as a distended or blocked gastrointestinal tract. Nervous system effects seen at dose levels near or above those causing lethality should not be considered, in isolation, to be evidence of neurotoxicity. In acute toxicity studies where high doses are administered, clinical signs are often observed which are suggestive of effects on the nervous system (e.g. observations of lethargy, postural or behavioural changes), and a distinction should be made between specific and non-specific signs of neurotoxicity.” “A consistent pattern of neurotoxic findings rather than a single or a few unrelated effects should be taken as persuasive evidence of neurotoxicity.”

or severe effects[117]. The examples provided in the legal text at REACH Annex IX/X, 8.7.3, Column 2 also provide guidance on the "severity level" of triggers for a particular concern with words such as "evidence of adverse effects" and findings "associated to adverse effects". There should be sufficient evidence, weighing all the information, to raise a reasonable expectation that the substance could be a developmental neurotoxicant (see Appendix R.7.6–5 of this Guidance for evaluation of triggers).

REACH specifies that an extended one-generation reproductive toxicity study including Cohorts 2A and 2B (developmental neurotoxicity cohorts) shall be proposed by the registrant or may be required by ECHA if a particular concern on (developmental) neurotoxicity.

Conditions for a particular concern for developmental neurotoxicity:

- existing information on the substance itself derived from relevant available *in vivo* or non-animal approaches, or
- specific mechanisms/modes of action of the substance with an association to (developmental) neurotoxicity, or
- existing information on effects caused by substances structurally analogous to the substance being studied, suggesting such effects or mechanisms/modes of action.

For the precise legal text see REACH regulation, Annexes IX and X, 8.7.3. The registrant must record the findings and justify the existence or nonexistence of the trigger(s) for the need to include the Cohorts 2A and 2B.

Examples of substance specific findings which may indicate a particular concern justifying inclusion of the developmental neurotoxicity cohort:

- abnormalities observed in the central nervous system or nerves
  - o changes in brain weight or in specific neural areas not secondary to body weight
  - o changes in brain volume or specific neural areas, obtained e.g. from morphometry/stereology measurements
  - o (histo)pathological findings in brain, spinal cord and/or nerves (e.g. sciatic nerve)
- any signs of behavioural or functional adverse effects on the nervous system in adult studies e.g. repeated-dose and acute toxicity studies and neurotoxicity studies, not likely to be secondary to general toxicity.
  - o clinical and/or behavioural signs (such as abnormal gait, narcosis, seizures or any other altered activity) if seen in absence of general toxicity
- specific mechanism/mode of action that has been closely linked to (developmental) neurotoxic effects (see Gupta RC (2011), pages 835-862),
  - o (adult) brain cholinesterase inhibition (by 20%);
  - o relevant changes in thyroid hormone levels or signs of thyroid toxicity indicating such changes,
  - o information on specific hormonal mechanisms/modes of action with clear association with the developing nervous system, such as oestrogenicity (Fryer *et al.*, 2012) and antiandrogenicity (Pallarés *et al*,. 2014)

[117] A serious or severe effect is an effect which has regulatory consequences, i.e. leads to a NOAEL values and/or contributes to hazard classification. Thus, a particular concern is an expectation that the substance has (developmental) neurotoxic properties contributing to the regulatory decision making. This also means that they are not secondary to other systemic toxicity.

- Information from (validated) non-animal approaches, such as from an *in vitro* developmental neurotoxicity test (see de Groot, 2013), predicting developmental neurotoxicity, e.g.:
  - Any sign of adverse neuronal differentiation *in vitro* e.g.:
    - Neurite outgrowth
    - Neural stem cell proliferation
    - Gene expression (mRNA and protein) biomarkers that are linked to neuronal differentiation, synaptogenesis and other neurodevelopmental differentiation
  - Functional endpoints, e.g. cell membrane potential, excitability, electrical activity
  - Specific modes of action that are linked to neurotoxic effects *in vivo* can be indicated *in vitro* by non-validated assays, e.g. cholinesterase inhibition, neuropathy target (neurotoxic) esterase inhibition.
- structurally analogue substances show (developmental) neurotoxic effects in *in vivo* or *in vitro* studies suggesting that similar effects or similar mechanisms/modes of action are likely to apply also for the registered substance (see the examples above for substance specific findings)
  - adequacy of an approach to use the trigger(s) from an analogous substance must be justified

The identified triggers should not be contradicted by other findings in the available data. The relevance and quality of triggers from the *in vivo* studies and non-animal approaches used must be adequately documented and justified. Evaluation of triggers is described in Appendix R.7.6–5 of this Guidance.

Further consideration related to adults vs developmental neurotoxicity is provided in Section R.7.6.4.2.3, of this Guidance under “General considerations related to investigation of (developmental) neurotoxicity and/or immunotoxicity”.

4) Inclusion of Cohort 3

The main concepts of the triggers (conditions) for Cohort 3 (developmental immunotoxicity, DIT) are based on a particular concern for (developmental) immunotoxicity[118]. A particular concern means that the concern should be specific to (developmental) immunotoxicity but also that the concern needs to reach a certain level of severity. Based on text in REACH Annex VIII, 8.6.1 for example, it can be understood that a particular concern is indicated, such as by serious or severe effects[119]. The examples provided in the legal text at REACH Annex IX/X, 8.7.3, Column 2 provides also guide on the “severity level” of triggers for a particular concern with wordings such as “evidence of adverse effects” and findings “associated to adverse effects”. There should be sufficient evidence, weighing all the information, to raise a

---

118 Both particular concerns for immunotoxicity as well as for developmental immunotoxicity may be addressed. See discussion in Section R.7.6.4.2.3 of this Guidance, under “General considerations related to investigation of (developmental) neurotoxicity and/or immunotoxicity” and “Proposals for developmental neurotoxicity or immunotoxicity studies”.

119 A serious or severe effect is an effect which has regulatory consequences, i.e. leads to a NOAEL values and/or contributes to hazard classification. Thus, a particular concern is an expectation that the substance has (developmental) immunotoxic properties contributing to the regulatory decision making. This also means that they are not secondary to other systemic toxicity.

reasonable expectation that the substance could be a developmental immunotoxicant (see Appendix R.7.6–5 of this Guidance for evaluation of triggers).

REACH specifies that an extended one-generation reproductive toxicity study including Cohort 3 (developmental immunotoxicity cohort) shall be proposed by the registrant or may be required by ECHA if a particular concern on (developmental) immunotoxicity.

Conditions for particular concern for developmental immunotoxicity:

- existing information on the substance itself derived from relevant available *in vivo* or non-animal approaches, or
- specific mechanisms/modes of action of the substance with an association to (developmental) immunotoxicity, or
- existing information on effects caused by substances structurally analogous to the substance being studied, suggesting such effects or mechanisms/modes of action.

For the precise legal text see REACH regulation, Annexes IX and X, 8.7.3. The registrant must record the findings and justify the existence or nonexistence of the trigger(s) for the need to include the Cohort 3.

Examples of substance specific findings which may indicate a particular concern justifying inclusion of the developmental immunotoxicity cohort:

- Combination of at least two (statistically significant and) biologically meaningful changes in haematology/clinical chemistry and/or organ weight associated with immunotoxicity, e.g. reduced leucocyte count in combination with reduced spleen weight.
- One severe (see footnote 43) statistically and/or biologically significant organ weight or histopathological finding related to an immunology organ, e.g. thymus atrophy.
- (respiratory) sensitisation (as a supportive factor only)
- Information on changes in immune function involving innate (e.g. NK-cell function, phagocytosis and oxidative burst) or acquired immunity (e.g. generation of immunological memory, cytotoxic T-cells and antibody production)
- Information on hormonal mechanisms/modes of action with clear association with the immune system, such as oestrogenicity (Adori *et al.*, 2010) and androgenicity (Trigunaite *et al.*, 2015) .
- Structural similarity with a substance causing structural or functional immunotoxicity or suggesting a similar mechanism/mode of action (see the examples above for substance specific findings)
    - adequacy of an approach to use the trigger(s) from an analogous substance must be justified

WHO Guidance document for immunotoxicity provides further examples of potential triggers for immunotoxicity testing (WHO, 2012). All effects on any immune-parameters found either *in vivo* (adult animals), or predicted *in vitro* or *in silico* may have impact on the developing immune system. These effects could be defined as quantitative or qualitative changes in cell counts or histopathology studying immune-specific organs or cell-populations in peripheral blood but may also include functional end-points such as antibody-production, delayed-type hypersensitivity test (to investigate cytotoxic T-cell activity), cytokine production, lymphocyte proliferation, NK-cell-function, phagocytosis, and oxidative burst.

The identified triggers should not be contradicted by other findings in the available data. The relevance and quality of triggers from the *in vivo* studies and non-animal approaches used must be adequately documented and justified. Evaluation of triggers is described in Appendix R.7.6–5 of this Guidance.

Appendix R.7.6–3 Premating exposure duration in the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443)

1. Importance of the premating exposure duration

The two main aspects in a reproductive toxicity study influencing how well fertility parameters and thus, the potential adverse effects on fertility can be evaluated are the length of the premating exposure duration and dose level setting.

The fertility part of the reproductive toxicity study should be capable of providing information on fertility that is adequate for both risk assessment and classification, including categorisation. For the classification purpose, it is important to produce and evaluate the full spectrum of effects on fertility. Just to detect a most sensitive effect may not be enough for deciding on classification categorisation because full information on magnitudes, incidences, severity and types of all effects (MIST information) should be evaluated together to assist the decision.

If the registrant applies ten weeks premating exposure duration in an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) no justification for premating exposure duration is needed. Substance specific justifications should be provided substantiated with data if shorter than ten weeks premating exposure duration is proposed.

Further insight to female and male reproduction toxicity can be obtained in relevant chapters of books for reproductive toxicology (such as Korach (1998) Reproductive and developmental toxicology; Gupta (2011) Reproductive and developmental toxicology)

1.1 Main parameters for evaluating effects on fertility

Mating/fertility

Mating and fertility are functional parameters which include effects on mating behaviour and fertility outcome. Parameters such as precoital interval, mating index, fertility index, preimplantation loss, post-implantation loss, number of corpora luteae, number of implantations, number of resorptions, dead foetuses, abortions, gestation length, litters size, and number of live pups are measuring effects on fertility (some of these parameters may also reflect developmental toxicity).

The length of the premating exposure may influence the mating and fertility parameters if the substance 1) causes adverse effects on primordial germ cell development, their migration and/or proliferation in embryo/foetus, 2) causes adverse effects on sperm development and maturation, 3) causes adverse effects on follicle development and/or development of ovum, 4) causes adverse effects on brain sexual development, 5) causes effects on hypothalamus-pituitary-gonad axis or other effects on hormonal control mechanisms.

The primordial germ cells already develop, migrate and proliferate during embryonic development. In addition to histopathological analysis of gonads, organ weight measurements and sperm parameter analysis, adverse effects on germ cell development/migration/proliferation during these early stages, as well as the other effects listed above, can be fully evaluated only by exposing the animals already *in utero* and then until adulthood and mating them. This full evaluation is possible if the mating and littering of the Cohort 1B animals is triggered in an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443).

An effect on fertility may be due to exposure *in utero,* postnatal period or during adulthood. In some cases it may be possible to conclude that effects on fertility are of developmental origin. For instance if there is information on fertility in both the parental animals and their offspring and effects on fertility are only seen in the mature offspring.

Sperm parameter analysis

Sperm parameter analysis includes for example, total cauda epididymal sperm number, percent progressively motile sperm, percent morphologically normal sperm and potentially percent of sperm with each identified abnormality for animals. The sperm count is measured by counting the number of sperm in cauda epididymis (sometimes also from testis as homogenisation resistant spermatid counts).

In the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) these parameters are to be reported for both the P and F1 males at termination. Other studies required in REACH as standard information requirements do not normally report results from sperm parameter analysis.

Sperm parameter analysis informs on the number of cauda epididymal sperm and their quality in terms of motility and morphological normality. The results from a sperm parameter analysis reflects the effects during the spermatogenic cycle in testes and during the epididymal maturation, if the exposure is long enough to cover both of these periods. The ability of sperm to fertilise eggs and produce alive and healthy offspring is examined in the reproductive toxicity studies by mating the animals and letting them litter. If the measurement of sperm parameters coincides close to mating, it assists and supports the evaluation of effects on fertility with the same exposure history through the same life stages. Sperm parameters may provide important information because in humans even a slight reduction in sperm quality/count may be critical for fertility.

Oestrous cycle

Oestrous cycle measurements reflect the normality of the hormonal level changes affecting the responsiveness of females. Direct measurements on function of hypothalamus-pituitary-gonad axis are not generally done in reproductive toxicity studies. In an extended one-generation reproductive toxicity study it is important to measure the oestrous cycle before mating and also after sexual maturity.

Organs weights of gonads and accessory sex organs

Organ weights of gonads and accessory sex organs, together with other parameters, can predict effects on fertility. These measurements can only be done at termination. Thus, this information can be obtained from the P males soon after mating but from the P females only after weaning. However, the measurements should be done as close as possible with the other information because information from various sources after the same exposure history allows combined and meaningful evaluation of effects on fertility based on all the data.

Histopathology of gonads and accessory sex organs

Histopathology of testes, ovaries and accessory sex organs can only be done at termination. Histopathological evaluation of testes allows assessment of the structural normality of testes including Leydig cells, Sertoli cells and seminiferous tubules with various developmental stages of sperm (e.g. Russell *et al.*, 1990). The information is generally qualitative, and quantitative measurements are not made and not required in test methods. Thus, it may not be possible to judge the amount of various cell types including the amount of various developmental stages of sperm. There may be a reduction of sperm at one developmental stage but it may be difficult to evaluate. Histopathological evaluation should reveal if multinuclear cells are present or another effect on sperm development if a significant reduction in the amount of certain cell types or their developmental stages is present. The information obtained is related to the morphological normality of testes but does not inform on the functional fertility and ability of the sperm to fertilise the eggs.

Histopathological evaluation of ovaries is complicated. The structure of an ovary is not organised and follicles at various developmental stages are distributed throughout the organ without a clear system. Thus, to count the number of follicles at different developmental stages requires several slices for histopathological examination. Quantitative evaluation of various cell types (e.g. granulosa cells and theca cells), indicative of toxicity is not generally done or required in the test methods. The number of primordial follicles (which can be combined with small growing follicles) is counted in the extended one-generation reproductive toxicity study in F1 animals which reflects the number of potential ova for future ovulations. The number of primordial and small growing follicles does not inform on actual functional fertility of the females but if the follicle number is reduced, it is a clear indication of gonad toxicity and should be taken into account in assessing effects on fertility.

Histopathology of accessory sex organs provides valuable information on how these organs have been developed and their morphological normality. The information should be evaluated together with other fertility findings.

It is important to be able to analyse the histopathological findings after the same exposure length and history as the other effects, including mating, to be able to understand a full picture of the spectrum of effects. Information on morphology is one important parameter in evaluation but as a stand alone measurement, to focus on morphology is too limited in order to provide a comprehensive picture of all the relevant aspects of fertility. However, it may be sufficient for classification (e.g. findings in histopathology alone from a repeated dose toxicity study may meet the classification criteria to category 1B for reproductive toxicity).

1.2 Ten weeks premating exposure duration

The full spermatogenesis, without sperm maturation, and folliculogenesis take 48-53 and 62 days in rats, respectively (e.g. Kerr *et al*., 2006; McGee and Hsueh, 2000). In addition to spermatogenesis, sperm maturation in rats takes around two weeks in epididymides. When the exposure is long enough, it covers both the sperm and follicle development through all the stages. Ten weeks premating exposure duration covers the full spermatogenesis and maturation meaning that the full cycle of development of sperm from spermatogonia into mature sperm is exposed. Thus, ten weeks premating exposure duration allows an assessment of the adverse effects on fertility by combining the information from all possible parameters in males evaluated at the same time. Similarly, the folliculogenesis, which lasts around 62 days, is fully covered only after a long exposure period, such as ten weeks. It is important to expose all the developmental stages of the sperm and follicles before the mating in order to be able to evaluate any potential adverse effect on fertility. Earlier stages of the spermatogenesis have been reported to be generally more sensitive than later stages to chemical and radiation exposure (Sjöblom *et al*., 1995) which also support that the exposure should cover all the stages before the mating.

For a comprehensive assessment of effects on fertility, which is often needed when deciding on classification for fertility effects, evaluation of the full spectrum of effects on fertility is necessary. Information from a limited number of parameters does not allow a conclusion on the absence of effects on fertility. The best outcome can be obtained when mating is allowed after an exposure covering one full spermatogenic cycle (including sperm maturation) and folliculogenesis, and an analysis of sperm parameters, organ weights and histopathology of gonads and accessory sex organs are conducted around the same time after the same exposure history.

In the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443), ten weeks premating exposure duration together with sperm parameter analysis, organ weights and histopathology of testis and accessory sex organs with the same exposure history is achievable for males. For females the organ weight measurements and histopathological analysis of gonads and accessory sex organs can only be made later and not near to mating. However, it is considered that the most important aspect is that the exposure duration for the female gonads covers the folliculogenesis before mating.

Organ weights (e.g. Bailey *et al*., 2004; Sellers *et al*., 2007; Hood *et al*., 2011)) and/or histopathology (e.g. Jacobson-Kram and Keller, 2006) of gonads may be among the most sensitive parameters for male fertility. For instance, testicular weight is quite a stable parameter because generally it is not influenced by small or moderate changes in body weight. Several studies have not established a correlation between testes-to-body weight and testes-to-brain weight (Bailey *et al*., 2004). Therefore, it could be concluded that variations on testicular weight will be linked to direct effects within the testes.

The most sensitive parameter showing an adverse effect is used to derive the NOAEL. However, the findings from the most sensitive parameters may not be sufficient for deciding on classification, including categorisation because the value of the NOAEL is not predictive for classification and (other) effects may be more relevant for classification purposes than the effect leading to a NOAEL. It is the clarity and the spectrum of the effects observed which counts for the classification and labelling. Thus, to address the fertility also for the classification and labelling purposes, including the categorisation, it is necessary to consider how well all the available parameters address the fertility endpoint. Information on magnitude, incidence, severity and type of all effects (MIST) influence on the classification, including categorisation. Evaluation of various parameters after the exposure length covering the critical reproductive aspects and after the same exposure history improves the quality of the assessment.

Environmental factors, such as chemical substances, pesticides, high temperatures and radiation have been associated with a reduction of sperm DNA integrity in infertile men (Evgeni *et al.*, 2014). It is to be noted that some effects on sperm, such as DNA fragmentation, may affect fertility and cannot be examined by routine gonadal histopathology (morphology) or sperm analysis. Several studies have attempted to investigate the possible correlation between human sperm DNA fragmentation and conventional sperm parameters. Most of them found an inverse correlation between DNA fragmentation rate and sperm quality (Evgeni *et al.*, 2014). In contrast, several authors have failed in finding a correlation between DNA fragmentation and standard sperm parameters, such as sperm concentration, motility and morphology (Evgeni *et al.*, 2014). The extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) does not contemplate sperm DNA damage assessment and consequently would not identify those cases where a reduction of sperm DNA integrity is not manifested in routine histopathology or in sperm parameter analysis.

The blood-testis barrier prevents free exchange of large proteins and some xenobiotics between the blood and the fluid within the seminiferous tubules (see Gupta, 2011, page 14). This may prolong time needed before the substance reaches the developing sperm supporting a long premating exposure duration. Thus, the blood level measurements of a substance may not reflect the exposure levels within the seminiferous tubules at a given time. All the different cell types representing various developmental stages of spermatogenesis are available in the testis at the same time and may allow detecting an adverse effect for a specific developmental stage or stages. However, a potential cumulative effect requiring exposure through several sequential stages cannot be detected with limited exposure duration.

In summary, the ten weeks premating exposure duration is one of the elements together with the appropriate dose level selection which allow production of data for an informed decision making for classification and labelling, including categorisation, for the hazard endpoint for sexual function and fertility according to CLP Regulation and for risk assessment.

1.3 Shorter than ten weeks premating exposure duration

Shorter than ten weeks premating exposure duration may be used based on substance specific justifications, but not shorter than 2 weeks. It is important to consider and document the reasoning why it is assumed that a longer premating exposure duration will not induce more or more severe effects.

A two weeks premating exposure duration is equivalent to the time for epidymal transit of maturing spermatozoa and thus, allows only the detection of post-testicular effects on sperm at mating (during the final stages of spermiation and epididymal sperm maturation). With a two-week premating exposure, the effects on functional fertility of exposure to the early stages of developing spermatozoa will not be covered as described above under heading 1.2.

The two weeks premating exposure duration is considered adequate to detect most of the male reproductive toxicants according to OECD GD 151. For females, two weeks premating exposure duration covers 2-3 oestrous cycles and effects on cyclicity may be detected. The detection of an effect may be adequate for NOAEL derivation but for classification and labelling purposes, including categorisation, information on magnitude, incidence, severity and all type of effects, i.e. full spectrum of effects is important (see text under heading 1.2).

Exposure during the full spermatogenic period and ovarian folliculogenesis are not covered at the time of mating, therefore, if only two weeks premating exposure duration has been selected, effects at earlier stages of spermatogenesis and folliculogenesis cannot be reflected in the functional fertility examination. This is a disadvantage and limited information may not allow adequate evaluation, including categorisation for classification, of potential adverse effects on fertility. It is to be noted that for the screening study (OECD TGs 421 or 422) the histopathological data will be limited also due to the limited duration of the whole study and limited statistical power as compared to the more comprehensive reproductive toxicity study such as the extended one-generation reproductive toxicity study.

A two-week premating period may be too short to produce results appropriate to conclude whether the substance meets the criteria for a category 1B reproductive toxicant, and thus may not be sufficient for classification and labelling purposes. Under point 2 below some considerations are presented on when a shorter than ten weeks premating exposure duration could be applied. In these cases substance specific justifications must be provided.

2. Considerations to be made in deciding if shorter than a ten weeks premating exposure duration could be adequate

2.1 Starting Point

To adequately assess the fertility endpoint, the best place to start considering the length of the premating exposure period should be ten weeks. Ten weeks cover the full spermatogenesis, sperm maturation and folliculogenesis before the mating allowing a meaningful assessment with the full spectrum of the effects after the same exposure history.

Based on substance specific justifications a shorter premating exposure duration may be proposed, but it should not be shorter than two weeks and sufficiently long to reach a steady-state (in reproductive organs) if such kinetic information is available.

2.2 Examples of cases where the existing information may support shorter than ten weeks premating exposure duration

If the registrant prefers another length of premating exposure duration than ten weeks, an acceptable substance-specific scientific justification substantiated with adequate data should be provided.

Such reasoning could be that *effects on fertility are already adequately addressed* and the extended one-generation reproductive toxicity study is used to address developmental toxicity. (It is, however, to be noted that the extended one-generation reproductive toxicity study does not provide equivalent information to the prenatal developmental toxicity study and thus cannot replace a prenatal development toxicity study)

There may be existing information from a good quality one-generation reproductive toxicity study (EU B.34, OECD TG 415) or similar addressing the fertility parameters. If information on a good quality one-generation reproductive toxicity study is available, then the fertility parameters are normally covered with adequate statistical power and the premating exposure duration may be shorter in a planned extended one-generation reproductive toxicity study.

An extended one-generation reproductive toxicity study is normally still needed to address the standard information requirement in REACH Annex IX/X, 8.7.3 because a one-generation reproductive toxicity study (EU B.34, OECD TG 415) does not cover the extended exposure period of F1 animals and the same parameters (e.g. sexual maturity and hormonal activity). In addition, the column 2 provisions of REACH Annex IX/X, 8.7.3 are not covered by one-generation reproductive toxicity study if triggered (for further details see Appendix R.7.6–2 of this Guidance for the extended one-generation reproductive toxicity study and Section R.7.6.4.2.3 of this Guidance, under “Proposals for developmental neurotoxicity or immunotoxicity studies” for separate developmental neurotoxicity and separate immunotoxicity studies).

There may be existing information from a good quality two-generation reproductive toxicity study (EU B.35, OECD TG 416) addressing the fertility parameters. If information on a good quality two-generation reproductive toxicity study is available, then the standard information requirement in REACH Annex IX/X, 8.7.3 is covered and an extended one-generation reproductive toxicity study may not be needed. However, the registrant must fulfil the column 2 provisions regarding developmental neurotoxicity and/or developmental immunotoxicity if the triggers are met. In these cases the registrant may consider fulfilling the adaptation requirements by proposing separate developmental neurotoxicity and/or developmental immunotoxicity study rather than an extended one-generation reproductive toxicity study. Similarly, if there are concerns related to the endocrine disrupting modes of action/properties not assessed in an existing two-generation reproductive toxicity study but which would have been measured in an extended one-generation reproductive toxicity study, the registrant may consider addressing these concerns in separate studies or add relevant parameters to other studies to be conducted (for further details see Appendix R.7.6–2 of this Guidance for the extended one-generation reproductive toxicity study and Section R.7.6.4.2.3, of this Guidance under "Proposals for developmental neurotoxicity or immunotoxicity studies" for separate developmental neurotoxicity and separate immunotoxicity studies).

There may also be cases where the fertility effects based on the existing information do meet the criteria for Reproductive toxicity Category 1B, but the column 2 adaptation (REACH Annex IX/X, 8.7) is not applicable due to further concerns on developmental toxicity. This information on fertility effects may stem for example, from good quality repeated dose toxicity studies (sex organ weights, histopathology of gonads and/or accessory sex organs, sperm parameters analysis), screening studies (OECD TGs 421 or 422; e.g. reduced fertility, litter size) or equivalent. In these cases, as the fertility is already addressed, shorter premating exposure duration in an extended one-generation reproductive toxicity study, if conducted to address the developmental toxicity, may be considered. If the effects from these studies only meet the classification criteria for Reproductive toxicity Category 2 for fertility, those should not be used as an argumentation to reduce the premating exposure length as the findings should be confirmed in a more comprehensive reproductive toxicity study (EU B.56, OECD TG 443).

There may be good quality information from existing repeated dose toxicity 90-day studies showing no effects in organ weights or histopathology of reproductive organs, and covering also the spermatogenesis and folliculogenesis. However, this information alone, or with the results from a screening study (OECD TGs 421 or 422) may not provide adequate confidence to shorten the premating exposure duration from ten weeks. This is because the information on mating and fertility from a screening study as well as the data from the repeated dose toxicity study is limited. Mating and fertility data from screening studies (OECD TGs 421 or 422) is after two weeks premating exposure duration not covering the full spermatogenesis and folliculogenesis and may also not be adequately long enough for detecting toxicity in hypothalamus-pituitary-gonad axis. In addition, the statistical power is low in these studies as they are not meant to provide comprehensive information on reproductive toxicity. Repeated dose toxicity 90-day studies may provide information on organ weights and histopathology but no mating data. The statistical power in the 90-day study is lower than that in the extended one-generation reproductive toxicity study also considering the data for histopathology. In addition, the exposure duration and exposure history are different in screening studies (OECD TGs 421 or 422) and 90-day studies. Thus, it may be difficult to conclude based on this information that a two weeks premating exposure duration is sufficient for a substance in question. However, the registrant may have additional information that may provide elements which together may support the justification, such as very low general toxicity (no effects up to the limit does of 1000 mg/kg bw/day in any of the existing studies), fast elimination, no distribution to sex organs, accessory sex organs and brain, and no concern on germ cell toxicity/mutagenicity (no effect in germ cell mutagenicity test). The substance specific justifications should be substantiated with adequate data.

Results showing no effects or some effects in reproductive organ weights and histopathology from a 28-day repeated dose toxicity study generally do not provide conclusive information to justify a shorter than ten weeks premating exposure duration. First of all, the length of the study is only 28-days and not covering the full spermatogenesis and folliculogenesis and the statistical power is low due to low number of animals.

Finally, if animals of Cohort 1B in an extended one-generation reproductive toxicity study are mated to produce the F2 generation, then the premating exposure duration will be ten weeks for these Cohort 1B animals and the fertility parameters will be covered allowing an evaluation of the full spectrum of effects on fertility. In these cases, shorter premating exposure duration for parental (P) animals may be considered. The consideration should take into account whether the findings from P animals (such as clinical signs, clinical chemistry, haematology) after a longer premating exposure would provide important information for interpretation of the findings in F1 animals, for example, when considering the potential developmental origin of such findings. It is to be noted that the results of the hazard class classification may differ depending on the interpretations of the origin of the results (differences in classification for specific target organ toxicity and developmental toxicity).

3. Summary

To fully evaluate effects on fertility, effects on all critical aspects and development stages should be covered; this can be done only by exposing the animals already *in utero* and then until adulthood and mating them. This full evaluation is possible if the extension (mating) of the Cohort 1B animals is triggered in an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443). The premating exposure duration of ten weeks is also covered in mated Cohort 1B animals.

If the extension of the Cohort 1B animals is not triggered, a ten -week premating exposure duration should be the starting point. This allows for assessing the consequences of early effects on the sex organs (spermatogenesis and folliculogenesis) assessor sex organs, hypothalamus-pituitary-gonad axis, and for example, prolonged distribution or any accumulation to relevant organs and tissues.

The registrant may prefer another length of premating exposure duration and substance specific justifications are needed to support shortened premating exposure duration.

Appendix R.7.6–4 Procedure for testing approaches and adaptation; Stage 3 - Stages 3.1.1 – 3.1.8

General adaptation rules of REACH Annex XI and certain specific adaptation rules in Column 2 provide possibilities for omitting the testing. These rules, except for those already passed at Stage 1, are presented here and the possibilities to omit the testing according to Stages 3.1.1 – 3.1.8 should be explored before conducting (REACH Annex VIII level test) or proposing (REACH Annexes IX and X level tests) the test.

Stage 3.1.1 Adaptation based on existing information not carried out according to GLP or the test methods indicated in the test method regulation (REACH Annex XI, 1.1.2)

Although the REACH standard information requirements refer to a specific series of reproductive studies, it is recognised that there may be other studies already performed that could address some of the endpoints covered by these standard protocols, reducing the need for new animal testing (adaptation according to REACH, Annex XI 1.1.2). The available data should be evaluated to assess their suitability for use, taking account of the robustness of design, and quality as outlined in Chapter R4 (*Guidance on IR&CSA*, *Chapter R4: "Evaluation of available information"*). The data from these studies (one or several together) are considered to be equivalent to data generated by the REACH standard test methods if the conditions of REACH Annex XI, Section 1.1.2 are met. An illustrative summary of these conditions is given below:

1) adequate for classification and labelling and/or risk assessment;
2) adequate and reliable coverage of key parameters;
3) exposure duration comparable or longer, if exposure duration is a relevant parameter;
4) adequate and reliable documentation;

a. adequate and reliable reporting of study design including dose levels tested.

Examples of other studies include: old studies conducted in other than preferred species; an NTP[120] modified one-generation study; non-GLP studies; or non-guideline investigations such as the NTP continuous breeding study (Chapin and Sloane, 1997).Such studies may be available and should be evaluated for fulfilling the criteria in REACH Annex XI, Section 1.1.2, in order to conclude that the information provided is equivalent to that foreseen to be the information provided by the EU test method. In addition, a study conducted according to a new test method not yet internationally acceptable may be valid and provide equivalent information.

It is to be noted that existing information on the two-generation reproductive toxicity study (EU B.35, OECD TG 416) is considered to fulfil the standard information requirement for REACH Annex IX/X, 8.7.3 (EU B.56, OECD TG 443), because this was the previous standard information requirement before the revision of the REACH Annexes to require an extended one-generation reproductive toxicity study. For further details see Section R.7.6.4.2.4 of this Guidance on the two-generation reproductive toxicity study (EU B.35, OECD TG 416).

Tests carried out according to old methods are evaluated case by case taking into account the toxicological properties of the substance. If the old study has for example, shorter exposure duration than the current test method, the registrant should justify using substance-specific arguments why the study with shorter exposure duration does not cause concern; for an example see Section R.7.6.4.2.2 of this Guidance. Similarly, if not all the

120 National Toxicology Program of NIEHS

key parameters are measured, but there are adequate substances-specific justifications to show that the missing information is of no concern, the old study may be acceptable. If the conditions summarised above for REACH Annex XI, 1.1.2 are not met, the study or test could still be of use for example, under REACH Annex XI, 1.2 as one element for weight of evidence adaptation.

Stage 3.1.2 Adaptation based on existing historical human data (REACH Annex XI, 1.1.3)

Epidemiological studies, conducted in the general population or in occupational cohorts, may provide information on possible associations between exposure to a chemical and adverse effects on reproduction. Clinical data and case reports (e.g. biomonitoring after accidental substance release) may also be available.

The criteria for assessing the adequacy of historical human data are listed in REACH Annex XI, Section 1.1.3. In exceptional cases human data may meet the classification criteria to Reproductive toxicity Category 1A and provide adequate information for risk assessment.

Stage 3.1.3 Adaptation based on existing information in a weight of evidence approach (REACH Annex XI, 1.2)

There are two possibilities to use the weight of evidence adaptation:

1) sufficient evidence from several independent sources of information; or
2) sufficient evidence from the use of newly developed test methods

leading to the conclusion that a substance has or has not a particular hazardous property.

It is to be noted that the weight of evidence approach described in REACH Annex XI, Section 1.2 needs to be substance and case specific and address the relevant standard information requirements of REACH Annex VII to X. Furthermore, it is hazard-based and therefore it has to be shown whether a substance has or has not a particular hazardous property. Because the weight of evidence approach is hazard-based, it means that exposure conditions or risk considerations are not part of the approach. To address the particular hazardous property of a substance, the key aspects/parameters of the study of the (standard) information requirement for which a weight of evidence approach is proposed need to be addressed to a sufficient extent.

In any case, adequate and reliable documentation of the information needs to be provided.

Adequate reporting of a weight of evidence approach is explained in the ECHA Practical Guide 2 (http://echa.europa.eu/documents/10162/13655/pg_report_weight_of_evidence_en.pdf).

Elements of a weight of evidence adaptation approach according to this adaptation rule for reproductive toxicity could be available from experimental studies addressing reproductive toxicity endpoints, reproductive toxicity studies performed with structurally similar substances and non-animal approaches, such as suitable validated *in vitro* methods, valid qualitative and quantitative structure-activity relationship models ((Q)SARs) or adverse outcome pathways (AOPs) (for further information on non-animal approaches see Stages 3.1.4 and 3.1.5).

Stage 3.1.4 Adaptation based on non-animal approaches such as QSAR approaches and in vitro methods (REACH Annex XI, 1.3 and 1.4)

REACH Annex XI, Sections 1.3 "Qualitative or Quantitative structure-activity relationship (QSAR) and Section 1.4 "*in vitro* methods" are potential adaptation possibilities. However, the available methods are currently not sufficient to address the complex endpoints on reproductive toxicity to replace an animal test. QSAR and *in vitro* methods may be used to support grouping and read-across approaches and may have a role in weight-of-evidence approach. For further details see Section R.7.6.4.1.1 of this Guidance.

Stage 3.1.5 Adaptation based on grouping and read-across (REACH Annex XI, 1.5)

The grouping of substances and read-across offer a possibility for adaptation of the standard information requirements of the REACH Regulation. If the read-across approach is adequate, unnecessary testing can be avoided. A read-across approach can also support a conclusion for a REACH endpoint using a weight of evidence approach.

The application of the grouping concept means that REACH information requirements for physicochemical properties, human health effects and/or environmental effects may be predicted from tests conducted on reference substance(s) within the group, referred to as source substance(s), by interpolation (extrapolation is generally not recommended for grouping) to other substances in the group, referred to as target substance(s) and this is called read-across.

The read-across approach has to be considered on an endpoint-by-endpoint basis due to the different complexities (e.g. key parameters, biological targets) of each endpoint. This means that read across (and category approach) is endpoint specific.

The term analogue approach is used when read-across is employed within a group of a very limited number of substances.

Read-across must, in all cases, be justified scientifically and documented thoroughly. There may be several lines of evidence used to justify the read-across, with the aim of strengthening the case.

Guidance on read-across is provided in the *Guidance on IR&CSA, Chapter R.6 "QSAR and grouping of chemicals"*. Further guidance can be found following this link: http://echa.europa.eu/support/grouping-of-substances-and-read-across.

Stage 3.1.6 Testing is technically not possible (REACH Annex XI, Section 2)

Tests do not need to be performed if it is not technically possible to do so. It may be that it is not possible to administer the substance for a particular reason. For example, the substance may be flammable in air or degrades explosively. It may also not be possible to produce sufficiently high enough exposure levels due to technical reasons. Justification for not performing tests is required and must be documented.

Stage 3.1.7 Substance-tailored exposure-driven testing (REACH Annex XI, Section 3)

The information requirements for reproductive toxicity at REACH Annex VIII, IX, and X levels may be omitted *if relevant human exposure can be excluded*. This clause states that tests may be omitted based on exposure scenarios developed in the Chemical Safety Report. The criteria defines three alternative sets of conditions that can, when justified and demonstrated, lead to an adaptation of standard information requirements (REACH Annex XI, 3.2.(a), (b) or (c)).

The adaptation according to REACH Annex XI Section 3.2.(a) is usually not applicable for REACH Annex IX and X reproductive toxicity studies as a DNEL derived from a reproduction/developmental toxicity screening test must not be considered appropriate to omit prenatal developmental toxicity study or an extended one-generation reproductive toxicity study (see REACH Annex XI, 3.2(a)(ii)footnote).

At REACH Annex IX level, the triggered prenatal developmental toxicity study on a second species may not need to be conducted based on a case-by-case justification. Such a justification may include the observation that triggers for the study on a second species are only at very high exposure levels compared with the identified and documented human exposure and that there are substance specific justifications that the second species would not be more sensitive/relevant to humans than the first species used. In such cases the DNEL derived based on the results from the first species may suffice although there were triggers for the study on second species.

For substances following strictly controlled conditions as described in REACH Annex XI, 3.2(b) or for substances rigorously permanently incorporated in an article according to REACH Annex XI, 3.2(c), the use of substance-tailored exposure-driven waiving may be possible.

In all cases, adequate justification and documentation must be provided (see REACH Annex XI, 3.2).

Stage 3.1.8 Adaptation based on column 2 rules others than CMR classification

(a) REACH Annex VIII (applicable for any registration of 10 tonnes or more per year)

The screening test for reproductive/developmental toxicity does not need to be conducted if a prenatal developmental toxicity study (OECD TG 414), an extended one-generation reproductive toxicity study (B.56, OECD TG 443) or a two-generation reproductive toxicity study (B.35, OECD TG 416) is available.

The screening test for reproductive/developmental toxicity provides initial information on reproduction toxicity. An extended one-generation reproductive toxicity study or a two-generation reproductive toxicity study provides more comprehensive information on the same and further key parameters with a higher statistical power. Thus, it is clear that these studies can cover the key parameters of the screening study and are superior to the screening study. However, if the prenatal developmental toxicity study is available, it provides information on embryonic and foetal development and the ability of the dam to maintain pregnancy, but not on fertility (or postnatal development). Thus, even though a prenatal developmental toxicity study is available, it is strongly recommended that the conduct of the screening study should be considered to obtain preliminary information on the fertility endpoint[121] and peri/early postnatal development.

(b) REACH Annexes IX and X (applicable for any registration of 100 tonnes or more per year)

The reproductive toxicity studies (prenatal developmental toxicity study(ies) and an extended one-generation reproductive toxicity study) do not need to be conducted if the following criteria are met:

1. The substance is of low toxicological activity (no evidence of toxicity seen in any of the tests available) <u>and</u>
2. It can be proven from toxicokinetic data that no systemic absorption occurs via relevant routes of exposure (e.g. plasma/blood concentrations below detection limit using a sensitive method and absence of the substance and of metabolites of the substance in urine, bile or exhaled air) <u>and</u>

[121] This position is supported by a relevant Ombudsman Case: "Hence it is strongly recommended in accordance with the endpoint specific REACH Guidance on information requirements and chemical safety assessment R.7, more specifically, paragraph 7.6.6.3 for reproductive toxicity that you consider conducting a screening reproductive/development toxicity study (OECD TGs 421 or 422) in addition to the pre-natal developmental toxicity study."

3. There is no or no significant human exposure[122].

It is necessary that all three criteria are fulfilled. The starting assumption is that substances with low toxicological activity may be less likely to be reproductive toxicants. The likelihood of the lack of reproductive toxicity potential is further increased and strengthened by requiring information proving no systemic absorption. When the substance has in addition no significant human exposure, it is considered safe to waive the reproductive toxicity study at REACH Annex IX and REACH Annex X levels.

[122] "No significant human exposure" must be considered in relation to the toxicity and amount and quality of available information.

Appendix R.7.6–5 Evaluation of triggers

Most of the triggers lead to information needs beyond the standard information requirements. For reproductive toxicity, the only standard information requirement (Column 1 requirement) which is triggered by toxicity and not only by a tonnage level is an extended one-generation reproductive toxicity which is triggered as indicated in Column 1 of REACH Annex IX, 8.7.3.

In this Appendix various aspects of triggers are discussed.

What is a trigger?

Triggers are findings which challenge the existing toxicity database. This means that due to existing triggers it is not possible to conclude on the potential for adverse health effects for a substance, and to address the concern, further information may be needed or is needed, depending on the condition. Before the concern is addressed with adequate information, the concern should be covered by applying (adequate) risk management measures.

In this document a general term of trigger is used. It is used instead of all the various possible terms used in the REACH Regulation or other places, such as an alert, condition, indication, indication of concern, serious concern, a particular concern.

A trigger is any factor present in the existing toxicological database, whether based on theoretical substance specific scientific considerations or from experimental or observational data that raises concerns that a substance may cause toxicity but information is not comprehensive enough to allow a conclusion to be drawn. It helps identifying where testing may need to go beyond the applicable standard information requirements. Where a standard information requirement applies, testing is required, unless an adaptation can be justified, irrespective of triggers. Case by case considerations are needed in evaluating triggers.

What needs to be done if there are triggers?

The term triggers is used as a general term. It depends also if there is legal text specifying what are the following actions needed. For example, if the legal text states "*if the conditions are met, the registrant shall…*" it means that in the existence of a trigger (condition) registrant must act accordingly. On the other hand, if the legal text states that the registrant may propose a test based on an indication or concern, then the registrant may act.

In the REACH Annex text for the information requirements the following terms are used as triggers:

1) **Condition**: if the conditions are met, the registrant **must** act. Condition may be e.g., **an (adverse)** effect, **an indication**, or **other relevant** existing information; thus, it may be e.g.:

   a. an effect which has (had) a regulatory consequence (NOAEL, classification; e.g. Muta 2), or

   b. a non-adverse effect (e.g. change in hormone level, *in vitro* results), other information (e.g. toxicokinetics), or

   c. indications of an effect inadequate for toxicological evaluation, or

   d. indications of modes of action from *in vivo* studies or non-animal approaches

   e. a combination of two or several indications (e.g. for a mode of action)

   f. a result of weighing all the relevant data for an endpoint (e.g. genotoxicity data)

2) **A particular concern**: if there is a particular concern, the registrant **must** act. A particular concern may be e.g. **serious/severe** effects, **adverse** effects, focused on a **specific type** of effects, or **other relevant** existing information; thus, it may be e.g.:

   a. an effect which has (had) a regulatory consequence (NOAEL, classification; e.g. STOT 1 or 2), or

b. existing information from non-animal approaches
c. specific mechanisms/modes of action
d. existing information on effects from various different data sources (in some cases also from structurally analogous substances)
e. information from one source may be sufficient when severe or considered adverse
f. a combination of two or several indications (e.g. for a mode of action)
g. a result of weighing all the relevant data (e.g. (developmental) neurotoxicity)

An exception: At REACH Annex VIII, 8.7.1, Column 2, based on a serious concern the registrant may act

3) Indications: may be

a. A condition
b. Adverse effects
c. Non-adverse effects, e.g. hormonal change
d. Mechanism/modes of action
e. From animal studies
f. From non-animal approaches
g. Indications are not the same as a particular concern, but may still require an action from the registrant, depending on the context

Sources for triggers

Triggers may stem from various sources of information including non-animal approaches, mechanistic studies, structurally analogous substances and *in vivo* studies and information from humans.

Findings observed in non-intact animals should generally be used as triggers unless there is evidence that the findings would not be also relevant for intact animals and/or humans. Experiments with non-intact animals may include animals with removal of an endocrine organ, such as ovary (ovariectomy). Another possibility is hormonal manipulation, for example, causing decrease or increase of organ weight. These animal models may be very sensitive to detect a change in for example, hormonal response; however, it should be considered whether the same applies in intact animals.

Classification and triggers

Adverse effects meeting the classification criteria for Category 1A or 1B reproductive toxicant are not triggers for further studies because they trigger the self-classification or harmonised classification and may allow omitting further reproductive toxicity studies according to REACH Annex VIII-X, point 8.7, Column 2 adaptation rules. However, effects meeting classification criteria for Category 2 reproductive toxicant may be triggers because they can raise concern that classification criteria for a higher category may be met.

Adverse effects not meeting classification criteria may be triggers. Whether findings which are considered non-adverse may serve as triggers depends on the parameter(s) in question and this is discussed below. The relevance and quality of triggers from the *in vivo* studies and non-animal approaches used should be adequately documented and justified.

Standard information requirements and triggers for further studies

The full (standard) information requirement in REACH Annex X, i.e. the extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) (or a two-generation reproductive toxicity study initiated before 15 March 2015; EU B.35, OECD TG 416), and prenatal development toxicity studies (EU B.31, OECD TG 414) performed in two species, when adequately conducted, should normally provide reliable information for conclusion on reproductive toxicity properties. If no conclusion can be drawn from the (standard) information requirement at the respective REACH Annex level, the registrant should address the remaining concern by proposing further studies to clarify the uncertainty over the reproductive potential of the substance.

For certain studies (e.g. the extended one-generation reproductive toxicity study, the study design is to be defined based on the existence/non-existence of the conditions/triggers.

Quality and relevance of the triggers

The generic guidance on the evaluation of available information gathered in the context of REACH Annexes VI-XI is provided in the *Guidance on IR&CSA*, *Chapter R.4: "Evaluation of available information"*.

Chapter R.4 applies for all kind of information; human, animal and non-animal sources and it is applicable also for information for reproductive toxicity endpoint. Principles described in Chapter R.4 apply to some extent also to the evaluation of triggers, although it is to be noted that a trigger is an indication of concern which challenges the available data as indicated in the definition of a trigger above and does not necessarily allow for conclusion on the hazardous properties to reproductive health – conclusion on classification or NOAEL values.

Certain general important aspects to assist the evaluation of triggers are presented below.

Consistency

It is important that the identified triggers are not contradicted by other findings in the available data. Consideration should be given to the statistical power and overall quality of the available data. Sometimes when the data is scarce it may not be possible to evaluate the consistency more than by noting if other data is contradicting with the potential trigger(s) or not.

When evaluating the consistency, differences in the existing studies must be taken into account. Apparent inconsistencies may be due to species/strain differences, different route and/or dose levels, different exposure duration, differences in methodology in measuring parameters, etc. Thus, whether the inconsistencies are likely due to methodological differences or differences in statistical power and not real inconsistencies in results, those must be analysed prior to weighing the results and deciding on the existence/non-existence of triggers.

Statistical significance and biological relevance

Dose responsiveness would provide more confidence and be more indicative of a chemically mediated effect rather than just a statistically significant finding in one dose group. The statistical power of the results from screening studies (OECD TGs 421 or 422) or 28-day study is quite low and there it may be more important to look at the ranges rather than statistical significances. It should also be remembered that statistical significance is not the same as biological relevance. There may be for example, 20% change in a parameter with biological relevance but without statistical significance. On the other hand there may be a statistically significant finding without a biological relevance. If the statistical power is high and biological variation is low for a parameter, the biological relevance of a change is high. It is necessary to evaluate if the statistical power is adequate in respect to the biological variation of a parameter. Historical data may provide guide for normal ranges but the control group of the study should generally be the main source of information in deciding on normal values and variation.

It should be also considered, case-by-case, the possibility of a non-monotonic dose-response curve.

Deciding on biological relevance of information from non-animal approaches may be challenging. Generally these predictive methods provide indication(s) and triggers rather than conclusions on hazardous properties of substances. If the non-animal approach is not reliable or the results are observed at extreme conditions (e.g. over 100x higher concentrations than the biologically plausible maximal concentration), the validity and relevance of such a single test result should be confirmed before conclusion. In best conditions results from two or more non-animal approaches are available supporting each other.

Human relevance

In the absence of further knowledge and proof, it is assumed that biologically relevant findings in animals are also relevant to humans. To justify that findings/modes of action/mechanisms of action are not relevant to human, information on humans is needed. It is not enough to state that there are no indications of the same findings/modes of action/mechanisms of action in humans than in animals, if the issue has not been adequately investigated.

Relationship of triggers with systemic toxicity

Clear triggers occur at dose levels without (other) systemic toxicity. However, the triggers have to be considered case-by-case as the relationship with the systemic toxicity may not be always clear although they may occur at the same dose level as the triggers. Generally triggers should be considered relevant even if observed at the same dose level than the (other) systemic toxicity findings if it cannot be justified why the triggers are secondary to (other) systemic toxicity.

Quality of the studies and tests

The quality of the studies or the reliability of the information should be considered. For example, triggers from *in vivo* and *in vitro* tests should have been tested with the biologically relevant material, in a robust system, and the data should be determined to be of adequate quality. Many non-animal approaches, for example, *in vitro* tests are not validated yet, but the result from them may be used if considered to be reliable case by case. For example, no *in vitro* tests for neuronal differentiation are validated but as triggers for motivating evaluation of developmental neurotoxicity, results from scientifically evaluated (peer reviewed) publications and reports may be used as triggers when considered relevant. The same goes for *in vitro* tests for other triggers such as for developmental immunotoxicity and endocrine disrupting modes of action/mechanisms.

When evaluating the results from non-animal approaches the predictivity and applicability domain and potential other limitations of the approaches need to be considered. Triggers from non-animal approaches such as QSAR predictions may be challenging to interpret especially when various methods show diverging results. Generally, consistent results from more than one non-animal approach are needed to increase the confidence of the existence or nonexistence of a trigger.

Triggers from structurally analogous substances

Triggers may also stem from structurally analogous substances. In that case, the adequacy to use the information as triggers should be considered and justified.

Evaluation of data for identification of triggers:

As part of the Stage 3.2.1 data review the following questions should be asked:

- Are there triggers for further studies/investigations specified in Column 2?
- Are there triggers for reproductive toxicity not specified in Column 2? (Considering also structurally analogous substances)

- Is there any knowledge of the substance, chemical groups or categories that would indicate special features related to reproductive toxicity to be included in the study design? If so, which?
- Are there triggers for mechanisms/modes of action relevant for reproductive toxicity? (Considering also structurally analogue substances)
- If Column 2 specific adaptation rules and REACH Annex XI general adaptation rules apply and the data is adequate for assessing and concluding the classification and labelling and risk assessment, evaluation of triggers is not needed. This means e.g. that if a substance meets the classification criteria for Category 1 for any of the CMR properties as defined at Stage 1 in Section R.7.6.2.3.2 of this Guidance and fulfils the adaptation criteria described in Column 2, then evaluation of triggers for further reproductive toxicity studies is not needed.

From a scientific perspective, it is not possible to generate an exhaustive and rigid list of triggers that would automatically trigger a particular study or have clearly defined implications for classification and risk assessment. However, certain conditions are specified in REACH Annexes and, when met, require a particular study or study design to be proposed.

A trigger (or triggers) may trigger:

- a study, which would fulfil a standard information requirement, which otherwise only applies at a higher tonnage level,; or
- a certain study design (or a particular independent study) when specified conditions are met (e.g. extension of Cohort 1B to include F2 or inclusion of Cohort 2 and/or 3 in an extended one-generation reproductive toxicity study); or
- inclusion of certain selected additional investigational parameters to a range-finding study or a study required in the (standard) information requirement (e.g. selected parameters for immunotoxicity under conditions where the trigger(s) need(s) to be confirmed before considering the need for further studies to address the concern; or
- special investigational studies/tests, e.g. studies on mechanisms/modes of action.

The following triggers are referred to in REACH Annex IX 8.7.3 and trigger the information requirement:

- **At REACH Annex IX level**, an extended one-generation reproductive toxicity study may be triggered by triggers from repeated dose toxicity studies (including screening studies) according to description in Column 1 (see further details in this Guidance, Section R.7.6.2.3.2, Stage 4.4 (iii) of this Guidance.

The following triggers are referred to in Column 2 adaptation rules for reproductive toxicity/developmental neurotoxicity/developmental immunotoxicity:

- **At REACH Annex VIII level**, based on trigger(s) for reproductive toxicity, either for developmental toxicity or for fertility, causing serious concern[123], the registrant **may** propose a prenatal developmental toxicity study or an extended one-generation reproductive toxicity study instead of a “screening for reproduction/developmental toxicity” test, as appropriate. The appropriate study depends on whether the concern is on prenatal developmental toxicity, prenatal developmental toxicity manifested

[123] Serious concern reflects a high likelihood for adverse effects on reproductive health.

postnatally, postnatal developmental toxicity or on fertility[124]. The triggers may stem for example from relevant non-animal approaches[125] or *in vivo* studies e.g. from 28-day repeated dose toxicity study which is required at this REACH Annex level or respective other information. A testing proposal is required for REACH Annex IX/X level studies.

- **At REACH Annex IX level**, trigger(s) for prenatal developmental toxicity should trigger a prenatal developmental toxicity study on a second species as a Column 2 requirement. Examples of triggers for this study are shown under Section R.7.6.2.3.2, Stage 4.4 (ii), prenatal developmental toxicity study of this Guidance.
- At REACH Annex IX level, if an extended one-generation reproductive toxicity study is triggered, triggers for extending the Cohort 1B, including Cohorts 2 and/or 3 are given in Column 2. The study design of an extended one-generation reproductive toxicity study and triggers to expand the study are described in Appendix R.7.6–2 of this Guidance.
- At the same REACH Annex level, an extended one-generation reproductive toxicity study on a second species or strain may be triggered at this REACH Annex (REACH Annex IX) or the next REACH Annex level (REACH Annex X). Examples of triggers are presented under Section R.7.6.2.3.2, Stage 4.4 (iii), extended one-generation reproductive toxicity study of this Guidance.
- **At REACH Annex X level**, an extended one-generation reproductive toxicity study is a (standard) information requirement. The triggers for extending the Cohort 1B, including Cohorts 2 and/or 3 are given in Column 2. The study design of the extended one-generation reproductive toxicity study and triggers to expand the study are described in Appendix R.7.6–2 of this Guidance.
- At REACH Annex X level, the full information requirements i.e. an extended one-generation reproductive toxicity study (EU B.56, OECD TG 443) (or a two-generation reproductive toxicity study initiated before 15 March 2015; EU B.35, OECD TG 416), and prenatal development toxicity studies (EU B.31, OECD TG 414) performed in two species, when adequately conducted, should normally provide reliable information for conclusion on reproductive toxicity properties as indicated above.

If no conclusion can be drawn from the (standard) information requirement, the registrant should address the remaining concern by proposing further studies to clarify the uncertainty over the reproductive potential of the substance.

[124] However, in case of proposing a prenatal developmental toxicity study it is strongly recommended that the registrant should consider conducting a screening study because a prenatal developmental toxicity study does not address the effects on the fertility endpoint and developmental toxicity manifested shortly after birth.

[125] In order to be considered providing "*serious concern*", information from non-animal approaches should be reliable, relevant and from validated studies with appropriate applicability domain (for QSAR models a formal validation process is not required). Based on case-by-case scientific justification results from non-validated and non-guideline tests may be acceptable. Generally several information sources may be needed.

Exposure triggers/conditions upgrading testing requirements

- Guidance on exposure-based adaptation and triggering of information requirements is provided in Section R.5.1 in the *Guidance on IR&CSA*, *Chapter R.5: Adaptation of information requirements.*
- The use pattern and the exposure to a substance may indicate a concern with the need for additional information requirements, on a case-by-case basis. For example, there may be serious concerns that human exposure, particularly to consumers, is close to the levels at which human health effects might be expected. Such concerns for human health need to be addressed by producing additional information on hazard. In very exceptional cases such concerns may be satisfactorily addressed by improved risk management measures.

Documentation and addressing the triggers/conditions

If the triggers for reproductive toxicity or the conditions described in Column 1 or 2 are met for further investigations, they must be described in the dossier as well as how they are addressed at the respective endpoint section.

R.7.7 Mutagenicity and carcinogenicity

R.7.7.1 Mutagenicity

R.7.7.1.1 Definition of mutagenicity

Mutagenicity refers to the induction of permanent transmissible changes in the amount or structure of the genetic material of cells or organisms. These changes may involve a single gene or gene segment, a block of genes or chromosomes. The term clastogenicity is used for agents giving rise to structural chromosome aberrations. A clastogen can cause breaks in chromosomes that result in the loss or rearrangements of chromosome segments. Aneugenicity (aneuploidy induction) refers to the effects of agents that give rise to a change (gain or loss) in chromosome number in cells. An aneugen can cause loss or gain of chromosomes resulting in cells that have not an exact multiple of the haploid number. For example, three number 21 chromosomes or trisomy 21 (characteristic of Down syndrome) is a form of aneuploidy.

Genotoxicity is a broader term and refers to processes which alter the structure, information content or segregation of DNA and are not necessarily associated with mutagenicity. Thus, tests for genotoxicity include tests which provide an indication of induced damage to DNA (but not direct evidence of mutation) *via* effects such as DNA strandbreaks, unscheduled DNA synthesis (UDS), sister chromatid exchange (SCE), DNA adduct formation or mitotic recombination, as well as tests for mutagenicity.

The chemical and structural complexity of the chromosomal DNA and associated proteins of mammalian cells, and the multiplicity of ways in which changes to the genetic material can be effected make it difficult to give more precise, discrete definitions.

In the risk assessment of substances it is necessary to address the potential effect of *mutagenicity*. It can be expected that some of the available data will have been derived from tests conducted to investigate potentially harmful effects on genetic material (*genotoxicity*). Hence, both the terms *mutagenicity* and *genotoxicity* are used in this document.

R.7.7.1.2 Objective of the guidance on mutagenicity

The aims of testing for genotoxicity are to assess the potential of substances to induce genotoxic effects which may lead to cancer or cause heritable damage in humans. Genotoxicity data are used in risk characterisation and classification of substances. Genotoxicity data are useful for the determination of the general mode of action of a substance (*i.e.* type(s) of genotoxic damage induced) and can provide some indication on the dose (concentration)-response relationship and on whether the observed effect can be reasonably assumed to have a threshold or not. Genotoxicity data are thus useful in deciding the best approach to use for the risk assessement. Expert judgement is necessary at each stage of the testing strategy to decide on the relevance of a result based on the data available for each endpoint.

Alterations to the genetic material of cells may occur spontaneously endogenously or be induced as a result of exposure to ionising or ultraviolet radiation, or genotoxic substances. In principle, human exposure to substances that are mutagens may result in increased frequencies of mutations above background.

Mutations in somatic cells may be lethal or may be transferred to daughter cells with deleterious consequences for the affected organism (*e.g.* cancer may result when they occur in proto-oncogenes, tumour suppressor genes and/or DNA repair genes) ranging from trivial to detrimental or lethal.

Heritable damage to the offspring, and possibly to subsequent generations, of parents exposed to substances that are mutagens may follow if mutations are induced in parental germ cells. To date, all known germ cell mutagens are also mutagenic in somatic cells *in vivo.* Substances that are mutagenic in somatic cells may produce heritable effects if they, or their active metabolites, have the ability to interact with the genetic material of germ cells. Conversely, substances that do not induce mutations in somatic cells *in vivo* would not be expected to be germ cell mutagens.

There is considerable evidence of a positive correlation between the mutagenicity of substances *in vivo* and their carcinogenicity in long-term studies with animals. Genotoxic carcinogens are substances for which the most plausible mechanism of carcinogenic action involves genotoxicity.

R.7.7.2 Information requirements on mutagenicity

The information requirements on mutagenicity are described by REACH Annexes VI-XI, that specify the information that must be submitted for registration and evaluation purposes. The information is thus required for substances produced or imported in quantities of >1 t/y (tons per annum). When a higher tonnage level is reached, the requirements of the corresponding Annex have to be considered. However, factors including not only production volume but also pre-existing toxicity data, information about the identified use of the substance and exposure of humans to the substance will influence the precise information requirements. The REACH Annexes must thus be considered as a whole, and in conjunction with the overall requirements of registration, evaluation and the duty of care.

Column 1 of REACH Annexes VII-X informs on the standard information requirements for substances produced or imported in quantities of >1 t/y, >10 t/y, >100 t/y, and >1000 t/y, respectively.

Column 2 of REACH Annexes VII-X lists specific rules according to which the required standard information may be omitted, replaced by other information, provided at a different stage or adapted in another way. If the conditions are met under which column 2 of these Annexes allows adaptations, the fact and the reasons for each adaptation should be clearly indicated in the registration dossier.

The standard information requirements for mutagenicity and the specific rules for adaptation of these requirements are presented in Table R.7.7–1.

Table R.7.7–1 REACH information requirements for mutagenicity

COLUMN 1 STANDARD INFORMATION REQUIRED	COLUMN 2 SPECIFIC RULES FOR ADAPTATION FROM COLUMN 1
Annex VII: 1. *In vitro* gene mutation study in bacteria.	Further mutagenicity studies shall be considered in case of a positive result.
Annex VIII: 1. *In vitro* cytogenicity study in mammalian cells or *in vitro* micronucleus study.	1. The study does not usually need to be conducted • if adequate data from an *in vivo* cytogenicity test are available or • the substance is known to be carcinogenic category 1A or 1B or germ cell mutagenic category 1A, 1B or 2.
2. *In vitro* gene mutation study in mammalian cells, if a negative result in Annex VII, 1 and Annex VIII, 1.	2. The study does not usually need to be conducted if adequate data from a reliable *in vivo* mammalian gene mutation test are available. Appropriate *in vivo* mutagenicity studies shall be considered in case of a positive result in any of the genotoxicity studies in Annex VII or VIII.
Annex IX:	If there is a positive result in any of the *in vitro* genotoxicity studies in Annex VII or VIII and there are no results available from an *in vivo* study already, an appropriate *in vivo* somatic cell genotoxicity study shall be proposed by the registrant. If there is a positive result from an *in vivo* somatic cell study available, the potential for germ cell mutagenicity should be considered on the basis of all available data, including toxicokinetic evidence. If no clear conclusions about germ cell mutagenicity can be made, additional investigations shall be considered.
Annex X:	If there is a positive result in any of the *in vitro* genotoxicity studies in Annex VII or VIII, a second *in vivo* somatic cell test may be necessary, depending on the quality and relevance of all the available data. If there is a positive result from an *in vivo* somatic cell study available, the potential for germ cell mutagenicity should be considered on the basis of all available data, including toxicokinetic evidence. If no clear conclusions about germ cell mutagenicity can be made, additional investigations shall be considered.

In addition to these specific rules, the required standard information set may be adapted according to the general rules contained in Annex XI. In this case as well, the fact and the reasons for each adaptation should be clearly indicated in the registration.

In some cases, the rules set out in Annex VII to XI may require certain tests to be undertaken earlier than or in addition to the tonnage-triggered requirements. Registrants should note that a testing proposal must be submitted for a test mentioned in Annex IX or X, independently from the registered tonnage. Following examination of such a testing proposal ECHA has to approve the test in its evaluation decision before it can be undertaken. See Section R.7.7.6 of this Guidance for further guidance on testing requirements.

R.7.7.3 Information and its sources on mutagenicity

To be able to evaluate the mutagenic potential of a substance in a comprehensive way, information is required on its capability to induce gene mutations, structural chromosome aberrations (clastogenicity) and numerical chromosome aberrations (aneugenicity). Many test methods are available by which such information can be obtained. Non-testing methods, such as SAR, QSAR and read-across approaches, may also provide information on the mutagenic potential of a substance.

Typically, *in vitro* tests are performed with cultured bacterial cells, human or other mammalian cells. The sensitivity and specificity of tests will vary with different classes of substances and, if adequate data are available for the class of substance to be tested, these data can guide the selection of the most appropriate test systems to be used. In order to detect mutagenic effects also of substances that need to be metabolically activated to become mutagenic, an exogenous metabolic activation system is usually added in *in vitro* tests. For this purpose the post-mitochondrial 9000 x *g* supernatant (S-9 fraction) of whole liver tissue homogenate containing a high concentration of metabolising enzymes and extracted from animals that have been induced to raise the oxidative P450 levels is most commonly employed. In the case when information is required on the mutagenic potential of a substance *in vivo,* several test methods are available. In *in vivo* tests whole animals are used, in which metabolism and toxicokinetic mechanisms in general exist as natural components of the test animal. It should be noted that species-specific differences in metabolism are known. Therefore, different genotoxic responses may be obtained. Some *in vivo* genotoxicity tests such as the Transgenic rodent (TGR) somatic and germ cell gene mutation assays and the comet assay employ methods by which any tissue (containing nucleated cells) of an animal can in theory be examined for effects on the genetic material. This gives the possibility to examine target tissues (including germ cells) and site-of-contact tissues (*i.e.* skin, epithelium of the respiratory or gastro-intestinal tract). However differences can exist regarding the number and type of tissues for which the use a specific test has been scientifically validated. For instance, the TGR assays can be used to examine germ cells whereas the comet assay as described in the OECD test guideline (TG) is, at present, not recommended for that purpose.

Some test methods, but not all, have an officially adopted EU and/or OECD TG for the testing procedure. In cases where no adopted EU or OECD TG is available for a test method, rigorous and robust protocols should be followed, such as those defined by internationally recognised groups of experts like the International Workshop on Genotoxicity Testing (IWGT) under the umbrella of the International Association of Environmental Mutagen Societies. Furthermore, modifications to OECD TGs have been developed for some classes of substances and may serve to enhance the accuracy of test results. Use of such modified protocols is a matter of expert judgement and will vary as a function of the chemical and physical properties of the substance to be evaluated. Similarly, use of standard test methods for the testing of tissue(s) not covered by those standard test methods should be scientifically justified and validity of the results will depend on the appropriateness of the acceptability criteria, which should have been specifically developed for this (these) tissue(s) based on sufficient experience and historical data.

R.7.7.3.1 Non-human data on mutagenicity

Non-testing data on mutagenicity

Non-test information about the mutagenicity of a substance can be derived in a variety of ways, ranging from simple inspection of the chemical structure through various read-across techniques, the use of expert systems, metabolic simulators, to *global* or *local* (Q)SARs. The usefulness of such techniques varies with the amount and nature of information available, as well as with the specific regulatory questions under consideration.

Regarding substances for which testing data exist, non-test information can be used in the *Weight of Evidence* approach, to help confirm results obtained in specific tests, or to help develop a better understanding of mutagenicity mechanisms. The information may be useful in deciding if, or what, additional testing is required. At the other extreme, where no testing data are available, similar alternative sources of information may assist in setting test priorities. In cases where no testing is likely to be done (low exposure, <1 t/y) they may be the only options available to establish a hazard profile.

Weight of Evidence approaches that use expert judgement to include test results for close chemical analogues are ways of strengthening regulatory positions on the mutagenicity of a substance. Methods that identify general *structural alerts* for genotoxicity such as the Ashby-Tennant super-mutagen molecule (Ashby and Tennant, 1988) may also be useful.

Prediction models for mutagenicity

There are hundreds of (Q)SAR models available in the literature for predicting test results for genotoxic endpoints for closely related structures (Naven *et al.*, 2012; Bakhtyari *et al.*, 2013). These are known as *local* (Q)SARs. When essential features of the information domain are clearly represented, these models may constitute the best predictive tools for estimating a number of mutagenic/genotoxic endpoints. However, quality of reporting varies from model to model and predictivity must be assessed case-by-case on the basis of clear documentation. Use of harmonised templates, such as the QSAR Model Reporting Format (QMRF) and the QSAR Prediction Reporting Format (QPRF) developed by the Joint Research Centre (JRC) of the European Commission (http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/qsar_tools/QRF), can help ensure consistency in summarising and reporting key information on (Q)SAR models and substance-specific predictions generated by (Q)SAR models. The JRC website also hosts the JRC (Q)SAR Model Inventory, which is an inventory of information on the validity of (Q)SAR models that have been submitted to the JRC (http://ihcp.jrc.ec.europa.eu/our_databases/jrc-qsar-inventory).

Generally, (Q)SAR models that contain putative mechanistic descriptors are preferred; however many models use purely structural descriptors. While such models may be highly predictive, they rely on statistical methods and the toxicological significance of the descriptors may be obscure.

(Q)SAR models for mutagenicity can apply to a limited set of congeneric substances (local models) or to a wide variety of non-congeneric substances (global models). Global (Q)SARs are usually implemented in computer programs and may comprise a set of local models; these global models first categorise the input molecule into the chemical domain it belongs to, and then apply the corresponding local prediction model. These are known as expert systems. Other global models apply the same mathematical algorithm on all input molecules without prior separation. It is generally observed that the concept of applicability domain is a useful one and the endpoints for substances inside the applicability domains of the models are better predicted than for substances falling outside.

Many global models for mutagenicity are commercial and some of the suppliers of these global models consider the data in their modelling sets to be proprietary. Proprietary means that the training set data used to develop the (Q)SAR model is hidden from the user. In other cases it means that it may not be distributed beyond use by regulatory authorities. The models do not always equal the software incorporating them, and the software often has flexible options for expert uses. Thus, the level of information available, from both (Q)SAR models and compiled databases, should be adequate for the intended purpose.

A list of the available (free and commercial) predictive software for ecotoxicological, toxicological and environmental endpoints, including mutagenicity models, has been compiled within the frame of the EU project Antares (http://www.antares-life.eu/).

The most common genotoxicity endpoint for global models has been to predict results of the Ames test. Some models for this endpoint include a metabolic simulator.

There are models for many other mutagenicity endpoints. For example, the Danish EPA and the Danish QSAR group at DTU Food (National Food Institute at the Technical University of Denmark) have developed a (Q)SAR database that contains predictions from a number of mutagenicity models. In addition to assorted Ames models, the database contains predictions of the following *in vitro* endpoints: chromosomal aberrations (CHO and CHL cells), mouse lymphoma/*tk*, CHO/hprt gene-mutation assays and UDS (rat hepatocytes); and the following *in vivo* endpoints: *Drosophila* SLRL, mouse micronucleus, rodent dominant lethal, mouse SCE in bone marrow and mouse comet assay data. The database is freely accessible *via* http://qsar.food.dtu.dk. The online database contains predictions for over 166,000 substances and includes a flexible system for chemical structure and parameter searching. A user manual with information on the individual models including training set information and validation results is available at the website. The database is also integrated into the OECD (Q)SAR Toolbox. A major update of the database with consensus predictions by use of different QSAR models for each of the modelled endpoints for more than 600,000 structures, including over 70,000 REACH pre-registered substances, and with an improved user interface is scheduled for the beginning of 2015.

Another example of a database with predictions on mutagenicity is the Enhanced NCI Database Browser (http://cactus.nci.nih.gov) sponsored by the U.S. National Cancer Institute. It contains predictions for over 250,000 substances for mutagenicity as well as other non-mutagenic endpoints, some of which may provide valuable mechanistic information (for example alkylating ability or microtubule formation inhibition). It is also searchable by a wide range of parameters and structure combinations.

Neither of these two examples is perfect, but they illustrate a trend towards predictions of multiple endpoints and may assist those making *Weight of Evidence* decisions regarding the mutagenic potential of untested substances. More detailed information on the strengths and limitations of the different (Q)SAR models can be found elsewhere (Serafimova *et al.*, 2010).

OECD QSAR Toolbox

To increase the regulatory acceptance of (Q)SAR models, the OECD has started the development of a QSAR Toolbox to make (Q)SAR technology readily accessible, transparent and less demanding in terms of infrastructure costs (http://www.qsartoolbox.org/). The OECD QSAR Toolbox facilitates the practical application of grouping and read-across approaches to fill gaps in (eco-)toxicity data, including genotoxicity and genotoxic carcinogenicity, for chemical hazard assessment. In particular, the OECD QSAR Toolbox covers the *in vitro* gene mutation (Ames test), *in vitro* chromosomal aberration, *in vivo* chromosomal aberration (micronucleus test), and genotoxic carcinogenicity endpoints. The predictions are based on the implementation of a range of profilers connected with genotoxicity and carcinogenicity (to quickly evaluate substances for common mechanisms or modes of action), and the incorporation of numerous databases with results from experimental studies (to support read-across and trend analysis) into a logical workflow. The Toolbox and guidance on its use are freely available. A user manual “Strategies for chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity” and various tutorials for categorisation of substances by use of the Toolbox in relation to protein- and DNA- binding and Ames test mutagenicity are also available on the OECD QSAR Toolbox web site.

The *Guidance on IR&CSA Chapter R.6: QSARs and grouping of chemicals* explains basic concepts of (Q)SARs and gives generic guidance on validation, adequacy and documentation for regulatory purposes. It also describes a stepwise approach for the use of read-across/grouping and (Q)SARs. Further information on the category formation and read-across approach for the prediction of toxicity can be found in Enoch (2010).

Testing data on mutagenicity

Test methods preferred for use are listed in Table R.7.7–2, Table R.7.7–3 and Table R.7.7–4. The introduction to the OECD TGs on genetic toxicity testing as well as some of the related OECD TGs are currently being revised under the OECD Test Guidelines Programme (TGP). In addition, an OECD Guidance Document on the selection and application of the assays for genetic toxicity is being developed. For further information, please see http://www.oecd.org/env/testguidelines.

In vitro data

Table R.7.7–2 *In vitro* test methods

Test method	GENOTOXIC ENDPOINTS measured / PRINCIPLE OF THE TEST METHOD	EU / OECD guideline[a]
Bacterial reverse mutation test	Gene mutations / The test uses amino-acid requiring strains of bacteria to detect (reverse) gene mutations (point mutations and frameshifts).	EU: B.13/14 OECD: 471
In vitro mammalian cell gene mutation test – *hprt* test	Gene mutations / The test identifies substances that induce gene mutations in the *hprt* gene of established cell lines.	EU: B.17 OECD: 476[b]
In vitro mammalian cell gene mutation test – Mouse lymphoma assay	Gene mutations and structural chromosome aberrations / The test identifies substances that induce gene mutations in the *tk* gene of the L5178Y mouse lymphoma cell line. If colonies in a *tk* mutation test are scored using the criteria of normal growth (large) and slow growth (small) colonies, gross structural chromosome aberrations (*i.e.* clastogenic effect) may be measured, since mutant cells that have suffered damage to both the *tk* gene and growth genes situated close to the *tk* gene have prolonged doubling times and are more likely to form small colonies.	EU: B.17 OECD: 476[b]
In vitro mammalian chromosome aberration test	Structural and numerical chromosome aberrations / The test identifies substances that induce structural chromosome aberrations in cultured mammalian established cell lines, cell strains or primary cell cultures. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this test is not optimal to measure numerical aberrations and is not routinely used for that purpose. Accordingly, this test guideline is not designed to measure numerical aberrations.	EU: B.10 OECD: 473[b]
In vitro micronucleus test	Structural and numerical chromosome aberrations / The test identifies substances that induce micronuclei in the cytoplasm of interphase cells. These micronuclei may originate from acentric fragments or whole chromosomes, and the test thus has the potential to detect both clastogenic and aneugenic substances.	EU: B.49 OECD: 487[b]

[a] For EU guidelines, see Regulation (EC) No 440/2008 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R0440:en:NOT) / for OECD guidelines see http://www.oecd.org/env/testguidelines
[b] OECD TGs 473, 476 and 487 are currently being revised (see http://www.oecd.org/env/testguidelines)

As noted earlier, accepted modifications to the standard test guidelines/methods have been developed to enhance test sensitivity to specific classes of substances. Expert judgement should be applied to judge whether any of these are appropriate for a given substance being registered. For example, protocol modifications for the Ames test might be appropriate for substances such as gases, volatile liquids, azo-dyes, diazo compounds, glycosides, and petroleum oil derived products, which should be regarded as special cases.

Animal data

- Somatic cells

Table R.7.7–3 *In vivo* test methods, somatic cells

Test method	GENOTOXIC ENDPOINTS measured / PRINCIPLE OF THE TEST METHOD	EU / OECD guideline[a]
In vivo mammalian bone marrow chromosome aberration test	Structural and numerical chromosome aberrations / The test identifies substances that induce structural chromosome aberrations in the bone-marrow cells of animals, usually rodents. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this test is not optimal to measure numerical aberrations and is not routinely used for that purpose. Accordingly, this test guideline is not designed to measure numerical aberrations.	EU: B.11 OECD: 475[b]
In vivo mammalian erythrocyte micronucleus test	Structural and numerical chromosome aberrations / The test identifies substances that cause micronuclei in erythroblasts sampled from bone marrow and/or peripheral blood cells of animals, usually rodents. These micronuclei originate from acentric fragments or whole chromosomes, and the test thus has the potential to detect both clastogenic and aneugenic substances.	EU: B.12 OECD: 474[b]
Unscheduled DNA synthesis (UDS) test with mammalian liver cells *in vivo*	DNA repair / The test identifies substances that induce DNA damage followed by DNA repair (measured as unscheduled "DNA" synthesis) in liver cells of animals, commonly rats. The test is usually based on the incorporation of tritium labelled thymidine into the DNA by repair synthesis after excision and removal of a stretch of DNA containing a region of damage.	EU: B.39 OECD: 486
Transgenic rodent (TGR) somatic and germ cell gene mutation assays	Gene mutations and chromosomal rearrangements (the latter specifically in the plasmid and Spi- assay models) / Since the transgenes are transmitted by the germ cells, they are present in every cell. Therefore, gene mutations and/or chromosomal rearrangements can be detected in virtually all tissues of an animal, including target tissues and specific site of contact tissues.	EU: B.58 OECD: 488
In vivo alkaline single-cell gel electrophoresis assay for DNA strand breaks (comet assay)	DNA strand breaks / The DNA strand breaks may result from direct interactions with DNA, alkali labile sites or as a consequence of incomplete excision repair. Therefore, the alkaline comet assay recognises primary DNA damage that would lead to gene mutations and/or chromosome aberrations, but will also detect DNA damage that may be effectively repaired or lead to cell death. The comet assay can be applied to almost every tissue of an animal from which single cell or nuclei suspensions can be made, including specific site of contact tissues.	EU: none OECD: 489

[a] For EU guidelines, see Regulation (EC) No 440/2008 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R0440:en:NOT) / for OECD guidelines see http://www.oecd.org/env/testguidelines
[b] OECD TGs 474 and 475 are currently being revised (see http://www.oecd.org/env/testguidelines)

A detailed review of transgenic animal model assays, including recommendations on how to perform such assays in somatic cells, has been produced for the OECD (Lambert *et al.,* 2005; OECD, 2009).

Validation studies and recommendations have been published in recent years, identifying experimental factors which are of importance for improved harmonisation of data obtained in the alkaline single-cell gel electrophoresis assay for DNA strand breaks (comet assay) (Ersson *et al*., 2013; Azqueta *et al*., 2013; Forchhammer *et al*., 2012; Azqueta *et al*., 2011a; Azqueta *et al*., 2011b; Forchhammer *et al*., 2009; Collins *et al*., 2008). Specifically, various international groups have proposed protocols and recommendations for performing the *in vivo* alkaline comet assay (Tice *et al.,* 2000; Hartmann *et al.,* 2003; McKelvey-Martin *et al.*, 1993; Brendler-Schwaab *et al.*, 2005; Burlinson et al., 2007; Smith *et al.*, 2008; Rothfuss *et al.*, 2010; Burlinson, 2012; Vasquez, 2012; Johansson *et al.*, 2010; Kirkland and Speit, 2008; EFSA, 2012). An international validation study on the *in vivo* alkaline single-cell gel electrophoresis assay was coordinated by the Japanese Centre for the Validation of Alternative Methods (JaCVAM) from 2006 to 2012. The validation study report was peer reviewed by the OECD and an OECD expert group drafted the comet OECD TG, which was approved by the OECD Working Group of National Coordinators of the Test Guidelines Programme (WNT) in April 2014. While awaiting the adoption of the comet OECD TG 489, the minimum criteria for acceptance of the comet assay published by EFSA (2012) can be used.

- Germ cells

Testing in germ cells has in the past been conducted only on very rare occasions (see Section R.7.7.6 of this Guidance).

Table R.7.7–4 *In vivo* test methods, germ cells

Test method	GENOTOXIC ENDPOINTS measured / PRINCIPLE OF THE TEST METHOD	EU / OECD guideline[a]
Mammalian spermatogonial chromosome aberration test	Structural and numerical chromosome aberrations / The test identifies substances that induce structural chromosome aberrations in mammalian, usually rodent, spermatogonial cells and is, therefore, expected to be predictive of induction of heritable mutations in germ cells. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this test is not optimal to measure numerical aberrations and is not routinely used for that purpose. Accordingly, this test guideline is not designed to measure numerical aberrations.	EU: B.23 OECD: 483[b]
Rodent dominant lethal test	Structural and numerical chromosome aberrations / The test identifies substances that induce dominant lethal effects causing embryonic or foetal death resulting from inherited dominant lethal mutations induced in germ cells of an exposed parent, usually the male. It is generally accepted that dominant lethals are due to structural and numerical chromosome aberrations. Rats or mice are recommended as the test species.	EU: B.22 OECD: 478[b]
Transgenic rodent (TGR) somatic and germ cell gene mutation assays	Gene mutations and chromosomal rearrangements (the latter specifically in the plasmid and Spi- assay models) / Since the transgenes are transmitted by the germ cells, they are present in every cell. Therefore, gene mutations and/or chromosomal rearrangements can be detected in virtually all tissues of an animal including specific site of contact tissues and germ cells. Delayed sampling times may need to be considered in order to detect mutations in different stages of spermatogenesis.	EU: none OECD: 488

[a] For EU guidelines, see Regulation (EC) No 440/2008 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R0440:en:NOT) / for OECD guidelines see http://www.oecd.org/env/testguidelines
[b] OECD TGs 478 and 483 are currently being revised (see http://www.oecd.org/env/testguidelines)

A detailed review of transgenic animal model assays, including recommendations on how to perform such assays in germ cells, has been produced for the OECD (Lambert *et al.,* 2005; OECD, 2009). The ability to include sampling of somatic and germ cells in a single study significantly reduces the need to perform additional studies to obtain such information, thereby conforming to the 3Rs principles. As specified in the OECD TG 488, additional sampling times may be needed to cover for the all the stages of spermatogenesis. The test can also be used to investigate transmission of mutations to the offspring since treatment of transgenic male mice can result in offspring carrying mutations (Barnett *et al*., 2002). An example of mutagenicity investigation in epididymal spermatozoa using a transgenic mouse model has been published (Olsen *et al.,* 2010).

The applicability of the standard alkaline comet assay to germ cells has been discussed by the OECD. The assay as described in the OECD TG 489 (see http://www.oecd.org/env/testguidelines) is not considered appropriate to measure DNA strand breaks in mature germ cells. Since high and variable background levels in DNA damage were reported in a literature review on the use of the comet assay for germ cell genotoxicity (Speit *et al.*, 2009), protocol modifications together with improved standardization and validation trials are deemed necessary before the comet assay on mature germ cells (*e.g.* sperm) can be included in the test guideline. In addition, the recommended exposure regimen described in this guideline is not optimal and longer exposures or sampling times would be necessary for a meaningful analysis of DNA strand breaks in mature sperm. Genotoxic effects as measured by the comet assay in testicular cells at different stages of differentiation have been described in the literature (Zheng *et al.*, 1997; Cordelli *et al.*, 2003). However, it should be noted that gonads contain a mixture of somatic and germ cells. For this reason, positive results in whole gonad (testis) are not necessarily reflective of germ cell damage, nevertheless, they suggest that tested chemicals have reached the gonad.

Databases with experimental data

There are several open-source databases with experimental information on mutagenicity and carcinogenicity (the two endpoints can often not easily be separated). A review of these databases can be found in Serafimova *et al.* (2010).

R.7.7.3.2 Human data on mutagenicity

Occasionally, studies of genotoxic effects in humans exposed by, for example, accident, occupation or participation in clinical studies (*e.g.* from case reports or epidemiological studies) may be available. Generally, cells circulating in blood are investigated for the occurrence of various types of genetic alterations.

R.7.7.4 Evaluation of available information on mutagenicity

Genotoxicity is a complex endpoint and requires evaluation by expert judgement. For both steps of the effects assessment, *i.e.* hazard identification and dose (concentration)-response (effect) assessment, it is very important to evaluate the data with regard to their adequacy and completeness. The evaluation of adequacy should address the reliability and relevance of the data in a way as outlined in the introductory chapter. The completeness of the data refers to the conclusion on the comparison between the available adequate information and the information that is required under the REACH provisions for the applicable tonnage level of the substance. Such a conclusion relies on *Weight of Evidence* approaches, which categorise available information based on the methods used: *guideline tests*, *non-guideline tests*, and other types of information which may justify adaptation of the standard testing regime. Such a *Weight of Evidence* approach also includes an evaluation of the available data as a whole, *i.e.* both *over and across* toxicological endpoints (for example, consideration of existing carcinogenicity data, repeated dose toxicity data and genotoxicity data all together can help understand whether a substance could be a genotoxic or non-genotoxic carcinogen).

This approach provides a basis to decide whether further information is needed on endpoints for which specific data appear inadequate or not available, or whether the requirements are fulfilled.

R.7.7.4.1 Non-human data on mutagenicity

Non-testing data for mutagenicity

In a more formal approach, documentation can include reference to a related substance or group of substances that leads to the conclusion of concern or lack of concern. This can either be presented according to scientific logic (read-across) or sometimes as a mathematical relationship of chemical similarity.

If well-documented and applicable (Q)SAR data are available, they should be used to help reach the decision points described in the section below. In many cases the accuracy of such methods will be sufficient to help, or allow either a testing or a specific regulatory decision to be made. In other cases the uncertainty may be unacceptable due to the severe consequences of a possible error. This may be driven by many factors including high exposure potential or toxicological concerns.

Substances for which no test-data exist or for which testing is technically not possible represent a special case in which reliance on non-testing data may be absolute. Many factors will dictate the acceptability of non-testing methods in reaching a conclusion based on no tests at all. It may be discussed whether *Weight of Evidence* decisions based on multiple genotoxicity and carcinogenicity estimates can equal or exceed those obtained by one or two *in vitro* tests, and whether general rules for adaptation of the standard testing regime as described in Annex XI to REACH may be invoked based on such estimates. This must be considered on a case-by-case basis.

Testing data on mutagenicity

Evaluation of genotoxicity test data should be made with care.

Regarding ***positive* findings**, particular points should be taken into account:

- are the testing conditions (*e.g.* pH, osmolality, precipitates) in *in vitro* mammalian cell assays relevant to the conditions *in vivo*?
- for studies *in vitro*, factors known to influence the specificity of mammalian cell assays such as the cell line used, the top concentration tested, the toxicity measure used or the metabolic activation system used, should be taken into consideration
- responses generated only at highly toxic/cytotoxic doses or concentrations should be interpreted with caution (*i.e.* taking into account the criteria defined in OECD guidelines)
- the presence or absence of a dose (concentration)-response relationship should be considered

Particular points to take into account when evaluating ***negative* test results** include:

- the doses or concentrations of test substance used (were they high enough? For studies *in vivo*, was a sufficienlty high dose level inducing signs of toxicity used? For studies *in vitro*, was a sufficient level of cytotoxicity reached?)
- was the test system used sensitive to the nature of the genotoxic changes that might have been expected? For example, some *in vitro* test systems will be sensitive to point mutations and small deletions but not to mutagenic events that create large deletions
- the volatility of the test substance (were concentrations maintained in tests conducted *in vitro*?)

- for studies *in vitro*, the possibility of metabolism not being appropriate in the test system including studies in extra-hepatic organs
- was the test substance taken up by the test system used for *in vitro* studies?
- were sufficient cells scored/sampled for studies *in vitro*? Has the appropriate number of samples/technical replicates been scored to support statistical significance of the putative negative result?
- for studies *in vivo*, did the substance reach the target organ? Or was the substance only in a position to act at the site of contact due to its high reactivity or insufficient systemic availability (taking also toxicokinetic data into consideration, *e.g.* rate of hydrolysis and electrophilicity may be factors that need to be considered)?
- for studies *in vivo*, was sampling appropriate? (Was a sufficient number of animals used? Were sufficient sampling times used? Was a sufficient number of cells scored/sampled?)

Different results between different test systems should be evaluated with respect to their individual significance. Examples of points to be considered are as follows:

- different results obtained in non-mammalian systems and in mammalian cell tests may be addressed by considering possible differences in substance uptake and metabolism, or in genetic material organisation and ability to repair. Although the results of mammalian tests may be considered of higher significance, additional data may be needed to explain differences
- if the results of indicator tests detecting putative DNA lesions (*e.g.* DNA binding, DNA damage, DNA repair; SCE) are not in agreement with results obtained in tests for mutagenicity, the results of mutagenicity tests are generally of higher significance provided that appropriate mutagenicity tests have been conducted. This is subject to expert judgement.
- if different findings are obtained *in vitro* and *in vivo*, in general, the results of *in vivo* tests indicate a higher degree of reliability. However, for evaluation of *negative* results *in vivo*, it should be considered whether the most appropriate tissues were sampled and whether there is adequate evidence of target tissue exposure
- the sensitivity and specificity of different test systems vary for different classes of substances. If available testing data for other related substances permit assessment of the performance of different assays for the class of substance under evaluation, the result from the test system known to produce more accurate responses would be given higher priority

Different results may also be available from the same test, performed by different laboratories or on different occasions. In this case, expert judgement should be used to evaluate the data and reach an overall conclusion. In particular, the quality of each of the studies and of the data provided should be evaluated, with special consideration of the study design, reproducibility of data, dose (concentration)-effect relationships, and biological relevance of the findings. The identity and purity of the test substance may also be a factor to take into account. In the case where an EU/OECD guideline is available for a test method, the quality of a study using the method is regarded as being higher if it was conducted in compliance with the requirements stated in the guideline, unless convincing scientific evidence can be provided to justify certain deviations from the standard test guideline for the specific substance evaluated. Furthermore, compared to non GLP-studies, studies compliant with GLP for the same assay generally provide more documentation and details of the study, which are important factors to consider when assessing study reliability/quality.

When making an assessment of the potential mutagenicity of a substance, or considering the need for further testing, data from various tests and genotoxic endpoints may be found. Both the strength and the weight of the evidence should be taken into account. The strongest evidence will be provided by modern, well-conducted studies with internationally established test guidelines/methods. For each test type and each genotoxic endpoint, there should be a separate *Weight of Evidence* analysis. It is not unusual for positive evidence of mutagenicity to be found in just one test type or for only one endpoint. In such cases the positive and negative results for different endpoints are not conflicting, but illustrate the advantage of using test methods for a variety of genetic alterations to increase the probability of identifying substances with mutagenic potential. Hence, results from methods testing different genotoxic endpoints should not be combined in an overall *Weight of Evidence* analysis, but should be subjected to such analysis separately for each endpoint. Based on the whole data set one has to consider whether there are data gaps: if there are data gaps further testing should be considered, otherwise an appropriate conclusion/assessment can be made.

R.7.7.4.2 Human data on mutagenicity

Human data have to be assessed carefully on a case-by-case basis. The interpretation of such data requires considerable expertise. Attention should be paid especially to the adequacy of the exposure information, confounding factors, co-exposures and to sources of bias in the study design or incident. The statistical power of the test may also be considered. It may be mentioned that, to date, no germ cell mutagen has been identified based on human data.

R.7.7.4.3 Remaining uncertainty on mutagenicity

Reliable data can be generated from well-designed and conducted studies *in vitro* and *in vivo*. However, due to the lack of human data available and the degree of uncertainty which is always inherent in testing, a certain level of uncertainty remains when extrapolating these testing data to the effect in humans.

R.7.7.5 Conclusions on mutagenicity

R.7.7.5.1 Concluding on Classification and Labelling

In order to conclude on an appropriate classification and labelling position with regard to mutagenicity, the available data should be considered using the criteria according to Annex I to the CLP Regulation (EC) No 1272/2008 (See also Section 3.5 of the *Guidance on the Application of the CLP criteria*).

R.7.7.5.2 Concluding on suitability for Chemical Safety Assessment

Considerations on dose (concentration)-response shapes and mode of action of mutagenic substances in test systems

Considerations on the dose (concentration)-response relationship and on possible mechanisms of action are important components of a risk assessment. The default assumption for genotoxic substances has for long been that they have a linear dose (concentration)-response relationship. However, this assumption has recently been challenged by experimental evidence showing that both direct and indirect acting genotoxins can possess non-linear or thresholded dose (concentration)-response curves.

Examples of non-DNA reactive mechanisms that may be demonstrated to lead to genotoxicity *via* non-linear or thresholded dose (concentration)-response relationships include inhibition of DNA synthesis, alterations in DNA repair, overloading of defence mechanisms (anti-oxidants or metal homeostatic controls), interaction with microtubule assembly leading to aneuploidy, topoisomerase inhibition, high cytotoxicity, metabolic overload and physiological perturbations (*e.g.* induction of erythropoeisis). The mechanisms underlying non-linear or thresholded dose (concentration)-response relationships for some DNA reactive genotoxic substances like alkylating agents seem linked to DNA repair capacity.

Assessment of the significance to be assigned to genotoxic responses mediated by such mechanisms would include an assessment of whether the underlying mechanism can be induced at substance concentrations that can be expected to occur under relevant *in vivo* conditions.

In general, several concentrations/doses are tested in genotoxicity assays. At least three experimental concentrations/doses have to be tested as recommended in the OECD test guidelines for genotoxicity. Determination of experimental dose (concentration)-effect relationships is one of several pieces of experimental information that are important to assess the genotoxic potential of a substance, and may be used as indicated below. It should be recognised that not all of these considerations may be applicable to *in vivo* data.

- the OECD introduction to the genotoxicity test guidelines lists the relevant criteria for identification of clear positive findings: (i) the increase in genotoxic response is concentration- or dose-related, (ii) at least one of the data points exhibits a statistically significant increase compared to the concurrent negative control, and (iii) the statistically significant result is outside the distribution of the historical negative control data (*e.g.* 95% confidence interval). In practice, the criterion for dose (concentration)-related increase in genotoxicity will be most helpful for *in vitro* tests, but care is needed to check for cytotoxicity or cell cycle delay which may cause deviations from a dose (concentration)-response related effect in some experimental systems

- genotoxicity tests are not designed in order to derive no effect levels. However, the magnitude of the lowest dose with an observed effect (*i.e.* the Lowest Observed Effect Dose or LOED) may, on certain occasions, be a helpful tool in risk assessment. This is true specifically for genotoxic effects caused by thresholded mechanisms, like, *e.g.* aneugenicity. Further, it can give an indication of the mutagenic potency of the substance in the test at issue. Modified studies, with additional dose or concentration points and improved statistical power may be useful in this regard. The Benchmark dose (BMD) approach presents several advantages over the NOED/LOED approach and can be used as an alternative strategy for dose (concentration)-response assessment (see the Guidance on IR&CSA, *Chapter R.8*)

- unusual shapes of dose (concentration)-response curves may contribute to the identification of specific mechanisms of genotoxicity. For example, extremely steep increases suggest an indirect mode of action or metabolic switching which could be confirmed by further investigation.

Considerations on genetic risks associated with human exposure to mutagenic substances

There are no officially adopted methods for estimating health risks associated with (low) exposures of humans to mutagens. In fact, most – if not all tests used today – are developed and applied to identify mutagenic properties of the substance, *i.e.* identification of the mutagenic hazard *per se*. In today's regulatory practice, the assessment of human health risks from exposure to mutagenic substances is considered to be covered by assessing and regulating the carcinogenic risks of these agents. The reason for this is that mutagenic events underlie these carcinogenic effects. Therefore, mutagenicity data is not used for deriving dose descriptors for risk assessment purposes and the reader is referred to this aspect in Section

R.7.7.8 (Carcinogenicity) for guidance on how to assess the chemical safety for mutagenic substances.

R.7.7.5.3 Information not adequate

A *Weight of Evidence* approach, comparing available adequate information with the tonnage-triggered information requirements by REACH, may result in the conclusion that the requirements are not fulfilled. In order to proceed in gathering further information, the following testing strategy can be adopted:

R.7.7.6 Integrated Testing Strategy (ITS) for mutagenicity

R.7.7.6.1 Objective / General principles

This testing strategy describes a flexible, stepwise approach for hazard identification with regard to the mutagenic potential of substances, so that sufficient data may be obtained for adequate risk characterisation including classification and labelling. It serves to help minimise the use of animals and costs as far as it is consistent with scientific rigour. A flow chart of the testing strategy is presented in Figure R.7.7–1 and recommendations on follow up procedures based on different testing data sets are given in Table R.7.7–5. As noted later in this section, deviations from this strategy may be considered if existing data for related substances indicate that alternate testing strategies yield results with greater sensitivity and specificity for mutagenicity *in vivo*.

The strategy defines a level of information that is considered sufficient to provide adequate reassurance about the potential mutagenicity of most substances. As described below, this level of information will be required for most substances at the Annex VIII tonnage level specified in REACH, although circumstances are described when the data may be required for substances at Annex VII.

For some substances, relevant data from other sources/tests may also be available (*e.g.* physico-chemical, toxicokinetic, and toxicodynamic parameters and other toxicity data; data on well-investigated, structurally similar, substances). These should be reviewed because, sometimes, they may indicate that either more or less genotoxicity studies are needed on the substance than defined by standard information requirements; *i.e.* they may allow tailored testing/selection of test systems. For example, bacterial mutagenesis assays of inorganic metal compounds are frequently negative due to limited capacity for uptake of metal ions and/or the induction of large DNA deletions by metals in bacteria potentially leading to an increased death rate in mutants. The high prevalence of false negatives for metal compounds might suggest that mutagenesis assays with mammalian cells, as opposed to bacterial cells, would be the preferred starting point for testing for this class of Annex VII substances.

In summary, a key concept of the strategy is that initial genotoxicity tests and testing guidelines/methods should be selected with due consideration to existing data that has established the most accurate testing strategy for the class of compound under evaluation. Even then, initial testing may not always give adequate information and further testing may sometimes be considered necessary in the light of all available relevant information on the substance, including its use pattern. Further testing will normally be required for substances which give rise to positive results in any of the *in vitro* tests.

If negative results are available from an adequate evaluation of genotoxicity from existing data in appropriate test systems, there may be no requirement to conduct additional genotoxicity tests.

Substances for which there is a harmonised classification in category 1A, 1B or 2 for germ cell mutagenicity and/or category 1A or 1B for carcinogenicity according to Annex VI to the CLP

Regulation (EC) No 1272/2008 will usually not require additional testing in order to meet the requirements of Annex VIII for the *in vitro* cytogenicity study in mammalian cells. Provided that appropriate risk management measures are implemented, the carcinogenicity study to meet the requirements of Annex X (see Section R.7.7.2 of this Guidance) and the reproductive toxicity studies to meet the requirements of Annexes VIII to X (see Section R.7.7.6 of this Guidance) may also be omitted for substances classified in category 1A or 1B for germ cell mutagenicity. In cases where a registrant is unsure of the formal position on the classification of a substance, or wishes to make a classification proposal himself, advice should be sought from an appropriate regulatory body before proceeding with any further testing.

In case additional testing is needed to meet the requirements of Annexes IX or X, the registrant must first submit a testing proposal to the European Chemicals Agency (ECHA) and obtain prior authorisation before any testing can be initiated.

It should also be noted that recommendations on a strategy for genotoxicity testing have also recently been published by other authoritative organisations (EFSA, 2011; EMA, 2012; UK COM, 2011). These strategies are based either on a step-wise approach or on a test-battery approach. Their principle is basically similar to the one detailed in this Guidance, *i.e.* the use of different pieces of information, including non-testing data and results from *in vitro* and *in vivo* testing, for a comprehensive assessment of the genotoxic potential a substance since no single test is capable of detecting all genotoxic mechanisms. However, as these strategies aim at serving different regulations and purposes, some differences can exist between them, in particular regarding the list of *in vitro* and *in vivo* tests recommended and the way to use them. For instance, while the UK COM and EFSA now both recommend the use of a core two-test battery (*i.e.* a bacterial reverse mutation test combined with an *in vitro* micronucleus test) for *in vitro* genotoxicity assessment, the REACH Regulation and this Guidance state the *in vitro* mammalian cell gene mutation test as a legal requirement in addition to the Ames test and the *in vitro* cytogenicity test if both are negative. Moreover, the *in vitro* chromosome aberration test is considered as a possible alternative option to the *in vitro* micronucleus test under REACH while it is now generally agreed that these tests are not equivalent since the *in vitro* chromosome aberration test is not optimal to measure numerical chromosome aberrations. Although this guidance aims at implementing the latest scientific developments in the field of genotoxicity testing, its main goal is to provide advice and support to the registrant in complying with the legal requirements under REACH and is thus in line with this Regulation.

R.7.7.6.2 Preliminary considerations

For a comprehensive coverage of the potential mutagenicity of a substance, information on gene mutations (base substitutions and deletions/additions), structural chromosome aberrations (breaks and rearrangements) and numerical chromosome aberrations (loss or gain of chromosomes, defined as aneuploidy) is required. This may be obtained from available data or tests on the substance itself or, sometimes, by prediction using appropriate *in silico* techniques (*e.g.* chemical grouping, read-across or (Q)SAR approaches).

It is important that whatever is known of the physico-chemical properties of the test substance is taken into account before devising an appropriate testing strategy. Such information may impact upon both the selection of test systems to be employed and/or modifications to the test protocols used. The chemical structure of a substance can provide information for an initial assessment of mutagenic potential. The need for special testing in relation to photomutagenicity may be indicated in some specific cases by the structure of a molecule, its light absorbing potential or its potential to be photoactivated. By using expert judgement, it may be possible to identify whether a substance, or a potential metabolite of a substance, shares or does not share structural characteristics with known mutagens. This can be used to justify a higher or lower level of priority for the characterisation of the mutagenic potential of a substance. Where the level of evidence for mutagenicity is particularly strong, it may be possible to make a conclusive hazard assessment in accordance with Annex I to REACH without additional testing on the basis of structure-activity relationships alone: in this case, the registrant still has to provide sufficient information to meet the requirements of Annexes VII to X but he may, if scientifically justified and duly documented in the registration dossier, invoke the general rules of Annex XI for adaptation of the standard testing regime by demonstrating, *inter alia*, that the results he wishes to use instead of testing in that context are adequate for the purpose of classification and labelling and/or risk assessment.

In vitro tests are particularly useful for gaining an understanding of the potential mutagenicity of a substance and they have a critical role in this testing strategy. They are not, however, without their limitations. Animal tests will, in general, be needed for the clarification of the relevance of positive findings and in case of specific metabolic pathways that cannot be simulated adequately *in vitro*.

The toxicokinetic and toxicodynamic properties of the test substance should be considered before undertaking, or appraising, animal tests. Understanding these properties will enable appropriate protocols for the standard tests to be developed, especially with respect to tissue(s) to be investigated, the route of substance administration and the highest dose tested. If little is understood about the systemic availability of a test substance at this stage, toxicokinetic investigations or modelling may be necessary.

Certain substances in addition to those already noted may need special consideration, such as highly electrophilic substances that give positive results *in vitro*, particularly in the absence of metabolic activation. Although these substances may react with proteins and water *in vivo* and thus be rendered inactive towards many tissues, they may be able to express their mutagenic potential at the initial site of contact with the body. Consequently, the use of test methods such as the comet assay or the gene mutation assays using transgenic animals that can be applied to the respiratory tract, upper gastrointestinal tract and skin may be appropriate. It is possible that specialised test methods will need to be applied in these circumstances, and that these may not have recognised, internationally valid, test guidelines. The validity and utility of such tests and the selection of protocols should be assessed by appropriate experts or authorities on a case-by-case basis.

Criteria for the evaluation and interpretation of results (e.g. how to define clear positive and clear negative results) are normally defined in the testing guidelines/methods. There is no requirement for verification of a clear positive or clear negative result. In cases where the response is neither clearly negative nor clearly positive and in order to assist in establishing the biological relevance of a result (e.g. a weak or borderline increase), the data should be evaluated by expert judgement and/or further investigations. A substance giving such a response should be reinvestigated immediately, normally using the same test method, but varying the conditions to obtain conclusive results. Only if, even after further investigations, the data set precludes coming to a conclusion of a positive or negative result, will the result be concluded as equivocal. Wherever possible, clear results should be obtained for one step in the strategic procedure before going on to the next. In cases where this does not prove to be possible and the study is inconclusive as a consequence of *e.g.* some limitation of the test or procedure, a further test should be conducted in accordance with the strategy.

Tests need not be performed if it is not technically possible to do so, or if they are not considered necessary in the light of current scientific knowledge. Scientific justifications for not performing tests required by the strategy should always be documented. It is preferred that tests as described in OECD Guidelines or Regulation (EC) No 440/2008 are used where possible. Alternatively, for other tests, up-to-date protocols defined by internationally recognised groups of experts, *e.g.* International Workshop on Genotoxicity Testing (IWGT, under the umbrella of the International Association of Environmental Mutagen Societies), may be used provided that the tests are scientifically justified. It is essential that all tests be conducted according to rigorous protocols in order to maximise the potential for detecting a mutagenic response, to ensure that negative results can be accepted with confidence and that results are comparable when tests are conducted in different laboratories. At the time of writing this guidance, a standard test guideline/method is still to be established for the *in vivo* comet assay described below. So if this test is to be conducted, and in waiting for the adoption of the comet OECD TG 489, consultation on the protocol with an appropriate expert or authority is advisable.

If a registrant wishes to undertake any tests for substances at the Annex IX or X tonnage levels that require the use of vertebrate animals, then there is a need to make a testing proposal to ECHA first. Testing may only be undertaken after ECHA has accepted the testing proposal in a formal decision.

R.7.7.6.3 Testing strategy for mutagenicity

Standard information requirement at Annex VII

A preliminary assessment of mutagenicity is required for substances at the REACH Annex VII tonnage level. All available information should be included but, as a minimum, there should normally be data from a gene mutation test in bacteria unless existing data for analogous substances indicates this would be inappropriate. For substances with significant toxicity to bacteria, not taken up by bacteria, or for which the gene mutation test in bacteria cannot be performed adequately, an *in vitro* mammalian cell gene mutation test may be used as an alternative test.

When the result of the bacterial test is positive, it is important to consider the possibility of the substance being genotoxic in mammalian cells. The need for further test data to clarify this possibility at the Annex VII tonnage level will depend on an evaluation of all the available information relating to the genotoxicity of the substance.

Standard information requirement at Annex VIII

For a comprehensive coverage of the potential mutagenicity of a substance, information on gene mutations, and structural and numerical chromosome aberrations is required for substances at the Annex VIII tonnage level of REACH.

In order to ensure the necessary minimum level of information is provided, at least one further test is required in addition to the gene mutation test in bacteria. This should be an *in vitro* mammalian cell test capable of detecting both structural and numerical chromosome aberrations.

There are essentially two different methods that can be viewed as alternative options according to REACH for this first mammalian cell test:

- An *in vitro* chromosome aberration test (OECD TG 473), *i.e.* a cytogenetic assay for structural chromosome aberrations using metaphase analysis. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this test is not optimal to measure numerical aberrations and is not routinely used for that purpose. Accordingly, this test guideline is not designed to measure numerical aberrations.

- An *in vitro* micronucleus test (OECD TG 487). This is a cytogenetic assay that has the advantage of detecting not only structural chromosomal aberrations but also aneuploidy. Use of a cytokinesis block, fluorescence *in situ* hybridisation with probes for centromeric DNA, or immunochemical labelling of kinetochore proteins can provide information on the mechanisms of chromosome damage and micronucleus formation. The labelling and hybridisation procedures can enable aneugens to be distinguished from clastogens. This may sometimes be useful for risk characterisation. If a substance is demonstrated to be an aneugen, it is assumed that its genotoxicity is thresholded, in contrast to non-thresholded genotoxicity. Both types of genotoxicity mechanisms trigger different ways to perform risk assessment.

Other *in vitro* tests may be acceptable as the first mammalian cell test, but care should be taken to evaluate their suitability for the substance being registered and their reliability as a screen for substances that cause structural and/or numerical chromosome aberrations. A supporting rationale should be presented for a registration with any of these other tests.

It is possible to present existing data from an *in vivo* cytogenetic test (*i.e.* a study or studies conducted previously) as an alternative to the first *in vitro* mammalian cell test. For instance, if an adequately performed *in vivo* micronucleus test is available already it may be presented as an alternative. There may however be specific cases where the *in vitro* mammalian cell test can still be justified even though *in vivo* cytogeneticity data exist. For example, in the *in vivo* micronucleus test, certain substances may not reach the bone marrow due to low bioavailability or specific tissue/organ distribution and would result negative. In addition, even if bioavailability of the parent compound in the bone marrow can be demonstrated, a clastogen requiring liver metabolism and for which the reactive metabolites formed are too short-lived to reach the bone marrow could give a negative result in the *in vivo* micronucleus test. In this case, *in vitro* testing could provide useful information on the mode of action of the substance, *e.g.* to understand whether the substance is clastogenic (or aneugenic) *in vitro*, and whether it requires a specific metabolism to be genotoxic. Justification of *in vitro* testing when *in vivo* data already exist should be considered on a case-by-case basis.

An *in vitro* gene mutation study in mammalian cells (OECD TG 476) is the second part of the standard information set required for registration at the Annex VIII tonnage level. For substances that have been tested already, this information should always be presented as part of the overall *Weight of Evidence* for mutagenicity with reference to induction of gene mutations in mammalian cells. For other substances, this second *in vitro* mammalian cell test will normally only be required when the results of the bacterial gene mutation test and the first study in mammalian cells (*i.e.* an *in vitro* chromosome aberration test or an *in vitro* micronucleus test) are negative. This is to detect *in vitro* mutagens that give negative results in the other two tests.

Under specific circumstances it may be possible to omit the second *in vitro* study in mammalian cells, *i.e.* if it can be demonstrated that this mammalian cell test will not provide any further useful information about the potential *in vivo* mutagenicity of a substance, then it does not need to be conducted. This should be evaluated on a case-by-case basis as there may be classes of compound for which conclusive data can be provided to show that the sensitivity of the first two *in vitro* tests cannot be improved by the conduct of the third test.

The *in vitro* mammalian cell gene mutation test will not usually be required if adequate information is available from a reliable *in vivo* study capable of detecting gene mutations. Such information may come from a TGR gene mutation assay. A comet assay or a liver UDS test may also be adequate. However, these two tests being indicator assays detecting putative DNA lesions, their use should be justified on a case-by-case basis, *e.g.* the UDS should be used only when it can be reasonably assumed that the liver is a target organ, since the UDS is restricted to the detection of primary DNA repair in liver cells.

Provided the *in vitro* tests have given negative results, normally, no *in vivo* tests will be required to fulfil the standard information requirements at Annex VIII. However, there may be rare occasions when it is appropriate to conduct testing *in vivo*, for example when it is not possible technically to perform satisfactory tests *in vitro*. Substances which, by virtue of, for example, their physico-chemical characteristics, chemical reactivity or toxicity cannot be tested in one or more of the *in vitro* tests should be considered on a case-by-case basis. In the same way, it may not always be possible with the S9 fraction used *in vitro* to mimic the *in vivo* metabolism of some substances, and the relevance of the *in vitro* negative results for those substances should be evaluated case by case. In addition, equivocal *in vitro* results or different results from different *in vitro* studies may require the consideration of further testing to reach a clear conclusion on mutagenicity. For those types of cases, expert judgement would be needed to determine whether *in vivo* testing is appropriate.

Requirement for testing beyond the standard levels specified for Annexes VII and VIII

Introductory comments

Concerns raised by positive results from *in vitro* tests usually require the consideration of further testing. The chemistry of the substance, data on analogous substances, toxicokinetic and toxicodynamic data, and other toxicity data will also influence the timing and pattern of further testing.

Unless there are appropriate results from an *in vivo* study already, testing beyond the standard set of *in vitro* tests is normally first directed towards investigating the potential for mutagenicity in somatic cells *in vivo*. Positive results in somatic cells *in vivo* constitute the trigger for consideration of investigation of potential expression of genotoxicity in germ cells. However, to avoid unnecessary testing of vertebrate animals and for cost reasons, as the TGR assays give the possibility to include sampling of somatic and male germ cells in a single study providing adapted sampling times (see OECD TG 488 for details), it is recommended to include such samples in the testing proposal for the TGR assays and to appropriately store the germ cell samples for later analysis in case there is a positive result in any of the somatic tissues tested.

Substances that are negative in the standard set of in vitro tests

In general, substances that are negative in the full set of *in vitro* tests specified in REACH Annexes VII and VIII are considered to be non-genotoxic. There are only a very limited number of substances that have been found to be genotoxic *in vivo*, but not in the standard *in vitro* tests. Most of these are pharmaceuticals designed to affect pathways of cellular regulation, including cell cycle regulation, and this evidence is judged insufficient to justify routine *in vivo* testing of industrial chemicals. However, occasionally, knowledge about the metabolic profile of a substance may indicate that the standard *in vitro* tests are not sufficiently reassuring and a further *in vitro* test, or an *in vivo* test, may be needed in order to ensure mutagenicity potential is adequately explored (*e.g.* use of an alternative to rat liver S9 mix, a reducing system, a metabolically active cell line, or genetically engineered cell lines might be judged appropriate).

Substances for which an in vitro test is positive

REACH Annex VII substances for which only a bacterial gene mutation test has been conducted and for which the result is positive should be studied further, according to the requirements of Annex VIII.

Regarding Annex VIII, when both the mammalian cell tests are negative but there was a positive result in the bacterial test, it will be necessary to decide whether any further testing is needed on a case-by-case basis. For example, suspicion that a unique positive response observed in the bacterial test was due to a specific bacterial metabolism of the test substance could be explored further by investigation *in vitro.* Alternatively, an *in vivo* test may be required (see below).

In REACH Annex VIII, following a positive result in an *in vitro* mammalian cell mutagenicity test, adequately conducted somatic cell *in vivo* testing is required to ascertain if this potential can be expressed *in vivo*. In cases where it can be sufficiently deduced that a positive *in vitro* finding is not relevant for *in vivo* situations (*e.g.* due to the effect of the test substances on pH or cell viability, *in vitro*-specific metabolism: see also Section R.7.7.4.1), or where a clear threshold mechanism coming into play only at high concentrations that will not be reached *in vivo* has been identified (*e.g.* damage to non-DNA targets at high concentrations), *in vivo* testing will not be necessary.

Annex VIII, Column 2 requires the registrant to consider appropriate mutagenicity *in vivo* studies already at the Annex VIII tonnage level, in cases where positive results in genotoxicity studies have been obtained. It should be noted that where this involves tests mentioned in Annexes IX or X, such as *in vivo* somatic cell genotoxicity studies, testing proposals must be submitted by the registrant and accepted by ECHA in a formal decision before testing can be initiated.

Standard information requirement according to Annexes IX and X

According to the requirements of Annexes IX and X, if there is a positive result in any of the *in vitro* studies from Annex VII or VIII and there are no appropriate results available from an *in vivo* study already, an appropriate *in vivo* somatic cell genotoxicity study should be proposed.

Before any decisions are made about the need for *in vivo* testing, a review of the *in vitro* test results and all available information on the toxicokinetic and toxicodynamic profile of the test substance is needed. A particular *in vivo* test should be conducted only when it can be reasonably expected from all the properties of the test substance and the proposed test protocol that the specific target tissue will be adequately exposed to the test substance and/or its metabolites. If necessary, a targeted investigation of toxicokinetics should be conducted before progressing to *in vivo* testing (*e.g*. a preliminary toxicity test to confirm that absorption occurs and that an appropriate dose route is used).

In the interest of ensuring that the number of animals used in genotoxicity tests is kept to a minimum, both males and females should not automatically be used. In accord with standard guidelines, testing in one sex only is possible when the substance has been investigated for general toxicity and no sex-specific differences in toxicity have been observed. If the test is performed in a laboratory with substantial experience and historical data, it should be considered whether a concurrent positive control and a concurrent negative control for all time points (*e.g*. for both the 24h and 48h time point in the micronucleus assay) will really be necessary (Hayashi *et al*., 2000).

For test substances with adequate systemic availability (*i.e*. evidence for adequate availability to the target cells) there are several options for the *in vivo* testing:

- A rodent bone marrow or mouse peripheral blood micronucleus test (OECD TG 474) or a rodent bone marrow chromosome aberration test (OECD TG 475). The micronucleus test has the advantage of detecting not only structural chromosomal aberrations (clastogenicity) but also numerical chromosomal aberrations (aneuploidy). Potential species-specific effects may also influence the choice of species and test method used.

- A transgenic rodent (TGR) mutation assay (OECD TG 488). TGR assays measure gene mutations and chromosomal rearrangements (the latter specifically in the plasmid and Spi- assay models) using reporter genes present in every tissue. In principle every tissue can be sampled, including target tissues and specific site of contact tissues.

- A comet (single cell gel electrophoresis) assay (OECD TG 489), which detects DNA strand breaks and alkali labile DNA lesions. In contrast to the above-mentioned *in vivo* micronucleus test and *in vivo* chromosome aberration test, this assay has the advantage of not being restricted to bone marrow cells. In principle every tissue from which single cell or nuclei suspensions can be prepared can be sampled, including specific site of contact tissues.

- Other DNA strand breakage assays may be presented as alternatives to the comet assay. All DNA strand break assays should be considered as surrogate tests, they do not necessarily detect permanent changes to DNA.

- A rat liver Unscheduled DNA synthesis (UDS) test (OECD TG 486). The UDS test is an indicator test measuring DNA repair of primary damage in liver cells but not a surrogate test for gene mutations *per se*. The UDS test can detect some substances that induce *in vivo* gene mutation because this assay is sensitive to some (but not all) DNA repair mechanisms. However not all gene mutagens are positive in the UDS test and it is thus useful only for some classes of substances. A positive result in the UDS assay can indicate exposure of the liver DNA and induction of DNA damage by the substance under investigation but it is not sufficient information to conclude on the induction of gene mutation by the substance. A negative result in a UDS assay alone is not a proof that a substance does not induce gene mutation.

Only the first two options for testing mentioned above can be used directly for providing evidence of *in vivo* chromosomal and gene mutagenicity, respectively. The other test methods require specific supporting information, for example results from *in vitro* mutagenicity studies, to be used for making definitive conclusions about *in vivo* mutagenicity and lack thereof.

In the framework of the 3Rs principles, the combination of *in vivo* genotoxicity studies or integration of *in vivo* genotoxicity studies into repeated dose toxicity studies, whenever possible and when scientifically justified, is strongly encouraged if this is to be performed to meet the requirements of the REACH Annex VIII tonnage level. All the above-mentioned *in vivo* tests for somatic cells are in principle amenable to such integration although sufficient experience is not yet available for all of the tests. It is possible for two or more endpoints to be combined into a single *in vivo* study, and thereby save on resources and numbers of animals used. The comet assay and the *in vivo* micronucleus test can be combined into a single acute study, although some modification of treatment and sampling times is needed (Hamada *et al*., 2001; Madrigal-Bujaidar *et al*., 2008; Pfuhler *et al.*, 2009; Bowen *et al*., 2011,). These same endpoints can be integrated into repeated dose (*e.g.* 28-day) toxicity studies (Pfuhler *et al*., 2009; Rothfuss *et al.*, 2011; EFSA, 2011).

Any one of these tests may be conducted, but this has to be decided using expert judgement on a case-by-case basis. The nature of the original *in vitro* response(s) (*i.e.* gene mutation, structural or numerical chromosome aberration) should be considered when selecting the *in vivo* study. For example, if the test substance showed evidence of *in vitro* clastogenicity, then it would be appropriate to follow this up with either a micronucleus test or chromosomal aberration test or a comet assay. However, if a positive result were obtained in the *in vitro* micronucleus test, the rodent micronucleus test would be appropriate to best address clastogenic and aneugenic potential.

For substances that appear preferentially to induce gene mutations, the TGR assays are the most appropriate and usually preferred tests to follow-up an *in vitro* gene mutation positive result and detect, *in vivo*, substances that induce gene mutation. With respect to the 3Rs principle and taking into account that a positive result in somatic cells triggers the need to consider the potential for germ cell testing, germ cells should always be collected, if possible, when a TGR study is performed. The rat liver UDS test has a long history of use and may in some specific cases be adequate to follow-up an *in vitro* gene mutation positive result, but not for tissues other than the liver. The sensitivity of the UDS test has been questioned (Kirkland and Speit, 2008) and the use of this test should be justified on a case-by-case basis, and take account of substance-specific considerations. The recommended use of the comet assay has been discussed at the OECD level and is indicated in the corresponding OECD TG (see http://www.oecd.org/env/testguidelines). The choice of any of these three assays can be justified only if it can be demonstrated that the tissue(s) studied in the assay is (are) sufficiently exposed to the test substance (or its metabolites). This information can be derived from toxicokinetic data or, in case no toxicokinetic data are available, from the observation of treatment-related effects in the organ of interest. Another type of data that can support evidence of organ exposure is knowledge on the target organ(s) of specific classes of substances (*e.g.* the liver for aromatic amines). In case the *in vivo* comet assay is used or proposed by the registrant, the test protocol followed or suggested should be described in detail and be in accordance with current scientific best practice, so as to ensure acceptability of the generated data. In waiting for the adoption of the comet OECD TG 489 the registrant should follow the EFSA guidance indicating the minimum criteria for acceptance of the comet assay (2012), as well as, for the combined comet-micronucleus test, the 3-day treatment schedule described by *e.g.* Bowen *et al.* (2011). The TGR and comet assays offer greater flexibility than the UDS test, most notably with regard to the possibility of selecting a range of tissues for study on the basis of what is known of the toxicokinetics and toxicodynamics of the substance. It should be realised that the UDS and comet tests are indicator assays: the comet assay detects DNA lesions whereas the UDS assay detects DNA repair patches (which depend on the DNA repair pathway involved and the proficiency of the cell type investigated), indirectly showing DNA lesions. In contrast, the TGR gene mutation assays measure mutations, *i.e.* permanent transmissible changes in the DNA.

Additionally, evidence for *in vivo* DNA adduct formation in somatic cells together with positive results from *in vitro* mutagenicity tests are sufficient to conclude that a substance is an *in vivo* somatic cell mutagen. In such cases, positive results from *in vitro* mutagenicity tests may not trigger further *in vivo* somatic tissue testing, and the substance would be classified at least as a category 2 mutagen. The possibility for effects in germ cells would need further investigation (see Section R.7.7.6.3, *Substances that give positive results in an in vivo test for genotoxic effects in somatic cells*).

Non-standard studies supported by published literature may sometimes be more appropriate and informative than established assays. Guidance from an appropriate expert or authority should be sought before undertaking novel studies. Furthermore, additional data that support or clarify the mechanism of action may justify a decision not to test further.

For substances inducing gene mutation or chromosomal aberration *in vitro*, and for which no indication of sufficient systemic availablity has been presented, or that are short-lived or reactive, an alternative strategy involving studies to focus on tissues at initial sites of contact with the body should be considered. Expert judgement should be used on a case-by-case basis to decide which tests are the most appropriate. The main options are the *in vivo* comet assay, TGR gene mutation assays, and DNA adduct studies. For any given substance, expert judgement, based on all the available toxicological information, will indicate which of these tests are the most appropriate. The route of exposure should be selected that best allows assessment of the hazard posed to humans. For insoluble substances, the possibility of release of active molecules in the gastrointestinal tract may indicate that a test involving the oral route of administration is particularly appropriate.

If the testing strategy described above has been followed and the first *in vivo* test is negative, the need for a further *in vivo* somatic cell test should be considered. The second *in vivo* test should only then be proposed if it is required to make a conclusion on the genotoxic potential of the substance under investigation; *i.e.* if the *in vitro* data show the substance to have potential to induce both gene and chromosome mutations and the first *in vivo* test has not addressed this comprehensively. In this regard, on a case-by-case basis, attention should be paid to the quality and relevance of all the available toxicological data, including the adequacy of target tissue exposure.

For a substance giving negative results in adequately conducted, appropriate *in vivo* test(s), as defined by this strategy, it will normally be possible to conclude that the substance is not an *in vivo* mutagen.

Substances that give positive results in an in vivo test for genotoxic effects in somatic cells

Substances that have given positive results in cytogenetic tests both *in vitro* and *in vivo* can be studied further to establish whether they specifically act as aneugens, and therefore whether thresholds for their genotoxic activity can be identified, if this has not been established adequately already. This should be done using *in vitro* methods and will be helpful in risk evaluation.

The potential for substances that give positive results in *in vivo* tests for genotoxic effects in somatic cells to affect germ cells should always be considered. The same is true for substances otherwise classified as category 2 mutagens under the CLP Regulation (EC) No 1272/2008 (for detailed information on the criteria for classification of substances for germ cell mutagenicity under the CLP Regulation (EC) No 1272/2008, see Section 3.5 of the *Guidance on the*

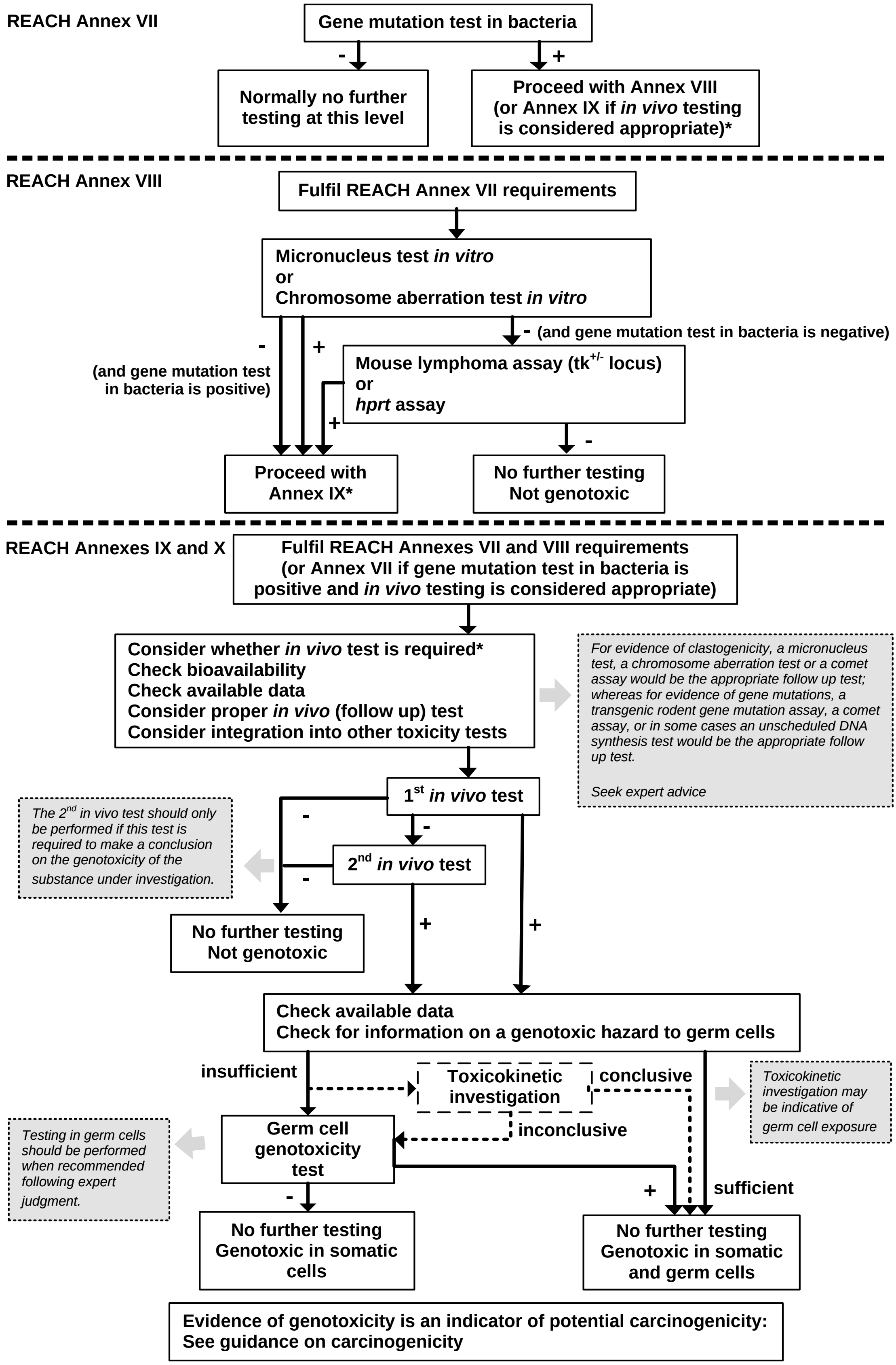
Application of the CLP criteria). The first step is to make an appraisal of all the available toxicokinetic and toxicodynamic properties of the test substance. Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified. Consequently, the substance is classified as a category 1B mutagen. If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation will be necessary. In the event that additional information about the toxicokinetics of the substance would resolve the problem, toxicokinetic investigation (*i.e.* not a full toxicokinetic study) tailored to address this should be performed. Although the hazard class for mutagenicity primarily refers to germ cells, the induction of genotoxic effects at site of contact tissues by substances for which no indication of sufficient systemic availability or presence in germ cells has been presented are also relevant and considered for classification. For such substances, at least one positive *in vivo* genotoxicity test in somatic cells can lead to classification in Category 2 germ cell mutagens and to the labelling as 'suspected of causing genetic defects' if the positive effect *in vivo* is supported by positive results of *in vitro* mutagenicity tests. Classification as Category 2 germ cell mutagen may also have implications for potential carcinogenicity classification.

If specific germ cell testing is to be undertaken, expert judgement should be used to select the most appropriate test strategy. Internationally recognised guidelines are available for investigating clastogenicity in rodent spermatogonial cells and for the dominant lethal test. Dominant lethal mutations are believed to be primarily due to structural or numerical chromosome aberrations.

Alternatively, other methods can be used if deemed appropriate by expert judgement. These may include the TGR gene mutation assays (with modified sampling times as indicated in the OECD TG 488 to detect effects at the different stages of spermatogenesis), or DNA adduct analysis. In principle, it is the potential for effects that can be transmitted to the progeny that should be investigated, but tests used historically to investigate transmitted effects (the heritable translocation test and the specific locus test) use very large numbers of animals. They are rarely used and should normally not be proposed for substances registered under REACH.

In order to minimise animal use, it is recommended to include cell samples from both relevant somatic and germ cell tissues (*e.g.* testes) in *in vivo* mutagenicity studies: the somatic cell samples can be investigated first and, if they are positive, germ cell tissues can then also be analysed. Finally, the possibility to combine reproductive toxicity testing with *in vivo* mutagenicity testing could be considered.

Figure R.7.7–1 Flow chart of the mutagenicity testing strategy



* Registrants should note that a testing proposal must be submitted for a test mentioned in Annex IX or X, independently from the registered tonnage. Following examination of such testing proposal ECHA has to approve the test in its evaluation decision before it can be undertaken.

Table R.7.7–5 Examples of different testing data sets and follow-up procedures to conclude on genotoxicity/mutagenicity according to the mutagenicity testing strategy.

Depending on the *in vitro* and *in vivo* test results available and the REACH Annex(es) of interest, further testing may be required to meet the standard information requirements for mutagenicity and allow for a conclusion on genotoxicity/mutagenicity to be reached. Recommendations on what should be done or particularly looked at in those different cases are mentioned in the table, together with specific rules for adaptation when applicable (for detailed guidance see also main text).

	GM bact	Cyt vitro	GM vitro	Cyt vivo	GM vivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation *[for detailed guidance, incl. timing of the tests, see main text]*	Comments
1	neg					Annex VII: no further tests are required. Annexes VIII, IX & X: conduct a CAbvitro or preferably a MNTvitro, and if this is negative, a GMvitro.	Annex VII: not genotoxic		Annexes VIII, IX & X: Select further tests in such a way that all the tests, together with other available information, enable thorough assessment for gene mutations and effects on chromosome structure and number.
2	neg	neg				Annex VII: no further tests are required. Annexes VIII, IX & X: conduct a GMvitro.	Annex VII: not genotoxic		Annexes VIII, IX & X: Select tests in such a way that all the tests, together with other available information, enable a thorough assessment for gene mutations and effects on chromosome structure and number.
3	neg		neg			Annex VII: no further tests are required. Annexes VIII, IX & X: conduct a CAbvitro or preferably a MNTvitro	Annex VII: not genotoxic		Annexes VIII, IX & X: Select tests in such a way that all the tests, together with other available information, enable a thorough assessment for gene mutations and effects on chromosome structure and number.

	GM bact	Cyt vitro	GM vitro	Cyt vivo	GM vivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation *[for detailed guidance, incl. timing of the tests, see main text]*	Comments
4	neg	neg	neg			Annexes VII, VIII, IX & X: no further tests are required.	not genotoxic		The available metabolic evidence may, on rare occasions, indicate that *in vitro* testing is inadequate*; in vivo* testing is needed. Seek expert advice. Annexes VIII, IX & X: Select tests in such a way that all the tests, together with other available information, enable a thorough assessment for gene mutations and effects on chromosome structure and number.
5	pos					Annexes VII, VIII, IX & X: Complete *in vitro* testing with a CAbvitro or preferably a MNTvitro.			Consider need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment, and to determine whether C&L is justified.

	GM bact	Cyt vitro	GM vitro	Cyt vivo	GM vivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation *[for detailed guidance, incl. timing of the tests, see main text]*	Comments
6	pos	neg				Annexes VII & VIII: *Complete in vitro* testing by conducting a GMvitro only under special conditions (see column 'Specific rules for adaption') Annexes IX & X: If systemic availability cannot be ascertained reliably, it should be investigated before progressing to *in vivo* tests. Select adequate somatic cell i*n vivo* test to investigate gene mutations *in vivo* (TGR, comet or if justified UDSvivo). If the TGR is to be conducted on somatic tissues, germ cell samples should be collected if possible, frozen and analysed for mutagenicity only in case of a positive result in somatic cells. If necessary seek expert advice.		Suspicion that a positive response observed in the GMbact was due to a specific bacterial metabolism of the test substance could be explored further by investigation *in vitro*.	Ensure that all tests together with other available information enable thorough assessment for gene mutations and effects on chromosome structure and number. Consider on a case-by-case basis need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment, and to determine whether C&L is justified.
7	neg	pos				Annexes VII, VIII, IX & X: If systemic availability cannot be ascertained reliably, it should be investigated before progressing to *in vivo* tests. Select adequate somatic cell *in vivo* test to investigate structural or numerical chromosome aberrations (MNTvivo or comet for *in vitro* clastogens and/or aneugens or CAbvivo for *in vitro*-clastogens) If necessary seek expert advice.			Ensure that all tests together with other available information enable thorough assessment for gene mutations and effects on chromosome structure and number. Consider need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment and to determine whether C&L is justified.

	GM bact	Cyt vitro	GM vitro	Cyt vivo	GM vivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation *[for detailed guidance, incl. timing of the tests, see main text]*	Comments
8	pos	pos				Annexes VII, VIII, IX & X: If systemic availability cannot be ascertained with acceptable reliability, it should be investigated before progressing to *in vivo* tests. Select adequate somatic cell *in vivo* tests to investigate both structural or numerical chromosome aberrations and gene mutations. If necessary seek expert advice.		Generally, both genotoxic endpoints should be investigated. If the first *in vivo* test is positive, a second *in vivo* test to confirm the other genotoxic endpoint need not be conducted. If the first *in vivo* test is negative, a second *in vivo* test is required if the first test did not address the endpoints comprehensively.	Ensure that all tests together with other available information enable thorough assessment for gene mutations and effects on chromosome structure and number. Consider need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment, and to determine whether C&L is justified.
9	neg	neg	pos			Annexes VII, VIII, IX & X: If systemic availability cannot be ascertained reliably, it should be investigated before progressing to *in vivo* tests. Select adequate somatic cell *in vivo* test to investigate gene mutations *in vivo* (TGR, comet or if justified UDSvivo). If the TGR is to be conducted on somatic tissues, germ cell samples should be collected if possible, frozen and analysed for mutagenicity only in case of a positive result in somatic cells. If necessary seek expert advice.			Ensure that all tests together with other available information enable thorough assessment for gene mutations and effects on chromosome structure and number. Consider on a case-by-case basis need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment, and to determine whether C&L is justified.
10	pos	neg			neg	Annexes VII, VIII, IX & X: no further tests are required.	not genotoxic		Further *in vivo* test may be necessary depending on the quality and relevance of available data.
	neg	pos		neg					

	GM bact	Cyt vitro	GM vitro	Cyt vivo	GM vivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation *[for detailed guidance, incl. timing of the tests, see main text]*	Comments
11	pos	neg			pos	Annexes VII, VIII, IX & X: No further testing in somatic cells is needed. Germ cell mutagenicity tests should be considered. If necessary seek expert advice on implications of all available data on toxicokinetics and toxicodynamics and on the choice of the proper germ cell mutagenicity test.	genotoxic	Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified.	If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation may be necessary. Risk assessment and C&L can be completed.
	neg	pos		pos					
	neg	neg	pos		pos				
12	pos	pos	(pos)	pos		Annexes VII, VIII, IX & X: No further testing in somatic cells is needed. Germ cell mutagenicity tests should be considered. If necessary seek expert advice on implications of all available data on toxicokinetics and toxicodynamics and on the choice of the proper germ cell mutagenicity test.	genotoxic	Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified.	If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation may be necessary. Risk assessment and C&L can be completed.
	pos	pos	(pos)		pos				
13	pos	pos	(pos)	neg		Annexes VII, VIII, IX & X: Select adequate somatic cell *in vivo* tests to investigate both structural or numerical chromosome aberrations and gene mutations. If necessary seek expert advice.			
	pos	pos	(pos)		neg				
14	pos	pos	(pos)	neg	neg	Annexes VII, VIII, IX & X: no further tests are required.	not genotoxic	Further *in vivo* test may be necessary pending on the quality and relevance of available data.	Risk assessment and C&L can be completed.

	GM bact	Cyt vitro	GM vitro	Cyt vivo	GM vivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation *[for detailed guidance, incl. timing of the tests, see main text]*	Comments
15	pos	pos	(pos)	neg	pos	Annexes VII, VIII, IX & X: No further testing in somatic cells is needed. Germ cell mutagenicity tests should be considered. If necessary seek expert advice on implications of all available data on toxicokinetics and toxicodynamics and on the choice of the proper germ cell mutagenicity test.	genotoxic	Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified.	If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation will be necessary. Risk assessment and C&L can be completed.
	pos	pos	(pos)	pos	neg				

Abbreviations: pos: positive; neg: negative; (pos): the follow up is independent from the result of this test; GM_{bact}: gene mutation test in bacteria (Ames test); Cyt_{vitro}: cytogenetic assay in mammalian cells; CAb_{vitro}: *in vitro* chromosome aberration test; MNT_{vitro}: *in vitro* micronucleus test; GM_{vitro}: gene mutation assay in mammalian cells; Cyt_{vivo}: cytogenetic assay in experimental animals; GM_{vivo}: gene mutation assay in experimental animals; CAb_{vivo}: *in vivo* chromosome aberration test (bone marrow); MNT_{vivo}: *in vivo* micronucleus test (erythrocytes); UDS_{vivo}: *in vivo* unscheduled DNA synthesis test; TGR: *in vivo* gene mutation test with transgenic rodent; comet: comet assay.

R.7.7.7 References on mutagenicity

Ashby J and Tennant RW (1988) Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. Mutat Res 204: 17-115.

Azqueta A, Gutzkow KB, Brunborg G and Collins AR (2011a) Towards a more reliable comet assay: optimising agarose concentration, unwinding time and electrophoresis conditions. Mutat Res 724: 41-5.

Azqueta A, Meier S, Priestley C, Gutzkow KB, Brunborg G, Sallette J, Soussaline F and Collins A (2011b) The influence of scoring method on variability in results obtained with the comet assay. Mutagenesis 26: 393-9.

Azqueta A, Gutzkow KB, Priestley CC, Meier S, Walker JS, Brunborg G and Collins AR (2013) A comparative performance test of standard, medium- and high-throughput comet assays. Toxicol In Vitro 27: 768-73.

Bakhtyari NG, Raitano G, Benfenati E, Martin T and Young D (2013) Comparison of in silico models for prediction of mutagenicity. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev 31: 45-66.

Bowen DE, Whitwell JH, Lillford L, Henderson D, Kidd D, Mc Garry S, Pearce G, Beevers C and Kirkland DJ (2011) Work conducted at Covance Laboratories Ltd., Harrogate. Evaluation of a multi-endpoint assay in rats, combining the bone-marrow micronucleus test, the Comet assay and the flow-cytometric peripheral blood micronucleus test. Mutat Res 722: 7-19.

Brendler-Schwaab S, Hartmann A, Pfuhler S and Speit G (2005) The *in vivo* comet assay: use and status in genotoxicity testing. Mutagenesis 20: 245-54.

Burlinson B, Tice RR, Speit G, Agurell E, Brendler-Schwaab SY, Collins AR, Escobar P, Honma M, Kumaravel TS, Nakajima M, Sasaki YF, Thybaud V, Uno Y, Vasquez M and Hartmann A (2007) Fourth International Workgroup on Genotoxicity Testing: result of the in vivo Comet assay workgroup. Mutat Res 627:31-5.

Burlinson B (2012) The *in vitro* and *in vivo* comet assays. Methods Mol Biol 817:143-63.

Collins AR, Oscoz AA, Brunborg G, Gaivão I, Giovannelli L, Kruszewski M, Smith CC, Stetina R (2008) The comet assay: topical issues. Mutagenesis 23: 143-51.

Cordelli E, Fresegna AM, Leter G, Eleuteri P, Spano M, Villani P (2003) Evaluation of DNA damage in different stages of mouse spermatogenesis after testicular X irradiation. Radiat Res 160:443-451.

EFSA, European Food Safety Authority (2011) Scientific Opinion of the Scientific Committee on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal 2011 9(9):2379 [69 pp.] Available online: www.efsa.europa.eu/efsajournal

EFSA, European Food Safety Authority (2012) Minimum Criteria for the acceptance of in vivo alkaline Comet Assay Reports. EFSA Journal 2012 10(11):2977 [12 pp.] Available online: www.efsa.europa.eu/efsajournal

EMA, European Medicines Agency (2012) International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use guideline S2 (R1) on genotoxicity testing and data interpretation for pharmaceuticals intended for human use - Step 5. Available online: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/12/WC500119604.pdf

Enoch SJ (2010) Chemical Category Formation and Read-Across for the Prediction of Toxicity *In:* Puzyn T, Leszczynski J and Cronin MTD (eds), Recent Advances in QSAR Studies, Volume 8, pp 209-219.

Ersson C, Møller P, Forchhammer L, Loft S, Azqueta A, Godschalk RW, van Schooten FJ, Jones GD, Higgins JA, Cooke MS, Mistry V, Karbaschi M, Phillips DH, Sozeri O, Routledge MN, Nelson-Smith K, Riso P, Porrini M, Matullo G, Allione A, Stepnik M, Ferlińska M, Teixeira JP, Costa S, Corcuera LA, López de Cerain A, Laffon B, Valdiglesias V, Collins AR and Möller L (2013) An ECVAG inter-laboratory validation study of the comet assay: inter-laboratory and intra-laboratory variations of DNA strand breaks and FPG-sensitive sites in human mononuclear cells. Mutagenesis 28: 279-86.

Johansson C, Møller P, Forchhammer L, Loft S, Godschalk RW, Langie SA, Lumeij S, Jones GD, Kwok RW, Azqueta A, Phillips DH, Sozeri O, Routledge MN, Charlton AJ, Riso P, Porrini M, Allione A, Matullo G, Palus J, Stepnik M, Collins AR and Möller L (2010) An ECVAG trial on assessment of oxidative damage to DNA measured by the comet assay. Mutagenesis 25: 125-32.

Forchhammer L, Johansson C, Loft S, Möller L, Godschalk RW, Langie SA, Jones GD, Kwok RW, Collins AR, Azqueta A, Phillips DH, Sozeri O, Stepnik M, Palus J, Vogel U, Wallin H, Routledge MN, Handforth C, Allione A, Matullo G, Teixeira JP, Costa S, Riso P, Porrini M and Møller P (2010) Variation in the measurement of DNA damage by comet assay measured by the ECVAG inter-laboratory validation trial. Mutagenesis 25: 113-23.

Forchhammer L, Ersson C, Loft S, Möller L, Godschalk RW, van Schooten FJ, Jones GD, Higgins JA, Cooke M, Mistry V, Karbaschi M, Collins AR, Azqueta A, Phillips DH, Sozeri O, Routledge MN, Nelson-Smith K, Riso P, Porrini M, Matullo G, Allione A, Stępnik M, Komorowska M, Teixeira JP, Costa S, Corcuera LA, López de Cerain A, Laffon B, Valdiglesias V and Møller P (2012) Inter-laboratory variation in DNA damage using a standard comet assay protocol. Mutagenesis 27: 665-72.

Hamada S, Sutou S, Morita T, Wakata A, Asanami S, Hosoya S, Ozawa S, Kondo K, Nakajima M, Shimada H, Osawa K, Kondo Y, Asano N, Sato S, Tamura H, Yajima N, Marshall R, Moore C, Blakey DH, Schechtman LM, Weaver JL, Torous DK, Proudlock R, Ito S, Namiki C and Hayashi M (2001) Evaluation of the rodent micronucleus assay by a 28-day treatment protocol: Summary of the 13th Collaborative Study by the Collaborative Study Group for the Micronucleus Test (CSGMT)/Environmental Mutagen Society of Japan (JEMS)-Mammalian Mutagenicity Study Group (MMS). Environ Mol Mutagen 37:93-110.

Hartmann A, Agurell E, Beevers C, Brendler-Schwaab S, Burlinson B, Clay P, Collins A, Smith A, Speit G, Thybaud V and Tice RR (2003) Recommendations for conducting the *in vivo* alkaline Comet assay. Mutagenesis 18: 45-51.

Hayashi M, MacGregor JT, Gatehouse DG, Adler ID, Blakey DH, Dertinger SD, Krishna G, Morita T, Russo A and Sutou S (2000) In vivo rodent erythrocyte micronucleus assay. II. Some aspects of protocol design including repeated treatments, integration with toxicity testing, and automated scoring. Environ Mol Mutagen 35: 234-52.

Kirkland D and Speit G (2008) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens III. Appropriate follow-up testing in vivo. Mutat Res 654:114-32.

Lambert IB, Singer TM, Boucher SE and Douglas GR (2005) Detailed review of transgenic rodent mutation assays. Mutat Res 590: 1-280.

Madrigal-Bujaidar E, Madrigal-Santillán EO, Alvarez-Gonzalez I, Baez R and Marquez P (2008) Micronuclei induced by imipramine and desipramine in mice: a subchronic study. Basic Clin Pharmacol Toxicol 103:569-73.

McKelvey-Martin VJ, Green MH, Schmezer P, Pool-Zobel BL, De Méo MP and Collins A (1993) The single cell gel electrophoresis assay (comet assay): a European review. Mutat Res 288:47-63.

Naven RT, Greene N and Williams RV (2012) Latest advances in computational genotoxicity prediction. Expert Opin Drug Metab Toxicol 8: 1579-87.

OECD (2009) Detailed Review Paper on Transgenic Rodent Mutation Assays, Series on Testing and Assessment, N° 103, ENV/JM/MONO(2009)7.

Olsen AK, Andreassen A, Singh R, Wiger R, Duale N, Farmer PB and Brunborg G (2010) Environmental exposure of the mouse germ line: DNA adducts in spermatozoa and formation of de novo mutations during spermatogenesis. PLoS One 5:e11349.

Pfuhler S, Kirkland D, Kasper P, Hayashi M, Vanparys P, Carmichael P, Dertinger S, Eastmond D, Elhajouji A, Krul C, Rothfuss A, Schoening G, Smith A, Speit G, Thomas C, van Benthem J and Corvi R (2009) Reduction of use of animals in regulatory genotoxicity testing: Identification and implementation opportunities—Report from an ECVAM workshop. Mutat Res 680: 31–42.

Rothfuss A, O'Donovan M, De Boeck M, Brault D, Czich A, Custer L, Hamada S, Plappert-Helbig U, Hayashi M, Howe J, Kraynak AR, van der Leede B-J, Nakajima M, Priestley C, Thybaud V, Saigo K, Sawant S, Shi J, Storer R, Struwe M, Vock E and Galloway S (2010) Collaborative study on fifteen compounds in the rat-liver Comet assay integrated into 2- and 4-week repeat-dose studies. Mutat Res 702: 40-69.

Rothfuss A, Honma M, Czich A, Aardema MJ, Burlinson B, Galloway S, Hamada S, Kirkland D, Heflich RH, Howe J, Nakajima M, O'Donovan M, Plappert-Helbig U, Priestley C, Recio L, Schuler M, Uno Y and Martus HJ (2011) Improvement of in vivo genotoxicity assessment: combination of acute tests and integration into standard toxicity testing. Mutat Res 723:108-20.

Serafimova R, Fuart Gatnik M and Worth A (2010) Review of QSAR models and software tools for predicting genotoxicity and carcinogenicity. JRC Technical Report EUR 24427 EN. Publications Office of the European Union, Luxembourg. Available online: http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/EUR_24427_EN.pdf

Smith CC, Adkins DJ, Martin EA and O'Donovan MR (2008) Recommendations for design of the rat comet assay. Mutagenesis 23: 233-40.

Somers CM, Yauk C, White PA, Parfett CLJ and Quinn JS (2002) Air pollution induces heritable DNA mutations. Proc Natl Acad Sci USA 99: 15904-15907.

Speit G, Vasquez M, Hartmann A (2009) The comet assay as an indicator test for germ cell genotoxicity. Mutat Res 681:3-12.

Tice RR, Agurell E, Anderson D, Burlinson B, Hartmann A, Kobayashi H, Miyamae Y, Rojas E, Ryu J-C and Sasaki F (2000) Single cell gel/Comet assay: guidelines for *in vitro* and *in vivo* genetic toxicology testing. Environ Mol Mutagen 35: 206-221.

UK COM, Committee on Mutagenicity of Chemicals in Food, Consumers Products, and the Environment (2011) Guidance on a strategy for genotoxicity testing of chemical substances. Available online: http://www.iacom.org.uk/guidstate/documents/COMGuidanceFINAL2.pdf

Vasquez MZ (2012) Recommendations for safety testing with the *in vivo* comet assay. Mutat Res 747:142-56.

Yauk C (2004) Advances in the application of germline tandem repeat instability for in situ monitoring. Mutat Res 566: 169-182.

Zheng H, Olive PL (1997).Influence of oxygen on radiation-induced DNA damage in testicular cells of C3H mice. Int J Radiat Biol 71:275-282.

R.7.7.8 Carcinogenicity

R.7.7.8.1 Definition of carcinogenicity

Chemicals are defined as carcinogenic if they induce tumours, increase tumour incidence and/or malignancy or shorten the time to tumour occurrence. Benign tumours that are considered to have the potential to progress to malignant tumours are generally considered along with malignant tumours. Chemicals can induce cancer by any route of exposure (e.g., when inhaled, ingested, applied to the skin or injected), but carcinogenic potential and potency may depend on the conditions of exposure (e.g., route, level, pattern and duration of exposure). Carcinogens may be identified from epidemiological studies, from animal experiments and/or other appropriate means that may include (Quantitative) Structure-Activity Relationships ((Q)SAR) analyses and/or extrapolation from structurally similar substances (read-across). Each strategy for the identification of potential carcinogens is discussed in detail later in this report. The determination of the carcinogenic potential of a chemical is based on a *Weight of Evidence* approach. Classification criteria are given in the (EU Directive 67/548/EEC).[126]

The process of carcinogenesis involves the transition of normal cells into cancer cells *via* a sequence of stages that entail both genetic alterations (i.e. mutations[127]) and non-genetic events. Non-genetic events are defined as those alterations/processes that are mediated by mechanisms that do not affect the primary sequence of DNA and yet increase the incidence of tumours or decrease the latency time for the appearance of tumours. For example; altered growth and death rates, (de)differentiation of the altered or target cells and modulation of the expression of specific genes associated with the expression of neoplastic potential (e.g. tumour suppressor genes or angiogenesis factors) are recognised to play an important role in the process of carcinogenesis and can be modulated by a chemical agent in the absence of genetic change to increase the incidence of cancer.

Carcinogenic chemicals have conventionally been divided into two categories according to the presumed mode of action: genotoxic or non-genotoxic[127]. Genotoxic modes of action involve genetic alterations caused by the chemical interacting directly with DNA to result in a change in the primary sequence of DNA. A chemical can also cause genetic alterations indirectly following interaction with other cellular processes (e.g., secondary to the induction of oxidative stress). Non-genotoxic modes of action include epigenetic changes, i.e., effects that do not involve alterations in DNA but that may influence gene expression, altered cell-cell communication, or other factors involved in the carcinogenic process. For example, chronic cytotoxicity with subsequent regenerative cell proliferation is considered a mode of action by which tumour development can be enhanced: the induction of urinary bladder tumours in rats may, in certain cases, be due to persistent irritation/inflammation, tissue erosion and regenerative hyperplasia of the urothelium following the formation of bladder stones. Other modes of non-gentoxic action can involve specific receptors (e.g., PPARα, which is associated with liver tumours in rodents; or tumours induced by various hormonal mechanisms). As with other nongenotoxic modes of action, these can all be presumed to have a threshold.

---

[126] Directive 67/548/EEC will be repealed and replaced with the EU Regulation on classification, labelling and packaging of substances and mixtures, implementing the Globally Harmonized System (GHS).

[127] For a definition and for background information on the terms mutagnicity and genotoxicity see Section R.7.7.1.1.

R.7.7.8.2 Objective of the guidance on carcinogenicity

The objective of investigating the carcinogenicity of chemicals is to identify potential human carcinogens, their mode(s) of action, and their potency.

With respect to carcinogenic potential and potency the most appropriate source of information is directly from human epidemiology studies (e.g. cohort, case control studies). In the absence of human data, animal carcinogenicity tests may be used to differentiate carcinogens from non-carcinogens. However, the results of these studies subsequently have to be extrapolated to humans, both in qualitative as well as quantitative terms. This introduces uncertainty, both with regard to potency for as well as relevance to humans, due to species specific factors such as differences in chemical metabolism and toxicokinetics and difficulties inherent in extrapolating from the high doses used in animal bioassays to those normally experienced by humans.

Once a chemical has been identified as a carcinogen, there is a need to elucidate the underlying mode of action, i.e. whether the chemical is directly genotoxic or not. In risk assessment a distinction is made between different types of carcinogens (see above).

For genotoxic carcinogens exhibiting direct interaction with DNA it is not generally possible to infer the position of the threshold from the *no-observed-effect level* on a dose-response curve, even though a biological threshold below which cancer is not induced may exist.

For non-genotoxic carcinogens, *no-effect-thresholds* are assumed to exist and to be discernable (e.g. if appropriately designed studies of the dose response for critical non-genotoxic effects are conducted). No effect thresholds may also be present for certain carcinogens that cause genetic alterations *via* indirect effects on DNA following interaction with other cellular processes (e.g. carcinogenic risk would manifest only after chemically induced alterations of cellular processes had exceeded the compensatory capacity of physiological or homeostatic controls). However, in the latter situation the scientific evidence needed to convincingly underpin this indirect mode of genotoxic action may be more difficult to achieve. Human studies are generally not available for making a distinction between the above mentioned modes of action; and a conclusion on this, in fact, depends on the outcome of mutagenicity/genotoxicity testing and other mechanistic studies. In addition to this, animal studies (e.g. the carcinogenicity study, repeated dose studies, and experimental studies with initiation-promotion protocols) may also inform on the underlying mode of carcinogenic action.

The cancer hazard and mode of action may also be highly dependent on exposure conditions such as the route of exposure. A pulmonary carcinogen, for example, can cause lung tumours in rats following chronic inhalation exposure, but there may be no cancer hazard associated with dermal exposure. Therefore, all relevant effect data and information on human exposure conditions are evaluated in a *Weight of Evidence* approach to provide the basis for regulatory decisions.

R.7.7.9 Information requirements on carcinogenicity

For the endpoint of carcinogenicity, standard information requirements are specifically described for substances produced or imported in quantities of ≥1000 t/y (Annex X). The precise information requirements will differ from substance to substance, according to the toxicity information already available and details of use and human exposure for the substance in question. The REACH Annexes VI to XI should be considered as a whole and in conjunction with the overall requirements of registration and evaluation.

Column 2 of Annex X lists specific rules according to which the required standard information may be omitted, replaced by other information, provided at a different stage or adapted in another way. If the conditions are met for adaptations under column 2 of this Annex, the fact and the reasons for each adaptation should be clearly indicated in the registration.

The standard information requirements for carcinogenicity and the specific rules for adaptation of these requirements are presented in Table R.7.7–6.

Table R.7.7–6 Standard information requirements for carcinogenicity and the specific rules for adaptation of these requirements

COLUMN 1 STANDARD INFORMATION REQUIRED	COLUMN 2 SPECIFIC RULES FOR ADAPTATION FROM COLUMN 1
Annexes VII-IX	
Annex X: 1. Carcinogenicity study.	1. A carcinogenicity study may be proposed by the registrant or may be required by the Agency in accordance with Articles 40 or 41 if: - the substance has a widespread dispersive use or there is evidence of frequent or long-term human exposure; and - the substance is classified as mutagen category 3 or there is evidence from the repeated dose study(ies) that the substance is able to induce hyperplasia and/or pre-neoplastic lesions. If the substance is classified as mutagen category 1 or 2, the default presumption would be that a genotoxic mechanism for carcinogenicity is likely. In these cases, a carcinogenicity test will normally not be required.

R.7.7.10 Information and its sources on carcinogenicity

There are many different sources of information that may permit inferences to be drawn regarding the potential of chemicals to be carcinogenic to humans. Clearly, these sources not only allow the identification of potential carcinogenic activity, but in case a substance is identified as a likely carcinogen they should also be informative with respect to the underlying mode of action as well as probable carcinogenic potency. The requirements of REACH call for proper classification and labelling, as well as for a quantitative assessment of risk that permits conclusions to be drawn regarding conditions under which safe use of the chemical may occur: i.e. the data should allow concluding on threshold or non-threshold mode of action, and on some dose descriptor (characterising the dose-response), preferably in quantitative terms.

It is noted (and indicated below), that the various sources inform differently on the aspects of hazard identification, mode of action, or carcinogenic potency.

R.7.7.10.1 Non-human data on carcinogenicity

Non-testing data on carcinogenicity

The capacity for performing the standard rodent cancer bioassay is limited by economic, technical and animal welfare considerations, such that an increased emphasis is being placed on the development of alternative, non-animal testing methods. However, carcinogenicity predictions through use of non-testing data currently represent an extreme challenge due to the multitude of possible mechanisms. Prediction of carcinogenicity in humans is especially problematic.

Although significant challenges remain, a broad spectrum of non-testing techniques exist for elucidating mechanistic, toxicokinetic or toxicodynamic factors important in understanding the carcinogenic process. These range from expert judgement, to the evaluation of structural similarities and analogues (i.e. read-across and grouping), to the use of (Q)SAR models for carcinogenicity. Such information may assist with priority setting, hazard identification, elucidation of the mode of action, potency estimation and/or with making decisions about testing strategies based on a *Weight of Evidence* evaluation.

Genotoxicity remains an important mechanism for chemical carcinogenesis and its definitive demonstration for a chemical is often decisive for the choice of risk assessment methodology. A commentary about non-testing options for genotoxicity is provided in Section R.7.7.1. It has long been known that certain chemical structures or fragments can be associated with carcinogenicity, often through DNA-reactive mechanisms. Useful guidance for structures and fragments that are associated with carcinogenicity *via* DNA reactive mechanisms has been provided by the US Food and Drug Administration's "Guideline for Threshold Assessment, Appendix I, Carcinogen Structure Guide" (US FDA, 1986); the Ashby-Tennant "super-mutagen model" (*e.g.*, Ashby and Tennant, 1988); and subsequent builds on this model (*e.g.*, Ashby and Paton, 1993; Munro *et al*., 1996). Additional information on structural categories can be found in the "IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Man" (IARC, 2006).

Models predicting test results for genotoxic endpoints for closely related structures are known as *local* or congeneric (Q)SARs. These congeneric models are less common for carcinogenicity than for mutagenicity. Franke *et al.* (2001) provide an example of such a model for a set of genotoxic carcinogens.

The situation is far more complex for non-genotoxic carcinogenicity due to the large number of different mechanisms that may be involved. However, progress is being made in predicting activity for classes of compounds that exert effect *via* binding to oestrogen receptors, induction of peroxisomal proliferation, and binding to tubulin proteins. Although many potentially useful models exist, their applicability will be highly dependant on the proposed mechanism and chemical class.

Several *global* (non-congeneric) models exist which attempt to predict (within their domain) the carcinogenic hazard of diverse (non-congeneric) groups of substances (*e.g.* Matthews and Contrera, 1998). These models may also assist in screening, priority-setting, deciding on testing strategies and/or the assessment of hazard or risk based on *Weight of Evidence*. Most are commercial and include expert systems such as Onco-Logic® (currently made available by US-EPA) and DEREK, artificial intelligence systems from MULTICASE, and the TOPKAT program. Historically, the performance of such models has been mixed and is highly dependent on the precise definition of carcinogenicity among those substances used to develop and test the model. These have been reviewed by ECETOC (2003) and Cronin *et al*. (2003).

Free sources of carcinogenicity predictions include the Danish EPA (Q)SAR database (accessible through the European Commission's Chemicals Bureau: ECB http://qsar.food.dtu.dk). Predictions in this database for 166,000 compounds include eight MULTICASE FDA cancer models, a number of genotoxicity predictions, rodent carcinogenic potency, hepatospecificity, oestrogenicity and aryl hydrocarbon (AH) receptor binding. Another source of carcinogenicity predictions is the Enhanced NCI database "*Browser*", which is sponsored by the US National Cancer Institute. This has 250,000 chemical predictions within it (http://cactus.nci.nih.gov), including general carcinogenicity, mutagenicity and additional endpoints, which may be of potential mechanistic interest in specific cases.

Further information on carcinogenicity models is available in the OECD Database on Chemical Risk Assessment Models where they are listed in an effort to identify tools for research and development of chemical substances. (http://www.olis.oecd.org/comnet/env/models.nsf/MainMenu?OpenForm).

The guidance on the Grouping of Chemicals and on (Q)SARs (see Sections R.6.2 and R.6.1, respectively) explains basic concepts of grouping and (Q)SARs and gives generic guidance on validation, adequacy and documentation for regulatory purposes. The guidance also describes a stepwise approach for the use of read-across/grouping and (Q)SARs.

It is noted that all the above mentioned sources may potentially inform on possible carcinogenic hazard and on the underlying mode of action, as well as on carcinogenic potency.

Testing data on carcinogenicity

In vitro data

The following *in vitro* data, which provide direct or indirect information useful in assessing the carcinogenic potential of a substance and (potentially) on the underlying mode(s) of action, may be available. No single endpoint or effect in and of itself possesses unusual significance for assessing carcinogenic potential but must be evaluated within the context of the overall toxicological effects of a substance under evaluation as described in Section R.7.7.11.1. Except as noted, standardised protocols do not exist for most of the *in vitro* endpoints noted. Rather, studies are conducted in accordance with expert judgement using protocols tailored to the specific substance, target tissue and cell type or animal species under evaluation.

genotoxicity studies: the ability of substances to induce mutations or genotoxicity (as defined in Section R.7.7.1) can be indicative of carcinogenic potential. However, correlations between mutagenicity/genotoxicity and carcinogenesis are stronger when effects are observed in appropriately designed *in vivo* as opposed to *in vitro* studies.

***in vitro* cell transformation assay results:** such assays assess the ability of chemicals to induce changes in the morphological and growth properties of cultured mammalian cells that are presumed to be similar to phenotypic changes that accompany the development of neoplastic or pre-neoplastic lesions *in vivo* (OECD, 2006). The altered cells detected by such assays may other targeted mechanisms of action

possess, or can subsequently acquire, the ability to grow as tumours when injected into appropriate host animals. As in vitro assays, cell transformation assays are restricted to the detection of effects of chemicals at the cellular level and will not be sensitive to carcinogenic activity mediated by effects exerted at the level of intact tissues or organisms.

mechanistic studies, e.g. on:

possess, or can subsequently acquire, the ability to grow as tumours when injected into appropriate host animals. As in vitro assays, cell transformation assays are restricted to the detection of effects of chemicals at the cellular level and will not be sensitive to carcinogenic activity mediated by effects exerted at the level of intact tissues or organisms.

mechanistic studies, e.g. on:

- cell proliferation: sustained cell proliferation can facilitate the growth of neoplastic/pre-neoplastic cells and/or create conditions conducive to spontaneous changes that promote neoplastic development.

- altered intercellular gap junction communication: exchange of growth suppressive or other small regulatory molecules between normal and neoplastic/pre-neoplastic cells through gap junctions is suspected to suppress phenotypic expression of neoplastic potential. Disruption of gap junction function, as assessed by a diverse array of assays for fluorescent dye transfer or the exchange of small molecules between cells, may attenuate the suppression of neoplastic potential by normal cells.

- hormone- or other receptor binding; a number of agents may act through binding to hormone receptors or sites for regulatory substances that modulate the growth of cells and/or control the expression of genes that facilitate the growth of neoplastic cells. Interactions of this nature are diverse and generally very compound specific.

other targeted mechanisms of action

- immunosuppressive activity: neoplastic cells frequently have antigenic properties that permit their detection and elimination by normal immune system function. Suppression of normal immune function can reduce the effectiveness of this *immune surveillance* function and permit the growth of neoplastic cells induced by exogenous factors or spontaneous changes.

- ability to inhibit or induce apoptosis: apoptosis, or programmed cell death, constitutes a sequence of molecular events that results in the death of cells, most often by the release of specific enzymes that result in the degradation of DNA in the cell nucleus. Apoptosis is integral to the control of cell growth and differentiation in many tissues. Induction of apoptosis can eliminate cells that might otherwise suppress the growth of neoplastic cells; inhibition of apoptosis can permit pre-neoplastic/neoplastic cells to escape regulatory controls that might otherwise result in their elimination.

- ability to stimulate angiogenesis or the secretion of angiogenesis factors: the growth of pre-neoplastic/neoplastic cells in solid tumours will be constrained in the absence of vascularisation to support the nutritional requirements of tumour growth. Secretion of angiogenesis factors stimulates the vascularisation of solid tumour tissue and enables continued tumour growth.

Animal data

A wide variety of study categories may be available, which may provide direct or indirect information useful in assessing the carcinogenic potential of a substance to humans. They include:

carcinogenicity studies (conventional long-term or life-time studies in experimental animals): Carcinogenicity testing is typically conducted using rats and mice, and less commonly in animals such as the Guinea pig, Syrian hamster and occasionally mini-pigs, dogs and primates. The standard rodent carcinogenicity bioassay would be conducted using rats or mice randomly assigned to treatment groups. Exposures to test substances may be *via* oral, inhalation or dermal exposure routes. The selection of exposure route is often dictated by *a priori* assumptions regarding the routes of exposure relevant to humans and/or other data sources (e.g. epidemiology studies or repeated dose toxicity studies in animals) that may indicate relevance of a given exposure route. Standardised protocols for such studies have been developed and are well validated (*e.g.* OECD TGs 451, 453 or US-EPA 870.4200).

short and medium term bioassay data (e.g., mouse skin tumour, rat liver foci model, neonatal mouse model): multiple assays have been developed that permit the detection and quantitation of putative pre-neoplastic changes in specific tissues. The induction of such *pre-neoplastic foci* may be indicative of carcinogenic potential. Such studies are generally regarded as adjuncts to conventional cancer bioassays, and while less validated and standardised, are applicable on a case-by-case basis for obtaining supplemental mechanistic and dose response information that may be useful for risk assessment (Enzmann *et al*., 1998).

genetically engineered (transgenic) rodent models (*e.g.*, *Xpa*$^{-/-}$, *p53*$^{+/-}$, *ras*H2 or Tg.AC): animals can be genetically engineered such that one or more of the molecular changes required for the multi-step process of carcinogenesis has been accomplished (Tennant *et al.*, 1999). This can increase the sensitivity of the animals to carcinogens and/or decrease the latency with which spontaneous or induced tumours are observed. The genetic changes in a given strain of engineered animals can increase sensitivity to carcinogenesis in a broad range of tissues or can be specific to the changes requisite for neoplastic development in one or only a limited number of tissues (Jacobson-Kram, 2004; Pritchard *et al.*, 2003; ILSI/HESI 2001). Data from these models may be used in a *Weight of Evidence* analysis of a chemical's carcinogenicity.

genotoxicity studies *in vivo*: the ability of substances to induce mutations or genotoxicity (as defined in Section R.7.7.1.1) can be indicative of carcinogenic potential. There is, in general, a good correlation between positive genotoxicity findings *in vivo* and animal carcinogenicity bioassay results

repeated dose toxicity tests: can identify tissues that may be specific targets for toxicity and subsequent carcinogenic effects. Particular significance can be attached to the observation of pre-neoplastic changes (e.g. hyperplasia or metaplasia) suspected to be conducive to tumour development and may assist in the development of dose-effect relationships (Elcombe *et al.*, 2002).

studies on the induction of sustained cell proliferation: substances can induce sustained cell proliferation *via* compensatory processes that continuously regenerate tissues damaged by toxicity. Some substances can also be tissue-specific mitogens, stimulating cell proliferation in the absence of overt toxic effects. Mitogenic effects are often associated with the action of tumour promoters. Both regenerative cell proliferation and mitogenic effects can be necessary, but not sufficient, for tumour development but have sufficiently different mechanistic basis that care should be exercised in assessing which is occurring (Cohen and Ellwein, 1991; Cohen *et al.*, 1991).

studies on immunosuppressive activity: as noted earlier, suppression of normal immune surveillance functions can interfere with normal immune system functions that serve to identify and eliminate neoplastic cells.

studies on toxicokinetics: can identify tissues or treatment routes that might be the targets for toxicity and can deliver data on exposure and metabolism in specific organs. Linkages to subsequent carcinogenic impacts may or may not exist, but such data can serve to focus carcinogenesis studies upon specific tissue types or animal species.

other studies on mechanisms/modes of action, e.g. OMICs studies (toxicogenomics, proteomics, metabonomics and metabolomics): carcinogenesis is associated with multiple changes in gene expression, transcriptional regulation, protein synthesis and other metabolic changes. Specific changes diagnostic of carcinogenic potential have yet to be validated, but these rapidly advancing fields of study may one day permit assessment of a broad array of molecular changes that might be useful in the identification of potential carcinogens.

It is noted that the above tests differently inform on hazard identification, mode of action or carcinogenic potency. For example, conventional bioassays are used for hazard identification and potency estimation (i.e. derivation of a dose descriptor), whereas studies using genetically engineered animals are informative on potential hazard and possibly mode of action, but less on carcinogenic potency as they are considered to be highly sensitive to tumour induction.

R.7.7.10.2 Human data on carcinogenicity

Human data may provide direct information on the potential carcinogenicity of the substance. Relevant human data of sufficient quality, if available, are preferable to animal data as no extrapolations between species, or from high to low dose are necessary. Epidemiological data will not normally be available for new substances but may well be available for substances that have been in use for many decades. For substances in common use prior to the implementation of modern occupational hygiene measures, the intensity of human exposures to some carcinogens was sufficient to produce highly significant, dose-dependent increases in cancer incidence.

A number of basic epidemiological study designs exist and include cohort, case-control and registry based correlational (e.g. ecological) studies. The most definitive epidemiological studies on chemical carcinogenesis are generally cohort studies of occupationally exposed populations, and less frequently the general population. Cohort studies evaluate groups of initially healthy individuals with known exposure to a given substance and follow the development of cancer incidence or mortality over time. With adequate information regarding the intensity of exposure experienced by individuals, dose dependent relationships with cancer incidence or mortality in the overall cohort can be established. Case-control studies retrospectively investigate individuals who develop a certain type of cancer and compare their chemical exposure to that of individuals who did not develop disease. Case control studies are frequently nested within the conduct of cohort studies and can help increase the precision with which excess cancer can be associated with a given substance. Correlational or ecological studies evaluate cancer incidence/mortality in groups of individuals presumed to have exposure to a given substance but are generally less precise since measures of the exposure experienced by individuals are not available. Observations of cancer clusters and case reports of rare tumours may also provide useful supporting information in some instances but are more often the impetus for the conduct of more formal and rigorous cohort studies.

Besides the identification of carcinogens, epidemiological studies may also provide information on actual exposures in representative (or historical) workplaces and/or the environment and the associated dose-response for cancer induction. Such information can be of much value for risk characterisation.

Although instrumental in the identification of known human carcinogens, epidemiology studies are often limited in their sensitivity by a number of technical factors. The extent and/or quality of information that is available regarding exposure history (e.g. measurements of individual exposure) or other determinants of health status within a cohort is often limited. Given the long latency between exposure to a carcinogen and the onset of clinical disease, robust estimates of carcinogenic potency can be difficult to generate. Similarly, occupational and environmentally exposed cohorts often have co-exposures to carcinogenic substances that have not been documented (or are incompletely documented). This can be particularly problematic in the study of long established industry sectors (e.g. base metal production) now known to entail co-exposures to known carcinogens (e.g. arsenic) present as trace contaminants in the raw materials being processed.. Retrospective hygiene and exposure analyses for such sectors are often capable of estimating exposure to the principle materials being produced, but data documenting critical co-exposures to trace contaminants may not be available. Increased cancer risk may be observed in such settings, but the source of the increased risk can be difficult to determine. Finally, a variety of lifestyle confounders (smoking and drinking habits, dietary patterns and ethnicity) influence the incidence of cancer but are often inadequately documented for purposes of adequate confounder control. Thus, modest increases in cancer at tissue sites known to be impacted by confounders (e.g. lung and stomach) can be difficult to interpret.

Techniques for biomonitoring and molecular epidemiology are developing rapidly. These newly developed tools promise to provide information on biomarkers of individual susceptibility, critical target organ exposures and whether effects occur at low exposure levels. Such ancillary information may begin to assist in the interpretation of epidemiology study outcomes and the definition of dose response relationships. For example, monitoring the formation of chemical adducts in haemoglobin molecules (Birner *et al.*, 1990; Albertini *et al*., 2006), the urinary excretion of damaged DNA bases (Chen and Chiu, 2005), and the induction of genotoxicity biomarkers (micronuclei or chromosome aberrations; Boffetta *et al.*, 2007) are presently being evaluated and/or validated for use in conjunction with classical epidemiological study designs. Such data are usually restricted in their application to specific chemical substances but such techniques may ultimately become more widely used, particularly when combined with animal data that defines potential mechanisms of action and associated biomarkers that may be indicative of carcinogenic risk. Monitoring of the molecular events that underlie the carcinogenic process may also facilitate the refinement of dose response relationships and may ultimately serve as early indicators of potential cancer risk. However, as a generalisation, such biomonitoring tools have yet to demonstrate the sensitivity requisite for routine use.

R.7.7.10.3 Exposure considerations for carcinogenicity

Information on exposure, use and risk management measures should be collected in accordance with Article 9 and Annex VI of REACH.

It is indicated in REACH Annex X a carcinogenicity study may be required by the European Chemicals Agency (or proposed by the registrant) when the substance has a widespread dispersive use or there is evidence of frequent or long-term human exposure. Preliminary toxicokinetic studies may be required first to address specific questions regarding potential target tissues and relevant exposure routes relevant for the chemical of concern.

On the other hand, investigations on the carcinogenic properties of a chemical can be deferred, if it can be demonstrated to the satisfaction of the Agency that the chemical is used only in a closed system and that human exposures are negligible (i.e. risk reduction measures on the substance are already equivalent to those applied to high potency carcinogenic substances of category 1 and 2. Reasons for this could include the presence of other substances for which strict exposure regimes are implemented or enforced). The rationale for exemption from testing, of course, needs to be clearly documented upon registration.

Also, considerations on exposure may influence the search for information, e.g. applicable to the actual route of exposure. For example, if from exposure scenarios it is clear that only a single specific route is involved, toxicity data for this route is of higher relevance in data gathering and evaluation than for the other routes. Also, the involvement of inhalation exposure to particles will prioritise toxicity information needs in order to allow a proper hazard evaluation and risk assessment.

R.7.7.11 Evaluation of available information on carcinogenicity

This particular endpoint is complex and requires evaluation by expert judgement.

Note that the objective of this evaluation is to acquire information on the carcinogenic potential of the substance: i.e. is the substance carcinogenic or not, and, if so, what is the underlying mode of action (thresholded or not), and what is its carcinogenic potency (i.e. there is a need to define a dose descriptor).

An evaluation on the above mentioned properties requires a combining of various types of information, as indicated in Section R.7.7.10 (and below). Such an evaluation needs a *Weight of Evidence* approach for arriving at conclusions, i.e. a careful gathering, sorting and weighing of the various pieces of information available. This exercise is particularly complex and, therefore, requires expert judgement input.

R.7.7.11.1 Non-human data on carcinogenicity

Non-testing data for carcinogenicity

To date little experience is available for the evaluation of substances on non-testing data, since the use of non-testing data for regulatory decisions is rather new. Therefore, at every stage in the assessment for potential chemical toxicity, specialist judgement is essential. It is recognised though, that non-testing data may potentially inform on all carcinogenic properties, i.e. including mode of action and potency.

Documentation should include reference to a related chemical or groups of chemicals that give rise to concern or lack of concern. This can either be presented according to scientific logic (read-across) or as a mathematical relationship of chemical similarity.

In some cases, the carcinogenic potential posed by a substance can be assessed based upon analysis of the relative concentrations of constituents believed to present a risk in a complex mixture. For example, the classification of certain complex coal- and oil-derived substances as carcinogens can varies as a function of the content of marker carcinogens (benzene, 1,3-butadiene and benzene), whereas for others it depends on the level of polycyclic aromatic hydrocarbons measured following DMSO solvent extraction. (see Annex I of EU Directive 67/548/EEC). When properly validated, such chemical extraction and analysis techniques are highly predictive of the outcomes that would be obtained in animal carcinogenicity studies.

If well documented and applicable, (Q)SARs can be used to help reach the decision points described in the section below. The accuracy of such methods may be sufficient to help or allow either a testing or a specific regulatory decision to be made. Expert judgement is needed to make this determination.

Chemicals for which no test-data exist present a special case in which reliance on non-testing methods may be absolute. Many factors will dictate the acceptability of non-testing methods in reaching a conclusion based on no tests at all. A *Weight of Evidence* evaluation of carcinogenicity based on multiple genotoxicity and carcinogenicity estimates (*e.g.* from (Q)SAR models) may in some cases equal or exceed the decision basis which could be obtained by experimentally testing a chemical in one or two *in vitro* tests. This must be considered on a case-by-case basis by the registrant.

Further guidance on the use of Grouping of Chemicals and on (Q)SARs both for a qualitative (i.e. classification and labelling) as well as a quantitative assessment (i.e. identifying some dose descriptor value) is provided in Sections R.4.3.2 and R.6.2, respectively, and also includes basic concepts used, validation status, adequacy and documentation needs for regulatory purposes.

Testing data on carcinogenicity

In vitro data

In vitro data can only give preliminary information about the carcinogenic potential of a substance and possible underlying mode(s) of action. For example, *in vitro* genotoxicity studies may provide information about whether or not the substance is likely to be genotoxic *in vivo*, and thus a potential genotoxic carcinogen (see Section R.7.7.1), and herewith on the potential mode of action underlying carcinogenicity: with or without a threshold.

Besides genotoxicity data other *in vitro* data (described in Section R.7.7.10.1) such as *in vitro* cell transformation can help to decide, in a *Weight of Evidence* evaluation, whether a chemical possesses a carcinogenic potential. Cell transformation results in and of themselves do not inform as to the actual underlying mode(s) of action, since they are restricted to the detection of effects exerted at the level of the single cell and may be produced by mechanistically distinct processes.

Studies can also be conducted to evaluate the ability of substances to influence processes thought to facilitate carcinogenesis. Many of these endpoints are assessed by experimental systems that have yet to be formally validated and/or are the products of continually evolving basic research. Formalised and validated protocols are thus lacking for the conduct of these tests and their interpretation. Although it is difficult to give general guidance on each test due to the variety and evolving nature of tests available, it is important to consider them on a case-by-case basis and to carefully consider the context on how the test was conducted.

A number of the test endpoints evaluate mechanisms that may contribute to neoplastic development, but the relative importance of each endpoint will vary as a function of the overall toxicological profile of the substance being evaluated. It should further be noted that there are significant uncertainties associated with extrapolating *in vitro* data to an *in vivo* situation. Such *in vitro* data will, in many instances, provide insights into the nature of the *in vivo* studies that might be conducted to define carcinogenic potential and/or mechanisms.

Animal data

In vivo data can give direct information about the carcinogenic potential of a substance, possible underlying mode(s) of action, and its potency.

Testing for carcinogenicity is conventionally carried out in groups of rats or mice according to standard test protocols or guidelines (e.g. OECD TGs 451, 453 or US-EPA 870.4200) and a conclusion is based on a comparison of the incidence, nature and time of occurrence of neoplasms in treated animals and controls.

Knowledge of the historic tumour incidence for the strain of animal used is important (laboratory specific data are preferable). Also attention to the study design used is essential because of the requirement for statistical analyses. The quality, integrity and thoroughness of the reported data from carcinogenicity studies are essential to the subsequent analysis and evaluation of studies. A qualitative assessment of the acceptability of study reports is therefore an important part of the process of independent evaluation. Sources of guidance in this respect can be found in IEH (2002), CCCF (2004) and OECD (2002). If the available study report does not include all the information required by the standard test guideline, judgement is required to decide if the experimental procedure is or is not acceptable and if essential information is lacking.

The final design of a carcinogenicity bioassay may deviate from OECD guidelines if expert judgement and experience in the testing of analogous substances supports the modification of protocols. Such modifications to standard protocols can be considered as a function of the specific properties of the material under evaluation.

Carcinogenicity data may sometimes be available in species other than those specified in standard test guidelines (e.g., Guinea pig, Syrian hamster and occasionally mini-pigs, dogs and primates). Such studies may be in addition to, or instead of, studies in rats and mice and they should be considered in any evaluation.

Data from non-conventional carcinogenicity studies, such as short- and medium-term carcinogenicity assays with neonatal or genetically engineered (transgenic) animals, may also be available (CCCF, 2004; OECD, 2002). Genetically engineered animals possess mutations in genes that are believed to be altered in the multi-step process of carcinogenesis, thereby enhancing animal sensitivity to chemically induced tumours. A variety of transgenic animal models exist and new models are continually being development. The genetic alteration(s) in a specific animal model can be those suspected to facilitate neoplastic development in a wide range of tissue types or the alterations can be in genes suspected to be involved in tissue specific aspects of carcinogenesis. The latter must be applied with recognition of both their experimental nature and the specific mechanistic pathways they are designed to evaluate. For example, a transgenic animal model sensitive to mesothelioma induction would be of limited value in the study of a suspected liver carcinogen. While such animal model systems hold promise for the detection of carcinogens in a shorter period of time and using fewer animals, their sensitivity and specificity remains to be determined. Due to a relative lack of validation, such assays have not yet been accepted as alternatives to the conventional lifetime carcinogenicity studies, but may be useful for screening purposes or to determine the need for a rodent 2-year bioassay. Several evaluations of these types of study have been published (e.g., Jacobson-Kram, 2004; Pritchard *et al.*, 2003; ILSI/HESI (2001).

When data are available from more than one study of acceptable quality, consistency of the findings should be established. When consistent, it is usually straightforward to arrive at a conclusion, particularly if the studies were in more than one species or if there is a clear treatment-related incidence of malignant tumours in a single study. If a single study only is available and the test substance is not carcinogenic, scientific judgement is needed to decide on whether (a) this study is relevant or (b) additional information is required to provide confidence that it should not be considered to be carcinogenic.

Study findings also may not clearly demonstrate a carcinogenic potential, even when approved study guidelines have been followed. For example, there may only be an increase in the incidence of benign tumours or of tumours that have a high background incidence in control animals. Although less convincing than an increase in malignant and rare tumours, and recognising the potential over-sensitivity of this model (Haseman, 1983; Ames and Gold, 1990), a detailed and substantiated rationale should be given before such positive findings can be dismissed as not relevant.

Repeated dose toxicity studies may provide helpful additional information to the *Weight of Evidence* gathered to determine whether a substance has the potential to induce cancer, and for potential underlying modes of action (Elcombe *et al*., 2002). For example, the induction of hyperplasia (either through cytotoxicity and regenerative cell proliferation, mitogenicity or interference with cellular control mechanisms) and/or the induction of pre-neoplastic lesions may contribute to the *Weight of Evidence* for carcinogenic potential. Toxicity studies may also provide evidence for immunosuppressive activity, a condition favouring tumour development under conditions of chronic exposure.

Finally, toxicokinetic data may reveal the generation of metabolites with relevant structural alerts. It may also give important information as to the potency and relevance of carcinogenicity and related data collected in one species and its extrapolation to another, based upon differences in absorption, distribution, metabolism and or excretion of the substance. Species specific differences mediated by such factors may be demonstrated through experimental studies or by the application of toxicokinetic modelling.

Positive carcinogenic findings in animals require careful evaluation and this should be done with reference to other toxicological data (e.g. *in vitro* and/or *in vivo* genotoxicity studies, toxicokinetic data, mechanistic studies, (Q)SAR evaluations) and the exposure conditions (e.g., route). Such comparisons may provide evidence for (a) specific mechanism(s) of action, a significant factor to take into account whenever possible, that may then be evaluated with respect to relevance for humans.

A conceptual framework that provides a structured and transparent approach to the *Weight of Evidence* assessment of the mode of action of carcinogens has been developed (see Sonich-Mullin *et al.*, 2001; Boobis *et al.*, 2006). This framework should be followed when the mechanism of action is key to the risk assessment being developed for a carcinogenic substance and can be particularly critical in a determination of whether a substance induces cancer *via* genotoxic or nongenotoxic mechanisms.

For example, a substance may exhibit limited genotoxicity *in vivo* but the relevance of this property to carcinogenicity is uncertain if genotoxicity is not observed in tissues that are the targets of carcinogenesis, or if genotoxicity is observed *via* routes not relevant to exposure conditions (e.g. intravenous injection) but not when the substance is administered *via* routes of administration known to induce cancer. In such instances, the apparent genotoxic properties of the substance may not be related to the mechanism(s) believed to underlie tumour induction. For example, oral administration of some inorganic metal compounds will induce renal tumours *via* a mechanism believed to involve organ specific toxicity and forced cell proliferation. Although genotoxic responses can be induced in non-target tissues for carcinogenesis *via* intravenous injection, there is only limited evidence to suggest that this renal carcinogenesis entails a genotoxic mechanism (IARC, 2006). The *burden of proof* in drawing such mechanistic inferences can be high but can have a significant impact upon underlying assumptions made in risk assessment.

In general, tumours induced by a genotoxic mechanism (known or presumed) are, in the absence of further information, considered to be of relevance to humans even when observed in tissues with no direct human equivalent. Tumours shown to be induced by a non-genotoxic mechanism are, in principle, also considered relevant to humans but there is a recognition that some non-genotoxic modes of action do not occur in humans (see OECD, 2002). This includes, for example, some specific types of rodent kidney, thyroid, urinary bladder, forestomach and glandular stomach tumours induced by rodent-specific modes of action, i.e., by mechanisms/modes of action not operating in humans or operative in humans under extreme and unrealistic conditions. Reviews are available for some of these tumour types providing a detailed characterisation that includes the key biochemical and histopathological events that are needed to establish these rodent-specific mechanisms that are not relevant for human health (see Technical Publication Series by IARC). Recently, the IPCS has developed a framework and provided some examples on how to evaluate the relevance to humans of a postulated mode of action in animals (ILSI RSI, 2003; Boobis *et al.*, 2006).

The information available for substances identified as carcinogenic based on testing and/or non-testing data should be further evaluated in an effort to identify underlying mode(s) of action and potency in order to subsequently allow a proper quantitative assessment of risk (see Section R.7.7.12.2). As already pointed out, the use of non-standard animal models (e.g. transgenic or neonatal animals) needs careful evaluation by expert judgement as to how to apply the results obtained for hazard and risk assessment purposes; it is not possible to provide guidance for such evaluations.

R.7.7.11.2 Human data on carcinogenicity

Epidemiological data may potentially be used for hazard identification, exposure estimation, dose response analysis, and risk assessment. The degree of reliability for each study on the carcinogenic potential of a substance should be evaluated using accepted causality criteria, such as that of Hill (1965). Particular attention should be given to exposure data in a study and to the choice of the control population. Often a significant level of uncertainty exists around identifying a substance unequivocally as being carcinogenic because of inadequate reporting of exposure data. Chance, bias and confounding factors can frequently not be ruled out. A clear identification of the substance, the presence or absence of concurrent exposures to other substances and the methods used for assessing the relevant dose levels should be explicitly documented. A series of studies revealing similar excesses of the same tumour type, even if not statistically significant, may suggest a positive association, and an appropriate joint evaluation (meta-analysis) may be used in order to increase the sensitivity, provided the studies are sufficiently similar for such an evaluation. When the results of different studies are inconsistent, possible explanations should be sought and the various studies judged on the basis of the methods employed.

Interpretation of epidemiology studies must be undertaken with care and include an assessment of the adequacy of exposure classification, the size of the study cohort relative to the expected frequency of tumours at tissue sites of special concern and whether basic elements of study design are appropriate (e.g. a mortality study will have limited sensitivity if the cancer induced has a high rate of successful treatment). A number of such factors can limit the sensitivity of a given study – unequivocal demonstration that a substance is not a human carcinogen is difficult and requires detailed and exact measurements of exposure, appropriate cohort size, adequate intensity and duration of exposure, sufficient follow-up time and sound procedures for detection and diagnosis of cancers of potential concern. Conversely, excess cancer risk in a given study can also be difficult to interpret if relevant co-exposures and confounders have not been adequately documented. Efforts are ongoing to improve the sensitivity and specificity of traditional epidemiological methods by combining cancer endpoints with data on established pre-neoplastic lesions or molecular indicators (biomarkers) of cancer risk.

Once identified as a carcinogenic substance on the basis of human data, well-performed epidemiology studies may be valuable for providing information on the relative sensitivity of humans as compared to animals, and/or may be useful in demonstrating an upper bound on the human cancer risk. Identification of the underlying mode(s) of action – needed for the subsequent risk assessment (see Section R.7.7.12.2) – quite often depends critically on available testing and/or non-testing information.

R.7.7.11.3 Exposure considerations for carcinogenicity

Exposure considerations may lead to adaptation of the evaluation of available information, and / or of the testing strategy.

As indicated before, waiving of carcinogenicity studies may apply, e.g. when it can be demonstrated that the substance is only produced and used in closed systems, which among other reasons may be due to the presence of other substances for which strict exposure regimes are implemented or enforced. On the other hand, a carcinogenicity study may be required (by the Agency or proposed by the registrant) when the substance has a widespread dispersive use or there is evidence of frequent or long-term human exposure, and information on its carcinogenic properties cannot be obtained by others means (from available effect information). Preliminary toxicokinetic studies may be required first to identify the potential target tissues and exposure routes that would guide the design of appropriate studies for the chemical of concern.

In the former case, i.e. when the substance is produced and used in closed systems only, conclusions on safe use and handling can be verified by use of read-across to risk assessments of structurally related carcinogens or to the so-called Threshold of Toxicological Concern (TTC) concept (see Appendix R.7-1): this concept identifies a *de minimis* exposure value for all chemicals, including genotoxic carcinogens, below which there is no appreciable risk to human health for any chemical. If it can be demonstrated that exposures are below these values, there is good reason for not performing the required tests. Clearly, good quality exposure information is essential in all these cases.

R.7.7.11.4 Remaining uncertainty on carcinogenicity

As indicated in the previous sections, adequate human data for evaluating the carcinogenic properties of a chemical are most often not available, and alternative approaches have to be used.

As also indicated in the previous sections and the Section R.7.7.1, test systems for identifying genotoxic carcinogens are reasonably well developed and adequately cover this property. There is also agreement that animal carcinogens which act by a genotoxic mode of action may reasonably be regarded as human carcinogens unless there is convincing evidence that the mechanisms by which mutagenicity and carcinogenicity are induced in animals are not relevant to humans. Unclear, however, and herewith introducing some uncertainty, is the relationship between carcinogenic potency in animals and in humans.

There is, on the other hand, a shortage of sensitive and selective test systems to identify non-genotoxic carcinogens, apart from the carcinogenicity bioassay. In the absence of non-testing information on the carcinogenicity of structurally related chemicals, indications for possible carcinogenic properties may come from existing repeated dose toxicity data, or from *in vitro* cell transformation assays. However, whereas the former source of data will have a low sensitivity (*e.g.* in case of a 28-day study), there is a possibility that the latter may lead to an over-prediction of carcinogenic potential. Insufficient data are available to provide further general guidance in this regard.

Non-genotoxic carcinogens may be difficult to identify in the absence of animal carcinogenicity test data. However, it could be argued that current conservative (cautious) risk assessment methodology should cover the risk for carcinogenic effects *via* this mode of action as well: i.e. current risk assessments for many non-genotoxic carcinogens are based on NOAELs for precursor effects or target organ toxicity with the application of conservative assessment factors to address uncertainty. For example, see the risk assessment for coumarin (EFSA, 2004; Felter *et al*., 2006). Such a risk assessment is not performed, though, in case this substance is not classified as dangerous for any other properties.

Once identified as a non-genotoxic carcinogen (from testing or non-testing data) there may be uncertainty as to the human relevance of this observation, i.e. to the human relevance of the underlying mode of action. In the absence of specific data on this, observations in the animal are taken as relevant to humans. However, additional uncertainty will exist for the relationship between carcinogenic potency in animals and in humans; this uncertainty, though, will be addressed in the procedure for deriving human standards (ILSI RSI, 2003).

Finally, conventional assays of carcinogenicity in animals have been found to be insensitive for some well-established human carcinogenic substances (e.g. asbestos and arsenic compounds). These substances can be shown to be carcinogenic when the test conditions are modified, thus illustrating that there will always be a possibility that a chemical could pose a carcinogenic hazard in humans but be missed in conventional animal studies. This is also true for other toxicological endpoints and should be taken into account by risk managers, especially when making decisions about the acceptability of scenarios showing particularly high exposures to workers and/or consumers.

R.7.7.12 Conclusions on carcinogenicity

R.7.7.12.1 Concluding on suitability for Classification and Labelling

In order to conclude on an appropriate classification and labelling position with regard to carcinogenicity, the available data should be considered using the criteria and guidance associated with the (EU Directive 67/548/EEC)[128].

R.7.7.12.2 Concluding on suitability for Chemical Safety Assessment

Besides the identification of a chemical as a carcinogenic agent from either animal data, epidemiological data or both, dose response assessment is an essential further step in order to characterise carcinogenic risks for certain exposure conditions or scenarios. A critical element in this assessment is the identification of the mode of action underlying the observed tumour-formation, as already explained in Section R.7.7.11.1: i.e. whether this induction of tumours is thought to be *via* a genotoxic mechanism or not.

In regulatory work, it is generally assumed that in the absence of data to the contrary an effect-threshold cannot be identified for genotoxic carcinogens exhibiting direct interaction with DNA, i.e., it is not possible to define a *no-effect level* for carcinogenicity induced by such agents. However, in certain cases even for these compounds a threshold for carcinogenicity may be identified in the low-dose region: e.g. it has in certain cases been clearly demonstrated that an increase in tumours did not occur at exposures below those associated with local chronic cytotoxicity and regenerative hyperplasia. It is also recognised that for certain genotoxic carcinogens causing genetic alterations, a practical threshold may exist for the underlying genotoxic effect. For example, this has been shown to be the case for aneugens (agents that induce aneuploidy – the gain or loss of entire chromosomes to result in changes in chromosome number), or for chemicals that

[128] Directive 67/548/EEC will be repealed and replaced with the EU Regulation on classification, labelling and packaging of substances and mixtures, implementing the Globally Harmonized System (GHS).

cause indirect effects on DNA that are secondary to another effect (e.g., through oxidative stress that overwhelms natural antioxidant defence mechanisms).

Non-genotoxic carcinogens exert their effects through mechanisms that do not involve direct DNA-reactivity. It is generally assumed that these modes of actions are associated with threshold doses, and it may be possible to define no-effect levels for the underlying toxic effects of concern. There are many different modes of action thought to be involved in non-genotoxic carcinogenicity. Some appear to involve direct interaction with specific receptors (e.g. oestrogen receptors), whereas appear to be non-receptor mediated. Chronic cytotoxicity with subsequent regenerative cell proliferation is considered a mode of action by which tumour development can be induced: the induction of urinary bladder tumours in rats, for example, may, in certain cases, be due to persistent irritation/inflammation/erosion and regenerative hyperplasia of the urothelium following the formation of bladder stones which eventually results in tumour formation. Specific cellular effects, such as inhibition of intercellular communication, have also been proposed to facilitate the clonal growth of neoplastic/pre-neoplastic cells.

The identification of the mode of action of a carcinogen is based on a combination of results in genotoxicity tests (both *in vitro* and *in vivo*) and observations in animal experiments, e.g. site and type of tumour and parallel observations from pathological and microscopic analysis. Epidemiological data seldom contribute to this.

Once the mode of action of tumour-formation is identified as having a threshold or not, a dose descriptor has to be derived for the purpose of allowing to conclude on chemical safety assessment. For threshold mechanisms the No-Observed-Adverse-Effect-Level (NOAEL) or Lowest-Observed-Adverse-Effect-Level (LOAEL) (see general introduction for definition and derivation of these descriptors) for tumour-formation or for the underlying (toxic) effect should be established to allow the derivation of a so-called Derived-No-Effect-Level (DNEL) (Chapter R.8 of the *Guidance on IR&CSA*), that subsequently is used in the safety assessment to establish safe exposure levels.

If the mode of action of tumour formation is identified as non-thresholded, dose descriptors such as T25, BMD10 or BMDL10 (general introduction for definition and derivation these descriptors) are to be established, that allow the derivation of a so-called Derived-Minimal-Effect-Level (DMEL; for guidance see Section R.8.5), that subsequently is used in the safety assessment to establish exposure levels of minimal concern.

Though mainly derived from animal data, epidemiological data may also occasionally provide dose descriptors that allow derivation of a DNEL or DMEL, e.g. Relative Risk (RR) or Odds Ratio (OR).

Substance-specific data for carcinogenicity normally will be absent, especially for the lower tonnage level substances. As indicated in Section R.7.7.11.1, non-testing data (read-across, grouping and/or (Q)SAR) may occasionally be considered sufficient to conclude on this endpoint, i.e. for classification, but also for establishing the underlying mode of action and for estimating the carcinogenic potency. This may introduce some additional uncertainty, especially with respect to the dose descriptor value, the addressing of which requires expert judgement; it is noted that experience to date on this is extremely limited. Guidance on read-across and/or grouping, and the use of (Q)SAR is provided in Sections R.6.2 and R.6.1.respectively.

R.7.7.12.3 Information not adequate

A *Weight of Evidence* approach comparing available adequate information with the tonnage-tiered information requirements by REACH may result in the conclusion that the information/data requirements are not fulfilled. In order to proceed in further information gathering, the following testing strategy can be adopted.

R.7.7.13 Integrated Testing Strategy (ITS) for carcinogenicity

R.7.7.13.1 Objective / General principles

The objective of this strategy is to describe where required how carcinogenicity should be assessed for all substances subject to registration under REACH: i.e. to identify substances with carcinogenic properties, their associated underlying mode of action, and their potency. Guidance is provided especially for those substances lacking pre-existing epidemiological or toxicological data on carcinogenicity.

The strategy provides the rationale for deciding whether or not a standard animal carcinogenicity study or any other further testing is required. It is recognised that standard carcinogenicity tests take considerable time to conduct and report, are expensive, and involve the use of a large number of animals. Consequently, it is preferable that decisions about the potential carcinogenicity of substances under REACH be taken as frequently as possible without the conduct of such tests.

The strategy recognises that the available information will differ from substance to substance. This may include various different types of toxicity information for the substance in question and/or for its analogues/structurally related chemicals. Details about the use and human exposure potential of the substance will also be available. All this will have an impact on the need for further data acquisition. Proposals for conducting a carcinogenicity test should be made with regard to the potential risk to human health and with consideration of the actual or intended production and/or use pattern.

REACH only specifies a carcinogenicity test for substances at the Annex X tonnage level (≥1000 t/y; see Section R.7.7.9). However, REACH also requires that carcinogenic substances at all tonnage levels be identified as substances of high concern, taking into account information from all available relevant sources (see Section R.7.7.10).

At the tonnage levels below 1000 t/y, the main concern is for those chemicals that are genotoxic. Chemicals may cause cancer secondary to other forms of toxicity, but protection of human health against the underlying toxicity (e.g., as identified from a repeat-dose toxicity study) will also protect against cancer that is secondary to that toxicity. It is noted, though, that some of these non-genotoxic carcinogens, when not classified for any other property and not identified as such in (limited) repeated dose toxicity studies will go unidentified; this also regards the risks associated with human exposures.

Finally, the strategy recognises that the carcinogenic process is a complex multi-step process. Chemically-induced cancer may be induced by any number of different pathways or modes of action and this allows for a variety of different approaches to carcinogenicity assessment. Substances that have the potential to act as genotoxic carcinogens can be identified by *in vitro* and *in vivo* mutagenicity tests, as described in Section R.7.7.1. Carcinogens that act by non-genotoxic modes of action are more difficult to identify because comparable, well-validated, short-term tests for the potentially numerous modes of actions involved are generally not available, and those tests that are available are not required as part of the standard information requirements of REACH.

A flow chart of the strategy is presented in Figure R.7.7–2.

R.7.7.13.2 Preliminary considerations

As a starting point, there will be the information collected with respect to mutagenicity. If they are available, test and non-test data from a literature search and, if possible, from members of an applicable chemical category or (Q)SAR analysis should be taken into account.

For substances for which there is no concern for mutagenic activity, and no other toxicological indicators of concern for carcinogenicity (i.e. for the substance itself or for structurally-related substances), there is no need for further consideration of its carcinogenic potential. This applies equally to those substances at the Annex X tonnage level as to those at lower tonnage levels.

If, however, for non-genotoxic substances toxicological indicators of concern are available (e.g. hyperplastic or pre-neoplastic lesions in repeated dose toxicity studies of the substance itself and/or of closely related substances), they should be investigated further on a case-by-case basis. Any decision on further testing is dependent upon the type and strength of the indications for carcinogenicity, the potential mechanism of action and their relevance to humans, and the type and level of human exposure (see Section R.7.7.10.2).

If no conclusion can be drawn regarding the potential genotoxicity of the substance then, in general, it will be determined on a case-by-case basis when and how the carcinogenic potential should be explored further. Again, this will then depend on the type and strength of the indications for carcinogenicity, the potential mechanism(s) of action, and the type and level of human exposure.

At least for substances at the higher tonnage levels, subchronic and/or chronic studies may provide additional important information on possible carcinogenic effects. There may, for example, be indications of peroxisomal proliferation or of hyperplastic or pre-neoplastic responses, including dose-response characteristics. These should be investigated further on the already indicated case-by-case basis, depending on the type and strength of the indications for carcinogenicity, the potential mechanism of action and relevance to humans, and the type and level of human exposure.

It may be appropriate on occasions to propose other tests to be undertaken, e.g. to test a read-across option with available non-testing data. These could include short-term tests, such as those for *in vitro* cell transformation or cell proliferation, or medium-term tests, like genetically engineered (transgenic) or neonatal models. It may well be that data generated in this way supports this read-across to available non-testing data, and herewith provides sufficient confidence in a read-across derived estimate of the carcinogenic potency for the substance and also for the magnitude of the risks associated with experienced exposure levels. The data generated may also weaken or even disprove the basis for read-across. It is noted that experience to date on this is very limited (as indicated in Section R.7.7.11.1). Guidance on read-across and/or grouping is provided in Section R.6.2 in Chapter R.6 of the *Guidance on IR&CSA*.

As validated testing procedures are not yet available and published in the OECD test guideline programme, it is essential that appropriate expert advice is sought regarding the application and suitability of any of these other tests.

Substances for which concern for carcinogenicity is solely based on positive genotoxicity data will, in a first step, be evaluated according to the approach outlined for identification of the genotoxicity hazard (see Section R.7.7.5).

Formally, for a substance classified as a category 1 or 2 mutagen, a carcinogenicity study will not normally be required (see Section R.7.7.9); *i.e.* it will be regarded as a genotoxic carcinogen. In order to allow an assessment of the magnitude of potential cancer risks associated with the prevailing human exposures, it may well be that available non-testing data (read-across, grouping, (Q)SAR) provide a sufficiently helpful estimate of the carcinogenic potency of the substance (i.e. by read-across) from which risks can be assessed. Guidance on read-across and/or grouping, and the use of (Q)SAR is provided in Sections R.6.2 and R.6.1, respectively.

In case such an approach is not possible, an estimate of acceptable exposure conditions may alternatively be obtained by use of the available data from animal toxicity studies: i.e. by identifying the minimal toxic dose in sub-chronic studies (if available, as some surrogate value for the dose descriptor) and by applying a large assessment factor; see for further guidance Gold *et al*. (2003). It is stressed that expert judgement is definitively needed here.

On very rare occasions, a case may be made to perform a carcinogenicity study in animals for substances that have been classified for mutagenicity in categories 1 or 2. Such a case would have to explain why the study was critically important; e.g. in the context of the clarification of carcinogenic risk associated with human exposures.

For substances classified as category 3 mutagens, and for which there is no carcinogenicity study, there should first be an evaluation of whether classification in category 2 for mutagenicity is possible. If such a classification is made, then the approach described above can be followed with regards to carcinogenicity. Occasionally, it may be established that classification as a category 2 mutagen is not appropriate. In such instances, it should not be assumed automatically that the substance has carcinogenic potential. However, unless there is clear evidence to indicate the contrary, it is expected that these substances will be regarded as genotoxic carcinogens.

As the previous paragraph implies, mutagenic potential *in vivo* is not always a reliable indicator of carcinogenic potential. If repeated dose toxicity studies indicate that pre-neoplastic changes (e.g. hyperplasia, precancerous lesions) occur, then the probability that carcinogenic activity will be expressed is increased. Non-testing data such as read-across and (Q)SAR may also contribute to this evaluation.

For substances at the REACH Annex X tonnage level, the need for or waiving of a standard animal test should be clearly explained, taking into account all the available toxicological and hygiene information on the substance and/or other relevant substances. For example, if it can be demonstrated that the substance is used only in a closed system and that human exposures are negligible, it is possible to propose no further testing for carcinogenicity.

It is recommended that when a carcinogenicity bioassay is required, study design and test protocol are well considered prior to delivering the test-proposal (e.g. OECD TG 453). Particular consideration, based on all the available data, should be given to the selection of the species and strain to be used in the carcinogenicity test, the route of exposure and dose level selection. It is also recommended that when a carcinogenicity test is to be conducted, an investigation of chronic toxicity should, whenever possible, form part of the study protocol. Finally, the limited value of a mouse assay as second species should be considered in this (Doe *et al.*, 2006).

The approaches outlined below may be used in the assessment of the potential carcinogenic risk of a substance to humans, and to help decide whether or not a carcinogenicity test will be required and, if so, when.

R.7.7.13.3 Testing strategy for carcinogenicity

As for other endpoints, the following three steps apply for the assessment of carcinogenicity (i.e. the hazard, underlying mode of action, and potency) for substances at each of the tonnage levels specified in Annexes VII to X of REACH.

i. Gather and assess all available test and non-test data from read-across/proper chemical category and suitable predictive models. Examine the Weight of Evidence that relates to carcinogenicity.

ii. Consider whether the standard information requirements are met.

iii. Ensure that the information requirements of Annexes VII and VIII are met; make proposals to conform with Annexes IX and X.

Further details about the procedures to follow at each of the different tonnage levels are described below.

Substances at Annexes VII, VIII and IX

A definitive assessment of carcinogenicity is usually not possible from the data available at the Annex VII, VIII and IX tonnage levels. However, for all substances, any relevant test data that are already available, together with information from predictive techniques such as read-across or chemical grouping, should be used to form a judgement about this important hazard endpoint.

The minimum information to be provided at the Annex VII, VIII and IX tonnage levels in relation to this endpoint is equivalent to that required for the mutagenicity endpoint (see Section R.7.7.2): positive results from *in vitro* mutagenicity studies provide an alert for possible carcinogenicity, and need confirmation *via* further testing *in vitro* and/or *in vivo* mutagenicity testing. As such, this will not lead to classification of a substance as a carcinogen, but this evidence should be taken into account in risk assessment: substances shown to be *in vivo* mutagens should be assumed to be potentially carcinogenic.

Furthermore, the results of repeated dose toxicity studies and /or reproductive/ developmental toxicity tests may be informative about a possible carcinogenic potential: hyperplasia or other pre-neoplastic effects may be observed in these studies. These observations may also be informative on potential mode(s) of action underlying the carcinogenic effect.

Although the criteria for carcinogenicity classification may not be met in the absence of substance-specific carcinogenicity data, the evidence from the available information alerting to possible carcinogenicity should be taken into account in the risk assessment for this endpoint: ways that allow an assessment of the magnitude of potential cancer risks associated with human exposures without performing the assay are indicated in indicated in Section R.7.7.13.2 (see Section for derivation of DMEL and DNEL values in Chapter R.8 of the *Guidance on IR&CSA*).

It is important to note that at the tonnage levels below 1000 t/y, the main concern is for those chemicals that are genotoxic. The repeated dose toxicity studies mentioned above may indicate cancers which are secondary to other forms of toxicity. For those the protection of human health against the underlying toxicity will also protect against cancer that is secondary to the toxicity. It is noted, though, that some of these non-genotoxic carcinogens, when not classified for any other property and not identified as such in (limited) repeated dose toxicity studies will go unidentified; this also regards the risks associated with human exposures.

Substances at Annex X

All substances at this tonnage should be evaluated for carcinogenicity.

All relevant data from all toxicity studies should be assessed to see whether a sufficiently reliable assessment about the carcinogenicity of the substance is possible, including alternative means, if needed: i.e. predictive techniques such as chemical grouping and read-across, and the use of (Q)SARs. On some occasions, it may be proposed to supplement these predictive approaches with *in vitro* or alternative shorter-term *in vivo* investigations in order to circumvent the need for a carcinogenicity study. This should usually be in the context of adding to the *Weight of Evidence* that a substance may be carcinogenic.

Formally, if the substance is classified as a category 1 or 2 mutagen (GHS category 1), a carcinogenicity study will not normally be required. For a substance classified as a category 3 mutagen (GHS category 2) it should first be established whether a case should be made for a higher level of classification.

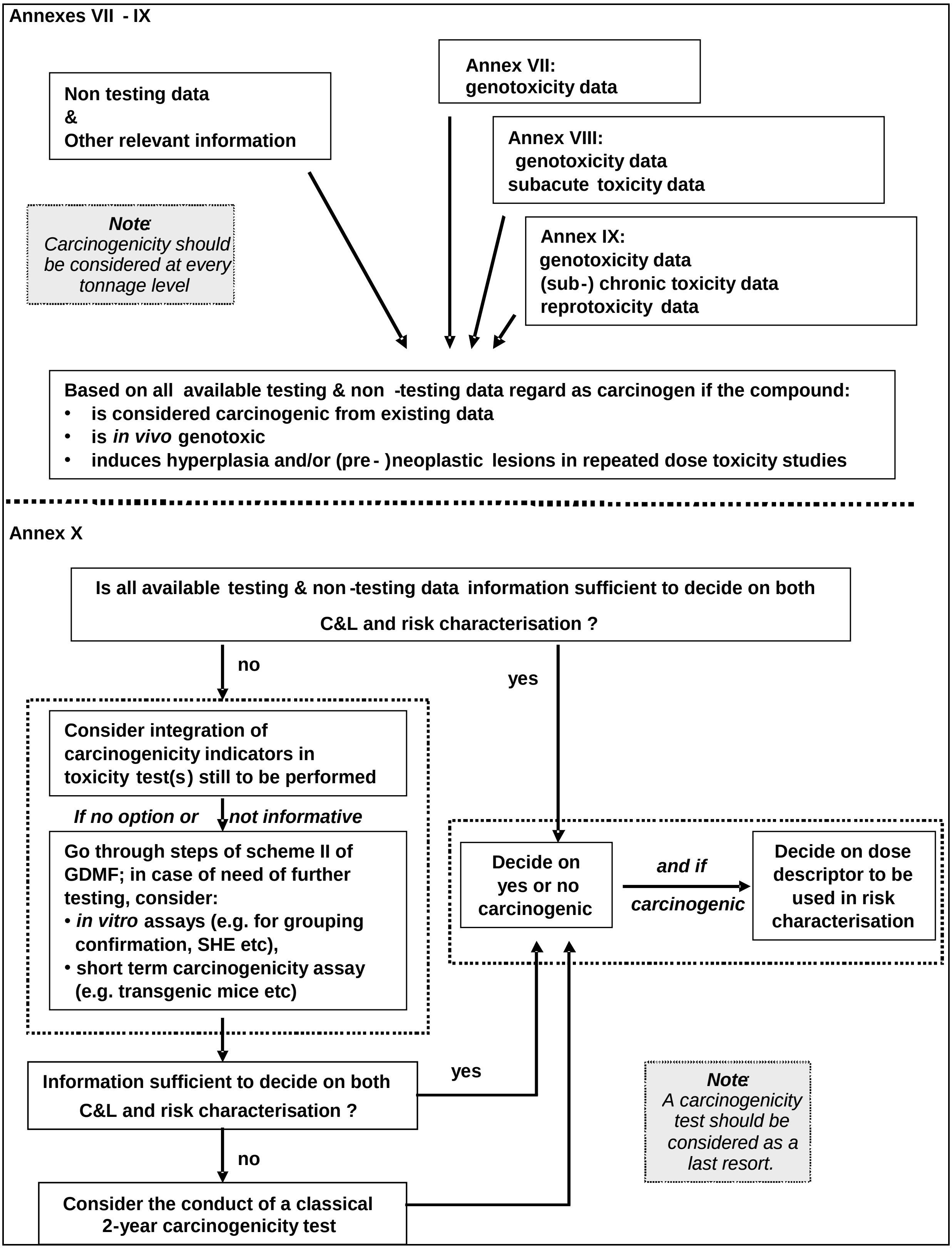
For risk assessment, all the substances are then regarded as genotoxic carcinogens unless there is scientific evidence to the contrary. Ways that allow an assessment of the magnitude of potential cancer risks associated with human exposures without performing the assay are indicated in Section R.7.7.13.2. (see Chapter R.8 of the *Guidance on IR&CSA* for derivation of DMEL and DNEL values).

A carcinogenicity study may, on occasion, be justified. If there are clear suspicions that the substance may be carcinogenic, and available information (from both testing and non-testing data) are not conclusive in this, both in terms of hazard and potency, then the need for a carcinogenicity study should be explored. In particular, such a study may be required for substances with a widespread, dispersive use or for substances producing frequent or long-term human exposures. However, it should be considered only as a last resort.

It is noted, though, that some of non-genotoxic carcinogens, i.e. when not classified for any other property and not identified as potential carcinogens in (limited) repeated dose toxicity studies will go unidentified; this also regards the risks associated with human exposures.

If, in any case there is a need for further testing, the registrant must prepare and submit a well-considered test proposal (see Section R.7.7.6.2), and a time schedule for fulfilling the information requirements.

Figure R.7.7–2 Integrated Testing Strategy for carcinogenicity



R.7.7.14 References on carcinogenicity

Albertini R, Bird M, Doerrer N, Needham L, Robison S, Sheldon L and Zenick H (2006) The use of biomonitoring data in exposure and human health risk assessments. Environ Health Perspect 114:1755–62.

Ames BN and Gold LS (1990) Chemical carcinogenesis: too many rodent carcinogens. Proc Natl Acad Sci USA 87: 7772-6.

Ashby J, Paton D (1993) The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures. Mutation Res 286:3–74.

Ashby J and Tennant RW (1988) Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. Mutat Res 204:17-115.

Boffetta P, van der Hel O, Norppa H, Fabianova E, Fucic A, Gundy S, Lazutka J, Cebulska-Wasilewska A, Puskailerova D, Znaor A, Kelecsenyi Z, Kurtinaitis J, Forni A, Vermeulen R, and Bonassie S (2007) Chromosomal aberrations and cancer risk: Results of a cohort study from Central Europe. Am J Epidemiol 165:36–43.

Boobis AR, Cophen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D and Farland W (2006) IPCS framework for analysing the relevance of a cancer mode of action for humans. Crit Rev Toxico. 36:781-92.

Chen HJ and Chiu WL (2005) Association between cigarette smoking and urinary excretion of 1,N2-ethenoguanine measured by isotope dilution liquid chromatography-electrospray ionization/tandem mass spectrometry. Chem Res Toxicol 18:1593–9.

Cohen SM and Ellwein LB (1991) Genetic errors, cell proliferation and carcinogenesis. Cancer Res 51:6493–505.

Cohen SM, Purtilo DT and Ellwein LB (1991) Pivotal role of increased cell proliferation in human carcinogenesis. Mod Pathol 4:371–5.

Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment (2004) Guidance on a Strategy for the Risk Assessment of Chemical Carcinogens.

Cronin MTD, Jaworska JS, Walker JD, Comber MHI, Watts CD and Worth AP (2003) Use of quantitative structure-activity relationships in international decision-making frameworks to predict health effects of chemical substances. Environl Health Perspect 22:1391-1401.

Doe JE, Lewis RW and Botham PA (2006) Comments on a scientific and animal welfare assessment of the OECD health effects test guidelines for the safety testing of chemicals under the European Union REACH system. Altern Lab Anim 34:111–4.

ECETOC (2003) QSARs: Evaluation of the commercially available software for human health and environmental endpoints with respect to chemical management applications. ECETOC Technical Report No. 89.

EFSA (European Food Safety Authority) (2004) Opinion of the Scientific Panel on Food Additives, Flavouring, Processing Aids and Materials in Contact with Food (AFC) on a Request from the Commission Related to Coumarin. Question Number EFSA-Q-2003-118. Adopted on 6 October 2004. The EFSA Journal 104, 1–36. Available from: http://www.efsa.europa.eu/en/efsajournal/pub/104.htm.

Enzmann H, Bomhard E, Iatropoulos M, Ahr HJ, Schleuter G and Williams GM (1998) Short- and intermediate-term carinogencitic testing – a review. Food Chem Toxicol 36:979–95.

Elcombe CR, Odum J, Foster JR, Stone S, Hasmall S, Soames AR, Kimber I and Ashby J (2002) Prediction of rodent non-genotoxic carcinogenesis: evaluation of biochemical and tissue changes in rodents following exposure to nine non-genotoxic NTP carcinogens. Env Health Perspect 110:363-70.

Felter SP, Vassallo JD, Carlton BD and Daston GP (2006) A safety assessment of coumarin taking into account species-specificity of toxicokinetics. Food Chem Toxicol 44:462–75.

Franke R, Gruska A, Giuliani A and Benigni R (2001) Prediction of rodent carcinogenicity of aromatic amines: a quantitative structure-activity relationships model. Carcinogenesis 22:1561-71.

Gold LS, Gaylor DW and Slone TH (2003) Comparison of cancer risk estimates based on a variety of risk assessment methodologies. Regul Toxicol Pharmacol 37:45-53.

Haseman JK (1983) A re-examination of false-positive rates for carcinogenesis studies. Fundam Appl Toxicol 3:334-9.

Hill AB (1965) The environment and disease: Association or causation? Proc R Soc Med 58:295-300.

IARC (International Agency for Research on Cancer) (2006). IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Preamble. Lyon. Available online: http://monographs.iarc.fr/ENG/Preamble/CurrentPreamble.pdf

IARC (International Agency for Research on Cancer) (1999). Species Differences in Thyroid, Kidney and Urinary Bladder Carcinogenesis. Capen CC, Dybing E, Rice JM & Wilbourn JD (Eds). IARC Scientific Publications 147, Lyon.

IARC (International Agency for Research on Cancer) (2006). Monograph 87: Inorganic and Organic Lead Compounds. IARC, Lyon.

ILSI/HESI (2001) Alternatives to carcinogenicity testing project. Toxicol Pathol 29(Suppl):1-351.

ILSI RSI (2003) A framework for human relevance analysis of information on carcinogenic modes of action. Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J and Patton DE. Crit Rev Toxicol 33:591-653.

Institute for Environment and Health (2002) Assessment of Chemical Carcinogens: Background to general principles of a weight of evidence approach, ISBN 1899110 37 2.

Jacobson-Kram D, Sistare FD, Jacobs AC (2004) Use of transgenic mice in carcinogenicity hazard assessment. Toxicol Pathol 32(Suppl 1):49-52.

Matthews EJ, Contrera JF (1998) A New Highly Specific Method for Predicting the Carcinogenic Potential of Pharmaceuticals in Rodents Using Enhanced MCASE QSAR-ES Software. Regul Toxicol Pharmacol 28:242-64.

Munro IC, Ford RA, Kennepohl E and Sprenger JG (1996) Thresholds of toxicological concern based on structure–activity relationships. Drug Metab Rev 28:209–17.

OECD (2002) Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies, OECD Environment, Health and Safety Publications, Series on Testing and Assessment No. 35 and Series on Pesticides No.14.

OECD (2006) draft OECD review paper on cell transformation assays for detection of cehmical carcinogens. OECD Environment, Health and Safety Publications, Series on Testing and Assessment No.31, Paris (France).

Pritchard JB, French JE, Davis BJ, Haseman JK (2003) The role of transgenic mouse models in carcinogen identification. Environ Health Perspect 111:444-54.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, Grant D, Hartley M, Knaap A, Kroese D, Mangelsdorf I, Meek E, Rice JM and Younes M (2001) IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. Reg Toxicol Pharmacol 34:146–52.

Tennant RW, Stasiewicz S, Mennear J, French JE and Spalding JW (1999) Genetically altered mouse models for identifying carcinogens. *In:* McGregor DB, Rice JM, Venitt S (Eds.). The use of short- and medium-term tests for carcinogens and data on genetic effects in carcinogenic hazard evaluation. Lyon, France: International Agency for Research on Cancer.

US FDA (1986) General principles for evaluating the safety of compounds used in foodproducing animals. Appendix 1. Carcinogen structure guide. Washington: U.S.Food and Drug Administration.

EUROPEAN CHEMICALS AGENCY
ANNANKATU 18, P.O. BOX 400,
FI-00121 HELSINKI, FINLAND
ECHA.EUROPA.EU