



**Committee of Experts on the Transport of Dangerous Goods
and on the Globally Harmonized System of Classification
and Labelling of Chemicals**

**Report of the Committee of Experts on the Transport of
Dangerous Goods and on the Globally Harmonized System of
Classification and Labelling of Chemicals on its eleventh
session**

held in Geneva on 9 December 2022

Addendum

Annex III

**Amendments to the ninth revised edition of the Globally Harmonized
System of Classification and Labelling of Chemicals (GHS)
(ST/SG/AC.10/30/Rev.9)**

Chapter 1.2

Insert the following definitions in the alphabetical order:

“Defined approach means an approach to testing and assessment that consists of a fixed data interpretation procedure used to interpret data generated with a defined set of information sources, that can either be used on its own, or together with other information sources within an overall weight of evidence, to satisfy a specific regulatory need;”

IATA means “Integrated Approach on Testing and Assessment”;

Chapter 1.3

1.3.2.4.8 Replace “determinations” with “assessments” in the last sentence.

1.3.2.4.9.2 Replace “determination” with “assessment” in the last sentence.

1.3.2.4.9.5 Replace “determination” with “assessment” in the first sentence.

Chapter 2.1

2.1.1.1 Amend the beginning of the definition of “pyrotechnic substance or mixture” to read: “*A pyrotechnic substance or mixture* is an explosive substance or mixture that is designed to produce an effect by heat...”.

Add the following new definition:

“Explosive or pyrotechnic effect in the context of 2.1.1.2.1 (c) means an effect produced by self-sustaining exothermic chemical reactions including shock, blast, fragmentation, projection, heat, light, sound, gas and smoke.”.

2.1.3 Add the following sentence at the end of the paragraph: “Table 2.1.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.2

2.2.3.1 Add the following sentence at the end of the paragraph: “Table 2.2.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.3

2.3.1.3 Add the following sentence at the end of the paragraph: “Table 2.3.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

2.3.2.3 Add the following sentence at the end of the paragraph: “Table 2.3.4 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.4

2.4.3 Add the following sentence at the end of the paragraph: “Table 2.4.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.5

- 2.5.3 Add the following sentence at the end of the paragraph: “Table 2.5.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.6

- 2.6.3 Add the following sentence at the end of the paragraph: “Table 2.6.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

- 2.6.4.2.4 In the first and second sentences replace “shall” with “should”.

Replace the last sentence with the following:

“Open-cup tests are acceptable for liquids which cannot be tested in closed-cup test methods (e.g. due to their viscosity) or when open-cup test data is already available. In these cases, 5.6°C should be subtracted from the measured value because open-cup test methods generally result in higher values than closed-cup methods.”

Chapter 2.7

- 2.7.1 Add the following definition under the definition of readily combustible solids: “*Metal powders* are powders of metals or metal alloys.”.

- 2.7.2.2 Replace “Powders of metals or metal alloys” with “Metal powders”.

- 2.7.3 Add the following sentence at the end of the paragraph: “Table 2.7.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.8

- 2.8.3 Add the following sentence at the end of the paragraph: “Table 2.8.1 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.9

- 2.9.3 Add the following sentence at the end of the paragraph: “Table 2.9.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.10

- 2.10.3 Add the following sentence at the end of the paragraph: “Table 2.10.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.11

- 2.11.3 Add the following sentence at the end of the paragraph: “Table 2.11.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.12

- 2.12.3 Add the following sentence at the end of the paragraph: “Table 2.12.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.13

- 2.13.3 Add the following sentence at the end of the paragraph: “Table 2.13.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.14

- 2.14.3 Add the following sentence at the end of the paragraph: “Table 2.14.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.15

- 2.15.3 Add the following sentence at the end of the paragraph: “Table 2.15.1 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.16

- 2.16.3 Add the following sentence at the end of the paragraph: “Table 2.16.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 2.17

- 2.17.1.1 Amend to read as follows:

“2.17.1.1 Desensitized explosives are substances and mixtures in the scope of chapter 2.1 which are phlegmatized to suppress their explosive properties in such a manner that they meet the criteria as specified in 2.17.2 and thus may be exempted from the hazard class “Explosives” (chapter 2.1; see paragraph 2.1.1.2.2).”.

- 2.17.2 Replace with the following:

“2.17.2 Classification criteria

2.17.2.1 A phlegmatized explosive should be considered for inclusion in this class if, in that state, the exothermic decomposition energy is ≥ 300 J/g.

NOTE 1: The exothermic decomposition energy may be estimated using a suitable calorimetric technique (see section 20, subsection 20.3.3.3 in Part II of the Manual of Tests and Criteria).

NOTE 2: Substances and mixtures with an exothermic decomposition energy < 300 J/g should be considered for inclusion in other physical hazard classes (e.g. as flammable liquids or flammable solids).

2.17.2.2 A phlegmatized explosive should be considered for inclusion in this class if, in that state, it meets the following criteria:

- (a) It is not intended to produce a practical explosive or pyrotechnic effect; and
- (b) it is phlegmatized to an extent that:
 - (i) it has no mass explosion hazard in accordance with test 6 (a) or 6 (b) of the *Manual of Tests and Criteria*; and
 - (ii) it is not too sensitive or thermally unstable in accordance with test series 3 of the *Manual of Tests and Criteria*;
 or that
 - (iii) it is too insensitive for inclusion into in the class of explosives in accordance with test series 2 of the *Manual of Tests and Criteria*; and
- (c) it presents no mass explosion hazard and has a corrected burning rate ≤ 1200 kg/min in accordance with the burning rate test of subsection 51.4 of the *Manual of Tests and Criteria*.

NOTE: *Phlegmatized explosives which do not meet the criteria of 2.17.2.2 should be classified as explosives (see chapter 2.1).*

2.17.2.3 In addition to the criteria in 2.17.2.1 and 2.17.2.2, nitrocellulose should be stable in accordance with appendix 10 of the *Manual of Tests and Criteria* in order to be used in nitrocellulose mixtures considered for this class.

NOTE: *Nitrocellulose mixtures containing no explosives other than nitrocellulose, do not need to meet the criterion of 2.17.2.2 (b) (ii).*

2.17.2.4 (new, former 2.17.2.2) Current paragraph 2.17.2.2 becomes new paragraph 2.17.2.4. Replace “using the test “burning rate test (external fire)”” with “determined using the burning rate (external fire) test”.

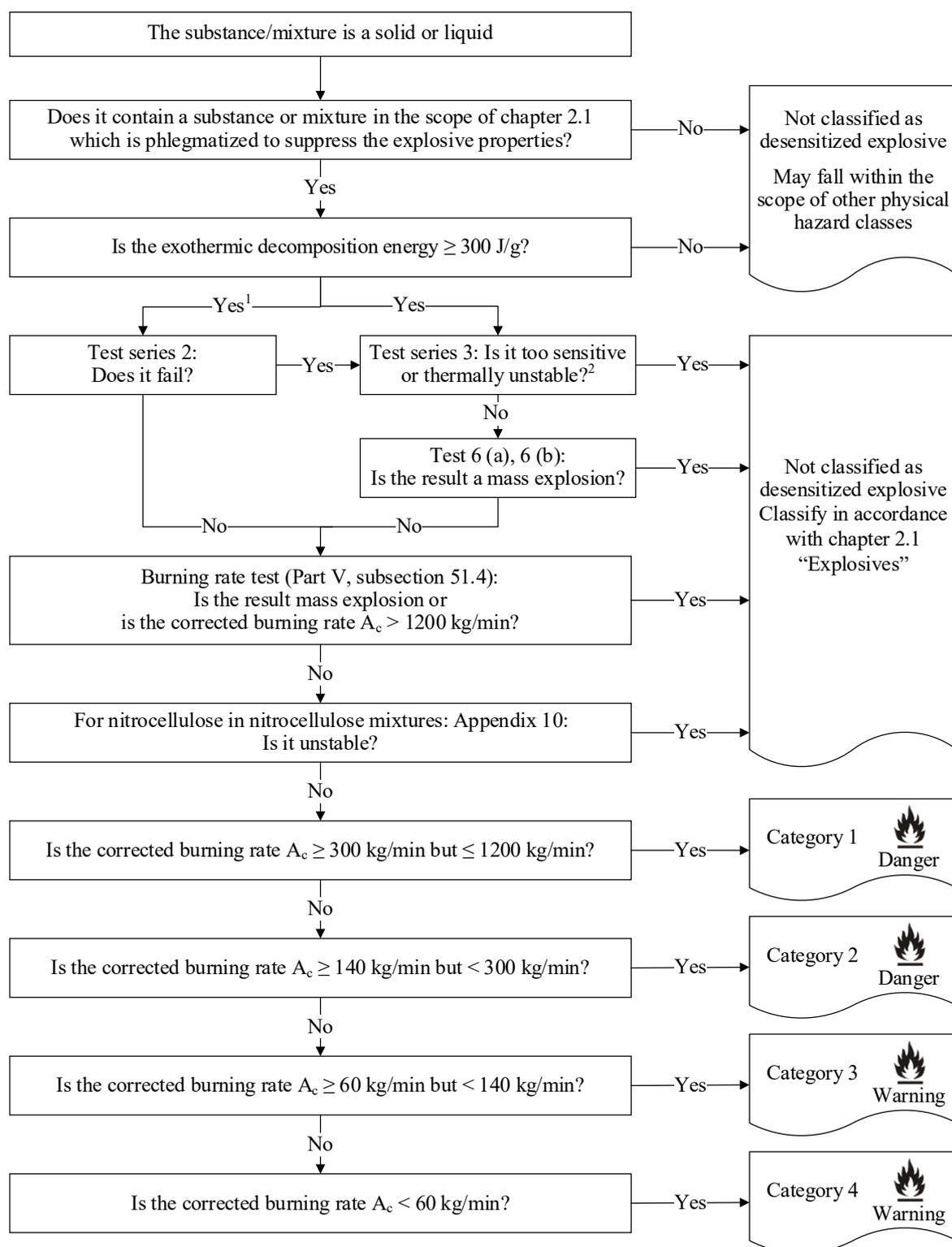
2.17.3 Add the following sentence at the end of the paragraph: “Table 2.17.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Current paragraph 2.17.4 remains unchanged.

2.17.4.1 Amend to read as follows:

“2.17.4.1 Decision logic

To classify desensitized explosives, data for the sensitivity, thermal stability, explosive potential and the corrected burning rate should be determined as described in Part I and Part V of the *Manual of Tests and Criteria*. Where a mixture contains nitrocellulose, additional data for the stability of the nitrocellulose as described in appendix 10 of the *Manual of Tests and Criteria* are needed in order to be used in nitrocellulose mixtures considered for this class. Classification is according to decision logic 2.17.1.”

Decision logic 2.17.1 for desensitized explosives

Insert the following footnotes:

¹ *Test series 2 is optional. The alternative route (via test 6 (a) and (b) and test series 3) may be taken directly without performing test series 2.*

² *Test series 3 is not applicable to nitrocellulose mixtures containing no explosives other than nitrocellulose.*

Current paragraph 2.17.4.2 remains unchanged.

Chapter 3.1

- 3.1.2.3 Replace “approach” with “assessment” in the last sentence.
- 3.1.2.6.1 Add the following sentence at the end of the existing paragraph:
 “Guidance on the conversion of experimental values for times other than a 1-hour exposure is provided in 3.1.5.3.”
- 3.1.3.5.6 Replace “assigned” with “classified in” in the last sentence.
- 3.1.3.5.7 At the beginning of the sentence, replace “aerosol” with “aerosolized”.
- 3.1.4.1 Amend the last sentence to read as follows: “Table 3.1.3 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”
- 3.1.5.3 Add a new section 3.1.5.3 to read as follows:

“3.1.5.3 *Guidance*

3.1.5.3.1 The ATE values used for inhalation toxicity classification in table 3.1.1 are based on a 4-hour experimental exposure in laboratory animals (3.1.2.6.1). Existing inhalation LC₅₀ values obtained in studies using exposure times other than 1 hour (3.1.2.6.1) can be adjusted to a 4-hour exposure using the ten Berge equation ($C^n \times t = k$) for gases and vapours and Haber’s rule ($C \times t = k$) for dusts and mists, as follows:

Formula for gases and vapours

$$LC_{50}(4 \text{ hours}) = \left(\frac{C^n \times t}{4} \right)^{1/n}$$

where:

- C = LC₅₀ concentration for exposure duration t
 n = chemical-specific exponent
 t = exposure duration, in hours, for C

Formula for dusts and mists

$$LC_{50}(4 \text{ hours}) = \frac{C \times t}{4}$$

where:

- C = LC₅₀ concentration for exposure duration t
 t = exposure duration, in hours, for C

3.1.5.3.2 A default value of 2 is used for n unless additional conclusive information is available to indicate that a different value is more appropriate. The accepted exposure times for conversion are from 30 minutes to 8-hour exposures. A competent authority may decide whether other exposure times are acceptable for conversion. Data from a long-term exposure should not be converted because this hazard class addresses acute toxicity. Guidance on the duration of short-term (i.e., acute) inhalation toxicity exposures can be found in OECD Guidance Document 39 (section 4.1: Outline of the exposure methodology).

Examples: classification using calculated 4-hour LC₅₀ values

Example 1: Substance (liquid)

- For the purpose of this example the substance has an experimental 6-hour vapour LC₅₀ = 13.6 mg/l
- No additional information on n is available so the default value (n = 2) will be used.

Criterion:

$$LC_{50}(4 \text{ hours}) = \left(\frac{C^n \times t}{4} \right)^{1/n}$$

Calculation

$$LC_{50}(4 \text{ hours}) = \left(\frac{C^n \times t}{4} \right)^{\frac{1}{n}} = \left(\frac{13.6^2 \times 6}{4} \right)^{\frac{1}{2}} = \mathbf{16.7}$$

3. Therefore, the substance is classified into Category 4 based on the vapours Category 4 criteria ($10.0 < ATE \leq 20.0$) from table 3.1.1.

Example 2: Substance (solid)

4. For this example, the substance has an experimental 2-hour dust $LC_{50} = 0.26$ mg/l

Criterion:

$$LC_{50}(4 \text{ hours}) = \frac{C \times t}{4}$$

Calculation

$$LC_{50}(4 \text{ hours}) = \frac{C \times t}{4} = \frac{0.26 \times 2}{4} = \mathbf{0.13}$$

5. Therefore, the substance is classified into Category 2 based on the dusts and mists Category 2 criteria ($0.05 < ATE \leq 0.5$) from table 3.1.1.”.

Chapter 3.2

- 3.2.1.2 Replace the second sentence with the following:

“Classification should be based on mutually acceptable data generated using methods that are validated according to international procedures. These include both OECD guidelines and equivalent methods (see 1.3.2.4.3).”.

In the last sentence replace “3.2.2.6” with “3.2.2.7”.

- 3.2.1.3 In the first sentence replace “3.2.2.7” with “3.2.2.8”.

In the last sentence replace “3.2.2.7.3” with “3.2.2.8.3” and “weight of evidence approach” with “weight of evidence assessment”. Insert “, 3.2.2.7” after “1.3.2.4.9” in the references between brackets at the end of the paragraph.

- 3.2.2.1 In the heading, add “(tier 1 in figure 3.2.1)” at the end.

- 3.2.2.2 In the heading, delete “test” and add “(tier 1 in figure 3.2.1)” at the end.

Amend the beginning of the first sentence to read: “OECD Test Guideline 404 is the currently available and internationally accepted animal test method...”.

- 3.2.2.3 In the heading, add “(tier 2 in figure 3.2.1)” at the end.

- 3.2.2.3.2 Replace the first sentence (“Wherever possible ... to be applied”) with the following text:

“The classification criteria for the currently available in vitro/ex vivo test methods adopted by the OECD in test guidelines 430, 431, 435, and 439 are described in tables 3.2.6 and 3.2.7 (see 3.2.5.3.4). Other validated in vitro/ex vivo test methods accepted by some competent authorities may also be considered. A competent authority may decide which classification criteria, if any, should be applied for other test methods to conclude on classification, including that a substance is not classified for effects on the skin.”.

3.2.2.3.3 (new) Place the two last sentences of current paragraph 3.2.2.3.2 (“In vitro/ex vivo...into consideration”) under a new paragraph 3.2.2.3.3 and replace “test methods used” with “test method(s) used”.

Current paragraphs 3.2.2.2.3 to 3.2.2.3.4.2 become 3.2.2.3.4 to 3.2.2.3.5.2.

3.2.2.3.4.1 (new, former 3.2.2.3.3.1) Add “(see 3.2.5.3.4)” at the end of the paragraph after “table 3.2.6”.

3.2.2.3.5.1 (new, former 3.2.2.3.4.1) Add “(see 3.2.5.3.4)” at the end of the paragraph after “table 3.2.7”.

3.2.2.3.5.2(new, former 3.2.2.3.4.2) Delete the last sentence (“In this situation...no classification.”).

3.2.2.3.6 (new, former 3.2.2.3.4.3) Place current paragraph 3.2.2.3.4.3 (“Where competent authorities do not adopt category 3 ... as not classified for skin irritation” under a new heading 3.2.2.3.6 and amend to read as follows:

“3.2.2.3.6 *No classification for effect on the skin*

Where competent authorities do not adopt Category 3, a negative result in an in vitro/ex vivo test method for skin irritation that is validated according to international procedures, e.g. OECD Test Guideline 439, can be used to conclude as not classified for skin irritation. Where competent authorities adopt Category 3, additional information is required to differentiate between Category 3 and no classification.”.

3.2.2.4 Amend the heading to read as follows:

“3.2.2.4 *Classification based on other existing animal skin data (tier 3 in figure 3.2.1)*”

3.2.2.5 Amend to read as follows:

“3.2.2.5 *Classification based on extreme pH (pH ≤ 2 or ≥ 11.5) and acid/alkaline reserve (tier 4 in figure 3.2.1)*

In general, substances with an extreme pH (pH ≤ 2 or ≥ 11.5) are expected to cause significant skin effects, especially when associated with significant acid/alkaline reserve. A substance with pH ≤ 2 or ≥ 11.5 is therefore considered to cause skin corrosion (Category 1) in this tier if it has a significant acid/alkaline reserve or if no data for acid/alkaline reserve are available. However, if consideration of acid/alkaline reserve suggests the substance may not be corrosive despite the extreme pH value, the result is considered inconclusive within this tier (see figure 3.2.1). A pH > 2 and < 11.5 is considered inconclusive and cannot be used for classification purposes. Acid/alkaline reserve and pH can be determined by different methods including those described in OECD Test Guideline 122 and Young et al. (1988), acknowledging that there are some differences between these methods (see 3.2.5.3.6). A competent authority may decide which criteria for significant acid/alkaline reserve can be applied.”

3.2.2.6 In the heading, add “*(tier 5 in figure 3.2.1)*” at the end.

3.2.2.6.1 In the last sentence, replace “Such methods” with “Non-test methods” and “(structural alerts, SAR); quantitative structure-activity relationships (QSARs); computer experts systems; and” with “(structural alerts, SAR) or quantitative structure-activity relationships (QSARs), computer experts systems, and”.

3.2.2.6.4 (new) Insert the following new paragraph:

“3.2.2.6.4 For conclusions on no classification from read-across and (Q)SARs the adequacy and robustness of the scientific reasoning and of the supporting evidence should be well substantiated and normally requires multiple negative substances with good structural and physical (related to toxicokinetics) similarity to the substance being classified, as well as a clear

absence of positive substances with good structural and physical similarity to the substance being classified.”.

3.2.2.7 (new) Insert a new section to read as follows:

“3.2.2.7 Classification based on an overall weight of evidence assessment (tier 6 in figure 3.2.1)”

3.2.2.7.1 An overall weight of evidence assessment using expert judgement is indicated where none of the previous tiers resulted in a definitive conclusion on classification. In some cases, where the classification decision was postponed until the overall weight of evidence, but no further data are available, a classification may still be possible.

3.2.2.7.2 A substance with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve (result considered inconclusive in tier 4; see 3.2.2.5) and for which no other information is available, should be classified as skin corrosion Category 1 in this tier. If inconclusive information is also available from other tiers but the overall weight of evidence assessment remains inconclusive, the extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) result should take precedence and the substance should be classified as skin corrosion Category 1 in this tier independently of its acid/alkaline reserve. For mixtures, the approach is different and is detailed in 3.2.3.1.3.”.

Current paragraphs 3.2.2.7 to 3.2.2.7.3 become 3.3.2.8 to 3.2.2.8.3.

3.2.2.8 (new, former 3.2.2.7) In the heading, add “**(figure 3.2.1)**” at the end.

3.2.2.8.1 (new, former 3.2.2.7.1) In the first sentence, delete “initial” and replace “elements” with “tiers as well as information within a tier”.

3.2.2.8.2 (new, former 3.2.2.7.2) Amend the first sentence to read as follows:

“In the tiered approach (figure 3.2.1), existing human and standard animal data form the highest tier, followed by in vitro/ex vivo data, other existing animal skin data, extreme pH and acid/alkaline reserve, and finally non-test methods.”.

In the second sentence, replace “weight of evidence approach” with “weight of evidence assessment”.

3.2.2.8.3 (new, former 3.2.2.7.3) Replace (twice) “weight of evidence approach” with “weight of evidence assessment”.

In the last sentence, replace “irritation” with “skin irritation” and add “are also available” at the end.

Figure 3.2.1 Amend as follows:

Text between tier 3 and tier 4 boxes: Replace “No data, conclusive for no classification, or inconclusive^b” with “No data, not classified for skin corrosion/irritation or inconclusive^b”.

Text between tier 4 and tier 5 boxes: Replace “data showing low/no acid/alkaline reserve” with “data showing non-significant acid/alkaline reserve”.

Text box for tier 6: replace “(see 3.2.2.7.3)” with “(see 3.2.2.7)”.

Exit box “Classification not possible”: amend the text to read: “Classification not possible for substances ^c”.

In the box on the right-hand side starting with “Assess consistency with lower tiers” replace “3.2.2.7.3” with “3.2.2.8.3”.

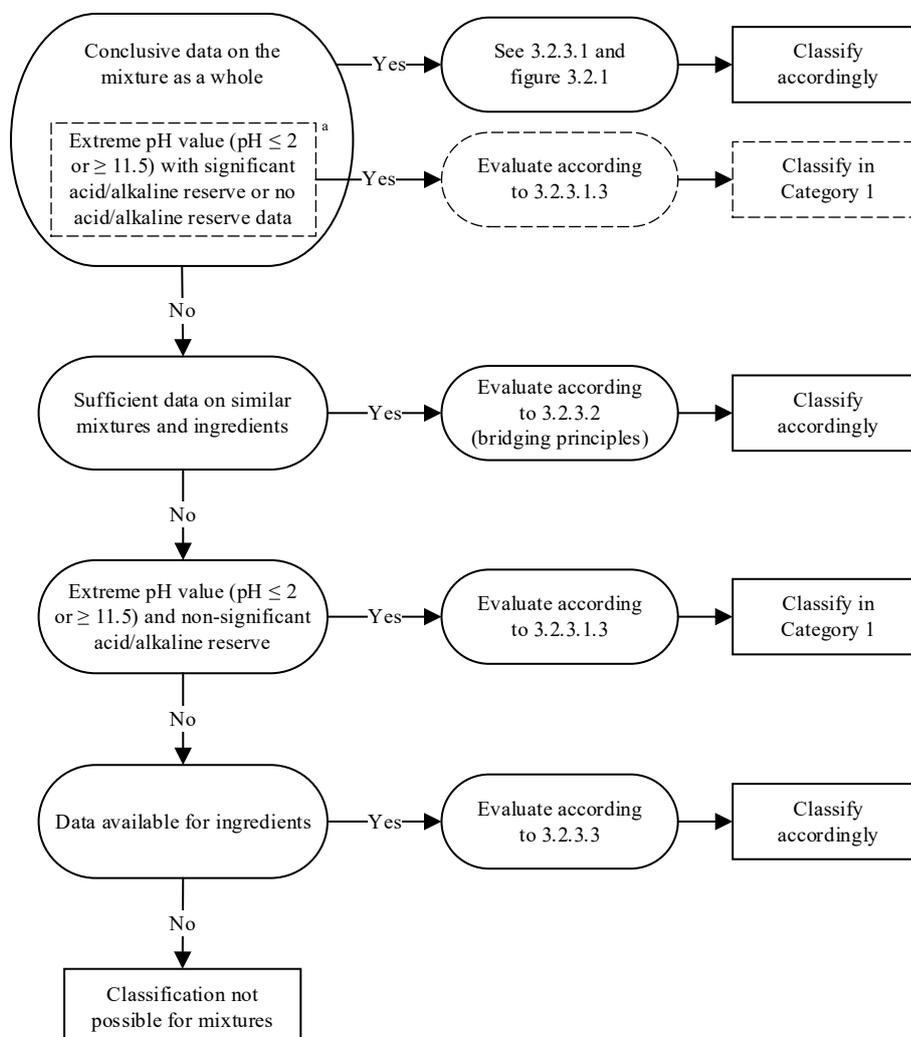
In note “a”, replace “3.2.2.7” with “3.2.2.8”.

Add a new note “c” to read as follows: “^c For mixtures, the flow chart in figure 3.2.2 should be followed.”.

3.2.3 Insert the following new text and figure under the current heading:

“The approach to classification for skin corrosion/irritation is tiered and is dependent upon the amount of information available for the mixture itself and for its ingredients. The flow chart of figure 3.2.2 below outlines the process to be followed.

Figure 3.2.2: Tiered approach to classification of mixtures for skin corrosion/irritation



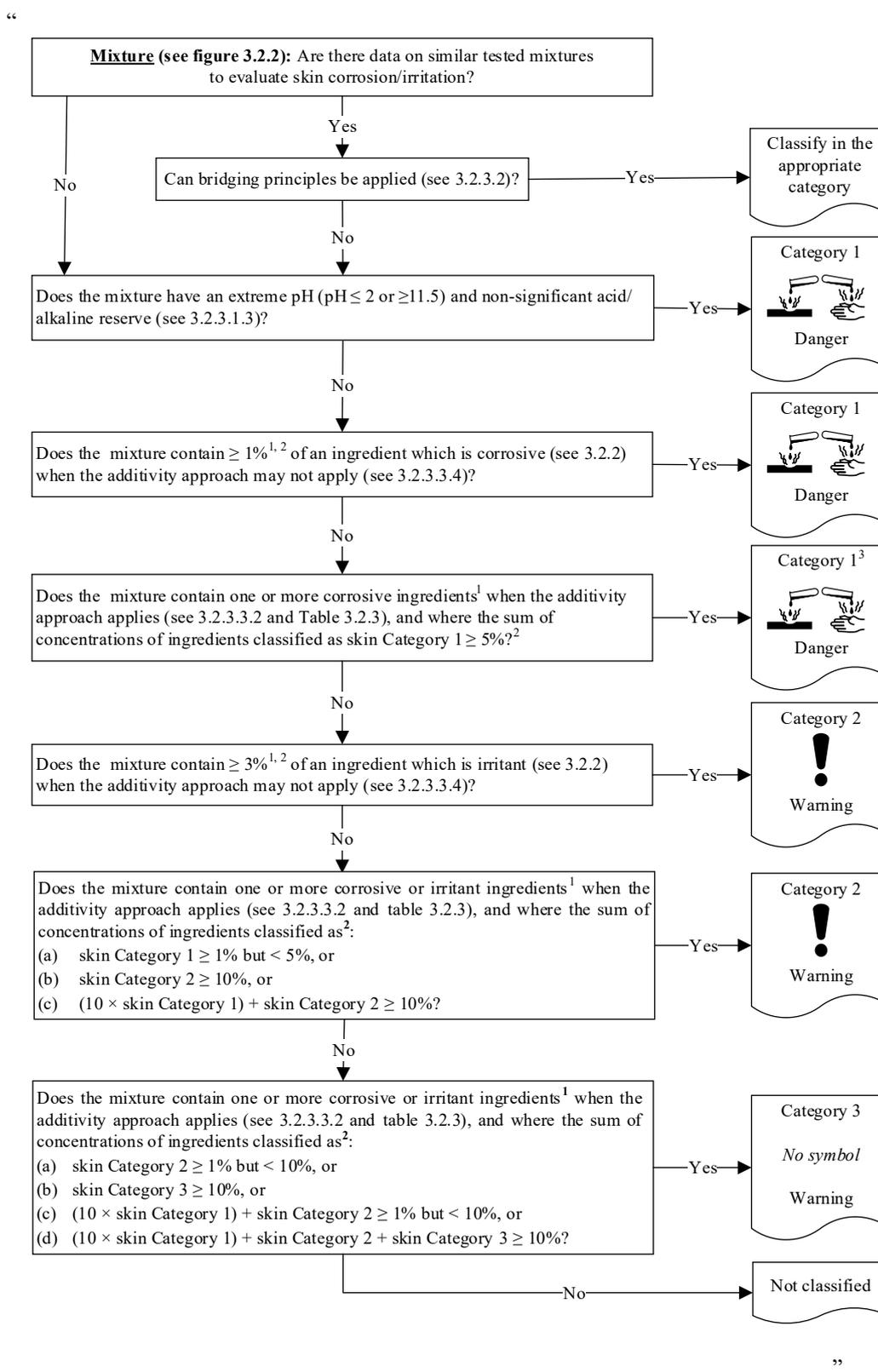
^a The dashed boxes represent an individual tier within conclusive data on the mixture as a whole. However, in contrast to substances, mixtures having an extreme pH value ($pH \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve but no other conclusive data on the mixture as a whole, or no conclusive weight of evidence assessment from all available data on the mixture as a whole, are not conclusive within the tiers for conclusive data on the mixture as a whole. Such mixtures should be first evaluated according to the bridging principles before the extreme pH value is considered as conclusive for classification.”

3.2.3.1.1 In the last sentence, replace “calculation method” with “classification based on ingredients”.

3.2.3.1.2 In the first sentence, replace with “data generated from validated test methods” with “test methods validated according to international procedures” and “test methods used” with “test method(s) used”.

- 3.2.3.1.3 Amend to read as follows:
- “A mixture with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) is considered corrosive (Category 1) in tier 4 if it has a significant acid/alkaline reserve or if no data for acid/alkaline reserve are available. However, if consideration of acid/alkaline reserve suggests the mixture may not be corrosive despite the extreme pH value, the result is considered inconclusive within tier 4 (see figure 3.2.1). If the overall weight of evidence assessment remains inconclusive or no data other than pH and acid/alkaline reserve are available, mixtures with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve should be assessed using the bridging principles described in 3.2.3.2. If the bridging principles cannot be applied, mixtures with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve should be classified as skin Category 1 (see figure 3.2.2). A $\text{pH} > 2$ and < 11.5 is considered inconclusive and cannot be used for classification purposes. Acid/alkaline reserve and pH can be determined by different methods including those described in OECD Test Guideline 122 and Young et al. (1988), acknowledging that there are some differences between these methods (see 3.2.5.3.6). A competent authority may decide which criteria for significant acid/alkaline reserve can be applied.”
- 3.2.3.2.5 In the heading, add “category” at the end.
- 3.2.3.2.7 Replace “aerosol” with “aerosolized” at the beginning of the sentence.
- 3.2.3.3.1 At the end of the first paragraph after “tiered approach” insert “for mixtures (see 1.3.2.3)”
- 3.2.3.3.4 In the first sentence, replace “when classifying certain types of chemicals” with “when classifying mixtures containing certain types of substances”.
- Amend the middle of the third sentence to read “...the pH should be used as the classification criterion (see 3.2.3.1.3) since extreme pH...”.
- 3.2.3.3.5 In the first sentence replace “concentration limits/cut-off values” with “cut-off values/concentration limits”.
- Delete “Classification of hazardous substances and mixtures – Use of cut-off values/Concentration limits” inside the parentheses in the second sentence and delete the parentheses around “1.3.3.2”.
- In the third sentence replace “concentration cut-off values” with “cut-off values/concentration limits”.
- Delete the last sentence of the paragraph (“In those cases...Figure 2.3.1”).
- 3.2.4 Amend the last sentence to read as follows: “Table 3.2.5 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”
- 3.2.5.1 In decision logic 3.2.1, amend the question starting with “Is the **substance or mixture**” to read as follows:
- “Is the **substance or mixture corrosive, irritant or mild irritant** (see 3.2.2 and 3.2.3.1) in accordance with the tiered approach (see 3.2.2.8 and figures 3.2.1 and 3.2.2)?”.

3.2.5.2 Replace decision logic 3.2.2 with the following:



In footnote 2, replace “see 3.2.3.3.6” with “see 3.2.3.3.5 and 3.2.3.3.6”.

3.2.5.3.1 Replace “weight of evidence approach” with “weight of evidence assessment”.

3.2.5.3.4 In the heading, replace “*in-vitro*” with “in vitro” and in the first sentence replace “or 439” with “and/or 439”.

3.2.5.3.5.2.6 In the second sentence replace “approach” with “assessment”.

3.2.5.3.6 Insert the following new paragraphs:

“3.2.5.3.6 *Guidance on the use of pH and acid/alkaline reserve for classification as skin corrosion/irritation*

3.2.5.3.6.1 Methods to determine the pH value such as OECD Test Guideline 122 and the method described by Young et al. (1988) differ in the concentration of the substance or mixture for which the pH is determined and include values of 1%, 10% and 100%. These methods also differ in the way the acid/alkaline reserve is determined, namely up to a pH of 7 for both acids and bases (OECD Test Guideline 122) or up to a pH of 4 for acids and a pH of 10 for bases (Young et al., 1988). Furthermore, there are differences between OECD Test Guideline 122 and Young et al. (1988) in the units used to express the acid/alkaline reserve.

3.2.5.3.6.2 Criteria to identify substances and mixtures requiring classification in Category 1 based on pH and acid/alkaline reserve have been developed for effects on the skin (Young et al., 1988). These criteria were developed using a combination of pH and acid/alkaline reserve values that were determined in a specific way (Young et al., 1988). Therefore, these criteria may not be directly applicable when other test concentrations or methods are used to measure pH and acid/alkaline reserve. Furthermore, the calibration and validation of these criteria was based on a limited dataset for effects on the skin. Thus, the predictive value of the combination of pH and acid/alkaline reserve for classification in Category 1 for effects on the skin is limited, especially for substances and mixtures with an extreme pH but a non-significant acid/alkaline reserve. The criteria developed by Young et al. (1988) for classification in Category 1 may be used as a starting point for determining whether a substance or a mixture has a significant acid/alkaline reserve or a non-significant acid/alkaline reserve. A competent authority may decide which criteria for significant acid/alkaline reserve can be applied.

* *References:*

Young, J.R., M.J. How, A.P. Walker, and W.M. Worth. 1988. Classification as corrosive or irritant to skin of preparations containing acidic or alkaline substances, without testing on animals. Toxicol. In Vitro, 2(1): 19-26. Doi: 10.1016/0887-2333(88)90032-x.”.

Chapter 3.3

3.3.1.2 Replace with the following:

“3.3.1.2 To classify, all available and relevant information on serious eye damage/eye irritation is collected and its quality in terms of adequacy and reliability is assessed. Classification should be based on mutually acceptable data/results generated using methods and/or defined approaches¹ that are validated according to international procedures. These include both OECD guidelines and equivalent methods/defined approaches (see 1.3.2.4.3). Sections 3.3.2.1 to 3.3.2.8 provide classification criteria for the different types of information that may be available.”.

Insert a new footnote 1 to read as follows:

¹ *According to OECD Guidance Document 255 on the reporting of defined approaches to be used within integrated approaches to testing and assessment, a defined approach to testing and assessment consists of a fixed data interpretation procedure (DIP) applied to data generated with a defined set of information sources to derive a result that can either be used on its own,*

or together with other information sources within an overall weight of evidence assessment, to satisfy a specific regulatory need.”.

3.3.1.3 and 3.3.1.4 Insert the following two new paragraphs:

“3.3.1.3 A tiered approach (see 3.3.2.10) organizes the available information into levels/tiers and provides for decision-making in a structured and sequential manner. Classification results directly when the information consistently satisfies the criteria. However, where the available information gives inconsistent and/or conflicting results within a tier, classification of a substance or a mixture is made on the basis of the weight of evidence within that tier. In some cases when information from different tiers gives inconsistent and/or conflicting results (see 3.3.2.10.3) or where data individually are insufficient to conclude on the classification, an overall weight of evidence assessment is used (see 1.3.2.4.9, 3.3.2.9 and 3.3.5.3.1).

3.3.1.4 Guidance on the interpretation of criteria and references to relevant guidance documents are provided in 3.3.5.3.”.

3.3.2 Delete “(see table 3.3.1)” in sub-paragraph (a) and “(see table 3.3.2)” in sub-paragraph (b) and in the last sentence.

3.3.2.1 Delete the heading “*Classification based on standard animal test data*”

3.3.2.1 and 3.3.2.2 (new) Insert the following new paragraphs:

“3.3.2.1 *Classification based on human data (tier 1 in figure 3.3.1)*

Existing reliable and good quality human data on serious eye damage/eye irritation should be given high weight where relevant for classification (see 3.3.5.3.2) and should be the first line of evaluation, as this gives information directly relevant to effects on the eye. Existing human data could be derived from single or repeated exposure(s), for example in occupational, consumer, transport or emergency response scenarios and epidemiological and clinical studies in well-documented case reports and observations (see 1.1.2.5 (c), 1.3.2.4.7 and 1.3.2.4.9). Although human data from accident or poison centre databases can provide evidence for classification, absence of incidents is not itself evidence for no classification, as exposures are generally unknown or uncertain.

3.3.2.2 *Classification based on standard animal data (tier 1 in figure 3.3.1)*

OECD Test Guideline 405 is the currently available and internationally accepted animal test method for classification as serious eye damage or eye irritant (see tables 3.3.1 and 3.3.2, respectively) and is the standard animal test. The current version of OECD Test Guideline 405 uses a maximum of three animals. Results from animal studies conducted under previous versions of OECD Test Guideline 405 that used more than three animals are also considered standard animal tests when interpreted in accordance with 3.3.5.3.3.”.

3.3.2.1.1 to 3.3.2.1.2.3 Current paragraphs 3.3.2.1.1 to 3.3.2.1.2.3 become new paragraphs 3.3.2.2.1 to 3.3.2.2.2.3.

Table 3.3.1 Delete note “a”. Current notes “b” and “c” become “a” and “b” respectively.

In note “b” replace “3.3.5.3” with “3.3.5.3.3”.

3.3.2.2.2.1 (new, former 3.3.2.1.2.1) In the last sentence replace “chemical” with “substance”.

3.3.2.2.2.2 (new, former 3.3.2.1.2.2) Replace “categories 2A and 2B” with “Category 2A and Category 2B”.

Table 3.3.2 Delete note “a”. Current notes “b” and “c” become “a” and “b” respectively.

In note “b”, replace “3.3.5.3” with “3.3.5.3.3”.

3.3.2.3 to 3.3.2.9 (new) Insert the following new paragraphs (and related footnotes 2 and 3), after table 3.3.2:

“3.3.2.3 Classification based on defined approaches (tier 2 in figure 3.3.1)

3.3.2.3.1 Defined approaches consist of a rule-based combination of data obtained from a predefined set of different information sources (e.g. *in vitro* methods, *ex vivo* methods, physico-chemical properties, non-test methods). It is recognized that most single *in vitro/ex vivo* methods are not able to replace *in vivo* methods fully for most regulatory endpoints. Thus, defined approaches can be useful strategies of combining data for classifying substances and mixtures. Results obtained with a defined approach validated according to international procedures, such as an OECD defined approach guideline or an equivalent approach, is conclusive for classification for serious eye damage/eye irritation if the criteria of the defined approach are fulfilled (see 3.3.5.3.4)². Data from a defined approach can only be used for classification when the tested substance is within the applicability domain of the defined approach used. Additional limitations described in the published literature should also be taken into consideration.

3.3.2.3.2 Where the results from defined approaches are assigned a level of confidence, a low confidence outcome of a defined approach cannot be used on its own to classify but may be considered in combination with other data.

3.3.2.3.3 Individual evidence used within a defined approach should not also be used outside of that defined approach.

3.3.2.4 Classification based on *in vitro/ex vivo* data (tier 2 in figure 3.3.1)

3.3.2.4.1 The classification criteria for the currently available *in vitro/ex vivo* test methods adopted by OECD in test guidelines 437, 438, 460, 491, 492, 494 and 496 are described in table 3.3.6 (see 3.3.5.3.5.1). When considered individually, these *in vitro/ex vivo* OECD test guidelines address serious eye damage and/or no classification for eye hazard, but do not address eye irritation. Therefore, data from a single *in vitro/ex vivo* OECD test guideline can only be used to conclude on either classification in Category 1 or no classification and cannot be used to conclude on classification in Category 2. When the result of a single *in vitro/ex vivo* method is “no stand-alone prediction can be made” (e.g. see table 3.3.6), a conclusion cannot be drawn on the basis of that single result and further data are necessary for classification (see 3.3.5.3.4.3 and 3.3.5.3.4.4).

3.3.2.4.2 *In vitro/ex vivo* methods in 3.3.2.4.1 with the result “no stand-alone prediction can be made” should within tier 2 only be used in combination with other types of data in defined approaches.

3.3.2.4.3 Other validated *in vitro/ex vivo* test methods accepted by some competent authorities are described in 3.3.5.3.5.2. Some of these *in vitro/ex vivo* test methods may be useful to classify in Category 2. A competent authority may decide which classification criteria, if any, should be applied for these test methods to conclude on classification, including that a substance is not classified for effects on the eye.

3.3.2.4.4 *In vitro/ex vivo* data can only be used for classification when the tested substance is within the applicability domain of the test method(s) used. Additional limitations described in the published literature should also be taken into consideration.

3.3.2.4.5 *Serious eye damage (Category 1)/Irreversible effects on the eye*

3.3.2.4.5.1 Where tests have been undertaken in accordance with OECD test guidelines 437, 438, 460, 491 and/or 496, a substance is classified for serious eye damage in Category 1 based on the criteria in table 3.3.6 (see 3.3.5.3.5.1).

3.3.2.4.5.2 Although the currently available OECD *in vitro/ex vivo* test guidelines and equivalent methods have not been developed to identify substances inducing discolouration of the eye, some comparable effects may be observed in these tests. Therefore, where, after washing, discolouration of the cornea or of the tested cells compared to the control is observed in OECD Test Guideline 437, 438, 492 or 494, or in other equivalent methods, suggesting a permanent effect, a competent authority may require classification of a substance for serious eye damage in Category 1.

3.3.2.4.6 *Eye irritation (Category 2)/Reversible effects on the eye*

3.3.2.4.6.1 A positive result in an *in vitro/ex vivo* test method that is validated according to international procedures for identification of substances inducing eye irritation can be used to classify for eye irritation in Category 2/2A³.

3.3.2.4.6.2 Where competent authorities adopt category 2A and category 2B, it is important to note that the currently validated *in vitro/ex vivo* test methods for effects on the eye do not allow discrimination between these two categories. In this situation, if the criteria for classification in Category 2 have been considered fulfilled, and no other relevant information is available, classification in Category 2/2A should be applied.

3.3.2.4.7 *No classification for effects on the eye*

OECD test guidelines 437, 438, 491, 492, 494 and 496 (see table 3.3.6 in 3.3.5.3.5.1) can be used to conclude that a substance is not classified for effects on the eye.

3.3.2.5 *Classification based on conclusive human data, standard animal data or in vitro/ex vivo data for skin corrosion (tier 3 in figure 3.3.1)*

Substances classified as corrosive to skin (skin Category 1) based on conclusive human data, standard animal data or *in vitro/ex vivo* data for skin corrosion according to the criteria in chapter 3.2 are also deemed as inducing serious eye damage (eye Category 1). Skin irritation (skin Category 2), mild skin irritation (skin Category 3) and no classification for skin irritation, as well as human patch data (as described in chapter 3.2), cannot be used alone to conclude on eye irritation or no classification for effects on the eye, but may be considered in an overall weight of evidence assessment.

3.3.2.6 *Classification based on other existing animal skin or eye data (tier 4 in figure 3.3.1)*

Other existing skin or eye data in animals may be used for classification, but there may be limitations regarding the conclusions that can be drawn (see 3.3.5.3.6). Substances classified as corrosive to skin (skin Category 1) based on other existing skin data according to the criteria in chapter 3.2 are also deemed as inducing serious eye damage (eye Category 1). Other existing skin data leading to classification in skin Category 2, 3 or no classification, cannot be used alone to conclude on eye irritation or no classification for effects on the eye, but may be considered in an overall weight of evidence assessment.”

3.3.2.7 *Classification based on extreme pH (pH ≤ 2 or ≥ 11.5) and acid/alkaline reserve (tier 5 in figure 3.3.1)*

In general, substances with an extreme pH (pH ≤ 2 or ≥ 11.5) are expected to cause significant eye effects, especially when associated with significant acid/alkaline reserve. A substance with pH ≤ 2 or ≥ 11.5 is therefore

considered to cause serious eye damage (Category 1) in this tier if it has a significant acid/alkaline reserve or if no data for acid/alkaline reserve are available. However, if consideration of acid/alkaline reserve suggests the substance may not cause serious eye damage despite the extreme pH value, the result is considered inconclusive within this tier (see figure 3.3.1). A pH > 2 and < 11.5 is considered inconclusive and cannot be used for classification purposes. Acid/alkaline reserve and pH can be determined by different methods including those described in OECD Test Guideline 122 and Young et al. (1988), acknowledging that there are some differences between these methods (see 3.3.5.3.7). A competent authority may decide which criteria for significant acid/alkaline reserve can be applied.

3.3.2.8 Classification based on non-test methods for serious eye damage/eye irritation or for skin corrosion (tier 6 in figure 3.3.1)

3.3.2.8.1 Classification, including the conclusion not classified, can be based on non-test methods, with due consideration of reliability and applicability, on a case-by-case basis. Non-test methods include computer models predicting qualitative structure-activity relationships (structural alerts, SAR) or quantitative structure-activity relationships (QSARs), computer expert systems, and read-across using analogue and category approaches.

3.3.2.8.2 Read-across using analogue or category approaches requires sufficiently reliable test data on similar substance(s) and justification of the similarity of the tested substance(s) with the substance(s) to be classified. Where adequate justification of the read-across approach is provided, it has in general higher weight than (Q)SARs.

3.3.2.8.3 Classification based on (Q)SARs requires sufficient data and validation of the model. The validity of the computer models and the prediction should be assessed using internationally recognized principles for the validation of (Q)SARs. With respect to reliability, lack of alerts in a SAR or expert system is not sufficient evidence for no classification.

3.3.2.8.4 Conclusive non-test data for skin corrosion may be used for classification for effects on the eye. Thus, substances classified as corrosive to skin (skin Category 1) according to the criteria in chapter 3.2 are also deemed as inducing serious eye damage (eye Category 1). Skin irritation (skin Category 2), mild skin irritation (skin Category 3) and no classification for skin irritation according to chapter 3.2 cannot be used alone to conclude eye irritation or no classification for effects on the eye, but may be considered in an overall weight of evidence assessment.

3.3.2.8.5 For conclusions on no classification from read-across and (Q)SARs the adequacy and robustness of the scientific reasoning and of the supporting evidence should be well substantiated and normally requires multiple negative substances with good structural and physical (related to toxicokinetics) similarity to the substance being classified, as well as a clear absence of positive substances with good structural and physical similarity to the substance being classified.

3.3.2.9 Classification based on an overall weight of evidence assessment (tier 7 in figure 3.3.1)

3.3.2.9.1 An overall weight of evidence assessment using expert judgement is indicated where none of the previous tiers resulted in a definitive conclusion on classification. In some cases, where the classification decision was postponed until the overall weight of evidence, but no further data are available, a classification may still be possible.

3.3.2.9.2 A substance with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve (result considered inconclusive in tier 5; see 3.3.2.7) and for which no other information is available, should be classified as serious eye damage Category 1 in this tier. If inconclusive

information is also available from other tiers but the overall weight of evidence assessment remains inconclusive, the extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) result should take precedence and the substance should be classified as serious eye damage Category 1 in this tier independently of its acid/alkaline reserve. For mixtures, the approach is different and is detailed in 3.3.3.1.3.”

Footnotes Insert the following new footnotes 2 and 3 at the bottom of the page in relation to paragraphs 3.3.2.3 (for footnote 2) and 3.3.2.4.6.1 (for footnote 3):

“² *Some defined approaches have been proposed for serious eye damage/eye irritation (Alépée et al., 2019a, b) but no classification criteria have yet been agreed internationally.*”

“³ *Although no classification criteria have yet been agreed internationally for some validated and/or accepted in vitro/ex vivo test methods proposed for identifying substances inducing eye irritation, these test methods may still be accepted by some competent authorities (see 3.3.2.4.2). If a defined approach (see 3.3.2.3) is not available or is not adequate for classification, data from these methods may be considered in a weight of evidence assessment within this tier.*”

3.3.2.2 and 3.3.2.2.1 Current paragraphs 3.3.2.2 and 3.3.2.2.1 become new paragraphs 3.3.2.10 and 3.3.2.10.1.

Delete existing paragraphs 3.3.2.2.2; 3.3.2.2.3, 3.3.2.2.4, 3.3.2.2.5 and 3.3.2.2.6.

3.3.2.10 and 3.3.2.10.1 (new, former 3.3.2.2 and 3.3.2.2.1) Amend to read as follows:

“3.3.2.10 Classification in a tiered approach (figure 3.3.1)”

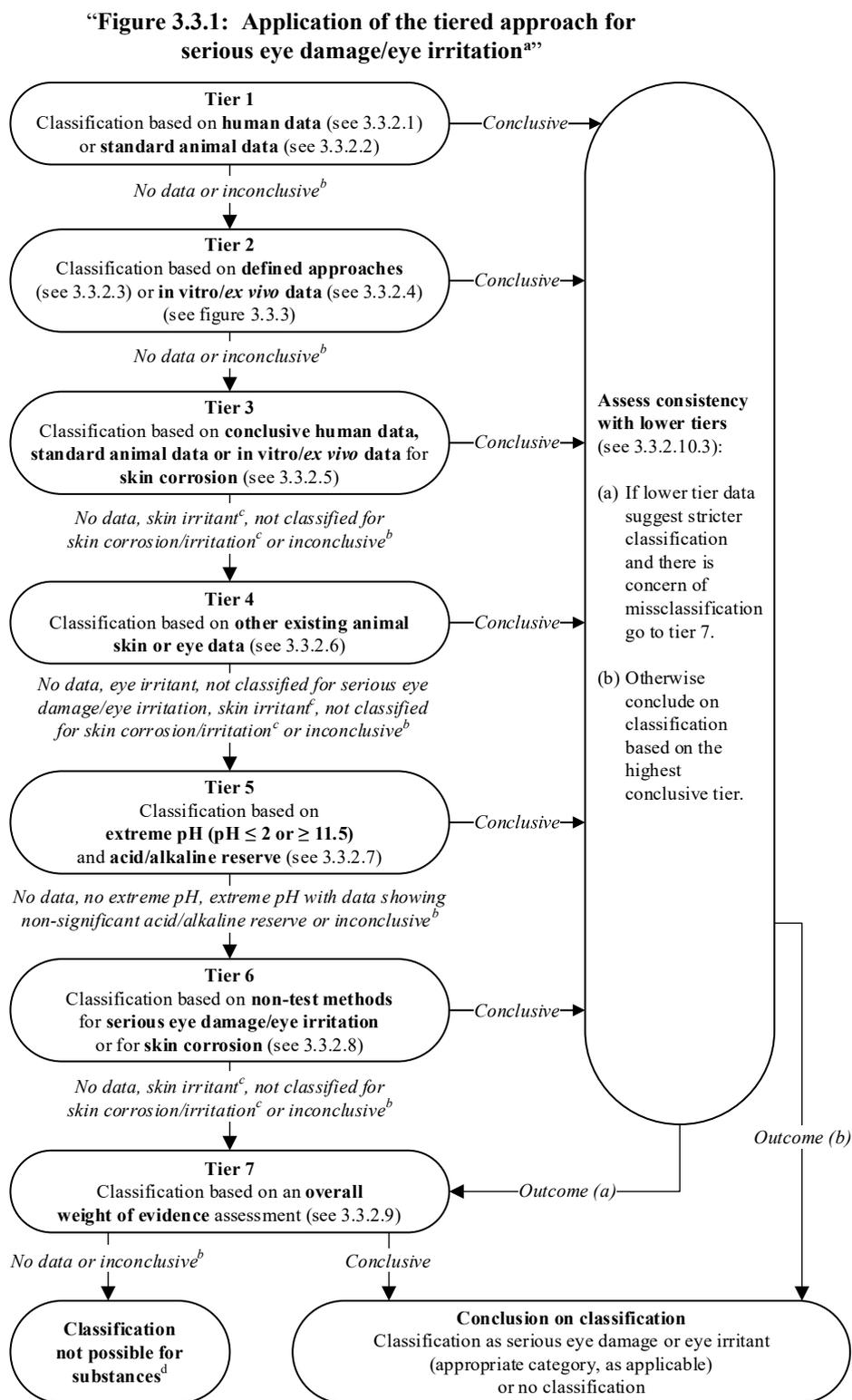
3.3.2.10.1 A tiered approach to the evaluation of information should be considered, where applicable (figure 3.3.1), recognizing that not all tiers as well as information within a tier may be relevant. However, all available and relevant information of sufficient quality needs to be examined for consistency with respect to the resulting classification.”

3.3.2.10.2 and 3.3.2.10.3 (new) Insert the following two new paragraphs:

“3.3.2.10.2 In the tiered approach (figure 3.3.1), existing human and standard animal data for eye effects form the highest tier, followed by defined approaches and in vitro/ex vivo data for eye effects, existing human/standard animal/in vitro/ex vivo data for skin corrosion, other existing animal skin or eye data, extreme pH and acid/alkaline reserve, and finally non-test methods. Where information from data within the same tier is inconsistent and/or conflicting, the conclusion from that tier is determined by a weight of evidence assessment.

3.3.2.10.3 Where information from several tiers is inconsistent and/or conflicting with respect to the resulting classification, information of sufficient quality from a higher tier is generally given a higher weight than information from a lower tier. However, when information from a lower tier would result in a stricter classification than information from a higher tier and there is concern for misclassification, then classification is determined by an overall weight of evidence assessment. For example, having consulted the guidance in 3.3.5.3 as appropriate, classifiers concerned with a negative result for serious eye damage in an in vitro/ex vivo study when there is a positive result for serious eye damage in other existing eye data in animals would utilise an overall weight of evidence assessment. The same would apply in the case where there is human data indicating eye irritation but positive results from an in vitro/ex vivo test for serious eye damage are also available.”

Figure 3.3.1: Replace with the following:



Replace current notes “a”, “b”, “c” and “d” to figure 3.3.1 with the following and delete notes “e” and “f”:

^a *Before applying the approach, the explanatory text in 3.3.2.10 as well as the guidance in 3.3.5.3 should be consulted. Only adequate and reliable data of sufficient quality should be included in applying the tiered approach.*

^b *Information may be inconclusive for various reasons, e.g.:*

- *The available data may be of insufficient quality, or otherwise insufficient/inadequate for the purpose of classification, e.g. due to quality issues related to experimental design and/or reporting;*
- *The available data may be insufficient to conclude on the classification, e.g. they might be indicative for absence of serious eye damage, but inadequate to demonstrate eye irritation;*
- *Where competent authorities make use of the eye irritation categories 2A and 2B, the available data may not be capable of distinguishing between category 2A and category 2B.*

^c *It is recognized that not all skin irritants are eye irritants and that not all substances that are non-irritant to skin are non-irritant to the eye (see 3.3.2.5, 3.3.2.6, 3.3.2.8.4 and 3.3.2.9.1).*

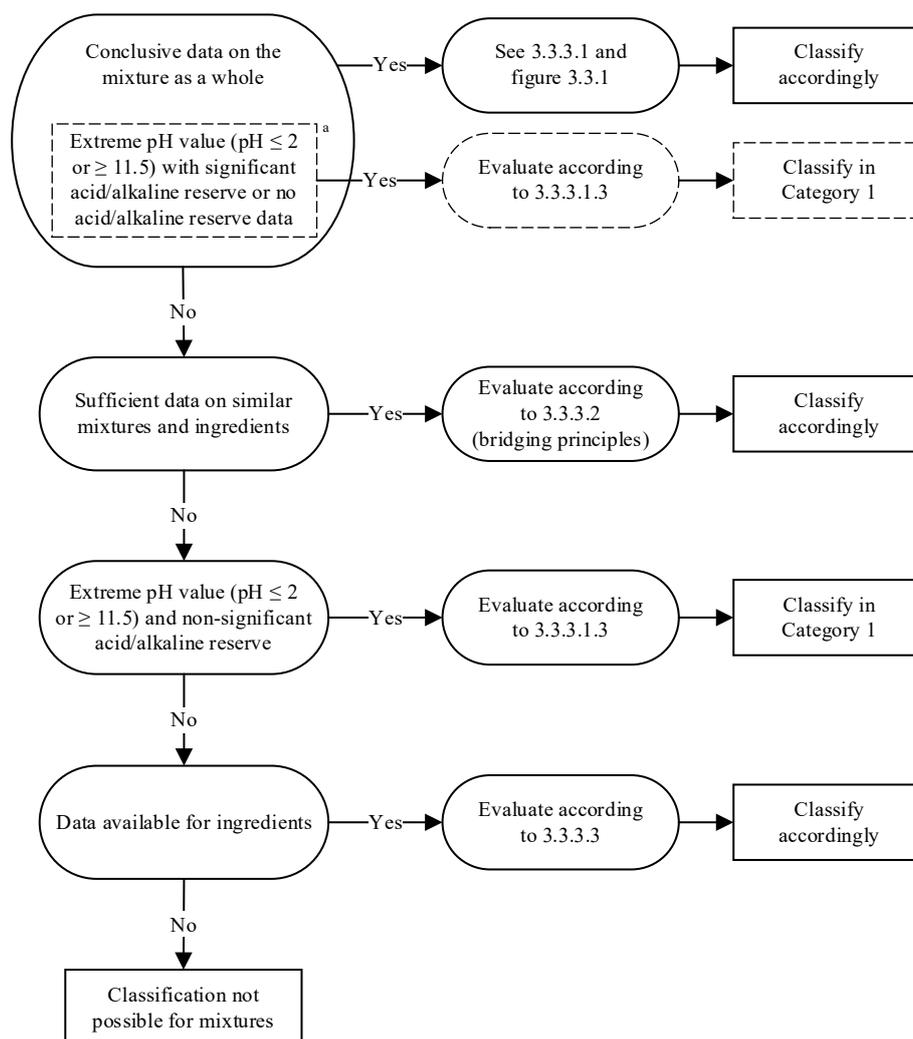
^d *For mixtures, the flow chart in figure 3.3.2 should be followed.”.*

3.3.3 Amend to read as follows:

“3.3.3 Classification criteria for mixtures

The approach to classification for serious eye damage/eye irritation is tiered and is dependent upon the amount of information available for the mixture itself and for its ingredients. The flow chart of figure 3.3.2 below outlines the process to be followed.

Figure 3.3.2: Tiered approach to classification of mixtures for serious eye damage/eye irritation



^a The dashed boxes represent an individual tier within conclusive data on the mixture as a whole. However, in contrast to substances, mixtures having an extreme pH value ($\text{pH} \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve but no other conclusive data on the mixture as a whole, or no conclusive weight of evidence assessment from all available data on the mixture as a whole, are not conclusive within the tiers for conclusive data on the mixture as a whole. Such mixtures should be first evaluated according to the bridging principles before the extreme pH value is considered as conclusive for classification.”

3.3.3.1.1 and 3.3.3.1.2 Amend to read as follows:

“3.3.3.1.1 In general, the mixture should be classified using the criteria for substances, taking into account the tiered approach to evaluate data for this hazard class (as illustrated in figure 3.3.1) and 3.3.3.1.2 and 3.3.3.1.3 below. If classification is not possible using the tiered approach, then the approach described in 3.3.3.2 (bridging principles), or, if that is not applicable, 3.3.3.3 (classification based on ingredients) should be followed.

3.3.3.1.2 Defined approaches and/or in vitro/ex vivo test methods validated according to international procedures may not have been validated using mixtures; although these approaches/methods are considered broadly applicable to mixtures, they can only be used for classification of mixtures when all ingredients of the mixture fall within the applicability domain of the defined approach or test method(s) used. Specific limitations regarding applicability domains are described in the respective defined approaches and

test methods and should be taken into consideration as well as any further information on such limitations from the published literature. Where there is reason to assume or evidence indicating that the applicability domain of a particular defined approach or test method is limited, data interpretation should be exercised with caution, or the results should be considered not applicable.”.

3.3.3.1.3 (new) Insert the following new paragraph:

“3.3.3.1.3 A mixture with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) is considered to cause serious eye damage (Category 1) in tier 5 if it has a significant acid/alkaline reserve or if no data for acid/alkaline reserve are available. However, if consideration of acid/alkaline reserve suggests the mixture may not cause serious eye damage despite the extreme pH value, the result is considered inconclusive within tier 5 (see figure 3.3.1). If the overall weight of evidence assessment remains inconclusive or no data other than pH and acid/alkaline reserve are available, mixtures with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve should be assessed using the bridging principles described in 3.3.3.2. If the bridging principles cannot be applied, mixtures with an extreme pH ($\text{pH} \leq 2$ or ≥ 11.5) and non-significant acid/alkaline reserve should be classified as eye Category 1 (see figure 3.3.2). A $\text{pH} > 2$ and < 11.5 is considered inconclusive and cannot be used for classification purposes. Acid/alkaline reserve and pH can be determined by different methods including those described in OECD Test Guideline 122 and Young et al. (1988), acknowledging that there are some differences between these methods (see 3.3.5.3.7). A competent authority may decide which criteria for significant acid/alkaline reserve can be applied.”.

3.3.3.2.6 In the last sentence, replace “by testing” with “based on test data” and “assigned” with “classified”.

3.3.3.2.7 Amend the beginning of the sentence to read “An aerosolized form of a mixture” and replace “non-aerosolized form of mixture” with “non-aerosolized form of the mixture”

Current footnote 1 becomes footnote 4.

3.3.3.3.1 In the introductory paragraph: replace “properties” with “hazards”, delete “the” before “mixtures” and amend the end to read “...where appropriate in the tiered approach for mixtures (see 1.3.2.3): ”

3.3.3.3.4 In the first sentence, insert “mixtures containing” after “classifying” and replace “chemicals” with “substances”.

In the third sentence, replace “the pH should be used as classification criterion (see 3.3.3.1.2) since pH” with “the pH should be used as the classification criterion (see 3.3.3.1.3) since extreme pH” and delete “(subject to consideration of acid/alkali reserve).”.

3.3.3.3.5 In the first sentence, replace “irreversible/reversible eye effects” with “serious eye damage/eye irritation”.

In the second sentence, delete “Use of cut-off values/Concentration limits” inside the parentheses.

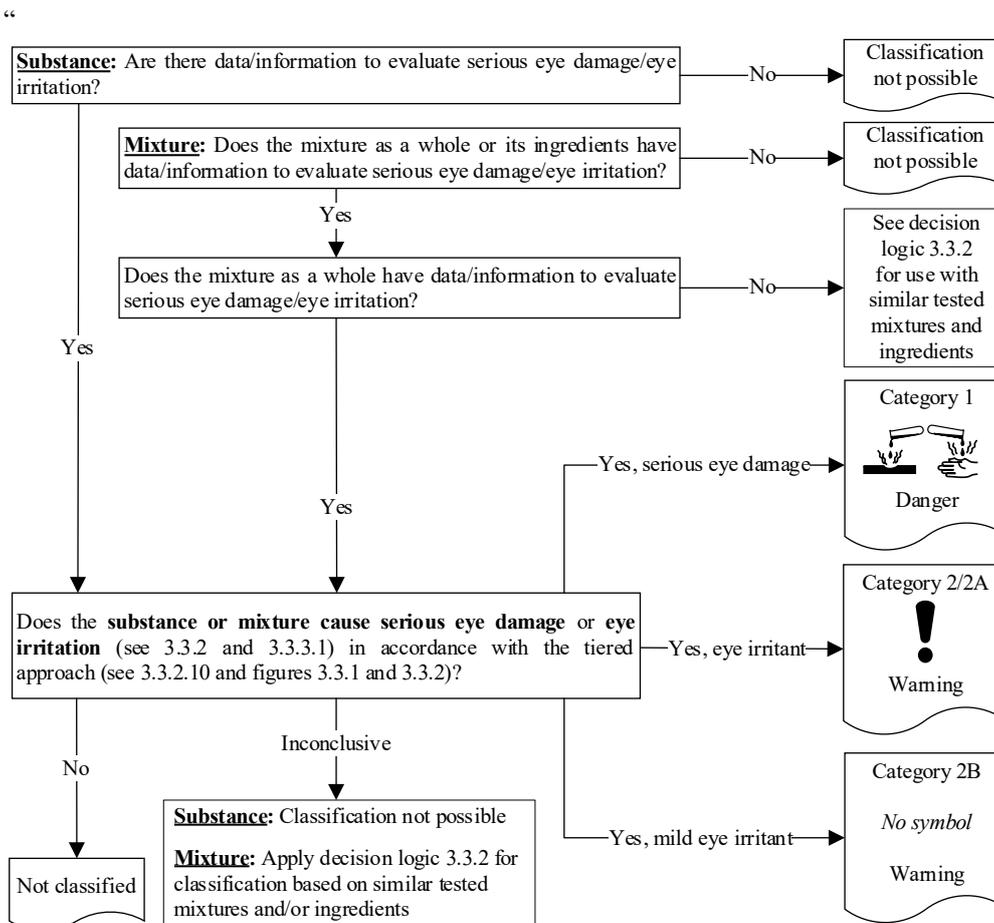
In the third sentence, replace “irreversible/reversible eye effects” with “serious eye damage/eye irritation” and “concentration/cut-off levels” with “cut-off values/concentration limits”.

Delete the last sentence of the paragraph (“In those cases...in this chapter”).

3.3.4 At the end of the paragraph, insert “Table 3.3.5 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

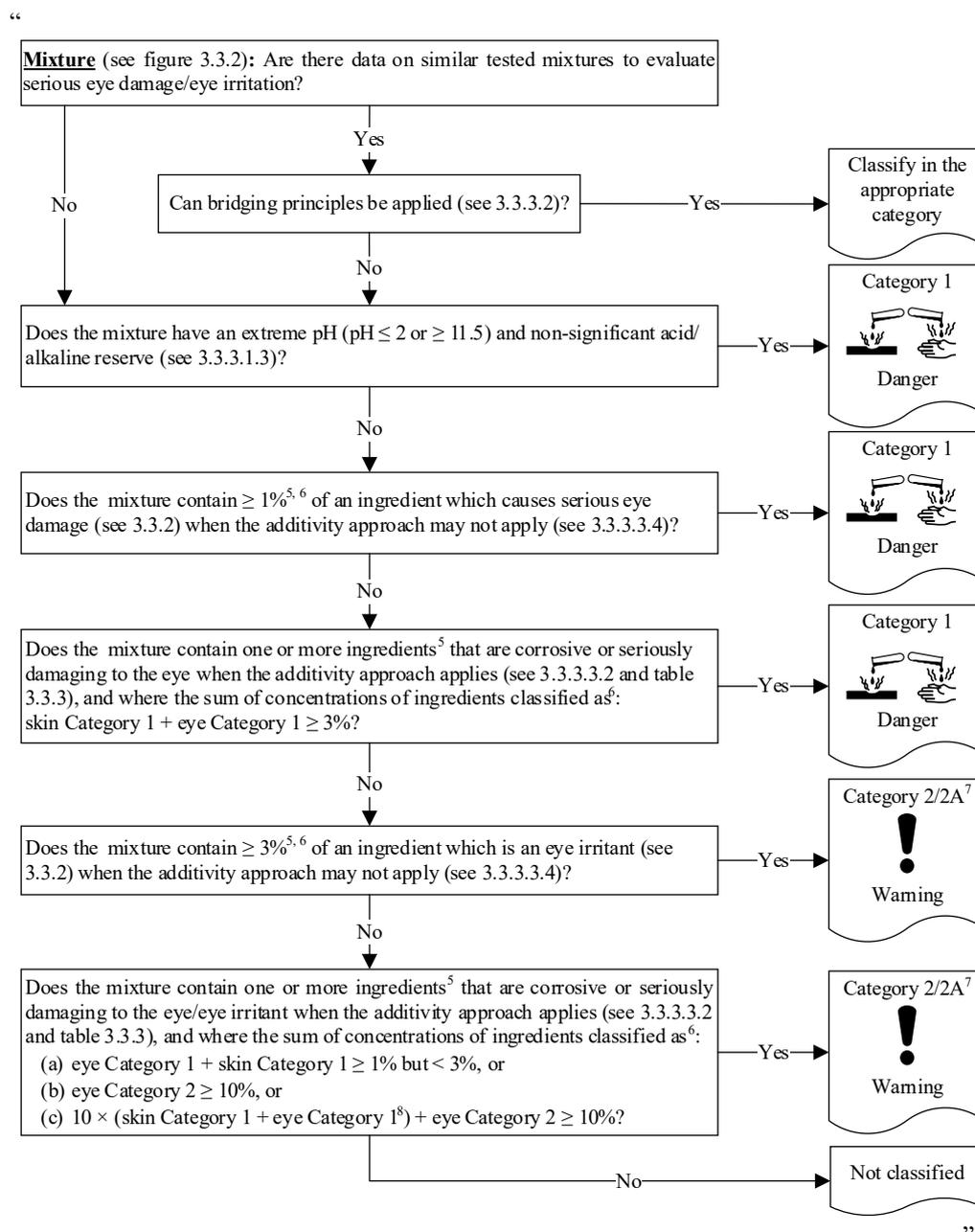
Table 3.3.5, third column Replace “Category 2A” with: “Category 2/2A”.

3.3.5.1 Replace decision logic 3.3.1 with the following and delete current footnotes 2 and 3:



”

3.3.5.2 Replace decision logic 3.3.2 with the following:



Current footnotes “4”, “5”, “6” and “7” become “5”, “6”, “7” and “8”.

3.3.5.3.1 and 3.3.5.3.2 (new) Insert the following two new paragraphs:

“3.3.5.3.1 *Relevant guidance documents*

Helpful information on the strengths and weaknesses of the different test and non-test methods, as well as useful guidance on how to apply a weight of evidence assessment, is provided in OECD Guidance Document 263 on an integrated approach on testing and assessment (IATA) for serious eye damage and eye irritation.

3.3.5.3.2 *Guidance on the use of human data for classification as serious eye damage/eye irritation*

The availability of human data for serious eye damage/eye irritation is limited and the data available may contain some uncertainty. However, where such data exist, they should be considered based on their quality. Human data may be obtained from epidemiological studies, human

experience (e.g. consumer experience), poison control centres, national and international home accident surveillance programs, case studies, or worker experience and accidents. Human case studies may have limited predictive value as often the presence of a substance or mixture in the eye will result in pain and quick washing of the eyes. Therefore, the effects observed may underestimate the intrinsic property of the substance or the mixture to affect the eye without washing. Further details on the strengths and limitations of human data for serious eye damage/eye irritation can be found in OECD Guidance Document 263 (section 4.1. Module 1: Existing human data on serious eye damage and eye irritation).”

3.3.5.3.3 Insert the following new heading:

“3.3.5.3.3 *Classification based on standard animal tests with more than three animals*”

3.3.5.3.1 to 3.3.5.3.5 Current paragraphs 3.3.5.3.1 to 3.3.5.3.5 become new paragraphs 3.3.5.3.3.1 to 3.3.5.3.3.5.

3.3.5.3.3.2 (new, former 3.3.5.3.2) Replace “3.3.2.1” with “3.3.2.2” and “done” with “performed”.

3.3.5.3.4 to 3.3.5.3.7.2 Insert the following new sections:

“3.3.5.3.4 *Guidance on the use of defined approaches and/or in vitro/ex vivo data for classification within tier 2 of figure 3.3.1*

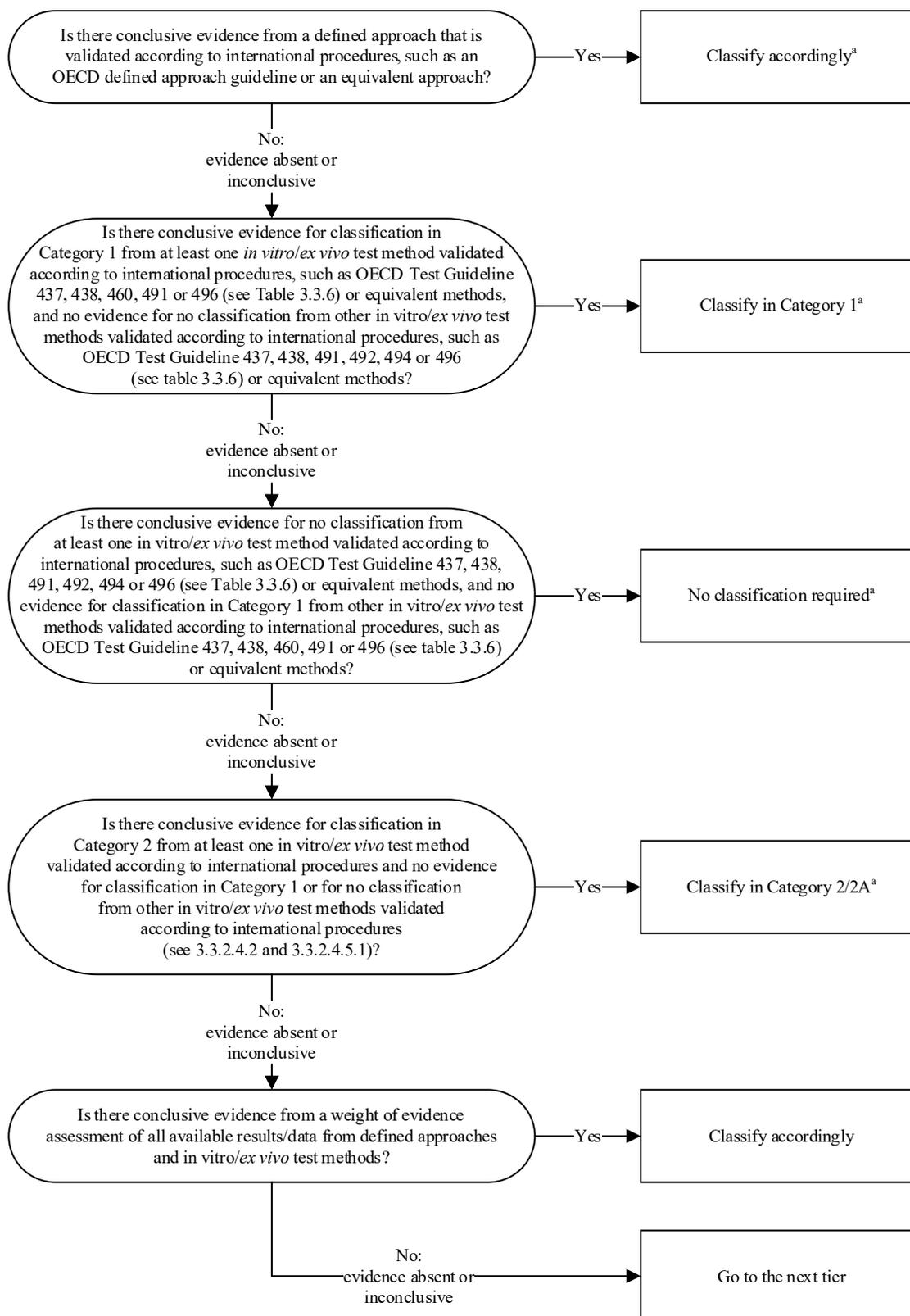
3.3.5.3.4.1 Defined approaches consist of a predefined set of different information sources (e.g. in vitro methods, *ex vivo* methods, physico-chemical properties, non-test methods) which, combined together through a fixed Data Interpretation Procedure (DIP) to convert input data into a prediction (or result), can provide a conclusion on the classification of a substance or mixture. A fixed DIP is defined as any fixed algorithm for interpreting data from one or typically several information sources and is rule-based in the sense that it is based, for example on a formula or an algorithm (e.g. decision criteria, rule or set of rules) that do not involve expert judgment. The output of a DIP generally is a prediction of a biological effect of interest or regulatory endpoint. Since in a defined approach the information sources are prescribed and the set of rules on how to integrate and interpret them is predetermined, the same conclusion will always be reached by different assessors on the same set of data as there is no room for subjective interpretation. In contrast, in a weight of evidence assessment, expert judgment is applied on an ad hoc basis to the available information, which may lead to different conclusions because there are no fixed rules for interpreting the data.

3.3.5.3.4.2 A stepwise approach to the evaluation of information derived from tier 2 of figure 3.3.1, i.e. defined approaches and/or in vitro/*ex vivo* test methods, should be considered where applicable (figure 3.3.3), recognizing that not all tiers as well as information within a tier may be relevant. However, all available and relevant information of sufficient quality needs to be examined for consistency with respect to the resulting classification. The outcome of a defined approach containing conclusive animal and/or human data may also eventually be considered during the overall weight of evidence in tier 7 (see figure 3.3.1). Where information from several steps is inconsistent and/or conflicting with respect to the resulting classification, information of sufficient quality from a higher step is generally given a higher weight than information from a lower step. However, when information from a lower step would result in a stricter classification than information from a higher step and there is concern for misclassification, then classification is determined by a within-tier weight of evidence assessment. For example, classifiers concerned with a negative result for serious eye damage in a defined approach when there is a positive result for serious eye damage in an in vitro/*ex vivo* method would utilise a within-tier weight of evidence assessment.

3.3.5.3.4.3 Current *in vitro/ex vivo* test methods are not able to distinguish between certain *in vivo* effects, such as corneal opacity, iritis, conjunctiva redness or conjunctiva chemosis, but they have shown to correctly predict substances inducing serious eye damage/eye irritation independently of the types of ocular effects observed *in vivo*. Many of the current *in vitro/ex vivo* test methods can thus identify substances or mixtures not requiring classification with high sensitivity but with limited specificity when used to distinguish not classified from classified substances or mixtures. This means that it is reasonably certain that a substance or mixture identified as not requiring classification by OECD Test Guideline 437, 438, 491, 492, 494 or 496 (see table 3.3.6) is indeed not inducing eye effects warranting classification, whereas some substances or mixtures not requiring classification will be over-predicted by these *in vitro/ex vivo* test methods when used in isolation. Furthermore, it should be considered that substances inducing serious eye damage are identified by many of these test methods with a high specificity but a limited sensitivity when used to distinguish Category 1 from Category 2 and not classified. This means that it is reasonably certain that a substance or mixture identified as Category 1 by OECD Test Guideline 437, 438, 460, 491 or 496 (see table 3.3.6) is indeed inducing irreversible eye effects, whereas some substances or mixtures inducing serious eye damage will be under-predicted by these *in vitro/ex vivo* test methods when used in isolation. As a consequence, a single *in vitro/ex vivo* OECD test guideline method is currently sufficient to conclude on either Category 1 or no classification according to the criteria defined in table 3.3.6, but not to conclude Category 2. When the result of an *in vitro/ex vivo* method is “no stand-alone prediction can be made” (e.g. see table 3.3.6), a conclusion cannot be drawn on the basis of that single result and further data are necessary for classification. Some *in vitro/ex vivo* test methods validated according to international procedures but not adopted as OECD test guidelines may be accepted by some competent authorities to classify in Category 2 (see 3.3.5.3.5.2). Moreover, combinations of *in vitro/ex vivo* methods in tiered approaches or their integration in defined approaches (see 3.3.2.3) may reduce the number of false predictions and show adequate performance for classification purposes.

3.3.5.3.4.4 In the absence of an adequate defined approach (see 3.3.2.3) or of conclusive *in vitro/ex vivo* data (see 3.3.2.4.1 and 3.3.2.4.2), a stand-alone prediction is not possible. In such cases, a within-tier weight of evidence assessment of data from more than one method would be needed to classify within tier 2. If a within-tier weight of evidence assessment is still not conclusive, then data from lower tiers may be required to reach a conclusion (see figure 3.3.1).

Figure 3.3.3: Classification based on defined approaches and/or in vitro/ex vivo data within tier 2 of figure 3.3.1



^a Evidence is considered conclusive if the data fulfil the criteria of the defined approach or of the method and there is no contradicting in vitro/ex vivo information. When information from a lower step would result in a stricter classification than information from a higher step and there is concern for misclassification, then classification is determined by a within-tier weight of evidence assessment.

3.3.5.3.5 *Classification criteria based on in vitro/ex vivo data*

3.3.5.3.5.1 Where in vitro/ex vivo tests have been undertaken in accordance with OECD test guidelines 437, 438, 460, 491, 492, 494 and/or 496, the criteria for classification in Category 1 for serious eye damage/irreversible effects on the eye and for no classification are set out in table 3.3.6.

Table 3.3.6: Criteria for serious eye damage/irreversible effects on the eye and for no classification^a for in vitro/ex vivo methods

Category	OECD Test Guideline 437 Bovine Corneal Opacity and Permeability test method	OECD Test Guideline 438 Isolated Chicken Eye test method	OECD Test Guideline 460 Fluorescein Leakage test method	OECD Test Guideline 491 Short Time Exposure test method	OECD Test Guideline 492 Reconstructed human Cornea-like Epithelium (RhCE)-based test methods: Methods 1, 2, 3 and 4 as numbered in Annex II of OECD Test Guideline 492	OECD Test Guideline 494 Vitrigel-Eye Irritancy Test Method	OECD Test Guideline 496 In vitro Macromolecular Test Method (test method 1)
	<p>Organotypic <i>ex vivo</i> assay using isolated corneas from the eyes of freshly slaughtered cattle. Test chemicals are applied to the epithelial surface of the cornea. Damage by the test chemical is assessed by quantitative measurements of:</p> <ul style="list-style-type: none"> - Corneal opacity changes measured using a light transmission opacitometer (opacitometer 1) or a laserlight-based opacitometer (LLBO, opacitometer 2) - Permeability (sodium fluorescein dye). <p>Both measurements are used to calculate an in vitro irritancy score (IVIS) when using opacitometer 1 or a LLBO Irritancy Score (LIS) when using opacitometer 2.</p> <p>Criteria based on IVIS or LIS.</p>	<p>Organotypic <i>ex vivo</i> assay based on the short-term maintenance of chicken eyes in vitro. Test chemicals are applied to the epithelial surface of the cornea. Damage by the test chemical is assessed by (i) a quantitative measurement of increased corneal thickness (swelling), (ii) a qualitative assessment of corneal opacity, (iii) a qualitative assessment of damage to epithelium based on application of fluorescein to the eye, and (iv) a qualitative evaluation of macroscopic morphological damage to the surface. Histopathology can be used to increase the sensitivity of the method for identifying Category 1 non-extreme pH ($2 < \text{pH} < 11.5$) detergents and surfactants. ^b</p> <p>Criteria based on the scores of corneal swelling, opacity and fluorescein retention, which are used to assign ICE classes (I, II, III or IV) to each endpoint, and on macroscopic and histopathology assessment ^b</p>	<p>Cytotoxicity and cell-function based in vitro assay that is performed on a confluent monolayer of Madin-Darby Canine Kidney (MDCK) CB997 tubular epithelial cells cultured on permeable inserts. The toxic effects of a test chemical are measured after a short exposure time (1 minute) by an increase in permeability of sodium fluorescein through the epithelial monolayer of MDCK cells. The amount of fluorescein leakage that occurs is proportional to the chemical-induced damage to the tight junctions, desmosomal junctions and cell membranes, and is used to estimate the ocular toxicity potential of a test chemical.</p> <p>Criteria based on mean percent fluorescein leakage following a defined exposure period</p>	<p>Cytotoxicity-based in vitro assay that is performed on a confluent monolayer of Statens Seruminstitut Rabbit Cornea (SIRC) cells. Each test chemical is tested at both 5 % and 0.05 % concentrations. Following five-minute exposure, cell viability is assessed by the enzymatic conversion in viable cells of the vital dye MTT into a blue formazan salt that is quantitatively measured after extraction from cells.</p> <p>Criteria based on mean percent cell viability following a defined exposure period</p>	<p>Three-dimensional RhCE tissues are reconstructed from either primary human cells or human immortalised corneal epithelial cells, which have been cultured for several days to form a stratified, highly differentiated squamous epithelium, consisting of at least 3 viable layers of cells and a non-keratinised surface, showing a cornea-like structure morphologically similar to that found in the human cornea. Following exposure and post-treatment incubation (where applicable), tissue viability is assessed by the enzymatic conversion in viable cells of the vital dye MTT into a blue formazan salt that is quantitatively measured after extraction from the tissues.</p> <p>Criteria based on mean percent tissue viability following defined exposure and post-exposure (where applicable) periods</p>	<p>In vitro assay using human corneal epithelium models fabricated in a collagen vitrigel membrane (CVM) chamber. The eye irritation potential of the test chemical is predicted by analysing time-dependent changes in transepithelial electrical resistance values using the value of three indexes. Resistance values are measured at intervals of 10 seconds for a period of three minutes after exposure to the test chemical preparation.</p> <p>Criteria based on the 3 measured indexes: time lag, intensity and plateau level of electrical resistance.</p>	<p>In vitro assay consisting of a macromolecular plant-based matrix obtained from jack bean <i>Canavalis ensiformis</i>. This matrix serves as the target for the test chemical and is composed of a mixture of proteins, glycoproteins, carbohydrates, lipids and low molecular weight components, which form a highly ordered and transparent gel structure upon rehydration. Test chemicals causing ocular damage lead to the disruption and disaggregation of the highly organized macromolecular reagent matrix, and produce turbidity of the macromolecular reagent. Such phenomena is quantified, by measuring changes in light scattering.</p> <p>Criteria based on a Maximum Qualified Score (MQS) derived from the Optical Density readings at different concentrations, calculated via a software.</p>

Table 3.3.6: Criteria for serious eye damage/irreversible effects on the eye and for no classification^a for in vitro/ex vivo methods (cont'd)

Category	OECD Test Guideline 437 Bovine Corneal Opacity and Permeability test method		OECD Test Guideline 438 Isolated Chicken Eye test method	OECD Test Guideline 460 Fluorescein Leakage test method	OECD Test Guideline 491 Short Time Exposure test method	OECD Test Guideline 492 Reconstructed human Cornea-like Epithelium (RhCE)-based test methods: Methods 1, 2, 3 and 4 as numbered in Annex II of OECD Test Guideline 492				OECD Test Guideline 494 Vitrigel-Eye Irritancy Test Method	OECD Test Guideline 496 In vitro Macromolecular Test Method (test method 1)
	Opacitometer 1 IVIS > 55	Opacitometer 2 LIS > 30 and lux/7 ≤ 145 and OD490 > 2.5, OR LIS > 30 and lux/7 > 145	At least 2 ICE class IV, OR Corneal opacity = 3 at 30 min (in at least 2 eyes), OR Corneal opacity = 4 at any time point (in at least 2 eyes), OR Severe loosening of the epithelium (in at least 1 eye), OR Certain histopathological effects ^b	Chemical concentration causing 20 % of Fluorescein Leakage (FL ₂₀) ≤ 100 mg/mL	Viability ≤ 70 % at 5 % and 0.05 %	No stand-alone prediction can be made				No stand-alone prediction can be made	MQS > 30.0
2/2A/2B	No stand-alone prediction can be made.	No stand-alone prediction can be made	No stand-alone prediction can be made	No stand-alone prediction can be made	No stand-alone prediction can be made	No stand-alone prediction can be made				No stand-alone prediction can be made	No stand-alone prediction can be made
Not classified	Opacitometer 1 IVIS ≤ 3	Opacitometer 2 LIS ≤ 30	ICE class I for all 3 endpoints, OR ICE class I for 2 endpoints and ICE class II for the other endpoint, OR ICE class II for 2 endpoints and ICE class I for the other endpoint	No stand-alone prediction can be made	Viability > 70 % at 5 % and 0.05 %	Test method 1 Liquids and Solids: Viability > 60 %	Test method 2 Liquids: Viability > 60 %; Solids: Viability > 50 %	Test method 3 Liquids and Solids: Viability > 40 %	Test method 4 Liquids: Viability > 35 %; Solids: Viability > 60 %	Time lag > 180 seconds and Intensity < 0.05 %/seconds and Plateau level ≤ 5.0 %	MQS ≤ 12.5

^a Grading criteria are understood as described in OECD test guidelines 437, 438, 460, 491, 492, 494 and 496.

^b For criteria, please consult OECD Test Guideline 438

3.3.5.3.5.2 A non-exhaustive list of other validated *in vitro/ex vivo* test methods accepted by some competent authorities but not adopted as OECD test guidelines are listed below. A competent authority may decide which classification criteria, if any, should be applied for these test methods:

- Time to Toxicity (ET₅₀) tests using the Reconstructed human Cornea-like Epithelia (RhCE) described in OECD Test Guideline 492 (Kandarova et al., 2018; Alépée et al., 2020);
- *Ex Vivo* Eye Irritation Test (EVEIT): an *ex vivo* assay that uses excised rabbit corneal tissues kept in culture for several days and monitors tissue recovery to model both reversible and non-reversible eye effects. Full-thickness tissue recovery is monitored non-invasively using optical coherence tomography (OCT) (Frentz et al., 2008; Spöler et al., 2007; Spöler et al., 2015);
- Porcine Ocular Cornea Opacity/Reversibility Assay (PorCORA): an *ex vivo* assay that uses excised porcine corneal tissues kept in culture for up to 21 days and monitors tissue recovery to model both reversible and non-reversible eye effects. The tissues are stained with fluorescent dye and effects on the corneal epithelia are visualised by the retention of fluorescent dye (Piehl et al., 2010; Piehl et al., 2011);
- EyeIRR-IS assay: a genomic approach applied to a RhCE model (Cottrez et al., 2021);
- *In vitro* Macromolecular Test Method (test method 2), similar to test method 1 described in OECD Test Guideline 496 (Choksi et al., 2020);
- Metabolic activity assay: *In vitro* assay consisting of measuring changes to metabolic rate in test-material treated L929 cell monolayer (Harbell et al., 1999; EURL ECVAM, 2004a; Hartung et al., 2010; Nash et al., 2014);
- Hen's Egg Test on the Chorio-Allantoic Membrane (HET-CAM): an organotypic assay that uses the vascularised membrane of fertile chicken eggs to assess a test material's potential to cause vascular changes (Spielmann et al., 1993; Balls et al., 1995; Spielmann et al., 1996; Brantom et al., 1997; ICCVAM, 2007; ICCVAM, 2010);
- Chorio-Allantoic Membrane Vascular Assay (CAMVA): an organotypic assay that uses the vascularised membrane of fertile chicken eggs to assess a test material's potential to cause vascular changes (Bagley et al., 1994; Brantom et al., 1997; Bagley et al., 1999; Donahue et al., 2011);
- Neutral Red Release (NRR) assay: *In vitro* assay that quantitatively measures a substance's ability to induce damage to cell membranes in a monolayer of normal human epidermal keratinocytes (NHEK) (Reader et al. 1989; Reader et al., 1990; Zuang, 2001; EURL ECVAM, 2004b; Settivari et al., 2016); and
- Isolated Rabbit Eye (IRE) test, similar to OECD Test Guideline 438 but using isolated rabbit eyes instead of isolated chicken eyes (Burton et al., 1981; Whittle et al. 1992; Balls et al., 1995; Brantom et al., 1997; ICCVAM, 2007; ICCVAM, 2010).

3.3.5.3.6 *Guidance on the use of other existing skin or eye data in animals for classification as serious eye damage or eye irritation*

3.3.5.3.6.1 The availability of other animal data for serious eye damage/eye irritation may be limited as tests with the eye as the route of exposure are not normally performed. An exception could be historical data from the Low Volume Eye Test (LVET) that might be used in a weight of evidence assessment. The LVET is a modification of the standard OECD Test Guideline 405 test method.

3.3.5.3.6.2 Existing data from the LVET test could be considered for the purpose of classification and labelling but must be carefully evaluated. The differences between the LVET and OECD Test Guideline 405 may result in a classification in a lower category (or no classification) based on LVET data, than if the classification was based on data derived from the standard in vivo test (OECD Test Guideline 405). Thus, positive data from the LVET test could be a trigger for considering classification in Category 1 on its own, but data from this test are not conclusive for a Category 2 classification or no classification (ECHA, 2017). Such data may, however, be used in an overall weight of evidence assessment. It is noted that the applicability domain of the LVET is limited to household detergent and cleaning products and their main ingredients (surfactants) (ESAC, 2009).

3.3.5.3.6.3 Effects on the eyes may be observed in acute or repeated dose inhalation studies with full body exposure. However, normally no scoring according to the Draize criteria is performed and the follow-up period may be shorter than 21 days. Also, the effects on the eyes will likely depend upon the concentration of the substance/mixture and the exposure duration. As there are no criteria for minimal concentration and duration, the absence of effects on the eyes or eye irritation may not be conclusive for the absence of serious eye damage. The presence of irreversible effects on the eye should be considered within a weight of evidence assessment.

3.3.5.3.7 *Guidance on the use of pH and acid/alkaline reserve for classification as serious eye damage*

3.3.5.3.7.1 Methods to determine the pH value such as OECD Test Guideline 122 and the method described by Young et al. (1988) differ in the concentration of the substance or mixture for which the pH is determined and include values of 1%, 10% and 100%. These methods also differ in the way the acid/alkaline reserve is determined, namely up to a pH of 7 for both acids and bases (OECD Test Guideline 122) or up to a pH of 4 for acids and a pH of 10 for bases (Young et al., 1988). Furthermore, there are differences between OECD Test Guideline 122 and Young et al. (1988) in the units used to express the acid/alkaline reserve.

3.3.5.3.7.2 Criteria to identify substances and mixtures requiring classification in Category 1 based on pH and acid/alkaline reserve have been developed for effects on the skin (Young et al., 1988) and the same criteria are applied for effects on the eye. These criteria were developed using a combination of pH and acid/alkaline reserve values that were determined in a specific way (Young et al., 1988). Therefore, these criteria may not be directly applicable when other test concentrations or methods are used to measure pH and acid/alkaline reserve. Furthermore, the calibration and validation of these criteria was based on a limited dataset for effects on the skin. Thus, the predictive value of the combination of pH and acid/alkaline reserve for classification in Category 1 for effects on the eye is limited, especially for substances and mixtures with an extreme pH but a non-significant acid/alkaline reserve. The criteria developed by Young et al. (1988) for classification in Category 1 may be used as a starting point for determining whether a substance or a mixture has a significant acid/alkaline reserve or a non-significant acid/alkaline reserve. A competent authority may decide which criteria for significant acid/alkaline reserve can be applied.

* *References:*

Alépée, N., E. Adriaens, T. Abo, D. Bagley, B. Desprez, J. Hibatallah, K. Mewes, U. Pfannenbecker, Á. Sala, A.R. Van Rompay, S. Verstraelen, and P. McNamee. 2019a. Development of a defined approach for eye irritation or serious eye damage for liquids, neat and in dilution, based on Cosmetics

Europe analysis of in vitro STE and BCOP test methods. *Toxicol. In Vitro*, 57: 154-163. Doi: 10.1016/j.tiv.2019.02.019.

Alépée, N., E. Adriaens, T. Abo, D. Bagley, B. Desprez, J. Hibatallah, K. Mewes, U. Pfannenbecker, A. Sala, A.R. Van Rompay, S. Verstraelen, and P. McNamee. 2019b. Development of a defined approach for eye irritation or serious eye damage for neat liquids based on Cosmetics Europe analysis of in vitro RhCE and BCOP test methods. *Toxicol. In Vitro*, 59: 100-114. Doi: 10.1016/j.tiv.2019.04.011.

Alépée, N., V. Leblanc, M.H. Grandidier, S. Tehuob, V. Tagliati, E. Adriaens, and V. Michaut. 2020. Development of the SkinEthic HCE Time-to-Toxicity test method for identifying liquid chemicals not requiring classification and labelling and liquids inducing serious eye damage and eye irritation. *Toxicol. In Vitro*, 69: 104960. Doi: 10.1016/j.tiv.2020.104960.

Bagley, D.M., D. Waters, and B.M. Kong. 1994. Development of a 10-day chorioallantoic membrane vascular assay as an alternative to the Draize rabbit eye irritation test. *Food Chem. Toxicol.*, 32(12): 1155-1160. Doi: 10.1016/0278-6915(94)90131-7.

Bagley, D.M., D. Cerven, and J. Harbell. 1999. Assessment of the chorioallantoic membrane vascular assay (CAMVA) in the COLIPA in vitro eye irritation validation study. *Toxicol. In Vitro*, 13(2): 285-293. Doi: 10.1016/s0887-2333(98)00089-7.

Balls, M., P.A. Botham, L.H. Bruner, and H. Spielmann. 1995. The EC/HO international validation study on alternatives to the draize eye irritation test. *Toxicol. In Vitro*, 9(6): 871-929. Doi: 10.1016/0887-2333(95)00092-5.

Brantom, P.G., L.H. Bruner, M. Chamberlain, O. De Silva, J. Dupuis, L.K. Earl, D.P. Lovell, W.J. Pape, M. Uttley, D.M. Bagley, F.W. Baker, M. Bracher, P. Courtellemont, L. Declercq, S. Freeman, W. Steiling, A.P. Walker, G.J. Carr, N. Dami, G. Thomas, J. Harbell, P.A. Jones, U. Pfannenbecker, J.A. Southee, M. Tchong, H. Argembeaux, D. Castelli, R. Clothier, D.J. Esdaile, H. Itigaki, K. Jung, Y. Kasai, H. Kojima, U. Kristen, M. Larnicol, R.W. Lewis, K. Marenus, O. Moreno, A. Peterson, E.S. Rasmussen, C. Robles, and M. Stern. 1997. A summary report of the COLIPA international validation study on alternatives to the draize rabbit eye irritation test. *Toxicol. In Vitro*, 11: 141-179. Doi:10.1016/S0887-2333(96)00069-0.

Burton, A.B., M. York, and R.S. Lawrence. 1981. The in vitro assessment of severe eye irritants. *Food Cosmet. Toxicol.*, 19(4): 471-480. Doi: 10.1016/0015-6264(81)90452-1.

Choksi, N., S. Lebrun, M. Nguyen, A. Daniel, G. DeGeorge, J. Willoughby, A. Layton, D. Lowther, J. Merrill, J. Matheson, J. Barroso, K. Yozzo, W. Casey, and D. Allen. 2020. Validation of the OptiSafe™ eye irritation test. *Cutan. Ocul. Toxicol.*, 39(3): 180-192. Doi: 10.1080/15569527.2020.1787431.

Cottrez, F., V. Leblanc, E. Boitel, H. Groux, and N. Alépée. 2021. The EyeIRR-IS assay: Development and evaluation of an in vitro assay to measure the eye irritation sub-categorization of liquid chemicals. *Toxicol. In Vitro*, 71: 105072. Doi: 10.1016/j.tiv.2020.105072.

Donahue, D.A., L.E. Kaufman, J. Avalos, F.A. Simion, and D.R. Cerven. 2011. Survey of ocular irritation predictive capacity using Chorioallantoic Membrane Vascular Assay (CAMVA) and Bovine Corneal Opacity and Permeability (BCOP) test historical data for 319 personal care products over fourteen years. *Toxicol. In Vitro*, 25(2): 563-572. Doi: 10.1016/j.tiv.2010.12.003.

ECHA. 2017. *Guidance on the Application of the CLP Criteria. Version 5.0. Reference ECHA-17-G-21-EN. Doi: 10.2823/124801. Available at: <https://echa.europa.eu/guidance-documents/guidance-on-clp>.*

ESAC. 2019. *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. Statement of the ECVAM Scientific Advisory Committee (ESAC) of 9th July 2009. Available at: https://ec.europa.eu/jrc/sites/jrcsh/files/esac31_lveta_20090922.pdf.*

EURL ECAM. 2004a. *Tracking System for Alternative Methods Towards Regulatory Acceptance (TSAR). Method TM2004-01. The cytosensor microphysiometer toxicity test. Available at: <https://tsar.jrc.ec.europa.eu/test-method/tm2004-01>.*

EURL ECAM. 2004b. *Tracking System for Alternative Methods Towards Regulatory Acceptance (TSAR). Method TM2004-03. Neutral Red Release Assay. Available at: <https://tsar.jrc.ec.europa.eu/test-method/tm2004-03>.*

Frentz, M., M. Goss, M. Reim, and N.F. Schrage. 2008. *Repeated exposure to benzalkonium chloride in the Ex Vivo Eye Irritation Test (EVEIT): observation of isolated corneal damage and healing. Altern. Lab. Anim., 36(1): 25-32. Doi: 10.1177/026119290803600105.*

Harbell, J.W., R. Osborne, G.J. Carr, and A. Peterson. 1999. *Assessment of the Cytosensor Microphysiometer Assay in the COLIPA In Vitro Eye Irritation Validation Study. Toxicol. In Vitro, 13(2): 313-323. Doi: 10.1016/s0887-2333(98)00090-3.*

Hartung, T., L. Bruner, R. Curren, C. Eskes, A. Goldberg, P. McNamee, L. Scott, and V. Zuang. 2010. *First alternative method validated by a retrospective weight of evidence approach to replace the Draize eye test for the identification of non-irritant substances for a defined applicability domain. ALTEX, 27(1): 43-51. Doi: 10.14573/altex.2010.1.43.*

ICCVAM. 2007. *ICCVAM test method evaluation report: in vitro ocular toxicity test methods for identifying ocular severe irritants and corrosives. NIH Publication No. 07-4517. National institute of environmental health sciences, research Triangle Park, North Carolina, USA.*

ICCVAM. 2010. *ICCVAM test method evaluation report: current validation status of in vitro test methods proposed for identifying eye injury hazard potential of chemicals and products. NIH Publication No. 10-7553. National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA.*

Kandarova, H., S. Letasiova, E. Adriaens, R. Guest, J.A. Willoughby Sr., A. Drzewiecka, K. Gruszka, N. Alépée, S. Verstraelen, and A.R. Van Rompay. 2018. *CON4EI: CONSortium for in vitro Eye Irritation testing strategy - EpiOcular™ time-to-toxicity (EpiOcular ET-50) protocols for hazard identification and labelling of eye irritating chemicals. Toxicol. In Vitro, 49: 34-52. Doi: 10.1016/j.tiv.2017.08.019.*

Nash, J.R., G. Mun, H.A. Raabe, and R. Curren. 2014. *Using the cytosensor microphysiometer to assess ocular toxicity. Curr. Protoc. Toxicol. 61: 1.13.1-11. Doi: 10.1002/0471140856.tx0113s61.*

Piehl, M., A. Gilotti, A. Donovan, G. DeGeorge, and D. Cerven. 2010. *Novel cultured porcine corneal irritancy assay with reversibility endpoint. Toxicol. In Vitro 24: 231-239. Doi:10.1016/j.tiv.2009.08.033.*

Piehl, M., M. Carathers, R. Soda, D. Cerven, and G. DeGeorge. 2011. *Porcine corneal ocular reversibility assay (PorCORA) predicts ocular damage and*

recovery for global regulatory agency hazard categories. *Toxicol. In Vitro*, 25: 1912-1918. Doi:10.1016/j.tiv.2011.06.008.

Reader, S.J., V. Blackwell, R. O'Hara, R.H. Clothier, G. Griffin, and M. Balls. 1989. A vital dye release method for assessing the short-term cytotoxic effects of chemicals and formulations. *Altern. Lab. Anim.*, 17: 28-33. Doi: 10.1177/026119298901700106.

Reader, S.J., V. Blackwell, R. O'Hara, R.H. Clothier, G. Griffin, and M. Balls. 1990. Neutral red release from pre-loaded cells as an in vitro approach to testing for eye irritancy potential. *Toxicol. In Vitro*, 4(4-5): 264-266. Doi: 10.1016/0887-2333(90)90060-7.

Settivari, R.S., R.A. Amado, M. Corvaro, N.R. Visconti, L. Kan, E.W. Carney, D.R. Boverhof, and S.C. Gehen. 2016. Tiered application of the neutral red release and EpiOcular™ assays for evaluating the eye irritation potential of agrochemical formulations. *Regul. Toxicol. Pharmacol.*, 81: 407-420. Doi: 10.1016/j.yrtph.2016.09.028.

Spielmann, H., S. Kalweit, M. Liebsch, T. Wirnsberger, I. Gerner, E. Bertram-Neis, K. Krauser, R. Kreiling, H.G. Miltenburger, W. Pape, and W. Steiling. 1993. Validation study of alternatives to the Draize eye irritation test in Germany: Cytotoxicity testing and HET-CAM test with 136 industrial chemicals. *Toxicol. In Vitro*, 7(4): 505-510. Doi: 10.1016/0887-2333(93)90055-a.

Spielmann, H., M. Liebsch, S. Kalweit, F. Moldenhauer, T. Wirnsberger, H.-G. Holzhütter, B. Schneider, S. Glaser, I. Gerner, W.J.W. Pape, R. Kreiling, K. Krauser, H.G. Miltenburger, W. Steiling, N.P. Luepke, N. Müller, H. Kreuzer, P. Mürmann, J. Spengler, E. Bertram-Neis, B. Siegemund, and F.J. Wiebel. 1996. Results of a validation study in Germany on two in vitro alternatives to the Draize eye irritation test, HET-CAM test and the 3T3 NRU cytotoxicity test. *Altern. Lab. Anim.*, 24: 741-858.

Spöler, F., M. Först, H. Kurz, M. Frentz, and N.F. Schrage. 2007. Dynamic analysis of chemical eye burns using high-resolution optical coherence tomography. *J. Biomed. Opt.*, 12: 041203. Doi:10.1117/1.2768018.

Spöler, F., O. Kray, S. Kray, C. Panfil, and N.F. Schrage. 2015. The Ex Vivo Eye Irritation Test as an alternative test method for serious eye damage/eye irritation. *Altern. Lab. Anim.*, 43(3): 163-179. Doi: 10.1177/026119291504300306.

Whittle, E., D. Basketter, M. York, L. Kelly, T. Hall, J. McCall, P. Botham, D. Esdaile, and J. Gardner. 1992. Findings of an interlaboratory trial of the enucleated eye method as an alternative eye irritation test. *Toxicol. Mech. Methods.*, 2: 30-41.

Young, J.R., M.J. How, A.P. Walker, and W.M. Worth. 1988. Classification as corrosive or irritant to skin of preparations containing acidic or alkaline substances, without testing on animals. *Toxicol. In Vitro*, 2(1): 19-26. Doi: 10.1016/0887-2333(88)90032-x.

Zuang, V. 2001. The neutral red release assay: a review. *Altern. Lab. Anim.*, 29(5): 575-599. Doi: 10.1177/026119290102900513.

Chapter 3.4

- 3.4.2.1.1.3 Replace “approach” with “assessment” twice.
- 3.4.2.2.1.2 Replace “3.4.2.2.1.3” with “3.4.2.2.2 to 3.4.2.2.6”.
- 3.4.2.2.1.3 Replace with the following:
 “3.4.2.2.1.3 For classification of skin sensitizers, all available and relevant information is collected and its quality in terms of adequacy and reliability is assessed. Classification should be based on mutually acceptable data/results generated using methods and/or defined approaches that are validated according to international procedures. These include both OECD guidelines and equivalent methods/defined approaches (see 1.3.2.4.3). Sections 3.4.2.2.2 to 3.4.2.2.6 provide classification criteria for the different types of information that may be available.”
- Table 3.4.2 Delete.
- 3.4.2.2.1.4 and 3.4.2.2.1.5 Insert the following two new paragraphs:
 “3.4.2.2.1.4 A tiered approach (see 3.4.2.2.7) organizes the available information on skin sensitization into levels/tiers and provides for decision-making in a structured and sequential manner. Classification results directly when the information consistently satisfies the criteria. However, where the available information gives inconsistent and/or conflicting results within a tier, classification of a substance or a mixture is made on the basis of the weight of evidence within that tier. In some cases when information from different tiers gives inconsistent and/or conflicting results (see 3.4.2.2.7.7) or where data individually are insufficient to conclude on the classification, an overall weight of evidence assessment is used (see 1.3.2.4.9 and 3.4.2.2.7.6).
 3.4.2.2.1.5 Guidance on the interpretation of criteria and references to relevant guidance documents are provided in 3.4.5.3.”
- 3.4.2.2.2 Amend the heading to read as follows: “*Classification based on human data (tier 1 in figure 3.4.1)*”
- 3.4.2.2.2.1 (new) Insert the following new paragraph and renumber subsequent paragraphs accordingly:
 “3.4.2.2.2.1 A substance is classified as a skin sensitizer in Category 1 if there is evidence in humans that the substance can lead to sensitization by skin contact in a substantial number of persons.”
- 3.4.2.2.2.2 (former 3.4.2.2.2.1) Amend the beginning of the paragraph to read as follows:
 “Substances showing a high frequency of occurrence in humans, can be presumed to have the potential to produce significant sensitization and are classified in sub-category 1A. Severity of reaction may also be considered. Human evidence for sub-category 1A can include”
- 3.4.2.2.2.3 (former 3.4.2.2.2.2) Amend the beginning of the paragraph to read as follows:
 “Substances showing a low to moderate frequency of occurrence in humans can be presumed to have the potential to produce sensitization and are classified in sub-category 1B. Severity of reaction may also be considered. Human evidence for sub-category 1B can include:”
- 3.4.2.2.3 Amend the heading to read as follows: “*Classification based on standard animal data (tier 1 in figure 3.4.1)*”
- 3.4.2.2.3.1 Amend the paragraph and insert a new table 3.4.2 to read as follows:
 “3.4.2.2.3.1 A substance is classified as a skin sensitizer if there are positive results from an appropriate animal test. For Category 1, when an adjuvant type

test method for skin sensitization is used, a response of at least 30% of the animals is considered as positive. For a non-adjuvant Guinea pig test method a response of at least 15% of the animals is considered positive. For Category 1, a stimulation index of three or more is considered a positive response in the radioisotopic local lymph node assay (LLNA). For the non-radioactive modifications to the LLNA, a stimulation index of 1.8 or more in the LLNA: DA, 1.6 or more in the LLNA: BrdU-ELISA, and 2.7 or more in the LLNA: BrdU-FCM are considered positive. Test methods for skin sensitization are described in the OECD Guideline 406 (the Guinea Pig Maximisation test and the Buehler guinea pig test) and in guidelines 429/442A/442B (Local Lymph Node Assays). Other methods may be used provided that they are well-validated and scientific justification is given. The Mouse Ear Swelling Test (MEST), appears to be a reliable screening test to detect moderate to strong sensitizers, and can be used as a first stage in the assessment of skin sensitization potential.

Table 3.4.2: Animal test results for Category 1

Assay	Criteria
Local lymph node assay	SI \geq 3
Local lymph node assay: DA	SI \geq 1.8
Local lymph node assay: BrdU-ELISA	SI \geq 1.6
Local lymph node assay: BrdU-FCM	SI \geq 2.7
Adjuvant Guinea pig test method	\geq 30% responding at any intradermal induction dose
Non-adjuvant Guinea pig test method	\geq 15% responding at any topical induction dose

3.4.2.2.3.2 Amend to read as follows (current table 3.4.3 remains unchanged):

“3.4.2.2.3.2 Substances showing a high potency in animals, can be presumed to have the potential to produce significant sensitization in humans and are classified in sub-category 1A. Severity of reactions may also be considered. Animal test results for sub-category 1A can include data with values indicated in table 3.4.3 below:”

Note to table 3.4.3 Add the following new note under current table 3.4.3:

“Note: For the LLNA: BrdU-ELISA, sub-categorization criteria (1A: EC1.6 value \leq 6%, 1B: EC1.6 value $>$ 6%, Maeda and Takeyoshi, 2019; Kobayashi et al., 2020) have been proposed and validated by OECD, but no sub-categorization criteria have yet been agreed internationally. Validated sub-categorization criteria may still be accepted by some competent authorities. A competent authority may decide which sub-categorization criteria, if any, should be applied for these test methods.

As for the LLNA: DA and LLNA: BrdU-FCM, there are currently no validated and internationally agreed criteria for subcategorization of skin sensitizers. Therefore, these test methods can only be used to conclude on either classification in category 1 or no classification.”

3.4.2.2.3.3 Amend to read as follows (current table 3.4.4 remains unchanged):

“3.4.2.2.3.3 Substances showing a low to moderate potency in animals, can be presumed to have the potential to produce significant sensitization in humans and are classified in sub-category 1B. Severity of reactions may also be considered. Animal test results for sub-category 1B can include data with values indicated in table 3.4.4 below:”

Note to table 3.4.4 Add the following new note under current table 3.4.4:

“Note: For the LLNA: BrdU-ELISA, sub-categorization criteria (1A: EC1.6 value \leq 6%, 1B: EC1.6 value $>$ 6%, Maeda and Takeyoshi, 2019; Kobayashi et al., 2020) have been proposed and validated by OECD, but no sub-categorization criteria have

yet been agreed internationally. Validated sub-categorization criteria may still be accepted by some competent authorities. A competent authority may decide which sub-categorization criteria, if any, should be applied for these test methods.

As for the LLNA: DA and LLNA: BrdU-FCM, there are currently no validated and internationally agreed criteria for subcategorization of skin sensitizers. Therefore, these test methods can only be used to conclude on either classification in category 1 or no classification.”.

- 3.4.2.2.4 Replace 3.4.2.2.4 (heading) and paragraphs 3.4.2.2.4.1 to 3.4.2.2.4.3 (including the related new footnotes 3 and 4) with the following:

“3.4.2.2.4 Classification based on defined approaches (tier 1 or tier 2 in figure 3.4.1)

3.4.2.2.4.1 Defined approaches consist of a rule-based combination of data obtained from a predefined set of different information sources (e.g. *in chemico* methods, *in vitro* methods, physico-chemical properties, non-test methods). It is recognized that most single non animal methods are not able to replace *in vivo* methods fully for most regulatory endpoints. Thus, defined approaches can be useful strategies of combining data for classifying substances and mixtures. Results obtained with a defined approach validated according to international procedures, such as OECD Guideline 497 or an equivalent approach, are conclusive for classification for skin sensitization if the criteria of the defined approach are fulfilled (see table 3.4.7)³. Data from a defined approach can only be used for classification when the tested substance is within the applicability domain of the defined approach used. Additional limitations described in the published literature should also be taken into consideration.

3.4.2.2.4.2 Where the results from defined approaches are assigned a level of confidence as for example in OECD Guideline 497, a low confidence outcome of a defined approach in tier 1 is inconclusive and thus cannot be used on its own to classify but may be considered in combination with other data in tier 2.

3.4.2.2.4.3 Individual evidence used within a defined approach should not also be used outside of that defined approach.

3.4.2.2.5 Classification based on *in chemico/in vitro* data (tier 1 or tier 2 in figure 3.4.1)

3.4.2.2.5.1 The currently available *in chemico/in vitro* methods address specific biological mechanisms leading to the acquisition of skin sensitization as described, for example, in the OECD Adverse Outcome Pathway for Skin Sensitisation (see OECD (2014)). Individual test methods that are validated according to international procedures and are accepted as stand-alone methods, can be used to conclude on the classification in tier 1. A competent authority may decide whether to use the method described in Appendix III to OECD Test Guideline 442C as a stand-alone method to discriminate between sub-category 1A and those not categorized as sub-category 1A (see 3.4.5.3.5).

3.4.2.2.5.2 Other non stand-alone *in chemico/in vitro* methods that are validated according to international procedures such as OECD test guidelines 442C (Annex I and II), 442D and 442E, are accepted as supportive evidence and should within tier 1 only be used in combination with other types of data in defined approaches. The use of these methods in tier 2 is described in 3.4.2.2.7.5.

3.4.2.2.5.3 Other validated *in chemico/in vitro* test methods accepted by some competent authorities are described in 3.4.5.3.6.2⁴. A competent

authority may decide which classification criteria, if any, should be applied for these test methods to conclude on classification.

3.4.2.2.5.4 *In chemico/in vitro* data can only be used for classification when the tested substance is within the applicability domain of the test method(s) used. Additional limitations described in the published literature should also be taken into consideration.

3.4.2.2.6 *Classification based on non-test methods (tier 2 in figure 3.4.1)*

3.4.2.2.6.1 Classification, including the conclusion not classified, can be based on non-test methods, with due consideration of reliability and applicability, on a case-by-case basis. Non-test methods include computer models predicting qualitative structure activity relationships (structural alerts, SAR) or quantitative structure-activity relationships (QSARs), computer expert systems, and read-across using analogue and category approaches.

3.4.2.2.6.2 Read-across using analogue or category approaches requires sufficiently reliable test data on similar substance(s) and justification of the similarity of the tested substance(s) with the substance to be classified. Where adequate justification of the read-across approach is provided, it has in general higher weight than (Q)SARs.

3.4.2.2.6.3 Classification based on (Q)SARs requires sufficient data and validation of the model. The validity of the computer models and the prediction should be assessed using internationally recognized principles for the validation of (Q)SARs. With respect to reliability, lack of alerts in a SAR or expert system is not sufficient evidence for no classification.

3.4.2.2.6.4 For conclusions on no classification from read-across and (Q)SARs the adequacy and robustness of the scientific reasoning and of the supporting evidence should be well substantiated and normally requires multiple negative substances with good structural and physical (related to toxicokinetics) similarity to the substance being classified, as well as a clear absence of positive substances with good structural and physical similarity to the substance being classified.

3.4.2.2.7 *Classification in a tiered approach (figure 3.4.1)*

3.4.2.2.7.1 A tiered approach to the evaluation of information should be considered, where applicable (figure 3.4.1), recognizing that not all tiers as well as information within a tier may be relevant. However, all available and relevant information of sufficient quality needs to be examined for consistency with respect to the resulting classification.

3.4.2.2.7.2 Tier 1 - Classification based on human data, standard animal data, defined approaches or stand-alone *in chemico/in vitro* methods

For classification of a substance, evidence in tier 1 may include data from any or all of the following lines of evidence. Where information from data within tier 1 is inconsistent and/or conflicting, the conclusion is determined in a weight of evidence assessment:

- (a) Experimental studies in humans (e.g., predictive patch testing, HRIPT, HMT (see paragraph 1.3.2.4.7, criteria in 3.4.2.2.2 (a) and 3.4.2.2.3 (a) and guidance in 3.4.5.3.2);
- (b) Epidemiological studies (e.g., case control studies, prospective studies) assessing allergic contact dermatitis (see paragraph 1.3.2.4.7, criteria in 3.4.2.2.2 (b) and (c) and 3.4.2.2.3 (b) and (c) and guidance in 3.4.5.3.2);

- (c) Well-documented cases of allergic contact dermatitis (see criteria in 3.4.2.2.2.2 (b) and 3.4.2.2.2.3 (b) and guidance in 3.4.5.3.2);
- (d) Appropriate animal studies (see criteria in 3.4.2.2.3, and guidance in 3.4.5.3.3);
- (e) Defined approaches validated according to international procedures (see 3.4.2.2.4, guidance in 3.4.5.3.4 and table 3.4.7);
- (f) Stand-alone *in chemico*/in vitro methods validated according to international procedures (see 3.4.2.2.5, guidance in 3.4.5.3.5 and table 3.4.8).

3.4.2.2.7.3 Tier 2 - Classification based on inconclusive data from tier 1, non stand-alone *in chemico*/in vitro methods or non-test methods.

In case a definitive conclusion on classification, including sub-categorization where required by a competent authority, cannot be derived from tier 1, additional lines of evidence shall be considered in a weight of evidence assessment in tier 2. These may include:

- (a) Data from non stand-alone *in chemico*/in vitro methods (see 3.4.2.2.5 and 3.4.5.3.5);
- (b) Data from non-test methods (see 3.4.2.2.6).

3.4.2.2.7.4 Evidence from non stand-alone *in chemico*/in vitro methods and from non-test methods should not be considered at this stage if this data is already used in a defined approach under 3.4.2.2.7.2.

3.4.2.2.7.5 Individual non stand-alone *in chemico*/in vitro methods validated according to international procedures and non-test methods (including read-across) can be applied in a weight of evidence assessment together with inconclusive data from tier 1 and should be used in this second tier because they can usually not be used as stand-alone (with the exception of good quality read-across). However, a competent authority may decide that a positive result with one of these non stand-alone *in chemico*/in vitro methods, may be used on its own to classify in Category 1 (see table 3.4.8).

3.4.2.2.7.6 Tier 3 - Classification based on overall weight of evidence assessment including additional indicators

In case a definitive conclusion on classification including sub-categorization where required by a competent authority cannot be derived from the previous tiers, an overall weight of evidence assessment using expert judgment should be used that may include a combination of two or more indicators of skin sensitization as listed below.

- (a) Isolated episodes of allergic contact dermatitis;
- (b) Epidemiological studies of limited power, e.g. where chance, bias or confounders have not been ruled out fully with reasonable confidence;
- (c) Data from animal tests, performed according to existing guidelines, which do not meet the criteria for a positive result described in 3.4.2.2.3, but which are sufficiently close to the limit to be considered significant;
- (d) Data from non-standard methods.

3.4.2.2.7.7 Where information from the various tiers is inconsistent and/or conflicting with respect to the resulting classification, information of sufficient quality from a higher tier is generally given a higher weight than information from a lower tier. However, when information from a lower tier would result in a stricter classification than information from a higher tier and there is concern for misclassification, then classification is determined by an overall weight of evidence assessment (i.e. in tier 3). For example, having consulted the guidance in 3.4.5.3 as appropriate, classifiers concerned with a negative result for skin sensitization in a Buehler study when there is a clear positive result in humans for very similar substances (from read-across) would utilise an overall weight of evidence assessment.”

3.4.2.2.4.4 Current paragraph 3.4.2.2.4.4 “Immunological contact urticaria” becomes new section 3.4.2.2.8. Renumber the two sub-paragraphs under this heading as follows:

3.4.2.2.8.1 Substances meeting the criteria for...”

3.4.2.2.8.2 There is no recognized animal model available...”

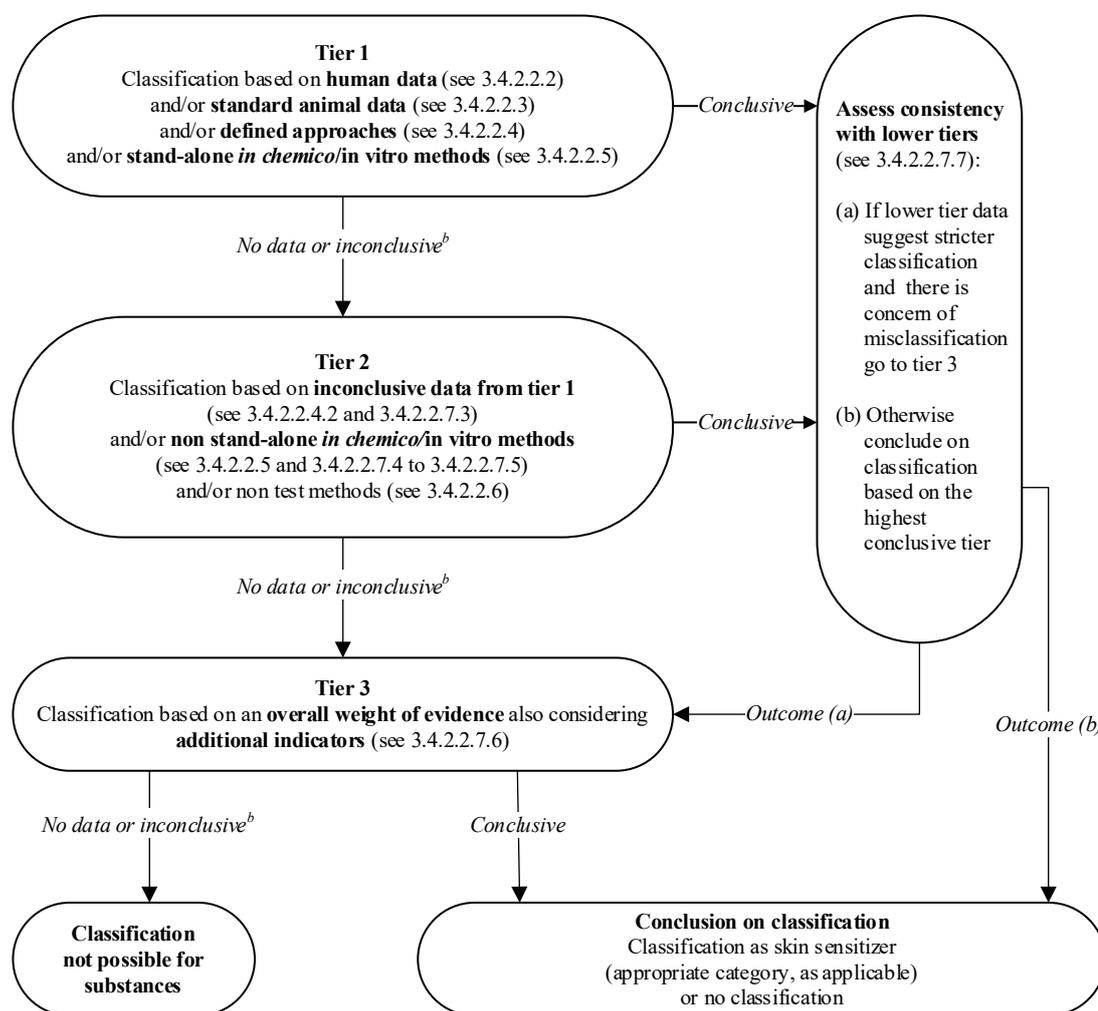
Footnotes 3 and 4 Insert the following new footnotes 3 and 4 at the bottom of the page in relation to paragraphs 3.4.2.2.4.1 (for footnote 3) and 3.4.2.2.5.3 (for footnote 4):

“³ Additional defined approaches have been proposed for skin sensitization (OECD 2017) but no classification criteria have yet been agreed internationally.”

“⁴ Additional in chemico/in vitro methods have been proposed for skin sensitization (see 3.4.5.3.6.2) but no classification criteria have yet been agreed internationally.”

Figure 3.4.1 Insert the following new figure 3.4.1 and related notes after section 3.4.2.2.8 (former 3.4.2.2.4.4):

Figure 3.4.1: Application of the tiered approach for skin sensitization^a



^a Before applying the approach, the explanatory text in 3.4.2.2.7 as well as the guidance in 3.4.5.3 should be consulted. Only adequate and reliable data of sufficient quality should be included in applying the tiered approach.

^b Information may be inconclusive for various reasons, e.g.:

- The available data may be of insufficient quality, or otherwise insufficient/inadequate for the purpose of classification, e.g. due to quality issues related to experimental design and/or reporting;
- Where competent authorities make use of the skin sensitization sub-categories 1A and 1B, the available data may not be capable of distinguishing between sub-category 1A and sub-category 1B.”

3.4.3.1 In the first sentence replace “evaluation” with “assessment”.

3.4.3.2.6 In the last sentence replace “by testing” with “based on test data” and “assigned” with “classified in”.

3.4.3.2.7 At the beginning of the sentence replace “aerosol form” with “aerosolized form”.

3.4.4.1 Amend the last sentence of the paragraph to read as follows: “Table 3.4.6 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”

- 3.4.5 Amend the heading to read as follows: “**Decision logic and guidance**”
- 3.4.5.1 Amend decision logic 3.4.1 as follows:
- In the second right hand text box from the top, replace “Category 1⁵” with “Category 1⁷”.
- In the bottom left hand text box starting with “Does the mixture contain...”, amend the footnote references “3, 4” to read: “5, 6”.
- Current footnotes “3”, “4” and “5” become “5”, “6” and “7” respectively.
- 3.4.5.2 Amend decision logic 3.4.2 as follows:
- Replace the text in the central box currently starting with “(a) Is there evidence...” with the following: “Is there evidence that the substance/mixture fulfils the criteria as described in 3.4.2.2.2.2 to 3.4.2.2.2.8 for substances and in 3.4.3.1 for mixtures?”.
- In the second right hand text box from the top, replace “Category 1⁶” with “Category 1⁸”.
- In the bottom left hand text box starting with “Does the mixture contain...”, amend footnote references “3, 4” to read: “5, 6”.
- Current footnotes “3”, “4” and “6” become “5”, “6” and “8”.
- 3.4.5.3 Insert the following new text:
- “3.4.5.3 *Background guidance***
- 3.4.5.3.1 *Relevant guidance documents*
- Mechanistic information on the process of skin sensitization is available in the OECD document on the Adverse Outcome Pathway for skin sensitization (see OECD (2014)). This information can be helpful in understanding the value of the individual *in chemico* and *in vitro* methods compared to the *in vivo* methods.
- 3.4.5.3.2 *Guidance on the use of human data*
- 3.4.5.3.2.1 The classification of a substance can be based on human evidence generated from a variety of sources. These sources include human predictive patch testing, epidemiological studies, case studies, case reports or histories, diagnostic patch testing and medical surveillance reports, and poison control centre information. This data may have been generated for consumers, workers, or the general population. When considering human evidence, consideration should be given to the size, exposure level, and exposure frequency of the exposed population. Guidance for evaluating human evidence and the criteria in 3.4.2.2.2 is provided by some competent authorities (e.g. ECHA Guidance on the Application of the CLP Criteria, 2017).
- 3.4.5.3.2.2 Positive data from predictive patch testing (HRIPT or HMT) conducted through human experimental and clinical studies, showing allergic contact dermatitis caused by the test substance can be used to classify substances for skin sensitization. These studies are generally conducted in controlled clinical settings and in general the larger the population size, the more reliable the study outcome is. Criteria for evaluating this data are provided in 3.4.2.2.2.1 and 3.4.2.2.2.
- 3.4.5.3.2.3 Positive data from well-run epidemiological studies (in accordance with WHO CIOMS guidelines, 2009) can be used for classifying substances for skin sensitization. Some examples of epidemiological studies may include case control studies, cohort studies, cross-sectional studies, or longitudinal studies. These studies should have large sample sizes with well-documented exposures to a substance.

3.4.5.3.2.4 A specific type of epidemiological study (such as randomized control studies or trials) may include information from diagnostic patch testing. Diagnostic patch testing is considered by some competent authorities to be the gold standard in diagnosing contact allergy in dermatitis patients (Johansen et al, 2015). Importantly, due consideration needs to be given to the appropriate selection of vehicle, substance and patch test concentrations for the purpose of not causing false negatives, false positives, irritant reactions or inducing contact allergy (skin sensitization). Positive data from experimental/clinical/diagnostic studies in humans and/or well-documented episodes of allergic contact dermatitis may be used to classify substances for skin sensitization, when it can be assumed with sufficient likelihood that the tested substance was indeed the most likely cause for induction of sensitization. Therefore, it should be established that there is at least a general likelihood that the respective patient(s) had been previously exposed to the substance. On the other hand, negative results from such tests are not sufficient to prove that the test substance should not be classified as a skin sensitizer.

3.4.5.3.2.5 Human data not generated in controlled experiments with volunteers for the purpose of hazard classification (e.g. case studies, case reports and case histories, and poison control centre information) can be used with caution. Consideration should be given to the frequency of cases, the inherent properties of the substances, as well as factors such as the exposure situation, bioavailability, individual predisposition, cross-reactivity and preventive measures taken.

3.4.5.3.2.6 Special consideration should be given to negative human data as full dose-response information is generally not available. For example, a negative result in an HRIPT or HMT at a low concentration may not allow for the conclusion that the substance does not have skin sensitizing properties as such effect at a higher concentration may not be excluded. In addition, negative human data should not necessarily be used to negate positive results from animal studies and/or defined approaches, but can be used as part of a weight of evidence assessment. For both animal and human data, consideration should be given to the impact of the vehicle (e.g. Wright et al, 2001 and Kligman, 1966).

3.4.5.3.2.7 For example, negative results from substances tested in a predictive patch test at a DSA (dose per skin area) $< 500 \mu\text{g}/\text{cm}^2$ imply that a classification for skin sensitization might not be needed at all, however, classification as sub-category 1A or 1B cannot be ruled out, because the concentration tested was not high enough to exclude these possibilities. The same holds for test results for which it is unknown whether the test concentration corresponded to a DSA $< 500 \mu\text{g}/\text{cm}^2$. Negative results from substances tested at a DSA $\geq 500 \mu\text{g}/\text{cm}^2$ suggest that classification might not be needed. However, while classification as sub-category 1A can be ruled out, classification as sub-category 1B cannot, because a higher test concentration might have resulted in a positive test result. However, a negative test result at a concentration of 100% can justify no classification (based on this test). Nevertheless, negative results at low concentrations may be informative for mixtures containing the substance at similar or lower concentrations.

3.4.5.3.3 *Guidance on the use of standard animal data*

3.4.5.3.3.1 The most common assays used for dermal sensitization testing in animals are the Local Lymph Node Assay (LLNA, OECD test guidelines 429 and 442A and 442B), the Guinea Pig Maximization Test (GPMT, OECD Test Guideline 406) and the Buehler test (OECD Test Guideline 406). When evaluating the quality of the study, consideration should be given, as relevant, to the strain of the mouse and guinea pig used, the number, age, and sex of the animals, and the test conditions used (e.g., preparation of patch test site, dose level selection, chemical preparation, positive and negative test controls).

3.4.5.3.3.2 OECD test guidelines for the LLNA include the radioactive assay (OECD Test Guideline 429) and non-radioactive assays (OECD Test Guideline 442A and 442B; LLNA:DA, LLNA:BrdU-ELISA, and LLNA:BrdU-FCM). In these tests, sensitizers are characterized by increasing the group mean stimulation index (“SI”, a measure of lymph node proliferation) in treated groups versus concurrent vehicle controls by more than a predefined critical value which is different for each form of the LLNA (e.g., $SI \geq 3$ for the radioactive LLNA, $SI \geq 1.6$ for the LLNA:BrdU-ELISA). For sensitizers, sub-categorization is performed based on the effective concentration (EC) causing an increase in SI of exactly the critical magnitude (e.g. the EC3 under OECD Test Guideline 429 is the concentration leading to an exactly threefold increase in group mean SI versus control).

3.4.5.3.3.3 The respective OECD test guidelines for the different LLNA variants specify that a pre-screen test should be undertaken to determine the highest concentration to be tested. If such a test has not been performed and the LLNA was carried out with a test concentration $< 100\%$, a rationale (e.g. based on solubility, local or systemic toxicity, see OECD test guidelines 429, and 442A and 442B) needs to be provided that the highest test concentration represents the maximum testable concentration. Otherwise, the reliability of a negative test result has to be considered compromised.

3.4.5.3.3.4 EC values are normally obtained by interpolation between adjacent test concentrations, i.e. between the highest test concentration causing an SI below, and the lowest test concentration causing an SI above the critical value. However, care must be taken when the EC value falls below the lowest concentration tested and can therefore only be estimated by extrapolation, which is associated with additional uncertainty. In some cases, the SI at the highest concentration tested falls only slightly below the critical SI value, which raises the question of upward extrapolation (unless the maximum testable concentration has been applied). These and other issues regarding the reliability of LLNA results are further discussed in Ryan et al. (2007) and Annex 3 of OECD Series on Testing and Assessment No. 336 (Supporting Document to OECD Guideline Document 497), which also provides a highly curated database of test guidelines 429 LLNA EC3 values.

3.4.5.3.3.5 Further limitations have been identified for the radioactive and non-radioactive LLNAs. For example, substances containing certain functional groups may interfere with the accuracy of the assay. These limitations as well as the possibility of borderline positive results are described in OECD test guidelines 429, and 442A and 442B. Variability in EC values for the same substance may also be the result of the vehicle used. For example, analysis has shown an underestimation of potency (i.e., higher EC3 values) with predominantly aqueous vehicles or propylene glycol (see Jowsey, 2008).

3.4.5.3.3.6 For OECD Test Guideline 406, the concentration of test chemical used for each induction exposure should be systemically well-tolerated using the highest dose to cause mild-to-moderate skin irritation. The concentration used for the challenge exposure should be the highest non-irritant dose. A positive result in a guinea pig test is defined as a grade above zero according to the applicable grading scale such as the Magnusson and Kligman grading scale for OECD Test Guideline 406 at one or more of the two observation time-points. A grade of 0.5, which is sometimes reported, is therefore also considered a positive result.

3.4.5.3.4 *Guidance on the use of defined approaches*

Defined approaches validated according to international procedures and described in OECD Guideline 497 have been characterized for the level of confidence that can be assigned to the predictions based on the applicability domain of the individual information sources used and the data interpretation procedure applied (see table 3.4.7). Other defined approaches

under consideration but not yet validated according to international procedures and described in OECD Guidance Document 256 according to internationally agreed criteria for their reporting (OECD Guidance Document 255) may be accepted by some competent authorities.

3.4.5.3.5 *Guidance on the use of non stand-alone in chemico/in vitro methods*

Individual *in chemico/in vitro* methods such as those reported in OECD test guidelines 442C, 442D and 442E, due to their limited mechanistic coverage, cannot be used on their own to conclude on Category 1 or no classification according to the criteria defined in table 3.4.8 and further data are necessary for classification in tier 2. In addition, although some of these methods provide quantitative information, these cannot be used for the purposes of subcategorization into sub-categories 1A and 1B since the criteria have not been validated according to international procedures. Nevertheless, such quantitative information may be accepted by a competent authority when used in a weight of evidence assessment under tier 2 for the purpose of subcategorization. This is also in line with the statement in these test guidelines that “*Depending on the regulatory framework, positive results generated with these methods may be used on their own to classify a chemical into UN GHS Category 1.*” Therefore, the GHS also allows a competent authority to decide that a positive result with one of these non stand-alone *in chemico/in vitro* methods, may be used on its own to classify in category 1 and whether Test Guideline 442C (Appendix III) kinetic Direct Peptide Reactivity Assay (kDPRA) can be used to differentiate between sub-category 1A and no sub-category 1A.

3.4.5.3.6 *Guidance on the use of non-standard data*

3.4.5.3.6.1 Validated but not yet adopted *in chemico/in vitro* methods such as those reported under 3.4.5.3.6.2 as well as *in vivo* test methods which do not comply with internationally agreed guidelines for the identification of skin sensitizers or the assessment of skin sensitizing potency may provide supportive evidence when used in an overall weight of evidence assessment (i.e. tier 3).

3.4.5.3.6.2 A non-exhaustive list of other validated *in chemico/in vitro* test methods accepted by some competent authorities but not adopted as OECD test guidelines is provided below. A competent authority may decide which classification criteria, if any, should be applied for these test methods:

- (a) The Genomic Allergen Rapid Detection (GARD) potency is a transcriptomics-based *in vitro* assay addressing the third key event of the skin sensitization Adverse Outcome Pathway (activation of dendritic cells) similar to the GARD_{skin} but uses a different gene signature that provides sub-categorization of skin sensitizers (Gradin et al., 2020; Zeller et al., 2017; Corsini et al. 2021).
- (b) The SENS-IS assay is a genomic approach applied to a Reconstructed Human Epidermis (RhE) (Cottrez et al., 2015; Cottrez et al., 2016).
- (c) The Epidermal Sensitization Assay (EpisensA) is based on the measurement of the upregulation of four genes in a reconstructed human epidermis (RhE) to discriminate between sensitizers and non-sensitizers (Saito et al., 2017).

3.4.5.3.7 *Guidance on the weight of evidence assessment for classifying substances and mixtures for skin sensitization*

3.4.5.3.7.1 There may be situations where results from tests and/or non-test methods are available but disagree with each other with respect to the classification. In these situations, the tiered approach to classification for skin sensitization requires a weight of evidence assessment consistent with the

principles elaborated in sections 1.3.2.4.2 and 1.3.2.4.9 on test data quality and weight of evidence, respectively. In addition, some guidance on the weight of evidence assessment specific for skin sensitization is provided below which can be applied when the general principles do not result in a conclusion on the classification. It should be noted that human and animal results for a substance obtained at low concentrations may still be informative for classifying a mixture containing the substance at similar or lower concentrations.

3.4.5.3.7.2 Mutual compatibility of study results

3.4.5.3.7.2.1 In cases where results are in disagreement with each other (e.g., not classified versus Category 1, sub-category 1A or 1B; sub-category 1A versus 1B), a weight of evidence assessment becomes necessary. However, less obvious situations may also occur such as where certain studies may point to not classified or sub-category 1B, while it cannot be excluded that a stricter classification might have resulted under a different dosing regime. For example, a negative HMT result at a dose per skin area of 100 µg/cm² cannot exclude that a positive result might have been obtained at e.g., 300 µg/cm² (sub-category 1A) or 700 µg/cm² (sub-category 1B). The same holds for LLNA test results obtained from tests which have not been carried out using the highest possible test concentration (see OECD Test Guideline 429 for details).

3.4.5.3.7.2.2 In the following ambiguous cases, study results for substances and mixtures would not be in disagreement with another study result pointing at that stricter classification:

- (a) A not classified result obtained at a lower test concentration does not exclude the possibility of a sub-category 1B outcome at a higher test concentration. Therefore, a not classified result obtained at a low concentration is compatible with other not classified outcomes, or with Category 1 and sub-category 1B outcomes obtained at higher test concentrations.
- (b) A not classified result at a very low-test concentration does not even exclude a possible outcome of sub-category 1A at a higher test concentration. Therefore, a not classified outcome obtained at a very low-test concentration is compatible with all possible classification outcomes (i.e., not classified, Category 1, sub-category 1A or 1B) obtained at higher test concentrations.
- (c) A sub-category 1B result at a higher test concentration does not exclude a sub-category 1A outcome at a lower test concentration. Therefore, a sub-category 1B classification tested at a high-test concentration is compatible with other outcomes of sub-category 1B, or even sub-category 1A, obtained at lower test concentrations.

3.4.5.3.7.2.3 If at least one unambiguous study result allows for sub-categorization of a substance or mixture and all other study results are not in disagreement (see above), then it can be classified into a sub-category. For example, if all study results are in the same sub-category (i.e., sub-category 1A or 1B), or with at least one study permitting sub-categorization (i.e., either sub-category 1A or 1B) and all other studies classified into Category 1 without sub-categorization, then the substance or mixture can be sub-categorized.

3.4.5.3.7.3 Weight of evidence considerations for giving one study result more weight than another.

3.4.5.3.7.3.1 Some classifiers or competent authorities may take various approaches to evaluate study results given the required level of expert judgement (see 1.3.2.4.8) required to perform a weight of evidence assessment. Competent authorities may specify their preferred approach in their own guidance. For example, through:

- (a) Applying a precautionary approach, giving more weight to studies resulting in the stricter classification outcome;
- (b) Giving human data higher weight than animal or non-test data;
- (c) Giving certain animal data (e.g., LLNA data) more weight than other animal data (e.g., Buehler test data).

3.4.5.3.7.3.2 Often, several results (of the same or different type) may have to be considered in the weight of evidence assessment. There are no generally recognized rules for this situation, however, possible solutions to integrating several results of the same type may include, for example:

- (a) A precautionary approach where the strictest classification outcome from all studies of sufficient quality is assigned as the overall classification outcome;
- (b) Averaging the obtained dose descriptors (e.g., LLNA EC3 values) or classification outcomes (no classification, Category 1, 1A, 1B). A detailed discussion of such approaches can be found in Annex 3 (on LLNA data) and Annex 4 (on HMT/HRIPT data) of OECD Series on Testing and Assessment No. 336 (Supporting document to OECD Guideline Document 497).

Table 3.4.7: Criteria for defined approaches

Category	OECD Guideline 497 on Defined Approaches for Skin sensitization "2 out of 3" (2o3) defined approach	OECD Guideline 497 on Defined Approaches for Skin sensitization Integrated testing strategy (ITSv1) defined approach and Integrated testing strategy (ITSv2 defined approach)
	<p>2o3 defined approach to skin sensitization hazard identification based on <i>in chemico</i> (key event 1 - Direct Peptide Reactivity Assay (KE1-DPRA)) and <i>in vitro</i> (key event 2-OECD 442D Appendix 1A, key event 3 - human Cell Line Activation Test (KE3-h-CLAT))</p> <p>Assays are run for two key events, and if these assays provide consistent results, then the chemical is predicted accordingly as sensitizer or non-sensitizer. If the first two assays provide discordant results, the assay for the remaining key event is run. The overall result is based on the two concordant findings taking into account the confidence on the obtained predictions as described in the Guideline</p>	<p>ITSv1 based on <i>in chemico</i> (KE1-DPRA) and <i>in vitro</i> (KE3-h-CLAT) data, and <i>in silico</i> (Derek Nexus) predictions. ITSv2 based on <i>in chemico</i> (KE1 -DPRA) and <i>in vitro</i> (KE3 -h-CLAT) data, and <i>in silico</i> (OECD QSAR Toolbox) predictions.</p> <p>Quantitative results of h-CLAT and DPRA are converted into a score from 0 to 3. For the <i>in silico</i> prediction (Derek or OECD QSAR ToolBox), a positive outcome is assigned a score of 1; a negative outcome is assigned a score of 0. When these scores have been assessed, a total battery score ranging from 0 to 7, calculated by summing the individual scores, is used to predict the sensitizing potential (hazard identification; GHS Category 1 versus no classification) and potency (GHS sub-category. 1A, Cat. 1B and no classification).</p>
1	2 out of 3 or 3 out of 3 positive predictions	Total battery score ≥ 2
1A	Not applicable	Total battery score 6-7
1B	Not applicable	Total battery score 2-5
Not classified	2 out of 3 or 3 out of 3 negative predictions	Total battery score < 2

Table 3.4.8: Criteria for individual *in chemico*/in vitro methods

Category	OECD Test Guideline 442C Key event-based Test Guideline for <i>in chemico</i> skin sensitization assays addressing the adverse outcome pathway (AOP) Key Event on covalent binding to proteins			OECD Test Guideline 442D Key event-based Test Guideline for in vitro skin sensitization assays addressing the AOP Key Event on keratinocyte activation antioxidant response element-nuclear factor-erythroid 2-related factor 2 (ARE-Nrf2) luciferase methods		OECD Test Guideline 442E In vitro skin sensitization assays addressing the AOP Key Event on activation of dendritic cells			
	Method described in Appendix I The Direct Peptide Reactivity Assay (DPRA) ^a	Method described in Appendix II The Amino acid Derivative Reactivity Assay (ADRA) ^a	Method described in Appendix III The kinetic Direct Peptide Reactivity Assay (kDPRA) ^b	Method described in Appendix 1A ^a	Method described in Appendix 1B Lusens ^a	Method described in Annex I human Cell Line Activation Assay (h-CLAT) ^a	Method described in Annex II U937 Cell Line Activation Test ^a	Method described in Annex III Interleukin-8 luciferase (IL-8 Luc) assay ^a	Method described in Annex IV Genomic Allergen Rapid Detection for assessment of skin sensitizers ^a
	<p>Methods: <i>in chemico</i> methods addressing the process of haptentation by quantifying the reactivity of test chemicals towards model synthetic peptides containing either lysine or cysteine (DPRA and kDPRA) or towards model synthetic amino acid derivatives containing either N-(2-(1-naphthyl) acetyl)-L-cysteine (NAC) or α-N-(2-(1-naphthyl) acetyl)-L-lysine (NAL) (ADRA).</p> <p>The criteria are based on the mean of cysteine and lysine peptides percent depletion (DPRA), kinetic rates of cysteine depletion (kDPRA) and mean NAC and NAL percent depletion value (ADRA). Predictions models based on the cysteine or NAC percent depletion value alone in case the unreacted lysine peptide or NAL cannot be reliably measured can be applied for the DPRA and ADRA.</p>			<p>Methods: cell-based methods addressing the process of keratinocytes activation, by assessing with the help of luciferase, the Nrf2-mediated activation of antioxidant response element (ARE)-dependent genes following exposure of the cells to the test chemical.</p> <p>Cell viability is quantitatively measured in parallel by enzymatic conversion of the dye 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT).</p> <p>The criteria are based on the induction of the luciferase gene above a given threshold, quantified at subtoxic concentrations. Criteria should be met in 2 of 2 or in 2 of 3 repetitions.</p>		<p>Methods: cell-based methods addressing the process of monocytes/dendritic cell activation by either quantifying the change in the expression of cell surface marker(s) (e.g. cluster of differentiation 54 (CD54), cluster of differentiation 86 (CD86)) or the change in IL-8 expression or the transcriptional patterns of an endpoint-specific genomic biomarker signature following exposure of the cells to the test chemical.</p> <p>Criteria should be met in 2 of 2 or in at least 2 of 3 repetitions for test methods described in Annexes I, II and III or in three valid biological replicates for test method described in Annex IV.</p>			

Table 3.4.8: Criteria for individual *in chemico*/in vitro methods (cont'd)

Category	OECD Test Guideline 442C Key event-based Test Guideline for <i>in chemico</i> skin sensitization assays addressing the adverse outcome pathway (AOP) Key Event on covalent binding to proteins			OECD Test Guideline 442D Key event-based Test Guideline for in vitro skin sensitization assays addressing the AOP Key Event on keratinocyte activation antioxidant response element-nuclear factor-erythroid 2-related factor 2 (ARE-Nrf2) luciferase methods		OECD Test Guideline 442E In vitro skin sensitization assays addressing the AOP Key Event on activation of dendritic cells			
	Method described in Appendix I The Direct Peptide Reactivity Assay (DPRA) ^a	Method described in Appendix II The Amino acid Derivative Reactivity Assay (ADRA) ^a	Method described in Appendix III The kinetic Direct Peptide Reactivity Assay (kDPRA) ^b	Method described in Appendix 1A ^a	Method described in Appendix 1B Lusens ^a	Method described in Annex I human Cell Line Activation Assay (h-CLAT) ^a	Method described in Annex II U937 Cell Line Activation Test ^a	Method described in Annex III Interleukin-8 luciferase (IL-8 Luc) assay ^a	Method described in Annex IV Genomic Allergen Rapid Detection for assessment of skin sensitizers ^a
1	The mean cysteine/lysine % depletion > 6.38% Or the mean cysteine % depletion > 13.89 %	The mean NAC and NAL % depletion ≥ 4.9% Or NAC% depletion ≥ 5.6%	Not applicable	The following 4 conditions are all met in 2 of 2 or in the same 2 of 3 repetitions: 1. Imax equal or higher than (≥) 1.5 fold and statistically significantly different to the solvent control 2. The cellular viability is higher than (>) 70% at the lowest concentration with induction of luciferase activity equal or above 1.5 fold 3. The EC _{1.5} value is less than (<) 1000 μM (or < 200 μg/mL for test chemicals with no defined molecular weight) 4. There is an apparent overall dose-dependent increase in luciferase induction	The following conditions are all met in 2 of 2 or in the same 2 of 3 repetitions: 1. A luciferase induction above or equal to (≥) 1.5 fold as compared to the solvent control is observed in at least 2 consecutive non-cytotoxic tested concentrations (i.e. cellular viability is equal or higher than (≥) 70%) 2. At least three tested concentrations should be non-cytotoxic (cellular viability equal or higher than (≥) 70%).	At least one of the following conditions is met in 2 of 2 or in at least 2 of 3 independent runs: The Relative Fluorescence Intensity of CD86 is equal to or greater than 150% at any tested concentration (with cell viability ≥ 50%) or the Relative Fluorescence Intensity of CD54 is equal to or greater than 200% at any tested concentration (with cell viability ≥ 50%).	The following condition is met in 2 of 2 or in at least 2 of 3 independent runs: The stimulation index of CD86 is equal or higher (≥) than 150% and/or interference is observed	The induction of normalised interleukin-8 luciferase activity (Ind-IL8LA) is equal or higher than (≥) 1.4 and the lower limit of the 95% confidence interval of Ind-IL8LA is equal or higher than (≥) 1.0 in at least 2 out of a maximum of 4 independent runs	The mean Decision Value (DV) is ≥ 0
1A	Not applicable		log kmax ≥ -2.0	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
1B	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable

Table 3.4.8: Criteria for individual *in chemico*/in vitro methods (cont'd)

Category	OECD Test Guideline 442C Key event-based Test Guideline for <i>in chemico</i> skin sensitization assays addressing the adverse outcome pathway (AOP) Key Event on covalent binding to proteins			OECD Test Guideline 442D Key event-based Test Guideline for in vitro skin sensitization assays addressing the AOP Key Event on keratinocyte activation antioxidant response element-nuclear factor-erythroid 2-related factor 2 (ARE-Nrf2) luciferase methods		OECD Test Guideline 442E In vitro skin sensitization assays addressing the AOP Key Event on activation of dendritic cells			
	Method described in Appendix I The Direct Peptide Reactivity Assay (DPRA) ^a	Method described in Appendix II The Amino acid Derivative Reactivity Assay (ADRA) ^a	Method described in Appendix III The kinetic Direct Peptide Reactivity Assay (kDPRA) ^b	Method described in Appendix 1A ^a	Method described in Appendix 1B Lusens ^a	Method described in Annex I human Cell Line Activation Assay (h-CLAT) ^a	Method described in Annex II U937 Cell Line Activation Test ^a	Method described in Annex III Interleukin-8 luciferase (IL-8 Luc) assay ^a	Method described in Annex IV Genomic Allergen Rapid Detection for assessment of skin sensitizers ^a
Not classified	The mean cysteine/lysine % depletion ≤ 6.38% or the mean cysteine % depletion ≤ 13.89 %	The mean NAC and NAL % depletion < 4.9% or NAC% depletion < 5.6%	Not applicable	At least one of the conditions for Category 1 is not met	At least one of the conditions for Category 1 is not met	None of the conditions for Category 1 is met	The stimulation index of CD86 is < 150% at all non-cytotoxic concentrations (cell viability ≥ 70%) and if no interference is observed	The Ind-IL8LA is less than (<) 1.4 and/or the lower limit of the 95% confidence interval of Ind-IL8LA is less than (<) 1.0 in at least 3 out of a maximum of 4 independent runs	The mean Decision Value (DV) is < 0

^a Data cannot be used as stand-alone to conclude on classification in Category 1 or on no classification in tier 1 but could be used to conclude on classification in Category 1 in tier 2 depending on the decision of the competent authority for their regulatory framework.

^b A competent authority may decide that data can be used as stand-alone to conclude on classification in sub-category 1A.”

Insert the following references at the end of chapter 3.4:

“

** References:*

Corsini, E., Clewell, R., Cotgreave, I., Eskes, C., Kopp-Schneider, A., Westmoreland, C., Alves, P.M., Navas, J.M. and Piersma, A., ESAC Opinion on the Scientific Validity of the GARDskin and GARDpotency Test Methods, Asturiol Bofill, D., Casati, S. and Viegas Barroso, J.F. editor(s), Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-40345-6, Doi:10.2760/626728, JRC125963.

Cottrez F, Boitel E, Auriault C, Aeby P, Groux H. Genes specifically modulated in sensitized skins allow the detection of sensitizers in a reconstructed human skin model. Development of the SENS-IS assay. *Toxicol In Vitro*. 2015 Jun;29(4):787-802. Doi: 10.1016/j.tiv.2015.02.012.

Cottrez F, Boitel E, Ourlin JC, Peiffer JL, Fabre I, Henaoui IS, Mari B, Vallauri A, Paquet A, Barbry P, Auriault C, Aeby P, Groux H. SENS-IS, a 3D reconstituted epidermis based model for quantifying chemical sensitization potency: Reproducibility and predictivity results from an inter-laboratory study. *Toxicol In Vitro* 2016 Apr;32:248-60. Doi: 10.1016/j.tiv.2016.01.007.

ECHA Guidance on the Application of the CLP Criteria Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures Version 5.0 July 2017

Gradin R., Johansson A., Forreryd A., Aaltonen E., Jerre A., Larne O., Mattson U., Johansson H. (2020) The GARDpotency assay for potency-associated subclassification of chemical skin sensitizers – Rationale, method development, and ring trial results of predictive performance and reproducibility. *Toxicol. Sci.* 176(2):423-432. Doi: 10.1093/toxsci/kfaa068

Johansson H., Lindstedt M., Albrekt A.S., Borrebaeck C.A. (2011) A genomic biomarker signature can predict skin sensitizers using a cell-based in vitro alternative to animal tests. *BMC Genomics* 12:399. Doi: 10.1186/1471-2164-12-399.

Johansson H., Rydnert F., Kühnl J., Schepky A., Borrebaeck C., Lindstedt M. (2014) Genomic allergen rapid detection in-house validation – A proof of concept. *Toxicol. Sci.* 139(2):362- 370. Doi: 10.1093/toxsci/kfu046.

Johansson H., Gradin R., Forreryd A., Agemark M., Zeller K., Johansson A., Larne O., van Vliet E., Borrebaeck C., Lindstedt M. (2017) Evaluation of the GARD assay in a blind Cosmetics Europe study. *ALTEX* 34(4):515-523. Doi: 10.14573/altex.1701121

Jowsey IR, Clapp CJ, Safford B, Gibbons BT, Basketter DA. (2008). The impact of vehicle on the relative potency of skin-sensitizing chemicals in the local lymph node assay. *Cutan Ocul Toxicol*: 27 (2); 67-75. Doi: 10.1080/15569520801904655.

Kligman A.M. (1966): The identification of contact allergens by human assay: II. Factors influencing the induction and measurement of allergic contact dermatitis. *Journal of Investigative Dermatology* 47 (5), 375-392. Doi: 10.1038/jid.1966.159

Kobayashi T., Maeda Y., Kondo H., Takeyoshi M. (2020) Applicability of the proposed GHS subcategorization criterion for LLNA:BrDU-ELISA (OECD TG442B) to the CBA/J strain mouse. *Journal of Applied Toxicology*. 40(10):1435-1439

Maeda Y., Takeyoshi M. (2019) Proposal of GHS sub-categorization criteria for LLNA: BrdU-ELISA (OECD TG442B). *Regulatory Toxicology and Pharmacology*. 107:104409.

OECD (2014). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins, OECD Series on Testing and Assessment, No. 168, OECD Publishing, Paris. Doi.org/10.1787/9789264221444-en

OECD (2017), Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation, OECD Series on Testing and Assessment, No. 256, OECD Publishing, Paris. Doi.org/10.1787/9789264279285-en.

Ryan CA et al. (2007): Extrapolating local lymph node assay EC3 values to estimate relative sensitizing potency. *Cutan Ocul Toxicol* 26(2), 135-45.

Saito K, Takenouchi O, Nukada Y, Miyazawa M, Sakaguchi H. An in vitro skin sensitization assay termed EpiSensA for broad sets of chemicals including lipophilic chemicals and pre/pro-haptens. *Toxicol In Vitro*. 2017 Apr;40:11-25. Doi: 10.1016/j.tiv.2016.12.005.

Wright ZM, Basketter PA, Blaikie L, Cooper KJ, Warbrick EV, Dearman RJ, Kimber I. Vehicle effects on skin sensitizing potency of four chemicals: assessment using the local lymph node assay. *Int J Cosmet Sci*. 2001 Apr;23(2):75-83. Doi: 10.1046/j.1467-2494.2001.00066.x.

Zeller K.S., Forreryd A., Lindberg T., Gradin R., Chawade A., Lindstedt M. (2017) The GARD platform for potency assessment of skin sensitizing chemicals. *ALTEX* 34(4):539-559. Doi: 10.14573/altex.1701101.”.

Chapter 3.5

- 3.5.3.2.4 In the last sentence replace “by testing” with “based on test data”.
- 3.5.4 Amend the last sentence of the paragraph to read as follows: “Table 3.5.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”
- 3.5.5.1.1 In decision logic 3.5.1, replace “approach” with “assessment” in the second and third text boxes down.

Chapter 3.6

- 3.6.3.2.4 In the last sentence replace “by testing” with “based on test data” and “assigned” with “classified in”.
- 3.6.4 Amend the last sentence of the paragraph to read as follows:
“Table 3.6.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”
- 3.6.5.3.2 In the second sentence replace “analysis” with “assessment”.
- 3.6.5.3.2.1 In the last sentence replace “evaluation” with “assessment”
- 3.6.5.1 In decision logic 3.6.1 replace “approach” with “assessment” in the second and third textboxes down.

Chapter 3.7

- 3.7.2.2.1 In the first sentence replace “an assessment of the total weight of evidence” with “a total weight of evidence assessment”.
- 3.7.2.3.1 In the first sentence replace “an assessment of the total weight of evidence” with “a total weight of evidence assessment”.
- 3.7.2.4.1 In the fourth sentence replace “approach” with “assessment”.
- 3.7.3.2.4 In the last sentence replace “by testing” with “based on test data” and “assigned” with “classified in”.
- 3.7.4 Add the following sentence at the end of the paragraph: “Table 3.7.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”
- 3.7.5.1.1 In decision logic 3.7.1 replace “approach” with “assessment” in the second and third textboxes down.

Chapter 3.8

- Table 3.8.1 In Note a, first sentence: replace “approach” with “assessment”.
In Note b, second sentence: replace “evaluation” with “assessment”.
- 3.8.2.1.10.1 Replace “approach” with “assessment”.
- 3.8.2.2.1 (d) In the last sentence replace “evaluation” with “assessment”.
- 3.8.3.2 In the first sentence replace “evaluation” with “assessment”.
- 3.8.3.3.6 In the last sentence replace “by testing” with “based on test data” and “assigned” with “classified in”.
- 3.8.3.3.7 At the beginning of the first sentence replace “aerosol form” with “aerosolized form”.
- 3.8.4.1 Insert the following sentence at the end of the paragraph: “Table 3.8.3 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”
- 3.8.5.1 In decision logic 3.8.1 replace “approach” with “assessment” in the fourth, fifth and sixth text boxes down.

Chapter 3.9

- 3.9.2.9.8 In the first sentence replace “approach” with “assessment”.
- 3.9.2.10.1 At the end of the paragraph replace “approach” with “assessment”.
- 3.9.3.2 In the first sentence replace “evaluation” with “assessment”.
- 3.9.3.3.6 In the last sentence replace “by testing” with “based on test data” and “assigned” with “classified in”
- 3.9.3.3.7 At the beginning of the paragraph replace “aerosol form” with “aerosolized form”.
- 3.9.4 Add the following sentence at the end of the paragraph: “Table 3.9.4 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”
- 3.9.5.1 In decision logic 3.9.1 replace “approach” with “assessment” in the fourth and fifth text boxes down.

Chapter 3.10

- 3.10.3.2.6 In the last sentence replace “assigned” with “classified in”.
- 3.10.4.1 Amend the last sentence of the paragraph to read as follows: “Table 3.10.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 4.1

- 4.1.2.5 In the fourth sentence replace “approach” with “assessment”.
- 4.1.3.4.6 In the last sentence replace “assigned” with “classified in”.
- 4.1.4 Insert the following sentence at the end of the paragraph: “Table 4.1.6 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Chapter 4.2

- 4.2.3 Insert the following sentence at the end of the paragraph: “Table 4.2.2 presents specific label elements for substances and mixtures classified into this hazard class based on the criteria in this chapter.”.

Annex 3

Section 1

- A3.1.2.3 Amend the last sentence to read as follows:
 “For example, H300 + H310 + H330 indicates that the text to appear on the label is “**Fatal if swallowed, in contact with skin or if inhaled.**”.”
- A3.1.2.4 Amend the last sentence to read as follows:
 “Also, where a combined hazard statement is permitted for two or more hazard statements (see A3.1.2.5), the competent authority may specify whether the combined hazard statement or the corresponding individual statements should appear on the label or may leave the choice to the manufacturer/supplier.”
- A3.1.2.5 (new) Insert the following new paragraph before current table A3.1.1:
 “A3.1.2.5 In addition to the combinations found in table A3.1.2, it is also permitted to combine more than one health hazard statement of equivalent severity if, for example, there is insufficient space on the label. When hazard statements are combined, all hazards must be clearly conveyed and only the repetitive text may be deleted. Statements can be combined by using the word “and”, additional punctuation, and changing the case of the initial letter of the word at the beginning of a statement. For example, H317 “May cause an allergic skin reaction” + H340 “May cause genetic defects” + H350 “May cause cancer” may all be combined because they are all for Category 1 health hazards (i.e., health hazard statements of equivalent severity) and have repetitive elements of the hazard statement (i.e., the statements begin with “may cause”). These statements may be combined to “May cause an allergic skin reaction, genetic defects, and cancer.” The competent authority may limit the types of combinations permitted to ensure comprehensibility (e.g., limit the number of hazard statements that can be combined).”.

Table A3.1.2

Insert the following note “a” under table A3.1.2:

“a Competent authorities may select the applicable hazard statement(s) depending on the serious eye damage/eye irritation hazard categories implemented in their jurisdiction (2/2A or 2A/2B).”

H317, column (3)

Replace “Sensitization, skin (chapter 3.4)” with “Skin sensitization (chapter 3.4)”.

H334, column (3)

Replace “Sensitization, respiratory (chapter 3.4)” with “Respiratory sensitization (chapter 3.4)”.

H315 + H319 (new)

Insert the following new entry after “H303 + H313 + H333”:

H315 + H319	Causes skin irritation and serious eye irritation^a	Skin corrosion/irritation (chapter 3.2) and serious eye damage/eye irritation (chapter 3.3)	2 (skin) + 2/2A (eye)
-------------------	--	---	--------------------------

H315 + H320

Under column (2), add a reference to note “a” at the end of the hazard statement, to read as follows: “Causes skin and eye irritation^a”.

Under column (4), replace “2 (skin)/2B (eye)” with “2 (skin) + 2B (eye)”.

Section 2

A3.2.2.4 Insert the following new paragraph:

“A3.2.2.4 Where square brackets [...] appear around a precautionary statement code, this indicates the precautionary statement is not appropriate in every case and should be used only in certain circumstances. In these cases, conditions for use explaining when the text should be used are given in column (5) of the tables.”

A3.2.4.4 Amend to read as follows:

“A3.2.4.4 Where square brackets [...] appear around some text in a precautionary statement, this indicates that the text in square brackets is not appropriate in every case and should be used only in certain circumstances. In these cases, conditions for use explaining when the text should be used are given in column (5) of the tables. For example, P264 states: “**Wash hands [and ...] thoroughly after handling.**”. This statement is given with the condition for use: “- text in square brackets to be used when the manufacturer/supplier or the competent authority specify other parts of the body to be washed after handling.”. The application of the condition for use should be interpreted as follows: If additional information is provided explaining what other part(s) of the body is to be washed following handling, then the text in brackets is to be used followed by the name of the relevant body part(s). However, if other part(s) of the body do not need to be specified, the text in square brackets should not be used, and the precautionary statement should read: “**Wash hands thoroughly after handling.**”.”

A3.2.5.2.2 Insert the following text after the first sentence:

“Precautionary statements can be combined by using the word “and”, additional punctuation, and changing the case of the initial letter of the word at the beginning of a statement. For example, P302 + P335 + P334 “**IF ON**

SKIN: Brush off loose particles from skin and immerse in cool water [or wrap in wet bandages].”

Table A3.2.2

P233

After the row applicable to “Acute toxicity, inhalation (chapter 3.1)” insert the following new row:

		Respiratory sensitization (chapter 3.4)	1, 1A, 1B	
--	--	---	-----------	--

(Note by the secretariat: the current condition for use under column (5) remains applicable and unchanged for “acute toxicity, inhalation” but does not apply to the new entry for respiratory sensitization).

Add the following condition of use under column (5) for the hazard classes “specific target organ toxicity, single exposure; respiratory tract irritation” and “specific target organ toxicity, single exposure narcotic effects”:

“– if the chemical is volatile and may generate a hazardous atmosphere.”.

P260

Insert the following new entry after the one applicable to acute toxicity, inhalation (chapter 3.1):

		Respiratory sensitization (chapter 3.4)	1, 1A, 1B	
--	--	---	-----------	--

In column (5), apply the following condition for use to the hazard classes “Acute toxicity, inhalation, respiratory sensitization”, “Specific target organ toxicity, single exposure” and “Specific target organ toxicity, repeated exposure”:

“Manufacturer/supplier or the competent authority to specify applicable physical state(s)”.

P261

Delete the entry for respiratory sensitization, categories 1, 1A, 1B.

P262, column (4)

Insert: “, 3” after: “1, 2”.

P264 and P270, column (4)

For acute toxicity (dermal), insert: “, 3” after: “1, 2”.

P271

In column (2) amend the precautionary statement to read as follows: **“Use only outdoors or with adequate ventilation.”**

Insert the following new entry after the one applicable to “Acute toxicity, inhalation (chapter 3.1)” :

		Respiratory sensitization (chapter 3.4)	1, 1A, 1B	
--	--	---	-----------	--

In column (5) insert the following condition for use applicable to all entries (including the one applicable to respiratory sensitization):

“Manufacturer/supplier to specify what type of ventilation would be adequate for safe use on the safety data sheet and in any supplemental safety instructions provided to consumers.”.

P280

Insert the following new row under the entry for “Eye irritation (chapter 3.3)”:

		Respiratory sensitization (chapter 3.4)	1, 1A, 1B	– <i>Specify protective gloves/clothing.</i> Manufacturer/supplier or the competent authority may further specify type of equipment where appropriate.
--	--	---	-----------	---

P284

In column (2), remove the square brackets around the text of the precautionary statement.

In column (5), amend the condition for use to read as follows:

“Manufacturer/supplier to specify on the safety data sheet what type of ventilation would be adequate for safe use and provide additional information with the chemical at the point of use that explains what type of respiratory equipment may also be needed.”

Table A3.2.3**P302 + P335 + P334, column (2)**

Amend the text to read as follows: “**IF ON SKIN: Brush off loose particles from skin and immerse in cool water [or wrap in wet bandages].**”

Table A3.2.4**P403**

Insert the following new row under the existing entry for “Acute toxicity, inhalation (chapter 3.1)”:

		Respiratory sensitization (chapter 3.4)	1, 1A, 1B	
--	--	---	-----------	--

In column (5) apply the following condition for use to “Acute toxicity, inhalation”; “Specific target organ toxicity, single exposure; respiratory tract irritation”; and “Specific target organ toxicity, single exposure narcotic effects”:

“– *if the chemical is volatile and may generate a hazardous atmosphere.*”

Section 3**Tables for flammable gases (chapter 2.2)**

Delete the note under the tables for pyrophoric gases and chemically unstable gases

Table for pyrophoric solids (chapter 2.10), hazard category 1, column “Response”, precautionary statement P302 + P335 + P334

Amend to read as follows:

“P302 + P335 + P334

IF ON SKIN: Brush off loose particles from skin and immerse in cool water or wrap in wet bandages.”

Table for “Acute toxicity - dermal (chapter 3.1)”, category 3, column “Prevention”

Insert the following entries:

“P262

Do not get in eyes, on skin, or on clothing.

P264

Wash hands [and ...] thoroughly after handling.

– text in square brackets to be used when the manufacturer/supplier or competent authority specify other parts of the body to be washed after handling.

P270

Do not eat, drink or smoke when using this product.”.

Table for “Acute toxicity, inhalation (chapter 3.1)”, categories 1, 2, column “Prevention”

Amend to read as follows:

“P260

Do not breathe dust/fume/gas/mist/vapours/spray.

Manufacturer/supplier or the competent authority to specify applicable physical state(s).

P271

Use only outdoors or with adequate ventilation.

Manufacturer/supplier to specify what type of ventilation would be adequate for safe use on the safety data sheet and in any supplemental safety instructions provided to consumers.

P284

In case of inadequate ventilation wear respiratory protection

Manufacturer/supplier to specify on the safety data sheet what type of ventilation would be adequate for safe use and provide additional information with the chemical at the point of use that explains what type of respiratory equipment may also be needed.”

Tables for “Acute toxicity, inhalation (chapter 3.1)”, categories 3 and 4, column “Prevention”, precautionary statement P271

Amend to read as follows:

“P271

Use only outdoors or with adequate ventilation.

Manufacturer/supplier to specify what type of ventilation would be adequate for safe use on the safety data sheet and in any supplemental safety instructions provided to consumers.”.

Table for “Sensitization – respiratory (chapter 3.4), heading

Amend to read as follows: “RESPIRATORY SENSITIZATION (CHAPTER 3.4)”.

Table for “Sensitization - Respiratory (chapter 3.4)”, categories 1, 1A, 1B

Column “Prevention”, amend to read as follows:

“P233

Keep container tightly closed.

P260

Do not breathe dust/fume/gas/mist/ vapours/spray.

Manufacturer/supplier or the competent authority to specify applicable physical state(s).

P271

Use only outdoors or with adequate ventilation.

Manufacturer/supplier to specify what type of ventilation would be adequate for safe use on the safety data sheet and in any supplemental safety instructions provided to consumers.

P280

Wear protective gloves/protective clothing.

Manufacturer/supplier or the competent authority may further specify type of equipment where appropriate.

P284

In case of inadequate ventilation wear respiratory protection.

Manufacturer/supplier to specify on the safety data sheet what type of ventilation would be adequate for safe use and provide additional information with the chemical at the point of use that explains what type of respiratory equipment may also be needed.”

Column “Storage”, add the following precautionary statement:

“P403

Store in a well-ventilated place.”**Table for “Sensitization – skin (chapter 3.4)”, heading**

Amend to read as follows: “SKIN SENSITIZATION (CHAPTER 3.4)”.

Table for “Specific target organ toxicity (single exposure) (chapter 3.8)”, categories 1 and 2, column “Prevention”, precautionary statement P260

Replace “to specify applicable conditions” with “to specify applicable physical state(s).”

Table for “Specific target organ toxicity (single exposure) (chapter 3.8)”, category 3, column “Prevention”, precautionary statement P271

Amend to read as follows:

“P271

Use only outdoors or with adequate ventilation.

Manufacturer/supplier to specify what type of ventilation would be adequate for safe use on the safety data sheet and in any supplemental safety instructions provided to consumers.”

Table for “Specific target organ toxicity (repeated exposure) (chapter 3.9)”, categories 1 and 2, column “Prevention”, conseil de prudence P260

Replace “to specify applicable conditions.” with “to specify applicable physical state(s).”

Annex 9

A9.4.3.5.1 In the second sentence, delete the quotation marks around “weight of evidence approach” and replace “approach” with “assessment”.

A9.4.3.6 Replace “approach” with “assessment in the fourth sentence.

A9.5.4.1 In the second sentence, delete the quotation marks around “weight of evidence approach” and replace “approach” with “assessment”.

A9.7.1.1 Amend the last but one sentence (“This section...or persistent hazards”) to read as follows: “This section does not take into account the non-metallic ion (e.g. CN-) of metal compounds, which may be toxic.”

A9.7.1.1.1 (new) Insert the following new paragraph:

“A9.7.1.1.1 Organometallic compounds (e.g. methyl mercury or tributyltin,...) and organometallic salts may also be of concern given that they may pose bioaccumulation or persistence hazards in case they do not quickly dissociate or dissolve in water. Unless they act as a significant source of the metal ion (as a result of the dissociation and/or degradation processes), the organic moieties and the inorganic components should be assessed individually (OECD 2015). They are therefore excluded from the guidance of this section and should be classified according to the general guidance provided in section 4. Alternatively, those metal compounds that contain an organic component but that dissociate or dissolve easily in water as the metal ion should be treated in the same way as metal compounds and classified according to this annex (e.g. Zinc acetate, ...).”

- A9.7.1.6 In the second sentence, replace the text between brackets with the following: “(e.g. partitioning or chemical speciation to a non-soluble and hence not-bioavailable form).”
- In the third sentence, replace “in assessing chronic classification” by “in assessing long-term (chronic) classification”
- A9.7.1.8 In the first sentence, replace “to cause toxicity at the level of the L(E)C₅₀,” by “to cause toxicity at the level of the ecotoxicity reference value (ERV), being the acute ERV (expressed as L(E)C₅₀), and/or the chronic ERV (expressed as the NOEC/EC_x),”
- A9.7.1.9 Amend the introductory paragraph to read as follows:
- “This section deals with metals and metal compounds. For how this guidance applies to organometallic compounds and organometallic salts, see A9.7.1.1.1. Within the context of this guidance document, metals and metal compounds are characterized as follows:”
- A9.7.2.1.1 Add the following new paragraphs:
- A9.7.2.1.1.1 Ecotoxicity data of soluble inorganic compounds are used and combined to derive the acute and chronic ecotoxicity reference value of the dissolved metal ion (ERV or ERV_{ion}). The ecotoxicity of soluble inorganic metal compounds is dependent on the physico-chemistry of the medium, irrespective of the original metal species released in the environment.
- A9.7.2.1.1.2 When evaluating ecotoxicity data and deriving ERVs, the general “weight of evidence” principle is also applicable to metals (see section A9.3.4).
- A9.7.2.1.1.3 The ecotoxicity data selected should be evaluated for their adequacy. Adequacy covers here both the reliability (inherent quality of a test relating to test methodology and the way that the performance and results of a test are described) and the relevance (extent to which a test is appropriate to be used for the derivation of an ecotoxicity reference value) of the available ecotoxicity data (see sections A9.2.6 and A9.3.6).
- (a) Under the reliability criteria, metal specific considerations include the description of some abiotic parameters in the test conditions for enabling the consideration of the bioavailable metal concentration and free metal ion concentration:
- (i) Description of the physical test conditions: in addition to the general parameters (O₂, T°, pH, ...), measurements of abiotic parameters such as dissolved organic carbon (DOC), hardness, alkalinity of the water that govern the speciation and hence the metal bioavailability are recommended.
 - (ii) Description of test materials and methods: to calculate the free metal ion concentration with speciation models the concentrations of dissolved major ions and cations (e.g. aluminium, iron, magnesium, and calcium) are recommended.
 - (iii) Concentration-effect relationship; hormesis: sometimes an increased performance in growth or reproduction is seen at low metal doses that exceed the control values, referred to as hormesis. Such effects can occur especially with major trace nutrients such as iron, zinc and copper but can also occur with a wide variety of non-essential substances. In such cases, positive effects should not be considered in the derivation of acute ERVs and especially chronic ERVs. Other models than the conventional log-logistic dose-response model should be used to fit the dose-response curve and consideration should be given to the adequacy of the control diet/exposure. Due to the essential nutritional needs, caution is

needed with regards to extrapolation of the dose-response curve (e.g. to derive an acute or chronic ERV) below the lowest tested concentration.

- (b) Under the relevancy criteria, certain considerations need to be made, related to the relevancy of the test substance and to acclimatisation/adaptation:
 - (i) Relevance of the test substance: tests conducted with soluble metal salts should be used for the purpose of deriving acute and chronic ERVs. The ecotoxicity adapted from organic metal compounds exposure should not be used.
 - (ii) Acclimatisation/adaptation: For essential metals, the culture medium should contain a minimal concentration not causing deficiency for the test species used. This is especially relevant for organisms used for chronic toxicity tests where the margin between essentiality and toxicity may become small. For this reason a proper description of culture conditions related to the level of essential metals is required.”

A9.7.2.1.2.1 Add the following text at the end of the existing paragraph:

“For the classification of metals and metal compounds, transformation/dissolution testing is carried out over a pH range (see A10.2.3.2). If evidence is available that the aquatic toxicity of the dissolved metal depends on pH, then transformation/dissolution data and aquatic toxicity are compared at a similar pH. If such evidence is not available, then the aquatic toxicity cannot be grouped according to pH. The highest aquatic toxicity observed is then compared to the transformation/dissolution data obtained at the pH which causes maximum transformation and dissolution.”

A9.7.2.1.2.2 Delete “or economic” at the end of the paragraph.

A9.7.2.1.2.3 In the second sentence, replace “(Tipping, 1994)” with “(Tipping, 1994; Tipping et al., 2011)”.

Amend the third and fourth sentence to read as follows: “Alternatively, the Biotic Ligand Model (BLM) allows for the calculation of the concentration of metal ion responsible for the toxic effect at the level of the organism, which may be affected by the DOC concentration, the pH, and the concentrations of competing ions such as calcium and magnesium. Such models may be investigated to better understand the impact of test medium composition on metal toxicity. The BLM model has at present been validated for specific metals, organisms, and end-points “(Santore and Di Toro, 1999; Garman et al., 2020).”

Add the following sentence at the end of the current paragraph: “In case a metal-specific BLM is available covering an appropriate pH range, a comparison of aquatic toxicity data can be made using the entire effects database for different reference pH values, relevant to the transformation/dissolution data.”

A9.7.2.2.2 Replace the last sentence with the following: “Where these are the only information available and the solubility data cannot provide a clear answer on the solubility rate and equilibrium, it is highly recommended that solubility data be generated using the Transformation/Dissolution Protocol (annex 10).”

A9.7.2.2.3 and A9.7.2.2.4 Replace with the following:

“A9.7.2.2.3 Screening test for assessing solubility of metal compounds

In the absence of solubility data for metal compounds, a screening test for assessing solubility should be performed as described in the Transformation/Dissolution Protocol (annex 10). The screening test is

conducted at the high loading rate (100 mg/l) and under rapid and vigorous agitation for 24 h. The function of the screening test is:

- (a) To identify those metal compounds which undergo either dissolution or rapid transformation such that their ecotoxicity potential is indistinguishable from soluble forms in that they may be classified based on the dissolved ion concentration.
- (b) To verify the pH dependency of the dissolution, in preparation of the full transformation/dissolution test. Where data at different pH are available from the screening test, then the full test should at least be conducted at the pH which maximizes the solubility. Where data are not available over the full pH range, a check should be made that this maximum solubility has been achieved by reference to suitable thermodynamic speciation models or other suitable methods (see A9.7.2.1.2.3). In the absence of suitable data or models, it is highly recommended that solubility data are generated to cover the full pH range. It should be noted that this screening test is only intended to be used for metal compounds. Metals should be assessed at the level of the full test (see A9.7.2.2.4).

A9.7.2.2.4 Full test for assessing solubility of metals and metal compounds

A9.7.2.2.4.1 The full test should at least be carried out at the pH⁶ that maximizes the concentration of dissolved metal ions in solution. The pH may be chosen following the same guidance as given for the screening test.

A9.7.2.2.4.2 Based on the data from the full test, it is possible to generate a concentration of the metal ions in solution after 7 days for each of the three loadings (i.e. 1 mg/l as “low”, 10 mg/l as “medium” and 100 mg/l as “high”) used in the test. If the purpose of the test is to assess the long-term (chronic) hazard of the substance, then the loadings⁷ should be 0.01 mg/l, 0.1 mg/l or 1 mg/l depending on the transformation rate, and the duration of the test should be extended to 28 days.”.

Insert the following new footnotes “6” and “7”:

⁶ *The Transformation/Dissolution Protocol specifies a pH range of 6 to 8.5 for the 7-day test and 5.5 to 8.5 for the 28-day test. Considering the difficulty in carrying out transformation/dissolution tests at pH 5.5, OECD only validated the test in the pH range of 6 to 8.*

⁷ *Lower loading rates than 1 mg/l may not be practically feasible for each case. While transformation/dissolution testing at lower loading rates is in principle the best way forward it is technically often not feasible. Extensive experience with the Transformation/Dissolution Protocol demonstrated that reliable predictions can be made for other loading rates. In order to make maximal use of existing transformation/dissolution data, the 28-day results for the lower loading rates (0.1 and 0.01 mg/l) can therefore often be derived by extrapolation from evidence at other loading rates. This approach should be justified on a case-by-case basis and supported by reliable information on the transformation/dissolution at different loading rates. It should be further noted that the relationship between loading rate and dissolved metal concentration may not be linear. Therefore, extrapolating transformation/dissolution data to lower loadings should preferably be made by using the equations of section A10.6.1 or alternatively by extrapolating in a precautionary way.”.*

A9.7.2.3 Amend to read as follows:

“A9.7.2.3 Comparison of aquatic toxicity data and solubility data

A decision on how to classify the substance will be made by comparing aquatic toxicity data and solubility data. Depending on the available data, two approaches can be followed:

- (a) If limited information on the transformation/dissolution at different pH levels is available, or if the aquatic toxicity of the dissolved metal does not depend on pH, then the lowest ERV and the highest transformation/dissolution result, each potentially derived at different pH levels, should provide the basis for classification (this should be used as the default approach).
- (b) If evidence is available that the aquatic toxicity of the dissolved metal depends on pH, and sufficient toxicity data are available at varying pH levels, then a split of the acute and chronic ERVs can be performed according to the pH band. If in addition transformation/dissolution data at different pH levels are available, then the classification may be derived by comparing transformation/dissolution data with the ERV at corresponding pH levels, meaning that toxicity data and transformation/dissolution data are in this case always compared at the same pH band. This split of the effects data into pH bands would apply in an equal way to the acute and the chronic effects data sets. The most stringent classification outcome across all pH bands should be used.”

A9.7.4.1 Amend to read as follows:

“A9.7.4.1 While $\log K_{ow}$ is a good predictor of BCF for certain types of organic compounds e.g. non-polar organic substances, it is irrelevant for inorganic substances such as inorganic metal compounds because metals, in contrast to organic substances, are not lipophilic and are generally not transported through cellular membranes by passive processes. Uptake of metal ions typically occurs through active processes.”

A9.7.4.3 Amend the end of the fourth sentence and insert a new fifth sentence as follows: “...are actively regulated in organisms in which the metal is essential (homeostasis). Removal and sequestration processes that minimize toxicity are complemented by an ability to up-regulate concentrations for essentiality.”

The rest of the paragraph remains unchanged (“Since nutritional requirement of the organisms... bioconcentration and environmental concern.”).

A9.7.4.4 (new) Insert a new paragraph A9.7.4.4 to read as follows:

“A9.7.4.4 For essential elements, measured BCFs decline as external concentrations increase because the internal concentrations are regulated by the organism. Non-essential metals are also actively regulated to some extent and therefore also for nonessential metals, an inverse relationship between the metal concentration and the external concentration may be observed (McGeer et al., 2003). When external concentrations are so high that they exceed a threshold level, or overwhelm the regulatory mechanism, this can cause harm to the organism. BCF and BAF may be used to estimate metal accumulation by:

- (a) Considering information on essentiality and homeostasis of metals/metal compounds. As a result of such regulation, the “bioaccumulative” criterion is not applicable to metals.
- (b) Assessing bioconcentration factors for non-essential metals, should preferably be done from BCF studies using environmentally relevant concentrations in the test media.”

A9.7.5.1.1 Amend to read as follows:

“A9.7.5.1.1 Short-term (acute) and long-term (chronic) hazards are assessed individually for metals and metal compounds. For long-term hazards preference should be given in applying the approach based on chronic toxicity data. Such evidence is often available for the readily soluble metal salts. The schemes for the determination of short-term and long-term aquatic hazards of metals and metal compounds are described below and summarized in the figures:

- (a) A9.7.1 (short-term hazard classification of metals);
- (b) A9.7.2 and A9.7.3 (long-term hazard classification of metals);
- (c) A9.7.4 (short-term hazard classification of metal compounds);
- (d) A9.7.5 (long-term hazard classification of metal compounds).

A9.7.5.1.1.1 There are several stages in these schemes where data are used for decision purposes. It is not the intention of the classification schemes to generate new data. In the absence of valid data, it will be necessary to use all available data and expert judgement.

A9.7.5.1.1.2 In the following sections, the reference to the acute and chronic ERVs refers to the data point(s) that will be used to select the hazard categories for the metal or metal compound.”.

A9.7.5.1.2 Amend to read as follows:

“A9.7.5.1.2 When considering acute and chronic ERVs for metal compounds (ERV_{compound}), it is important to ensure that the data point to be used as the justification for the classification is expressed in the weight of the molecule of the metal compound to be classified. This is known as correcting for molecular weight. Thus while most metal data are expressed in, for example, mg/l of the dissolved metal ion (abbreviated ERV_{ion}), this value will need to be adjusted to the corresponding weight of the metal compound. Thus:

$$ERV_{\text{compound}} = ERV_{\text{ion}} \times$$

(Molecular weight of metal compound / Σ atomic weight of the atom(s) of the metal in the compound)

where:

$ERV_{\text{compound}} = ERV$ of the metal compound

$ERV_{\text{ion}} = ERV$ of the dissolved metal ion”

A9.7.5.2 to A9.7.5.2.4.2 Replace with the following text:

“A9.7.5.2 *Classification strategy for metals*

A9.7.5.2.1 *Short-term (acute) aquatic hazard of metals*

A9.7.5.2.1.1 The scheme for determining the short-term (acute) aquatic hazard of metals is described in this section and summarized in figure A9.7.1.

A9.7.5.2.1.2 Where the acute ERV of the dissolved metal ion is greater than 100 mg/l, the metals need not be considered further in the classification scheme.

A9.7.5.2.1.3 Where the acute ERV of the dissolved metal ions is less than or equal to 100 mg/l, consideration must be given to the data available on the rate and extent to which these ions can be generated from the metal. Such data, to be valid and useable, should have been generated using the Transformation/Dissolution Protocol (annex 10).

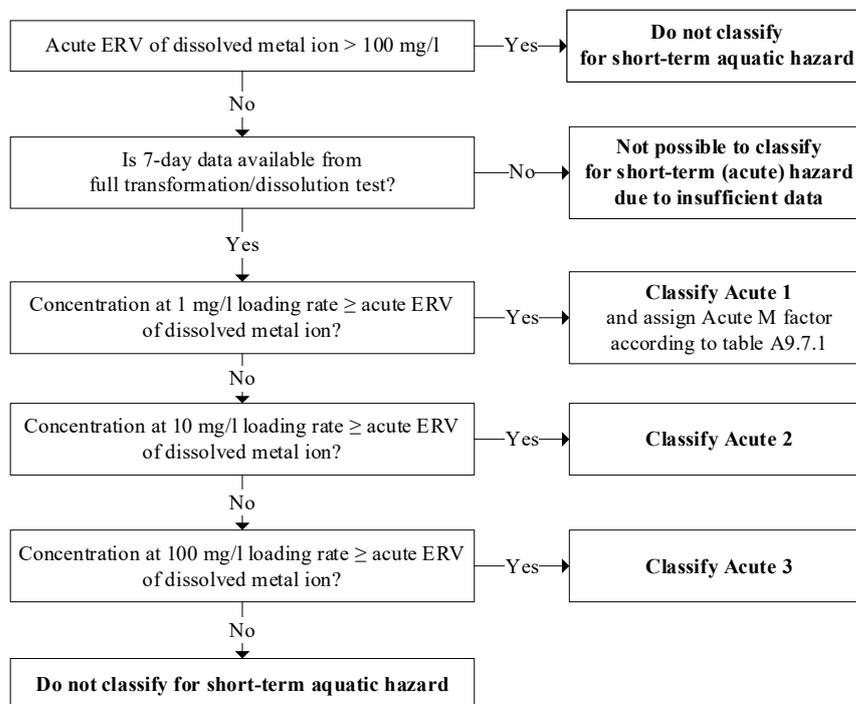
A9.7.5.2.1.4 Where 7-day data from the Transformation/Dissolution Protocol are available, then, the results should be used to aid classification according to the following rules. Classify the metal as:

- (a) Category Acute 1 if the dissolved metal ion concentration at the low loading rate is greater than or equal to the acute ERV. Assign an Acute M factor according to table A9.7.1;
- (b) Category Acute 2 if the dissolved metal ion concentration at the low loading rate is less than the acute ERV, but at the medium loading rate it is greater than or equal to the acute ERV;

- (c) Category Acute 3 if the dissolved metal ion concentration at the low and the medium loading rates is less than the acute ERV, but at the high loading rate it is greater than or equal to the acute ERV.

Do not classify the metal for short-term aquatic hazard if the dissolved metal concentration at all loading rates is below the acute ERV.

Figure A9.7.1: Classification strategy for determining the short-term (acute) aquatic hazard of metals



A9.7.5.2.2 Long-term (chronic) aquatic hazard of metals

The scheme for determining the long-term (chronic) aquatic hazard of metals is described in this section and summarized in figures A9.7.2 and A9.7.3. Metals can be classified for long-term aquatic hazard using chronic toxicity data when available or using the surrogate approach in absence of appropriate chronic toxicity data.

A9.7.5.2.2.1 Approach based on available chronic toxicity data

A9.7.5.2.2.1.1 Where the chronic ERV of the dissolved metal ion is greater than 1 mg/l, the metal need not be considered further in the classification scheme.

A9.7.5.2.2.1.2 Where the chronic ERV of the dissolved metal ion is less than or equal to 1 mg/l, consideration must be given to the available data on the rate and extent to which these ions can be generated from the metal. To be valid and useable, such data should have been generated or calculated using the Transformation/Dissolution Protocol (annex 10) for 28 days (see A9.7.2.2.4). If such data are unavailable, the surrogate approach should be used (see A9.7.5.2.2.2). Where 28-day transformation/dissolution data are available, then classify the metal as:

- (a) Category Chronic 1 if the dissolved metal ion concentration obtained at a loading rate of 0.1 mg/l (0.01 mg/l if there is evidence of rapid environmental transformation) is greater than or equal to the chronic ERV. Assign a Chronic M factor according to table A9.7.1;
- (b) Category Chronic 2 if the dissolved metal ion concentration obtained at a loading rate of 1 mg/l (0.1 mg/l if there is evidence of rapid

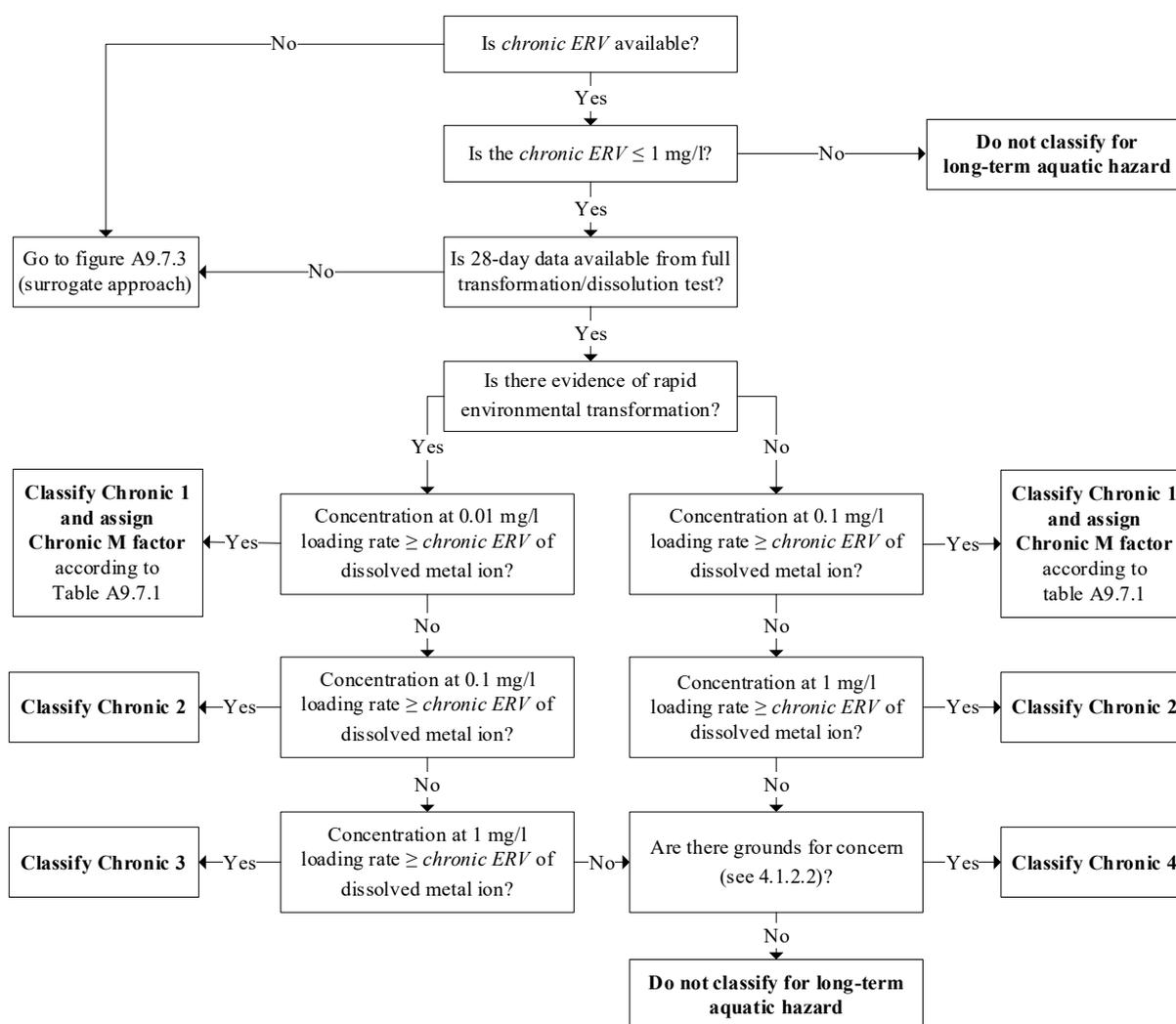
environmental transformation) is greater than or equal to the chronic ERV;

- (c) Category Chronic 3 if the dissolved metal ion concentration obtained at a loading rate of 1 mg/l is greater than the chronic ERV and there is evidence of rapid environmental transformation.

A9.7.5.2.2.1.3 Classify the metal as category Chronic 4 if the data available do not allow classification under the formal criteria but there are nevertheless some grounds for concern (see 4.1.2.2).

A9.7.5.2.2.1.4 Do not classify the metal for long-term aquatic hazard if the dissolved metal ion concentration obtained from the 28-day transformation/dissolution test at a loading rate of 1 mg/l is less than the chronic ERV of the dissolved metal ion.

Figure A9.7.2: Classification strategy for determining long-term aquatic hazard of metals on the basis of chronic data



A9.7.5.2.2.2 The surrogate approach

A9.7.5.2.2.2.1 Where appropriate chronic toxicity data and/or transformation/dissolution data are not available, but the metal is classified for short-term (acute) aquatic hazard, then classify the metal as (unless there is evidence of rapid environmental transformation and no bioaccumulation):

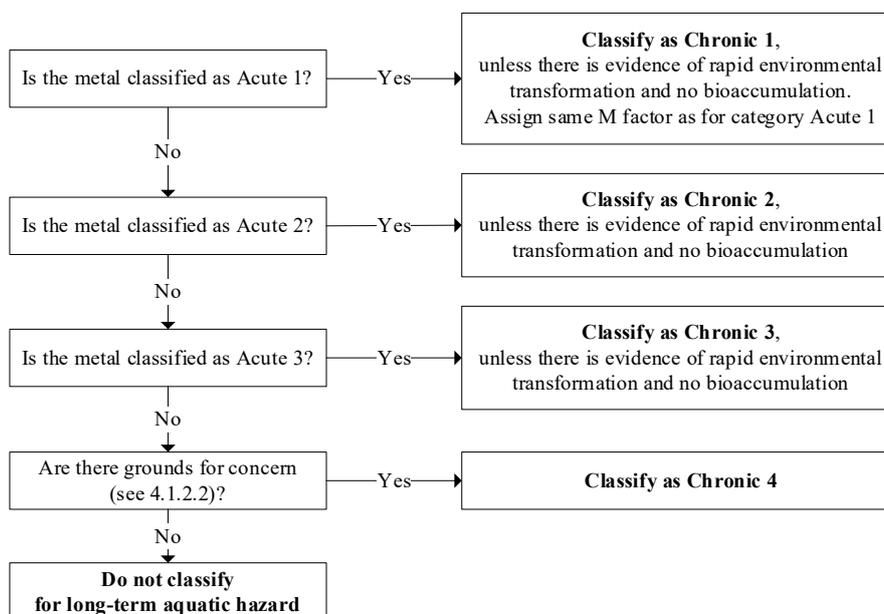
- (a) Category Chronic 1 if the metal is classified for short-term (acute) aquatic hazard as category Acute 1. Assign the same M factor as for category Acute 1.

- (b) Category Chronic 2 if the metal is classified for short-term (acute) aquatic hazard as category Acute 2.
- (c) Category Chronic 3 if the metal is classified for short-term (acute) aquatic hazard as category Acute 3.

A9.7.5.2.2.2.2 In the lack of a short-term aquatic hazard classification due to missing transformation/dissolution data, and there is no clear data of sufficient validity to show that the transformation of metal ions will not occur, the safety net classification (Chronic 4) should be applied when the known classifiable toxicity of these soluble forms is considered to produce sufficient concern. For example, this is the case when the acute ERV_{ion} is equal to or below 100 mg/l, and/or if the chronic ERV_{ion} is equal to or below 1 mg/l. In these cases, testing according to the Transformation/Dissolution Protocol may be considered.

A9.7.5.2.2.2.3 Do not classify the metal for long-term aquatic hazard if the metal is not classified for short-term aquatic hazard and if there are no grounds for concern.

Figure A9.7.3: Classification strategy for determining long-term aquatic hazard of metals in absence of appropriate chronic toxicity reference data and/or 28-day transformation/dissolution data



”

A9.7.5.3 to A9.7.5.3.3.3 Replace with the following text:

“A9.7.5.3 Classification strategy for metal compounds

Metal compounds will be considered as readily soluble if the water solubility (measured e.g. through a screening test according to the Transformation/Dissolution Protocol, or estimated e.g. from the solubility product), expressed as the concentration of dissolved metal ion, is greater than or equal to the acute ERV_{ion} . In the context of the classification criteria, metal compounds will also be considered as readily soluble if such data are unavailable, i.e. there are no clear data of sufficient validity to show that the transformation to metal ions will not occur. Care should be exercised for compounds whose solubility is close to the acute ERV as the conditions under which solubility is measured could differ significantly from those of the acute toxicity test. In these cases the results of the screening test are preferred. Metal compounds will be considered as poorly soluble if the water solubility (measured e.g. through a screening test, or estimated e.g. from the solubility

product), expressed as the concentration of dissolved metal ion, is less than the acute ERV_{ion} .

A9.7.5.3.1 *Short-term (acute) aquatic hazard of metal compounds*

A9.7.5.3.1.1 Readily soluble metal compounds are classified on the basis of the acute $ERV_{compound}$. Classify the readily soluble metal compound as:

- (a) Category Acute 1 if the acute $ERV_{compound}$ is equal to or less than 1 mg/l column. Assign an Acute M factor according to table A9.7.1;
- (b) Category Acute 2 if the acute $ERV_{compound}$ is greater than 1 mg/l but less than or equal to 10 mg/l;
- (c) Category Acute 3 if the acute $ERV_{compound}$ is greater than 10 mg/l but less than or equal to 100 mg/l.

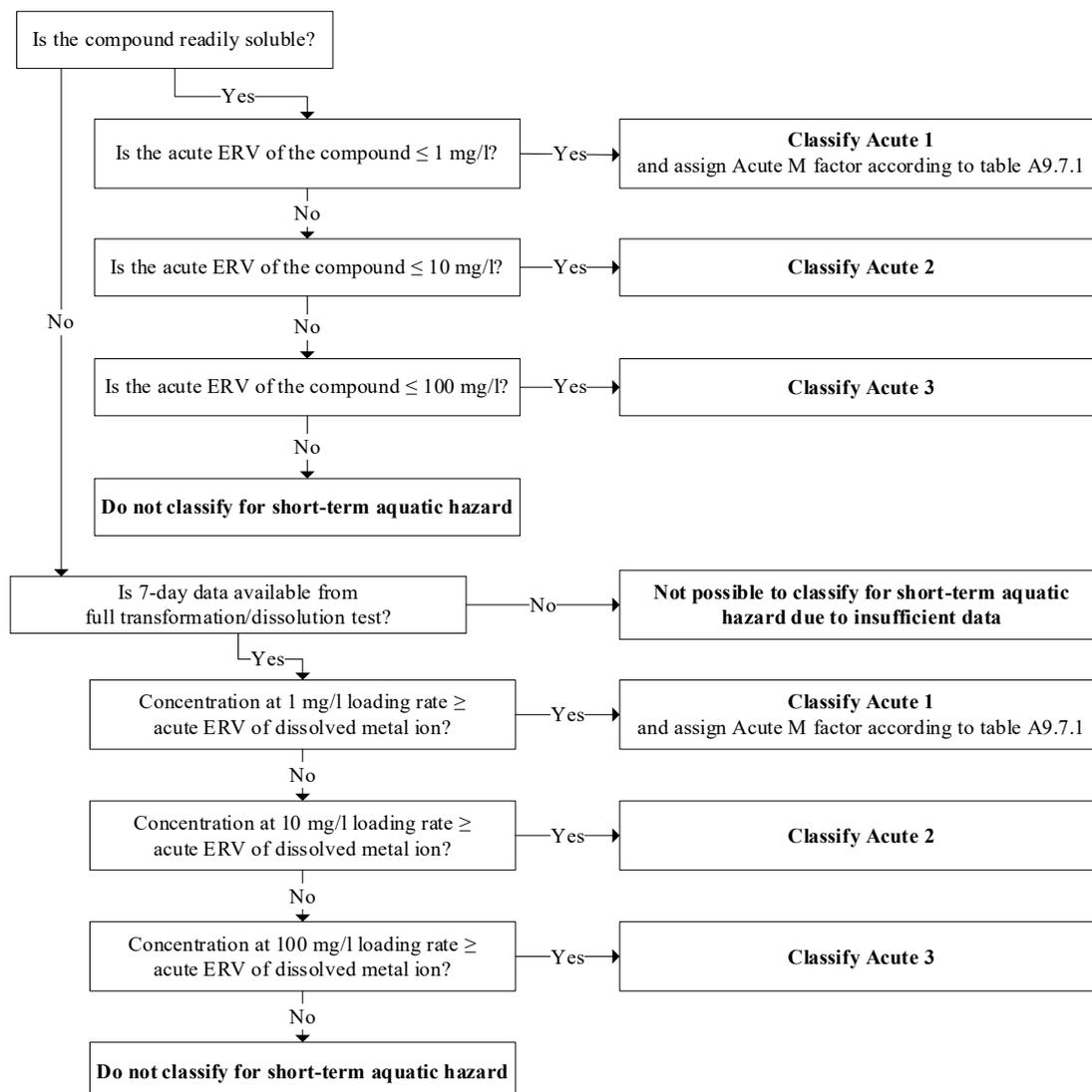
Do not classify the readily soluble metal compound for short-term aquatic hazard if the acute $ERV_{compound}$ is greater than 100 mg/l.

A9.7.5.3.1.2 Poorly soluble metal compounds are classified on the basis of the acute ERV of the dissolved metal ion and 7-day transformation/dissolution data. Classify the poorly soluble metal compound as:

- (a) Category Acute 1 if the dissolved metal ion concentration at the low loading rate is equal to or greater than the acute ERV_{ion} , and assign Acute M factor according to table A9.7.1;
- (b) Category Acute 2 if the dissolved metal ion concentration at the medium loading rate is equal to or greater than the acute ERV_{ion} ;
- (c) Category Acute 3 if the dissolved metal ion concentration at the high loading rate is equal to or greater than the acute ERV_{ion} .

Do not classify the poorly soluble metal compound for short-term (acute) aquatic hazard if the dissolved metal ion concentration is below the acute ERV of the dissolved metal ion at all loading rates.

Figure A9.7.4: Classification strategy for determining the short-term (acute) aquatic hazard of metal compounds



A9.7.5.3.2 Long-term (chronic) aquatic hazard of metal compounds

The scheme for determining the long-term (chronic) aquatic hazard of metal compounds is described in this section and summarised in figure A9.7.5. Metal compounds can be classified for long-term aquatic hazard using chronic toxicity data when available, or using the surrogate approach in absence of appropriate chronic toxicity data.

A9.7.5.3.2.1 Approach based on available chronic toxicity data

A9.7.5.3.2.1.1 Where the chronic ERV_{Compound} is greater than 1 mg/l, the metal compound need not to be considered further in the classification scheme for long-term hazard.

A9.7.5.3.2.1.2 Readily soluble metal compounds are classified on the basis of the chronic ERV_{compound} . If there is no evidence of rapid environmental transformation, then classify the readily soluble metal compound as:

- (a) Category Chronic 1 if the chronic ERV_{compound} is equal to or less than 0.1 mg/l (0.01 mg/l if there is evidence of rapid environmental transformation). Assign a chronic M factor according to table A9.7.1;

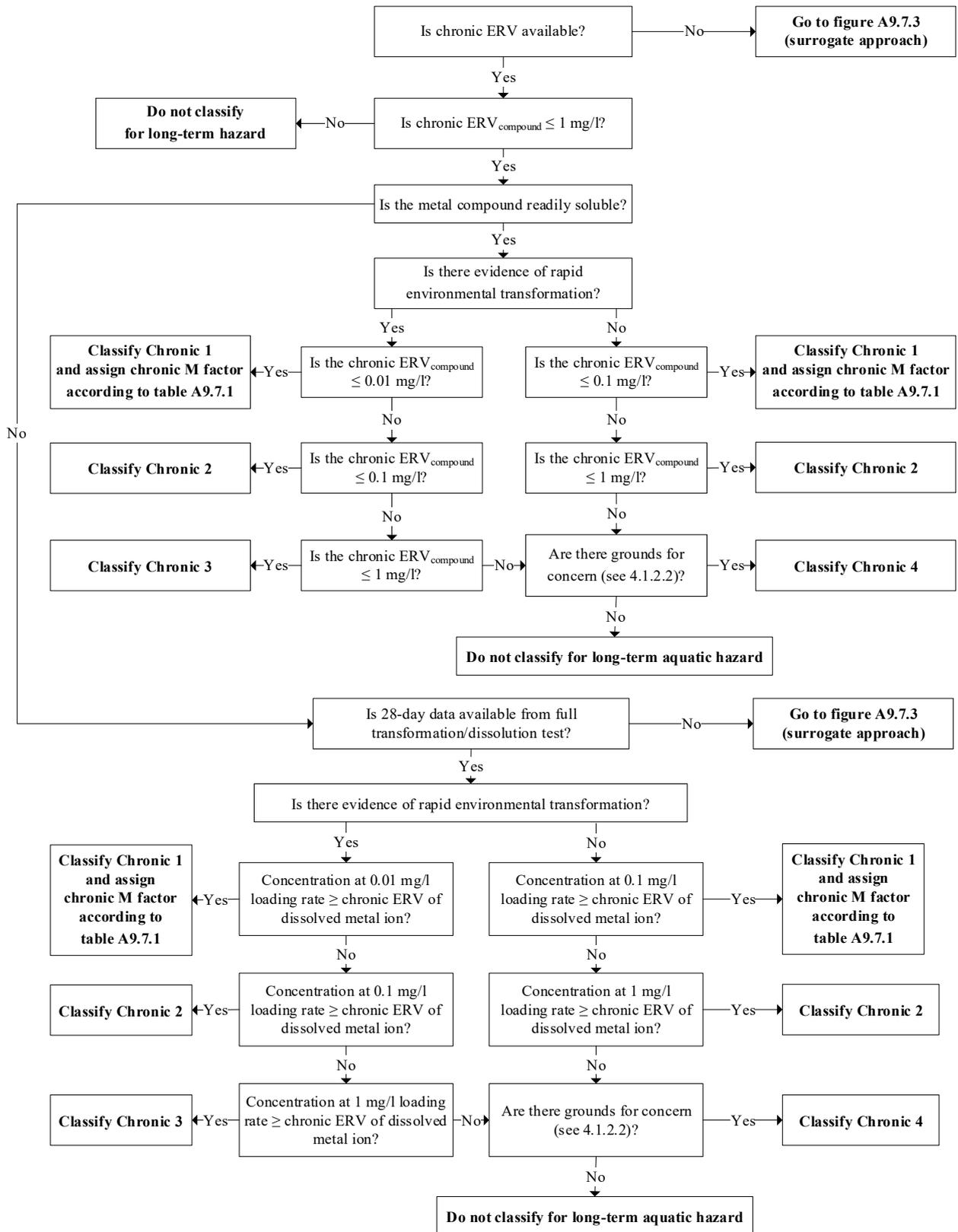
- (b) Category Chronic 2 if the chronic ERV_{compound} is equal to or less than 1 mg/l (0.1 mg/l if there is evidence of rapid environmental transformation);
- (c) Category Chronic 3 if the chronic ERV_{compound} is equal to or less than 1 mg/l and there is evidence of rapid environmental transformation;
- (d) Category Chronic 4 if the data available do not allow classification under the formal criteria but there are nevertheless some grounds for concern (see 4.1.2.2)

A9.7.5.3.2.1.3 Poorly soluble metal compounds: Consideration must be given to the data available on the rate and extent to which these ions can be generated from the metal compound. For such rate and extent data, to be valid and useable, they should have been generated using the Transformation/Dissolution Protocol for a 28-day period. Where such 28-day transformation/dissolution data are unavailable, the surrogate approach should be used (see A9.7.5.3.2.2). Where 28-day transformation/dissolution data are available, then classify the poorly soluble metal compound as:

- (a) Category Chronic 1 if the dissolved metal ion concentration obtained at a loading rate of 0.1 mg/l (0.01 mg/l if there is evidence of rapid environmental transformation) is greater than or equal to the chronic ERV of the dissolved metal ion. Assign a chronic M factor according to table A9.7.1;
- (b) Category Chronic 2 if the dissolved metal ion concentration obtained at a loading rate of 1 mg/l (0.1 mg/l if there is evidence of rapid environmental transformation) is greater than or equal to the chronic ERV of the dissolved metal ion;
- (c) Category Chronic 3 if the dissolved metal ion concentration obtained at a loading rate of 1 mg/l is greater than or equal to the chronic ERV of the dissolved metal ion and there is evidence of rapid environmental transformation;
- (d) Category Chronic 4 if the data available do not allow classification under the formal criteria but there are nevertheless some grounds for concern (see 4.1.2.2)

Do not classify the poorly soluble metal compound for long-term (chronic) aquatic hazard if the dissolved metal ion concentration obtained from the 28-day transformation/dissolution test at a loading rate of 1 mg/l is less than the chronic ERV of the dissolved metal ion.

Figure A9.7.5: Classification strategy for determining long-term aquatic hazard of metal compounds on the basis of chronic data



A9.7.5.3.3.2 The surrogate approach

Where appropriate chronic toxicity data and/or transformation/dissolution data are not available, but the metal compound is classified for short-term (acute) aquatic hazard, then the metal compound is classified according to the surrogate approach. The surrogate approach for metal compounds is identical to that for metals (see A9.7.5.2.2.2)."

A9.7.5.4.3 Add the following paragraph after the current table in A9.7.5.4.3:

"Massive forms will usually be tested as 1 mm particles. Alternatively, the transformation/dissolution testing of materials with different surface areas may result in highly reliable dissolution kinetic equations that allow to define the critical particle diameter for appropriate loadings for the acute and long-term hazard assessment."

A9.7.5.4.4 Amend to read as follows:

"A9.7.5.4.4 For some forms of metals, it may be possible, using the Transformation/Dissolution Protocol (OECD 2001), to obtain a correlation between the concentration of the metal ion after a specified time interval as a function of the surface area loadings of the forms tested. Such correlations should be established for the relevant pH ranges as specified in the Transformation/Dissolution Protocol. In such cases, it could then be possible to estimate the level of dissolved metal ion concentration of the metal with different particles, using the critical surface area approach (Skeaff et. al., 2000) (See reference in appendix VI, part 5, Metals and metal compounds). From this correlation and a linkage to the appropriate toxicity data at corresponding pH level, it is possible to determine a critical surface area of the substance that delivers the L(E)C₅₀ to the dissolution medium and then to convert the critical surface area to a critical particle diameter (see example). This critical particle diameter at appropriate mass loadings for acute and long-term hazard assessment can then be used to:

- (a) determine the classification category of powders based on the finest representative powder on the market and
- (b) determine an accurate classification of the massive metal by applying a 1 mm (default) diameter.

A9.7.5.4.4.1 Within the critical surface area approach an equation is developed to predict metal ion release (based on previously measured metal ion release from different loadings of the metal), which is correlated to measured surface area, and a corresponding calculated equivalent particle diameter. The basis of the critical surface area approach is that the release of metal ions is dependent on the surface area of the substance, with this release being predictable once the relationship has been established. The critical surface area as the surface area loading (mm²/l) to a medium that delivers a selected ecotoxicity reference value to that medium. The term SA is the measured specific surface area (m²/g) of the metal sample. The measured specific critical surface area (SA_{crit}) (m²/g) is the measured specific surface area for the corresponding low, medium and high loadings which are associated with the respective acute and long-term aquatic toxicity classification categories in the classification scheme for metals and metal compounds. A typical equation for this relationship for a given substance, aquatic medium, pH and retention time is:

$$\log(C_{Me(aq)}) = a + b \log(A_{meas})$$

C_{Me(aq)} = total dissolved concentration of metal ion (in mg/l) at a particular length of test time (i.e. 168 hours for short-term hazard assessment) under certain conditions (i.e. pH, specified medium, etc.), as determined by transformation/dissolution testing of different surface area loadings

a, b = regression coefficients

A_{meas} = initial surface area loading (in mm^2/l), calculated as follows:

$$A_{\text{meas}} = SA \times \text{substance mass loading} \times 10^6$$

where:

SA = specific surface area (in m^2/g) measured with the Brunauer-Emmet Teller (BET) nitrogen adsorption-desorption technique.

Substance mass loading in g/l .”

Figure A9.7.1 Delete

A9.7.5.5 Insert a new subsection A9.7.5.5 to read as follows:

“A9.7.5.5 Setting M factors for metals and inorganic metal compounds

A9.7.5.5.1 For the hazard class “Hazardous to the Aquatic Environment”, where the application of the normal cut-off values or concentrations limits may lead to an “under-classification” of the mixture, the M factor concept is used. The M factors are used in application of the summation method for the classification of mixtures containing substances that are classified as very ecotoxic. The concept of M factors has been established to give an increased weight to very toxic substances when classifying mixtures. This ensures that the magnitude of their toxicity is not lost in the derivation of the mixtures classification. M factors are only applicable to the concentration of a substance classified as hazardous to the aquatic environment (categories Acute 1 and Chronic 1) and are used to derive by the summation method the classification of a mixture in which the substance is present. They are, however, substance specific and it is important that they are established when classifying substances. It is important to note that separate Acute and Chronic M factors should be derived and these may not necessarily be of the same value, depending how each was determined (e.g. the basis of the separate acute and chronic ERV values).

A9.7.5.5.2 For readily soluble metal compounds M factors are applied as for organic substances (see table A9.7.1).

A9.7.5.5.3 For poorly soluble metal compounds and metals M factors are applied based on the ratio of the dissolved metal ion concentration (obtained from transformation/dissolution testing after respectively 7 and 28 days for the loading that was used to establish the classification of Category Acute 1 or Category Chronic 1) and the ERV of the dissolved metal ion. If that ratio is below 10 then an M factor of 1 is applied; if that ratio is ≥ 10 and < 100 then an M factor of 10 is applied; if that ratio is ≥ 100 and < 1000 then an M factor of 100 is applied... (continue this approach in factor 10 intervals).

Table A9.7.1: M factors for readily soluble metal compounds

Acute ERV _{compound} (mg/l)	Acute Multiplication factors (M)
$0.1 < \text{Acute ERV} \leq 1$	1
$0.01 < \text{Acute ERV} \leq 0.1$	10
$0.001 < \text{Acute ERV} \leq 0.01$	100
$0.0001 < \text{Acute ERV} \leq 0.001$	1000
Continue in factor 10 intervals	

Chronic ERV _{compound} (mg/l)	Chronic Multiplication factors (M)	
	No rapid environmental transformation	Rapid environmental transformation
0.01 < Chronic ERV ≤ 0.1	1	-
0.001 < Chronic ERV ≤ 0.01	10	1
0.0001 < Chronic ERV ≤ 0.001	100	10
0.00001 < Chronic ERV ≤ 0.0001	1000	100
Continue in factor 10 intervals		

”

Annex 9, Appendix VI

Insert the following references under section 5 “Metals and metal compounds”:

“Garman, E.R., Meyer, J.S., Bergeron, C.M., Blewett, T.A., Clements, W.H., Elias, M.C., Farley, K.J., Gissi, F. and Ryan, A.C. (2020), Validation of Bioavailability-Based Toxicity Models for Metals. *Environmental Toxicology & Chemistry*, 39: 101-117.

OECD (2015). Guidance on selecting a strategy for assessing the ecological risk of organometallic and organic metal salt substances based on their environmental fate. OECD Series on Testing and Assessment nr. 212. OECD, Paris, France.

Tipping, E., Lofts, S., and Sonke, J.E. (2011). Humic Ion-Binding Model VII: a revised parameterisation of cation-binding by humic substances. *Environmental Chemistry* 8 225—235.

Annex 10

A10.1 In the third sentence delete “(SSIMs).

Amend the two last sentences of the paragraph to read as follows: “The experimental work on several metals and metal compounds upon which this Test Guidance is based has been conducted and reported (references 5 to 15, this annex). This test guidance has subsequently also been published as a guidance document by OECD (reference 16).”.

A10.1.2 Replace the term “dissolution/transformation” in the last sentence with Transformation/Dissolution”.

A10.1.4 Insert the following new paragraph:

“A10.1.4 This test guidance is not applicable to organometallic compounds.”

A10.2.2.1 Amend the end of the last sentence to read as follows: “...is indistinguishable from soluble forms, and to verify the pH dependency of the dissolution, in preparation of the full transformation/dissolution test (see A9.7.2.3).”

A10.2.3.1 Amend the end of the paragraph to read as follows: “... using a loading of 1 mg/l, 0.1 mg/l, or 0.01 mg/l depending on the transformation rate.”.

A10.2.3.2 Amend the last sentence to read as follows:

“...except for the 28-day full test where the pH range of 5.5 to 8.5 is recommended if technically feasible to take into consideration possible long-term effects on acidic lakes.”

A10.2.3.3 In the first sentence, replace “while massive are tested” with “while massive forms are tested”.

Add the following sentence at the end of the existing paragraph: “The tested material should also be free from oxidation/corrosion layers due to storage, given the latter may disturb the transformation rate. Appropriate pre-treatment of the samples is recommended.”

- A10.4 (f) Amend the end of the sentence to read “or equivalent technique, and particle size distribution;”.
- A10.5 In sub-paragraph (d), delete “radial impeller”.
- In sub-paragraph (e) replace “(A10.5.1.7) acrodisc filter” with “(A10.5.1.10) filter”
- In sub-paragraph (k) replace “coupled axial plasma spectrometry” with “coupled plasma mass spectrometry” and amend the end to read as follows: “...the lowest chronic ecotoxicity reference value or the lowest acute ecotoxicity reference value if only a 7-day test is conducted;”.
- A10.5.1.2 Amend the text between brackets in the first sentence to read as follows: “(e.g. HCl or aqua regia)”
- Insert the following new second sentence: “Specific attention to the type of glassware is required for metals that can be released from the glass.”
- In the third sentence insert “e.g.” before “one or two two-litre reaction kettles”
- Footnote 2 Replace “dissolution/transformation protocol” with “Transformation/Dissolution Protocol”.
- A10.5.1.4 Amend to read as follows:
- “A10.5.1.4 The concentration of total organic carbon in the medium before adding the substance, should not exceed 2.0 mg/l.”
- A10.5.1.5 In the first sentence, replace “transformation of the metal compound” with “transformation of the metal or metal compound”.
- A10.5.1.7 Insert the following new note 3 under table A10.1:
- NOTE 3:** *Equilibration via headspace is recommended given CO₂ gas bubbling does not guarantee equal distribution between different test vessels.*”
- A10.5.1.9 Amend to read as follows:
- “A10.5.1.9 During the full transformation/dissolution test, agitation should be used which is sufficient to maintain the flow of aqueous medium over the test substance while maintaining the integrity of the surface of the test substance and of any solid reaction product coatings formed during the test. For 1 l of aqueous medium, this may be accomplished by the use of a 1.0 to 3.0 l flask capped with a rubber stopper and placed on an orbital or laboratory shaker set at 100 r.p.m. Other methods of gentle agitation may be used provided they meet the criteria of surface integrity and homogeneous solution.”.
- A10.5.1.10 Amend the end of the first sentence to read as follows: “...which will in turn depend on particle size distributions, the shape of the particles and particle density.”
- Replace the last two sentences (“Hence, filtration...may be useful”) with:
- “Alternative techniques may be considered in case of finer particles. If there is concern that particles will remain in suspension, then filtration efficiency should be checked prior to any testing. Options that could be considered to increase filtration efficiency include centrifugation followed by filtration, or waiting for about 5 minutes for the suspension to settle prior to taking a solution sample.”.

- A10.5.2.1 Amend the first paragraph to read as follows:
 “A suitable validated analytical method for the total dissolved metal analysis is essential to the study. The analytical detection limit should preferably be 5 times lower than the appropriate chronic ecotoxicity reference value, or the acute ecotoxicity reference value in case a 7-day test is conducted.”
- A10.5.2.3.1 Delete “(e.g. 37-44 μm)” and add the following sentence at the end of the paragraph: “This variability may be higher at the lower loadings.”
- A10.5.2.3.3 Amend the end of the final sentence in the second paragraph to read as follows: “with target pH 1 and analysed for total dissolved metal concentration.”
- A10.5.2.3.5 Amend the final sentence to read as follows: “It is a requirement to check the specific surface area of powder samples.”
- A10.5.3.1 Amend the heading to read as follows: “*Screening transformation/dissolution test – sparingly soluble metal compounds*”.
- A10.5.3.1.1 In the last sentence, in the text between brackets, replace “addition of the solids” with “addition of the test material”.
- A10.5.3.1.2 Amend the end of the first sentence to read as follows: “rapidly and vigorously (e.g. on an orbital shaker at 200 rpm, if feasible).”
- A10.5.3.2 Amend the heading to read: “*Full transformation/dissolution test - metals and metal compounds*”
- A10.5.3.2.2 Amend to read as follows:
 “For 7-day test, substance loadings of 1, 10 and 100 mg/l, respectively, are added to the test vessels (number of which depends on the reproducibility as established in subsection A10.5.2.3), containing the aqueous medium. The test vessels are closed (but allowing for equilibration with air if required) and agitated as described in A10.5.1.9. If a 28-day test is to be conducted, then the loading may be 0.01 mg/l, 0.1 mg/l or 1 mg/l depending on the transformation rate. The test with 1 mg/l loading may be extended to 28 days, provided that the same pH value is to be chosen for both 7 day and 28-day tests. The 7-day tests are only conducted at pH ranges of 6 up to 8.5, while a somewhat broader pH range of 5.5 and 6 to 8.5 is recommended if technically feasible for the 28-day tests. A concurrent control test with no substance loaded (i.e. a blank test solution) is required. At established time intervals (e.g. 2 hours, 6 hours, 1, 4 and 7 days for the short-term test and additionally at e.g. 14, 21 and 28 days for the long-term test), the temperature, pH and dissolved O₂ concentrations are measured in each test vessel, and at least two samples (e.g. 10 - 15 ml) are drawn by syringe from each test vessel. The solid and dissolved fractions....”
 [*the rest of the text remains unchanged*].
 At the end of the last sentence replace “(long term test)” with “(the long-term test)”.
- A10.6.2.1 Insert the following new paragraph at the end of the section:
 “The release rate may also be expressed relative to the surface area of the test substance (e.g. $\mu\text{g}/\text{mm}^2$) to allow for a comparison of the release rates between different surface loadings or particle sizes.”
- A10.6.2.2 Amend to read as follows:
 “A10.6.2.2 *Long-term test*
 The dissolved metal concentrations, measured from the 1 mg/l loading during the 28-day test, are plotted versus time and the transformation/dissolution kinetics determined, if possible, as described in A10.6.2.1.”.

Annex 10, Appendix

Insert the following new references:

12. Skeaff, J.M., Hardy, D.J. and King, P. (2008), A new approach to the hazard classification of alloys based on transformation/dissolution. *Integr Environ Assess Manag*, 4: 75-93. https://doi.org/10.1897/IEAM_2007-050.1
 13. Skeaff, J., Adams, W.J., Rodriguez, P., Brouwers, T. and Waeterschoot, H. (2011), Advances in metals classification under the United Nations globally harmonized system of classification and labeling. *Integr Environ Assess Manag*, 7: 559-576. <https://doi.org/10.1002/ieam.194>
 14. Skeaff, J.M. and Beaudoin, R. (2015), Transformation/dissolution characteristics of a nickel matte and nickel concentrates for acute and chronic hazard classification. *Integr Environ Assess Manag*, 11: 130-142. <https://doi.org/10.1002/ieam.1573>
 15. Huntsman-Mapila, P., Skeaff, J.M., Pawlak, M. and Beaudoin, R. (2016), Addressing aquatic hazard classification for metals, metal compounds and alloys in marine systems, *Marine Pollution Bulletin* 109:550-557. <https://doi.org/10.1016/j.marpolbul.2016.03.055>
 16. OECD Environment Health and Safety Publications; Series on Testing and Assessment n° 29. Guidance document on Transformation Dissolution of Metals and Metal Compounds in Aqueous media, July 2001.
-