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

26 November 2018

Sub-Committee of Experts on the Globally Harmonized System of Classification and Labelling of Chemicals

Thirty-sixth session

Geneva, 5-7 December 2018

Item 3 (d) of the provisional agenda

Classification criteria and related hazard communication: use of non-animal testing methods for classification of health hazards

Revision of Chapter 3.2 to fully incorporate non-animal test methods

Transmitted by the experts from the United Kingdom and the Netherlands on behalf of the informal working group on the use of non-animal test methods for classification of health hazards

Introduction

1. This informal document revises and replaces UN/SCEGHS/36/INF.6 to accommodate a limited number of editorial changes to the text of the revised Chapter 3.2, and updates the Sub-Committee on one specific aspect of the work of the Informal Working Group regarding classification based on pH, as reported in ST/SG/AC.10/C.4/2018/29.

Discussion

2. As planned, the informal working group held a webinar on 30 October 2018 after ST/SG/AC.10/C.4/2018/29 and informal document INF.6 had been submitted for discussion at the thirty-sixth session. The purpose was to undertake a final review of the revised text of Chapter 3.2, which now properly reflects the contribution of non-animal test methods in classifying substances and mixtures for skin corrosion and skin irritation.

3. A number of minor editorial changes and clarifications were agreed in the following sections and paragraphs: 3.2.1.2, 3.2.1.3, 3.2.2.1, 3.2.2.2.2.5, 3.2.2.3.1, 3.2.2.3.2, 3.2.2.3.3.1, 3.2.2.3.4.1, 3.2.2.6.1, Fig 3.2.1, 3.2.3.1.1, 3.2.3.1.2, 3.2.3.1.3, Decision logic 3.2.1, 3.2.5.3.2.4, 3.2.5.3.2.5, 3.2.5.3.4, 3.2.5.3.5.2.3 and 3.2.5.3.5.2.5.

4. In addition, the text on classification of substances and mixtures on the basis of chemical properties (3.2.2.5 and 3.2.3 respectively) was reviewed, and possible amendments considered. As indicated in paragraph 21 of ST/SG/AC.10/C.4/2018/29, this had been the subject of active discussion in the working group in the final stages of finalising the text. It was concluded that the discussions on the details, including the classification outcome when there is extreme pH and low acid/alkaline reserve, were not sufficiently mature and some working group members felt there had been insufficient time to consult their stakeholders. The relevant paragraphs of the text (3.2.2.5 and 3.2.3.1.3) in informal document INF.6/Rev.1 therefore revert to the corresponding existing text in GHS 7th revised edition (3.2.2.2.5 and 3.2.3.1.2) with minor editorial changes already agreed by the working group. Further, Figure

3.2.1 was adapted to retain the approach in Figure 3.2.1 of GHS 7th revised edition with regard to classification in case of with extreme pH and low buffer capacity.

5. Subject to the Sub-Committee agreeing proposed terms of reference for further work on non-animal test methods in the next biennium, the issue will be revisited when the informal working group addresses its next priority, Chapter 3.3 on serious eye damage and eye irritation. In this chapter paragraphs 3.3.2.2.4 and 3.3.3.1.2 mirror the corresponding paragraphs on pH in Chapter 3.2. The United Kingdom and the Netherlands, who lead this work, are confident that agreement can be reached in the next biennium, leading to consistent text in both Chapters 3.2 and 3.3.

Outcome

6. Annex 1 to this paper updates the Annex to ST/SG/AC.10/C.4/2018/29 by listing the additional changes introduced in Chapter 3.2 in line with the further input by the informal working group on 30 October and subsequent discussions. Annex 2 sets out the revised text of Chapter 3.2 as proposed by the Informal Working Group. As before, text in black is unchanged from GHS 7th revised edition, and new text relative to GHS 7th revised edition is shown in red; changes relative to INF.6 are shown by ~~strikethrough~~ where text is deleted and in blue where new text has been introduced. However, presentational improvements in the lower part of Decision logic 3.2.1, including deletion of a blank horizontal box are not shown.

Annex 1

Amend the revised chapter 3.2 as presented in Working Document 2018/29 as follows:

3.2.1 Replace 3.2.1.2 with the following:

"3.2.1.2 To classify, all available and relevant information on skin corrosion/irritation is collected and its quality in terms of adequacy and reliability is assessed. Wherever possible classification should be based on data generated using internationally validated and accepted methods, such as OECD Test Guidelines (TG) or equivalent methods. Sections 3.2.2.1 to 3.2.2.6 provide classification criteria for the different types of information that may be available."

Replace 3.2.1.3, as introduced in Working Document 2018/29, with the following:

"3.2.1.3 A tiered approach (see 3.2.2.7) 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 3.2.2.7.3) or where data individually are insufficient to conclude on the classification, an overall weight of evidence approach is used (see 1.3.2.4.9 and 3.2.5.3.1)."

3.2.2.1 Replace the text of 3.2.2.1, as introduced in Working Document 2018/29, with the following:

"Existing reliable and good quality human data on skin corrosion/irritation should be given high weight where relevant for classification (see 3.2.5.3.2) and should be the first line of evaluation, as this gives information directly relevant to effects on the skin. 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."

In paragraph 3.2.2.2.2.5 as renumbered in Working Document 2018/29, replace the text with:

"An irritation category (Category 2) is presented in Table 3.2.2 using the results of animal testing. Authorities (e.g. for classifying pesticides) also have available a less severe mild irritation category (Category 3). Several criteria distinguish the two categories (Table 3.2.2). They mainly differ in the severity of skin reactions. The major criterion for the irritation category is that at least 2 of 3 tested animals have a mean score of ≥ 2.3 and ≤ 4.0. For the mild irritation category, the mean score cut-off values are ≥ 1.5 and < 2.3 for at least 2 of 3 tested animals. Test materials in the irritation category are excluded from the mild irritation category.

Replace the text in Table 3.2.2 regarding "Mild irritation", in the second column, with the following:

"Mean score of ≥ 1.5 and < 2.3 for erythema/eschar or for oedema in at least 2 of 3 tested animals from gradings at 24, 48 and 72 hours after patch removal or, if reactions are delayed, from grades on 3 consecutive days after the onset of skin reactions (when not included in the irritant category above)".

3.2.2.3 Replace the text of 3.2.2.3.1, as introduced in Working Document 2018/29, with the following:

"3.2.2.3.1 The currently available individual *in vitro/ex vivo* test methods address either skin irritation or skin corrosion, but do not address both endpoints in one single test. Therefore, classification based solely on *in vitro/ex vivo* test results may require data from more than one method. For authorities implementing category 3 it is important to note that the currently available internationally validated and accepted *in vitro/ex vivo* test methods do not allow identification of substances classified as category 3."

Replace the text of 3.2.2.3.2, as introduced in Working Document 2018/29, with the following:

"3.2.2.3.2 Wherever possible classification should be based on data generated using internationally validated and accepted *in vitro/ex vivo* test methods, and the classification criteria provided in these test methods needs to be applied. *In vitro/ex vivo* data can only be used for classification when the tested substance is within the applicability domain of the test methods used. Additional limitations described in the published literature should also be taken into consideration."

Replace the text of 3.2.2.3.3.1, as introduced in Working Document 2018/29, with the following:

"3.2.2.3.3.1 Where tests have been undertaken in accordance with OECD Test Guidelines (TGs) 430, 431, or 435, a substance is classified for skin corrosion in category 1 (and, where possible and required into sub-categories 1A, 1B or 1C) based on the criteria in Table 3.2.6."

Replace the text of 3.2.2.3.4.1, as introduced in Working Document 2018/29, with the following:

"3.2.2.3.4.1 Where classification for corrosivity can be excluded and where tests have been undertaken in accordance with OECD Test Guideline 439, a substance is classified for skin irritation in category 2 based on the criteria in Table 3.2.7."

3.2.2.5 Replace the text of 3.2.2.5, as inserted in Working Document 2018/29, with the following:

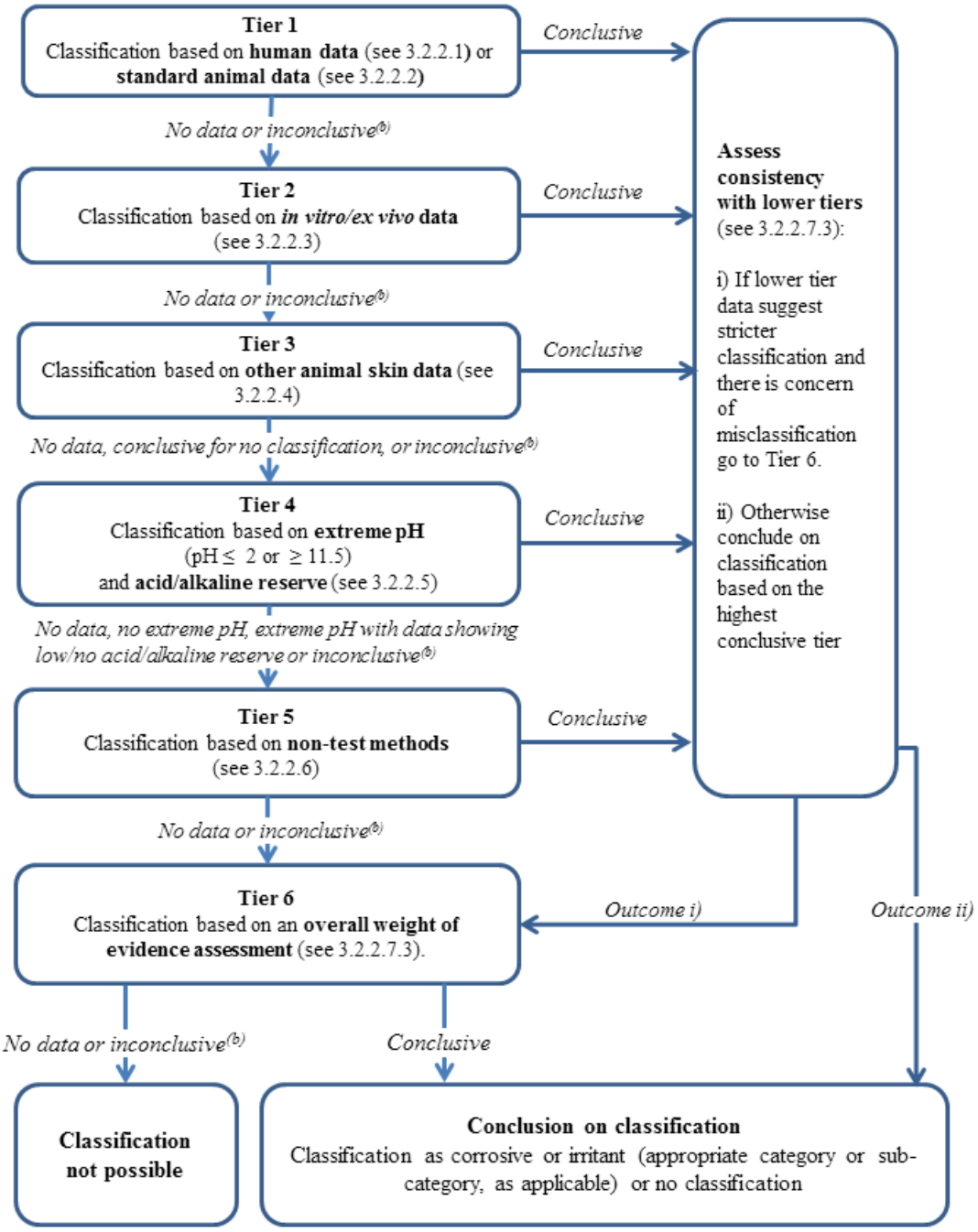
"Skin effects may be indicated by pH extremes such as ≤ 2 and ≥ 11.5 especially when associated with significant acid/alkaline reserve (buffering capacity). Generally, such substances are expected to produce significant effects on the skin. In the absence of any other information, a substance is considered corrosive (Skin Category 1) if it has a pH ≤ 2 or a pH ≥ 11.5. However, if consideration of acid/alkaline reserve suggests the substance may not be corrosive despite the low or high pH value, this needs to be confirmed by other data, preferably from an appropriate validated *in vitro/ex vivo* test. Buffering capacity and pH can be determined by test methods including OECD TG 122."

3.2.2.6 Replace the text of 3.2.2.6.1, as inserted in Working Document 2018/29, with the following:

"3.2.2.6.1 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. Such methods include computer models predicting qualitative structure-activity relationships (structural alerts, SAR);

quantitative structure-activity relationships (QSARs); computer expert systems; and read-across using analogue and category approaches."

3.2.2.7 Replace the new Figure 3.2.1, as introduced in Working Document 2018/29 as follows:



3.2.3 Replace the text of 3.2.3.1.1, as inserted in Working Document 2018/29, with the following:

"3.2.3.1.1 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 3.2.1). If classification is not possible using the tiered approach, then the approach described in 3.2.3.2 (bridging principles), or, if that is not applicable 3.2.3.3 (calculation method) should be followed."

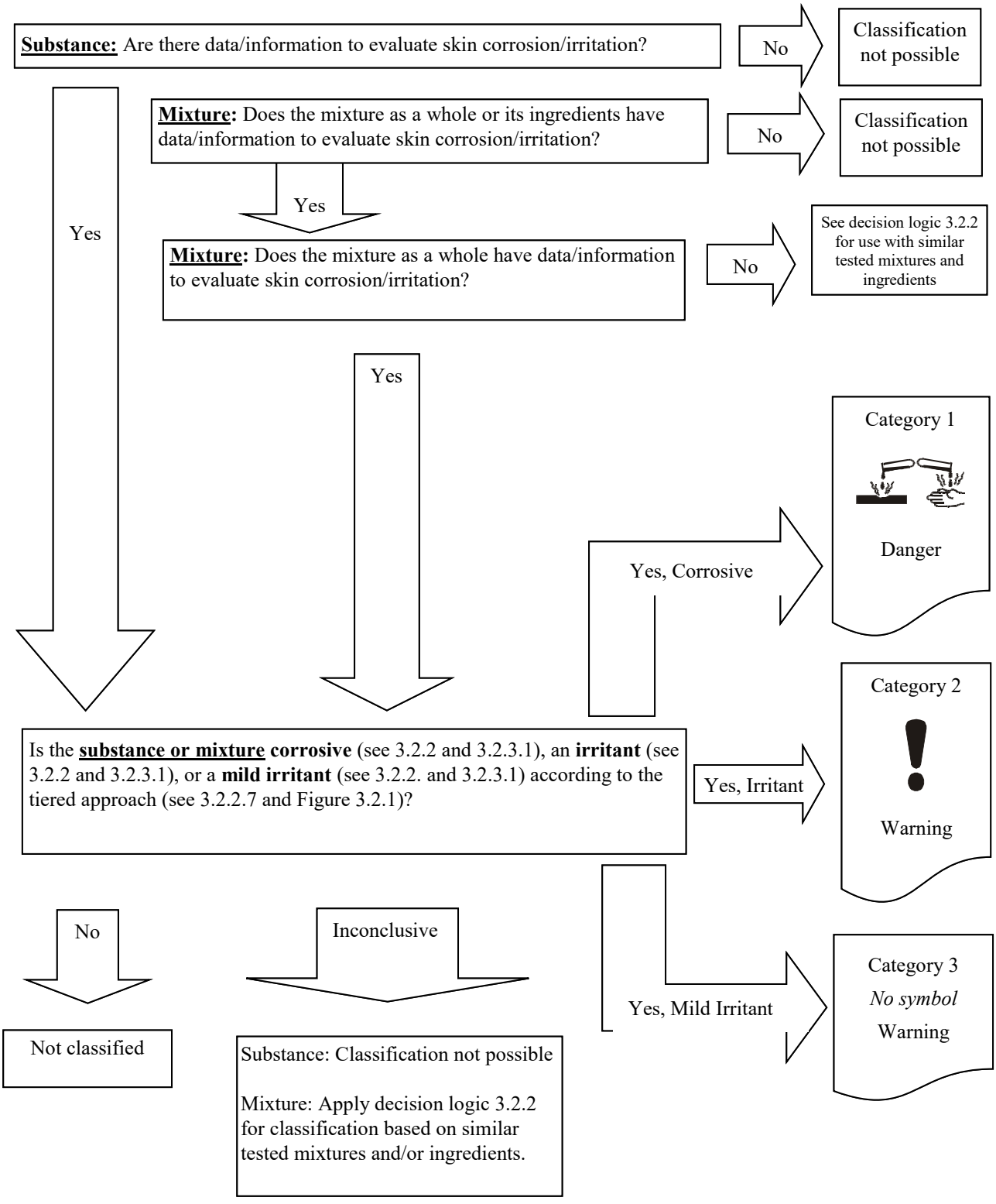
Replace the text of 3.2.3.1.2, as inserted in Working Document 2018/29, with the following:

"3.2.3.1.2 *In vitro/ex vivo* data generated from validated test methods, may not have been validated using mixtures, however these methods are considered broadly applicable to mixtures, but can only be used for classification of mixtures when all ingredients of the mixture fall within the applicability domain of the test methods used. Specific limitations regarding applicability domains are described in the respective 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 test method is limited, data interpretation should be exercised with caution, or the results should be considered not applicable."

Add a new paragraph 3.2.3.1.3 using text from the current GHS Rev 7 as follows:

"3.2.3.1.3 In the absence of any other information, a mixture is considered corrosive (Skin Category 1) if it has a pH $\leq$ 2 or a pH $\geq$ 11.5. However, if consideration of acid/alkaline reserve suggests the substance may not be corrosive despite the low or high pH value, this needs to be confirmed by other data, preferably from an appropriate validated *in vitro/ex vivo* test."

3.2.5.1 Replace Decision logic 3.2.1 and notes with the Decision logic as follows:


Substance: Are there data/information to evaluate skin corrosion/irritation?
No
Classification not possible
Mixture: Does the mixture as a whole or its ingredients have data/information to evaluate skin corrosion/irritation?
No
Classification not possible
Yes
Yes
Mixture: Does the mixture as a whole have data/information to evaluate skin corrosion/irritation?
No
See decision logic 3.2.2 for use with similar tested mixtures and ingredients
Yes
Category 1
Danger
Yes, Corrosive
Is the substance or mixture corrosive (see 3.2.2 and 3.2.3.1), an irritant (see 3.2.2 and 3.2.3.1), or a mild irritant (see 3.2.2. and 3.2.3.1) according to the tiered approach (see 3.2.2.7 and Figure 3.2.1)?
Yes, Irritant
Category 2
Warning
No
Inconclusive
Yes, Mild Irritant
Category 3
No symbol
Warning
Not classified
Substance: Classification not possible
Mixture: Apply decision logic 3.2.2 for classification based on similar tested mixtures and/or ingredients.

3.2.5.2 In Decision logic 3.2.2, replace footnote number "3" by "1" and replace footnote number "4" by "2" on the second page of the Decision logic in both text boxes, leading to the following text:

"Does the mixture contain one or more corrosive or irritant ingredients[1] when the additivity approach applies (see 3.2.3.3.2 and Table 3.2.3), and where the sum of concentrations of ingredients classified as[2]:".

3.2.5.3 Replace the text of 3.2.5.3.2.4, as inserted in Working Document 2018/29, with the following:

"3.2.5.3.2.4 Specific criteria for HPT results leading to classification as category 2 (skin irritation), category 3 (mild irritation) or not classified, have not been established at the international level. Therefore, the results of an HPT are generally used within a weight of evidence assessment. However, some competent authorities may provide specific guidance. A clearly negative result in an HPT with a sufficient number of volunteers after exposure to the undiluted substance for 4 hours can justify no classification."

Replace the text of 3.2.5.3.2.5, as inserted in Working Document 2018/29, with the following:

"3.2.5.3.2.5 Human case reports may be used for classification as corrosive if irreversible damage to the skin was observed. There are no internationally accepted classification criteria for irritation. Therefore, where competent authorities have not provided specific guidance on this matter, expert judgement may be required to evaluate whether the exposure duration and any available long-term follow-up information are sufficient to allow for a conclusion on classification. Cases resulting in irritation or no effects may not be conclusive on their own but can be used in a weight of evidence assessment."

Replace the text of 3.2.5.3.4, as inserted in Working Document 2018/29, with the following:

"Where *in vitro/ex vivo* tests have been undertaken in accordance with OECD Test Guidelines (TGs) 430, 431, 435 or 439, the criteria for classification in category 1 (and, where possible and required into sub-categories 1A, 1B or 1C) for skin corrosion and in category 2 for skin irritation are set out in Tables 3.2.6 and 3.2.7."

Replace the header of the second column in Table 3.2.6 from "OECD TG 430 (Transcutaneous Electrical Resistance test method)" to "OECD TG 430 Transcutaneous Electrical Resistance test method."

Replace the text of 3.2.5.3.5.2.3, as inserted in Working Document 2018/29, with the following:

"3.2.5.3.5.2.3 Repeated dose dermal studies (e.g. OECD TG 410 and 412) can be used to classify as corrosive when destruction of the skin is observed after the initial exposures. However, normally such exposures are avoided and corrosive effects may only be observed in the range-finding studies. Moreover, sub-categorisation for corrosion will rarely be possible due to a longer time period between start of exposure and first observation. The observation of skin irritation or the absence of skin irritating effects should be considered as not conclusive. Skin effects only observed after multiple exposures may indicate skin sensitisation rather than skin irritation."

Replace the text of 3.2.5.3.5.2.5, as inserted in Working Document 2018/29, with the following:

"3.2.5.3.5.2.5 Irritation data from the Local Lymph Node Assay (e.g. OECD TG 429, 442A and 442B) should normally not be used for classification as the test substance is applied to the dorsum of the ear by open topical application,

and in some cases specific vehicles for enhancement of skin penetration are used. Further, due to the proportional increase of skin thickness associated with increased body weight, the skin thickness of mice deviates significantly from that of rabbits and humans."

Annex 2

Chapter 3.2

Skin corrosion/irritation

3.2.1 Definitions and general considerations

3.2.1.1 *Skin corrosion* refers to the production of irreversible damage to the skin; namely, visible necrosis through the epidermis and into the dermis occurring after exposure to a substance or mixture.

Skin irritation refers to the production of reversible damage to the skin occurring after exposure to a substance or mixture.

3.2.1.2 To classify, all available and relevant information on skin corrosion/irritation is collected and its quality in terms of adequacy and reliability is assessed. Wherever possible classification should be based on data generated using internationally validated and accepted methods, such as OECD Test Guidelines (TG) or equivalent methods. Sections 3.2.2.1 to 3.2.2.6 provide classification criteria for the different types of information that may be available.

3.2.1.3 A *tiered approach* (see 3.2.2.7) 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 3.2.2.7.3) or where data individually are insufficient to conclude on the classification, an overall weight of evidence approach is used (see 1.3.2.4.9 and 3.2.5.3.1).

3.2.1.4 Guidance on the interpretation of criteria and references to relevant guidance documents are provided in 3.2.5.3.

3.2.2 Classification criteria for substances

Substances can be allocated to one of the following three categories within this hazard class:

(a) Category 1 (skin corrosion)

This category may be further divided into up to three sub-categories (1A, 1B and 1C) which can be used by those authorities requiring more than one designation for corrosivity.

Corrosive substances should be classified in Category 1 where sub-categorization is not required by a competent authority or where data are not sufficient for sub-categorization.

When data are sufficient, and where required by a competent authority, substances may be classified in one of the three sub-categories 1A, 1B or 1C.

(b) Category 2 (skin irritation)

(c) Category 3 (mild skin irritation)

This category is available for those authorities that want to have more than one skin irritation category (e.g. for classifying pesticides).

3.2.2.1 ***Classification based on human data***

Existing reliable and good quality human data on skin corrosion/irritation should be given high weight where relevant for classification (see 3.2.5.3.2). ~~Information from human exposure~~ and should be the first line of evaluation, as this gives information directly relevant to effects on the skin. 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 ~~Chapter 1.1 paragraph~~ 1.1.2.5 (c), ~~and Chapter 1.3, paragraphs~~ 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.

3.2.2.2 ***Classification based on standard animal test data***

OECD TG 404 is the currently available internationally validated and accepted animal test for classification as skin corrosive or irritant (See Tables 3.2.1 and 3.2.2, respectively) and is the standard animal test. The current version of OECD TG 404 uses a maximum of 3 animals. Results from animal studies conducted under previous versions of OECD TG 404 that used more than 3 animals are also considered standard animal tests when interpreted in accordance with 3.2.5.3.3.

3.2.2.2.1 *Skin corrosion*

3.2.2.2.1.1 A substance is corrosive to skin when it produces destruction of skin tissue, namely, visible necrosis through the epidermis and into the dermis, in at least one tested animal after exposure for up to 4 hours.

3.2.2.2.1.2 For those authorities wanting more than one designation for skin corrosion, up to three sub-categories are provided within the corrosion category (Category 1, see Table 3.2.1): sub-category 1A, where corrosive responses are noted following up to 3 minutes exposure and up to 1 hour observation; sub-category 1B, where corrosive responses are described following exposure greater than 3 minutes and up to 1 hour and observations up to 14 days; and sub-category 1C, where corrosive responses occur after exposures greater than 1 hour and up to 4 hours and observations up to 14 days.

Table 3.2.1: Skin corrosion category and sub-categories

	Criteria
Category 1	Destruction of skin tissue, namely, visible necrosis through the epidermis and into the dermis, in at least one tested animal after exposure ≤ 4 h
Sub-category 1A	Corrosive responses in at least one animal following exposure ≤ 3 min during an observation period ≤ 1 h
Sub-category 1B	Corrosive responses in at least one animal following exposure > 3 min and ≤ 1 h and observations ≤ 14 days
Sub-category 1C	Corrosive responses in at least one animal after exposures > 1 h and ≤ 4 h and observations ≤ 14 days

3.2.2.2.2 *Skin irritation*

3.2.2.2.2.1 A substance is irritant to skin when it produces reversible damage to the skin following its application for up to 4 hours.

3.2.2.2.2.2 An irritation category (Category 2) is provided that:

(a) recognizes that some test materials may lead to effects which persist throughout the length of the test; and

(b) acknowledges that animal responses in a test may be variable.

An additional mild irritation category (Category 3) is available for those authorities that want to have more than one skin irritation category.

3.2.2.2.2.3 Reversibility of skin lesions is another consideration in evaluating irritant responses. When inflammation persists to the end of the observation period in 2 or more test animals, taking into consideration alopecia (limited area), hyperkeratosis, hyperplasia and scaling, then a material should be considered to be an irritant.

3.2.2.2.2.4 Animal irritant responses within a test can be variable, as they are with corrosion. A separate irritant criterion accommodates cases when there is a significant irritant response but less than the mean score criterion for a positive test. For example, a test material might be designated as an irritant if at least 1 of 3 tested animals shows a very elevated mean score throughout the study, including lesions persisting at the end of an observation period of normally 14 days. Other responses could also fulfil this criterion. However, it should be ascertained that the responses are the result of chemical exposure. Addition of this criterion increases the sensitivity of the classification system.

3.2.2.2.2.5 An irritation category (Category 2) is presented in Table 3.2.2 using the results of animal testing. Authorities (e.g. for classifying pesticides) also have available a less severe mild irritation category (Category 3). Several criteria distinguish the two categories (Table 3.2.2). They mainly differ in the severity of skin reactions. The major criterion for the irritation category is that at least 2 of 3 tested animals have a mean score of ≥ 2.3 and ≤ 4.0. For the mild irritation category, the mean score cut-off values are ≥ 1.5 and < 2.3 for at least 2 of 3 tested animals. Test materials in the irritation category are excluded from the mild irritation category.

Table 3.2.2: Skin irritation categories [a,b]

Categories	Criteria
Irritation (Category 2) (applies to all authorities)	(1) Mean score of ≥ 2.3 and ≤ 4.0 for erythema/eschar or for oedema in at least 2 of 3 tested animals from gradings at 24, 48 and 72 hours after patch removal or, if reactions are delayed, from grades on 3 consecutive days after the onset of skin reactions; or (2) Inflammation that persists to the end of the observation period normally 14 days in at least 2 animals, particularly taking into account alopecia (limited area), hyperkeratosis, hyperplasia, and scaling; or (3) In some cases where there is pronounced variability of response among animals, with very definite positive effects related to chemical exposure in a single animal but less than the criteria above.
Mild irritation (Category 3) (applies to only some authorities)	Mean score of ≥ 1.5 and < 2.3 for erythema/eschar or for oedema in at least 2 of 3 tested animals from gradings at 24, 48 and 72 hours after patch removal or, if reactions are delayed, from grades on 3 consecutive days after the onset of skin reactions (when not included in the irritant category above).

[a] *Grading criteria are understood as described in OECD Test Guideline 404.*

[b] *Evaluation of a 4, 5 or 6-animal study should follow the criteria given in 3.2.5.3.3*

3.2.2.3 *Classification based on in vitro/ex vivo data*

3.2.2.3.1 The currently available individual *in vitro/ex vivo* test methods address either skin irritation or skin corrosion, but do not address both endpoints in one single test. Therefore, classification based solely on *in vitro/ex vivo* test results may require data from more than one method. For authorities implementing category 3 it is important to note that the currently available internationally validated and accepted *in vitro/ex vivo* test methods do not allow identification of substances classified as category 3.

3.2.2.3.2 Wherever possible classification should be based on data generated using internationally validated and accepted *in vitro/ex vivo* test methods, and the classification criteria provided in these test methods needs to be applied. *In vitro/ex vivo* data can only be used for classification when the tested substance is within the applicability domain of the test methods used. Additional limitations described in the published literature should also be taken into consideration.

3.2.2.3.3. *Skin corrosion*

3.2.2.3.3.1 Where tests have been undertaken in accordance with OECD Test Guidelines (TGs) 430, 431, or 435, a substance is classified for skin corrosion in category 1 (and, where possible and required into sub-categories 1A, 1B or 1C) based on the criteria in Table 3.2.6.

3.2.2.3.3.2 Some *in vitro/ex vivo* methods do not allow differentiation between sub-categories 1B and 1C (See Table 3.2.6). Where sub-categories are required by competent authorities and existing *in vitro/ex vivo* data cannot distinguish between the sub-categories, additional information has to be taken into account to differentiate between these two sub-categories. Where no or insufficient additional information is available, category 1 is applied.

3.2.2.3.3.3 A substance identified as not corrosive should be considered for classification as skin irritant.

3.2.2.3.4 *Skin irritation*

3.2.2.3.4.1 Where ~~a conclusion of~~ classification for corrosivity can be excluded and where tests have been undertaken in accordance with OECD Test Guideline 439, a substance is classified for skin irritation in category 2 based on the criteria in Table 3.2.7.

3.2.2.3.4.2 Where competent authorities adopt category 3, it is important to note that currently available *in vitro/ex vivo* test methods for skin irritation (e.g. OECD TG 439) do not allow for classification of substances in category 3. In this situation, if the classification criteria for either category 1 or 2 are not fulfilled, additional information is required to differentiate between category 3 and no classification.

3.2.2.3.4.3 Where competent authorities do not adopt category 3, a negative result in an internationally accepted and validated *in vitro/ex vivo* test for skin irritation, e.g. OECD TG 439, can be used to conclude as not classified for skin irritation.

3.2.2.4 *Classification based on other, existing skin data in animals*

Other existing skin data in animals may be used for classification, but there may be limitations regarding the conclusions that can be drawn (see 3.2.5.3.5). If a substance is highly toxic via the dermal route, an *in vivo* skin corrosion/irritation study may not have been conducted since the amount of test substance to be applied would considerably exceed the toxic dose and, consequently, would result in the death of the animals. When observations of skin corrosion/irritation in acute toxicity studies are made, these data may be used for classification, provided that the dilutions used and species tested are relevant. Solid substances (powders) may become corrosive or irritant when moistened or in contact with moist skin or mucous membranes. This is generally indicated in the standardised test methods. Guidance on the use of other existing skin data in animals including acute and repeated dose toxicity tests as well as other tests is provided in 3.2.5.3.5.

3.2.2.5 *Classification based on chemical properties*

Skin effects may be indicated by pH extremes such as ≤ 2 and ≥ 11.5 especially when associated with significant acid/alkaline reserve (buffering capacity). Generally, such substances are expected to produce significant effects on the skin. *In the absence of any other information, a* substance is considered corrosive (Skin Category 1) if it has a pH ≤ 2 or a pH ≥ 11.5. However, if consideration of acid/alkaline reserve suggests the substance may not be corrosive despite the low or high pH value, this needs to be confirmed by other data, preferably from an appropriate validated *in vitro/ex vivo* test. ~~If no additional data are available in case of extreme pH in combination with low buffer capacity the situation is considered non-conclusive.~~ Buffering capacity and pH can be determined by test methods including OECD TG 122.

3.2.2.6 *Classification based on non-test methods*

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

3.2.2.6.2 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(s) to be classified. Where adequate justification of the read-across approach is provided, it has in general higher weight than (Q)SARs.

3.2.2.6.3 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 recognised 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.

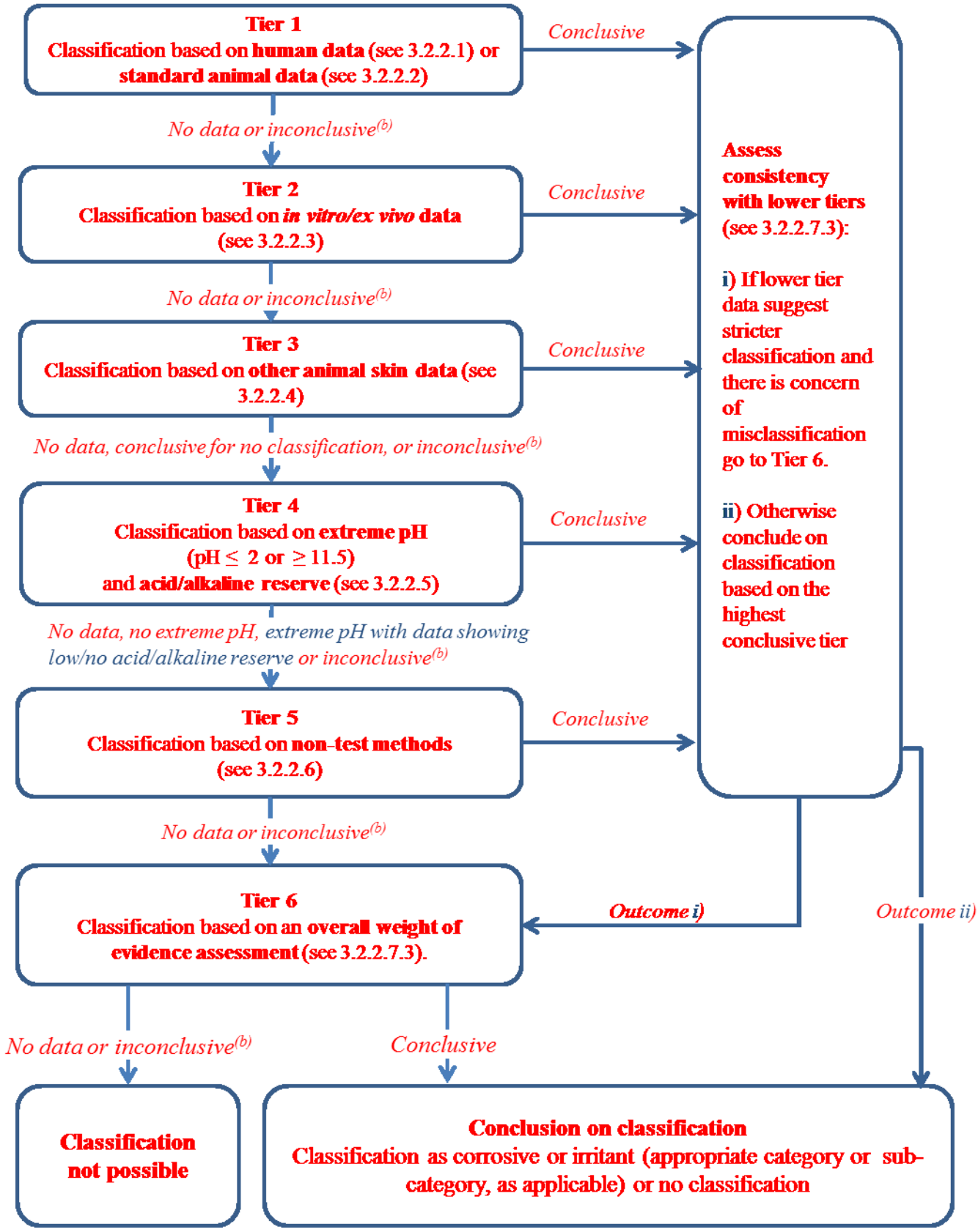
3.2.2.7 Classification in a tiered approach

3.2.2.7.1 A tiered approach to the evaluation of initial information should be considered, where applicable (Figure 3.2.1), recognising that not all elements may be relevant. However, all available and relevant information of sufficient quality needs to be examined for consistency with respect to the resulting classification.

3.2.2.7.2 In the tiered approach (Figure 3.2.1), existing human and animal data form the highest tier, followed by *in vitro/ex vivo* data, other existing skin data in animals, and then other sources of information. 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 approach.

3.2.2.7.3 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 approach. For example, having consulted the guidance in 3.2.5.3 as appropriate, classifiers concerned with a negative result for skin corrosion in an *in vitro/ex vivo* study when there is a positive result for skin corrosion in other existing skin data in animals would utilise an overall weight of evidence approach. The same would apply in the case where there is human data indicating irritation but positive results from an *in vitro/ex vivo* test for corrosion.

Figure 3.2.1: Application of the tiered approach for skin corrosion and irritation [a]



(a) Before applying the approach, the explanatory text in 3.2.2.7 as well as the guidance in 3.2.5.3 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 adequate to demonstrate irritancy, but inadequate to demonstrate absence of corrosivity
- Where competent authorities make use of the mild skin irritation category 3, the available data may not be capable of distinguishing between category 3 and category 2, or between category 3 and no classification.
- The method used to generate the available data may not be suitable for concluding on no classification (see 3.2.2. and 3.2.5.3 for details). Specifically, *in vitro/ex vivo* and non-test methods need to be validated explicitly for this purpose.

3.2.3 Classification criteria for mixtures

3.2.3.1 *Classification of mixtures when data are available for the complete mixture*

3.2.3.1.1 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 3.2.1). If classification is not possible using the tiered approach, then ~~apply~~ the approach described in 3.2.3.2 (bridging principles), or, if that is not applicable 3.2.3.3 (calculation method) ~~as appropriate~~ should be followed.

3.2.3.1.2 *In vitro/ex vivo* data generated from validated test methods may not have been validated using mixtures; although ~~they~~ these 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 test methods used. Specific limitations regarding applicability domains are described in the respective test methods, and should be taken into consideration as well as any further information on such limitations from the published literature. Where there ~~are~~ is reason~~s~~ to assume or evidence indicating that the applicability domain of a particular test method is limited, data interpretation should be exercised with caution, or the results should be considered not applicable.

3.2.3.1.3 In the absence of any other information, a mixture is considered corrosive (Skin Category 1) if it has a pH ≤ 2 or a pH ≥ 11.5. However, if consideration of acid/alkaline reserve suggests the substance may not be corrosive despite the low or high pH value, this needs to be confirmed by other data, preferably from an appropriate validated *in vitro/ex vivo* test.

[**Note:** the following sections are unchanged from GHS Rev 7:

3.2.3.2 Classification of mixtures when data are not available for the complete mixture: bridging principles

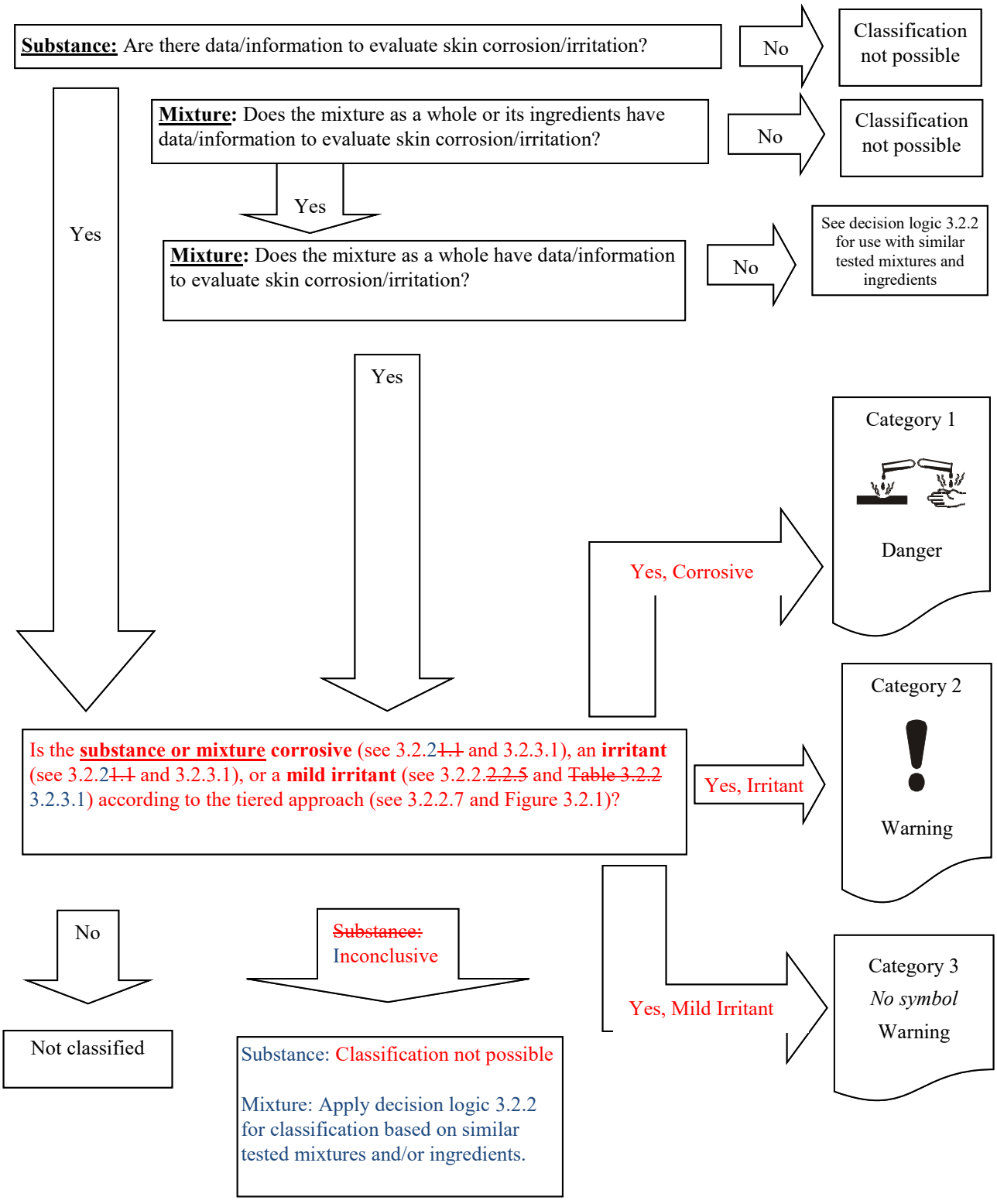
3.2.3.3 Classification of mixtures when data are available for all ingredients, or only for some ingredients of the mixture

3.2.4 Hazard communication]

3.2.5 Decision logics and guidance

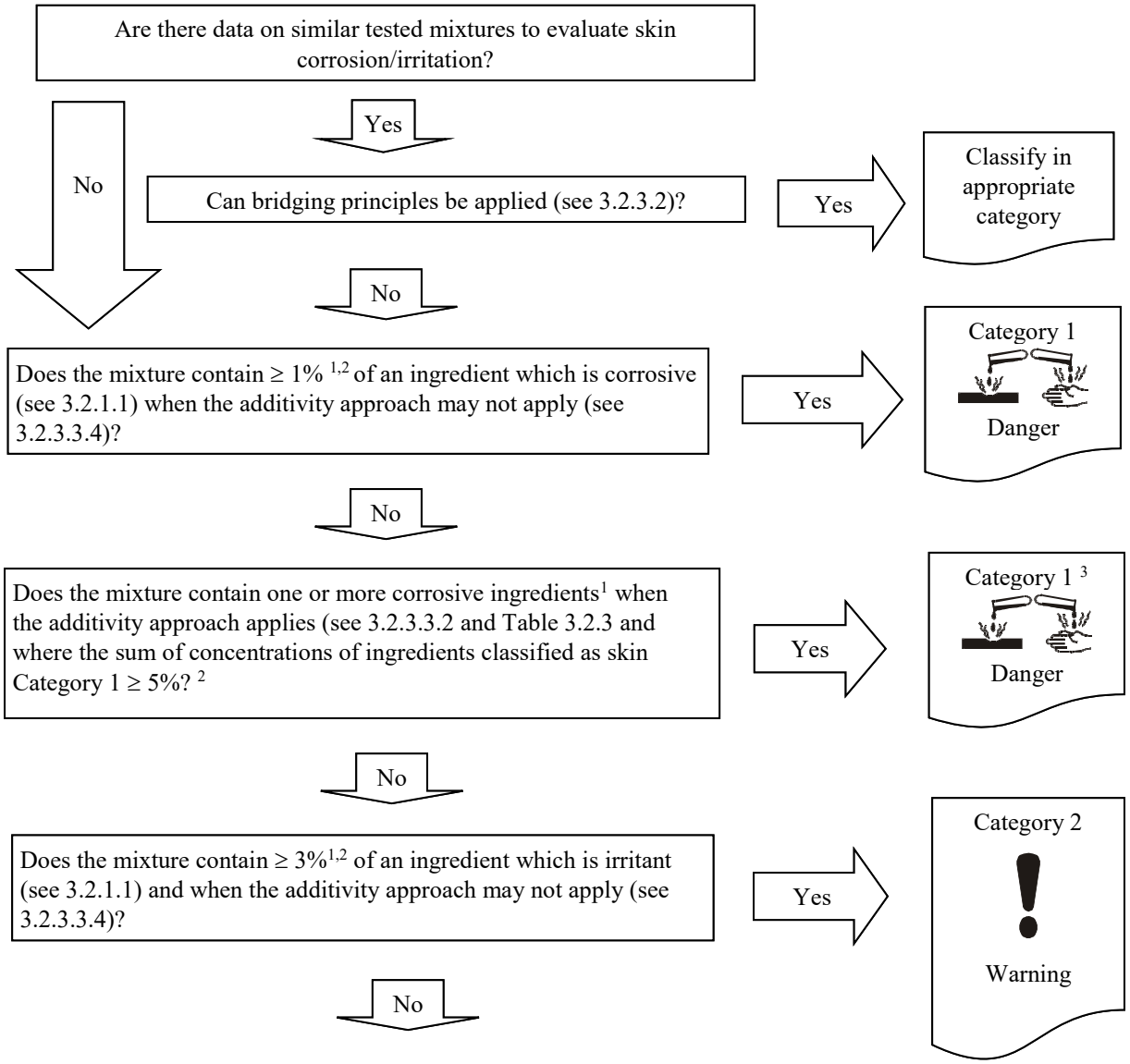
The decision logics which follow are not part of the harmonized classification system but are provided here as additional guidance. It is strongly recommended that the person responsible for classification study the criteria before and during use of the decision logics.

3.2.5.1 ***Decision logic 3.2.1 for skin corrosion/irritation***


Substance: Are there data/information to evaluate skin corrosion/irritation?
No
Classification not possible
Mixture: Does the mixture as a whole or its ingredients have data/information to evaluate skin corrosion/irritation?
No
Classification not possible
Yes
Yes
Mixture: Does the mixture as a whole have data/information to evaluate skin corrosion/irritation?
No
See decision logic 3.2.2 for use with similar tested mixtures and ingredients
Yes
Yes
Category 1
Danger
Yes, Corrosive
Is the **substance or mixture corrosive** (see 3.2.2~~1.1~~ and 3.2.3.1), an **irritant** (see 3.2.2~~1.1~~ and 3.2.3.1), or a **mild irritant** (see 3.2.2.~~2.2.5~~ and ~~Table 3.2.2~~ 3.2.3.1) according to the tiered approach (see 3.2.2.7 and Figure 3.2.1)?
Yes, Irritant
Category 2
Warning
No
~~Substance:~~ Inconclusive
Yes, Mild Irritant
Category 3
No symbol
Warning
Not classified
Substance: Classification not possible
Mixture: Apply decision logic 3.2.2 for classification based on similar tested mixtures and/or ingredients.

3.2.5.2 Decision logic 3.2.2 for skin corrosion/irritation [1 2 3]

Classification of mixtures on the basis of information/data on similar tested mixtures and/or ingredients

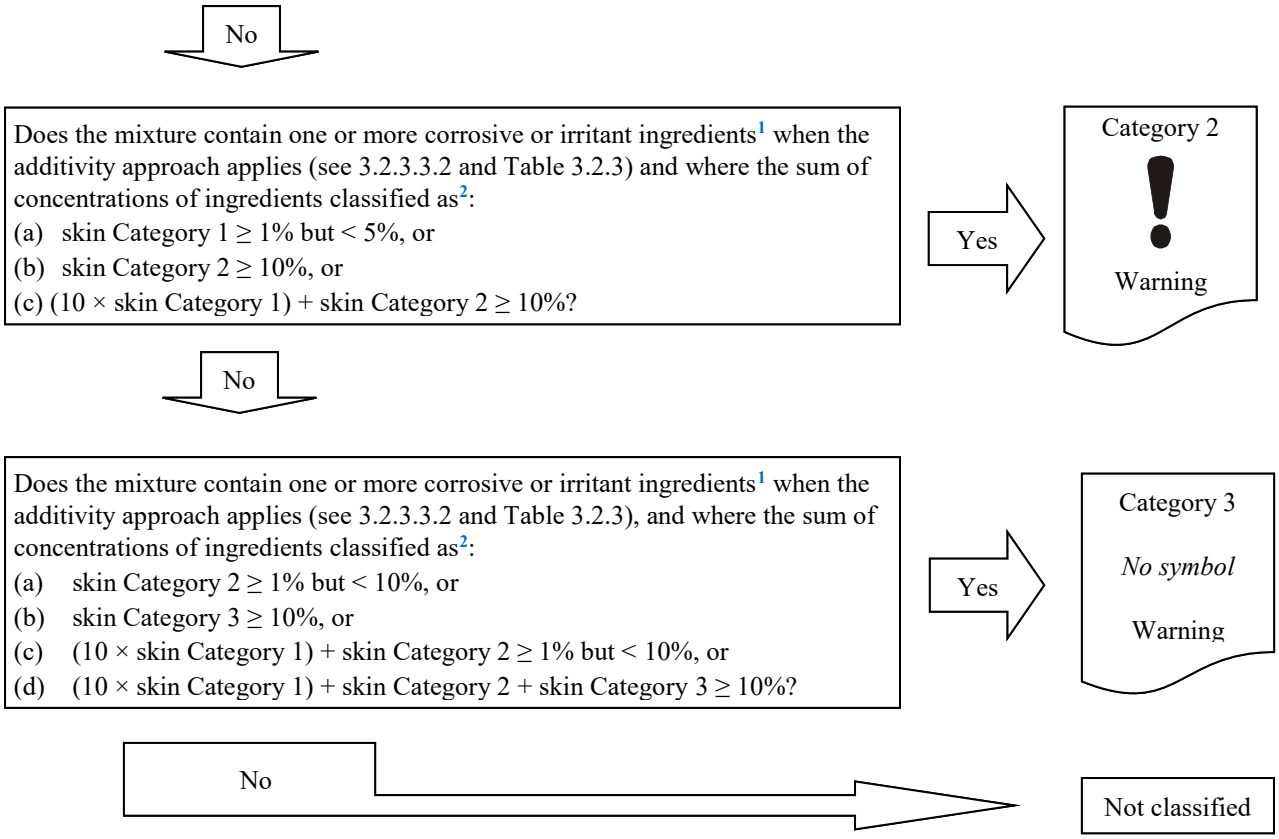


(Cont'd on next page)

[1] *Where relevant < 1%, see 3.2.3.3.1.*

[2] *For specific concentration limits, see 3.2.3.3.6. See also Chapter 1.3, para. 1.3.3.2 for "Use of cut-off values/concentration limits".*

[3] *See note to Table 3.2.3 for details on use of Category 1 sub-categories.*



3.2.5.3 Background guidance

3.2.5.3.1 *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 approach, is provided in OECD Guidance Document 203, An Integrated Approach on Testing and Assessment (IATA) for Skin Corrosion and Irritation.

3.2.5.3.2 *Guidance on the use of human data for classification as skin corrosion or skin irritation*

3.2.5.3.2.1 Human data generally refers to two types of data: prior human experience (e.g. published case studies from occupational, consumer, transport, emergency response scenarios, epidemiological studies) or from human tests (e.g. clinical trials, dermal patch test). Relevant, reliable and good quality human data is generally given high weight for classification. However, human data may have limitations. Further details on the strengths and limitations of human data for skin irritation/corrosion can be found in OECD guidance document 203 (section III. A, Part 1, Module 1).

3.2.5.3.2.2 Generally, Human Patch Tests (HPT) are performed to discriminate between irritant and non-irritant substances. Application of a corrosive substance to human skin is generally avoided. Therefore, another test is normally performed in advance to exclude corrosivity. The HPT alone does not normally discriminate between irritant and corrosive substances. In rare circumstances, there may be HPT data that can be used for classification as corrosive (e.g. application of an HPT after a false negative *in vitro* test). However, the combination of an HPT and sufficient other information on skin corrosion can be used for classification within a weight of evidence assessment.

3.2.5.3.2.3 Some competent authorities do not allow HPT testing solely for hazard identification (see 1.3.2.4.7) while some competent authorities recognize the use of HPT for classification as skin irritant.

3.2.5.3.2.4 Specific criteria for HPT results leading to classification as category 2 (skin irritation), category 3 (mild irritation) or not classified, have not been established at the international level. Therefore, the results of an HPT are generally used within a weight of evidence assessment. However, some competent authorities may provide specific guidance. A clearly negative result in an HPT with a sufficient number of volunteers after exposure to the undiluted substance for 4 hours can justify no classification.

3.2.5.3.2.5 Human case reports may be used for classification as corrosive if irreversible damage to the skin was observed. There are no internationally accepted classification criteria for irritation. Therefore, where competent authorities have not provided specific guidance on this matter, expert judgement may be required to evaluate ~~the sufficiency of~~ whether the exposure duration and ~~the~~ any available ~~of sufficient~~ long-term follow-up information are sufficient ~~and~~ to allow for a conclusion on ~~the~~ classification. Cases resulting in irritation or no effects may not be conclusive on their own but can be used in a weight of evidence assessment.

3.2.5.3.3 *Classification based on standard animal tests with more than 3 animals*

3.2.5.3.3.1 Classification criteria for the skin and eye hazard classes are detailed in the GHS in terms of a 3-animal test. It has been identified that some older test methods may have used up to 6 animals. However, the GHS criteria do not specify how to classify based on existing data from tests with more than 3 animals. Guidance on how to classify based on existing data from studies with 4 or more animals is given in the following paragraphs.

3.2.5.3.3.2 Classification criteria based on a 3-animal test are detailed in 3.2.2.2. Evaluation of a 4, 5 or 6-animal study should follow the criteria in the following paragraphs, depending on the number of animals tested. Scoring for erythema/eschar and oedema should be performed at 24, 48 and 72 hours after exposure or, if reactions are delayed, from grades on 3 consecutive days after the onset of skin reactions.

3.2.5.3.3.3 In the case of a study with 6 animals the following principles apply:

(a) The substance or mixture is classified as skin corrosion Category 1 if destruction of skin tissue (that is, visible necrosis through the epidermis and into the dermis) occurs in at least one animal after exposure up to 4 hours in duration;

(b) The substance or mixture is classified as skin irritation Category 2 if at least 4 out of 6 animals show a mean score per animal of ≥ 2.3 and ≤ 4.0 for erythema/eschar or for oedema;

(c) The substance or mixture is classified as skin irritation Category 3 if at least 4 out of 6 animals show a mean score per animal of ≥ 1.5 and < 2.3 for erythema/eschar or for oedema.

3.2.5.3.3.4 In the case of a study with 5 animals the following principles apply:

(a) The substance or mixture is classified as skin corrosion Category 1 if destruction of skin tissue (that is, visible necrosis through the epidermis and into the dermis) occurs in at least one animal after exposure up to 4 hours in duration;

(b) The substance or mixture is classified as skin irritation Category 2 if at least 3 out of 5 animals show a mean score per animal of ≥ 2.3 and ≤ 4.0 for erythema/eschar or for oedema;

(c) The substance or mixture is classified as skin irritation Category 3 if at least 3 out of 5 animals show a mean score per animal of ≥ 1.5 and < 2.3 for erythema/eschar or for oedema.

3.2.5.3.3.5 In the case of a study with 4 animals the following principles apply:

(a) The substance or mixture is classified as skin corrosion Category 1 if destruction of skin tissue (that is, visible necrosis through the epidermis and into the dermis) occurs in at least one animal after exposure up to 4 hours in duration;

(b) The substance or mixture is classified as skin irritation Category 2 if at least 3 out of 4 animals show a mean score per animal of ≥ 2.3 and ≤ 4.0 for erythema/eschar or for oedema;

(c) The substance or mixture is classified as skin irritation Category 3 if at least 3 out of 4 animals show a mean score per animal of ≥ 1.5 and < 2.3 for erythema/eschar or for oedema.

3.2.5.3.4 *Classification criteria based on in-vitro/ex vivo data*

Where *in vitro/ex vivo* tests have been undertaken in accordance with OECD Test Guidelines (TGs) 430, 431, 435 or 439, the criteria for classification in category 1 (and, where possible and required into sub-categories 1A, 1B or 1C) for skin corrosion and in category 2 for skin irritation are set out in Tables 3.2.6 and 3.2.7.

Table 3.2.6: Skin corrosion criteria for *in vitro/ex vivo* methods

Category	OECD TG 430 (Transcutaneous Electrical Resistance test method)	OECD TG 431 Reconstructed human Epidermis test methods: Methods 1, 2, 3, 4 as numbered in Annex 2 of OECD TG 431				OECD TG 435 Membrane barrier test method	
	Using rat skin discs corrosive chemicals are identified by their ability to produce a loss of normal *stratum corneum* integrity. Barrier function of the skin is assessed by recording the passage of ions through the skin. The electrical impedance of the skin is measured using transcutaneous electrical resistance (TER). A confirmatory test of positive results using a dye-binding step that assesses if an increase in ionic permeability is due to the physical destruction of the *stratum corneum* is performed in case of a reduced TER (less than or around 5 kΩ) in the absence of obvious damage. The criteria are based on the mean TER value in kΩ and sometimes on dye content.	Four similar methods where the test chemical is applied topically to a three-dimensional reconstructed human epidermis (RhE) which closely mimics the properties of the upper parts of human skin. The test method is based on the premise that corrosive chemicals are able to penetrate the *stratum corneum* by diffusion or erosion, and are cytotoxic to the cells in the underlying layers. Tissue viability is assessed by enzymatic conversion of the dye MTT into a blue formazan salt that is quantitatively measured after extraction from the tissues. Corrosive chemicals are identified by their ability to decrease tissue viability below defined threshold values. The criteria are based on the percent tissue viability following a defined exposure period.				An in vitro membrane barrier test method comprising a synthetic macromolecular bio-barrier and a chemical detection system (CDS). Barrier damage is measured after the application of the test chemical to the surface of the synthetic membrane barrier. The criteria are based on the mean penetration/breakthrough time of the chemical through the membrane barrier.	
						Type 1 chemicals (high acid/alkaline reserve)	Type 2 chemicals (low acid/alkaline reserve)
1	(a) mean TER value ≤ 5 kΩ and the skin discs are obviously damaged (e.g. perforated), or (b) mean TER value ≤ 5 kΩ, and (i) the skin discs show no obvious damage (e.g. perforation), but (ii) the subsequent confirmatory testing of positive results using a dye binding step is positive.	Method 1 < 35% after 3, 60 or 240 min exposure	Methods 2, 3, 4 < 50% after 3 min exposure; or ≥ 50% after 3 min exposure and < 15% after 60 min exposure			≤ 240 min	≤ 60 min
1A	Not applicable	Method 1 < 35% after 3 min exposure	Method 2 < 25% after 3 min exposure	Method 3 < 18% after 3 min exposure	Method 4 < 15% after 3 min exposure	0-3 min.	0-3 min
1B		≥ 35% after 3 min exposure and < 35% after 60 min exposure or ≥ 35% after 60 min exposure and < 35% after 240 min exposure	≥ 25% after 3 min exposure and fulfilling criteria for category 1	≥ 18% after 3 min exposure and fulfilling criteria for category 1	≥ 15% after 3 min exposure and fulfilling criteria for category 1	> 3 to 60 min.	> 3 to 30 min
1C						> 60 to 240 min.	> 30 to 60 min
Not classified as skin corrosive	(a) the mean TER value > 5 kΩ, or (b) the mean TER value ≤ 5 kΩ, and (i) the skin discs show no obvious damage (e.g. perforation), and (ii) the subsequent confirmatory testing of positive results using a dye binding step is negative	≥ 35% after 240 min exposure	≥ 50% after 3 min exposure and ≥ 15% after 60 min exposure			> 240 min.	> 60 min

Table 3.2.7 Skin irritation criteria for *in vitro* methods

Category	TG 439 Reconstructed Human Epidermis test methods
	Four similar methods (1-4) where the test chemical is applied topically to a three-dimensional reconstructed human epidermis (RhE) which closely mimics the properties of the upper parts of human skin. Tissue viability is assessed by enzymatic conversion of the dye MTT into a blue formazan salt that is quantitatively measured after extraction from the tissues. Positive chemicals are identified by their ability to decrease tissue viability below defined threshold levels. The criteria are based on mean percent tissue viability after exposure and post-treatment incubation.
1 or 2	Mean percent tissue viability (≤) 50%. Note: The RhE test methods covered by this TG cannot resolve between GHS categories 1 and 2. Further information on skin corrosion will be required to decide on its final classification [see also the OECD Guidance Document No. 203].
2	Mean percent tissue viability ≤ 50% and the test chemical is found to be noncorrosive (e.g., based on TG 430, 431 or 435)
Not classified as skin irritant or category 3	Mean percent-tissue viability > 50% Note: The RhE test methods covered by this TG cannot resolve between GHS optional category 3 and not classified as skin irritant. Further information on skin irritation is required for those authorities that want to have more than one skin irritation category.

3.2.5.3.5 *Guidance on the use of other existing skin data in animals for classification as skin corrosion or skin irritation*

3.2.5.3.5.1 General approach

All existing other animal data should be carefully reviewed and only used if they are conclusive for classification. In evaluating other existing skin data in animals, however, it should be recognised that the reporting of dermal lesions may be incomplete, testing and observations may be made in a species other than the rabbit, and species may differ in sensitivity in their responses. In general skin thickness decreases with body weight. However, other factors also affect species variability. In addition, for most of these tests, irritating and corrosive effects need to be avoided. Therefore, these effects may only be observed in range finding studies using a small number of animals with limited observations and reporting.

3.2.5.3.5.2 Other data limitations and consequences for classification

3.2.5.3.5.2.1 Acute dermal toxicity tests, repeated dose animal studies, skin sensitisation studies and skin absorption studies may all differ from the standard *in vivo* acute dermal irritation/corrosion test (e.g. OECD TG 404) with regard to exposure duration, area dose, the use of dissolved substances, level of occlusion, patch type, scoring and follow-up of the skin lesions and the test species.

3.2.5.3.5.2.2 Destruction of the skin in any acute dermal toxicity test (e.g. OECD TG 402) should be considered for classification as corrosive (category 1 or sub-category 1A, 1B or 1C where possible and required). Skin irritation in an acute dermal study in rabbits fulfilling the criteria in Table 3.2.2, should be considered for classification as irritant if the exposure conditions are such that corrosive effects can be excluded. Skin irritation in an acute dermal study in other species should be considered as not conclusive, as these species may be less or more sensitive than rabbits. Such data should be taken into account in a weight-of-evidence assessment. The absence of skin irritation should also be considered as not conclusive and taken into account in a weight-of-evidence assessment.

3.2.5.3.5.2.3 Repeated dose dermal studies (e.g. OECD TG 410 and 412) can be used to classify as corrosive when destruction of the skin is observed after the initial exposures. However, normally such exposures are avoided and ~~such~~ corrosive effects may only be observed in the range-finding studies. Moreover, sub-categorisation for corrosion will rarely be possible due to a longer time period between start of exposure and first observation. The observation of skin irritation or the absence of skin irritating effects should be considered as not conclusive. Skin effects only observed after multiple exposures may indicate skin sensitisation rather than skin irritation.

3.2.5.3.5.2.4 In skin sensitisation studies in guinea pigs (e.g. OECD TG 406), severely irritating and corrosive exposure must be avoided. Therefore, such effects are normally only observed in range-finding studies. The range-finding results, with the exception of intradermal exposure in the maximisation test, can be used to classify as corrosive when destruction of the skin is observed.

The presence or absence of skin irritation in a skin sensitisation study should be considered as not conclusive by itself as the species tested may be more or less sensitive than rabbits, but signs of irritation should be taken into account in a weight of evidence assessment.

3.2.5.3.5.2.5 Irritation data from the Local Lymph Node Assay (e.g. OECD TG 429, 442A and 442B) should normally not be used for classification as the test substance is applied to the dorsum of the ear by open topical application, and in some cases specific vehicles for enhancement of skin penetration are used. Further, due to the proportional increase of skin thickness associated with increased body weight, the ~~mouse~~ skin thickness of mice deviates ~~the most~~ significantly from that of rabbits and humans.

3.2.5.3.5.2.6 In skin absorption studies (e.g. OECD TG 427), corrosive exposure conditions are generally avoided as this affects the absorption. Therefore, information on skin effects from these studies does not allow classification directly but may be considered within a weight of evidence approach. However, information on the dermal absorption may be taken into account in a weight-of-evidence assessment as a high dermal absorption in combination with additional evidence for high cytotoxicity may indicate irritation or corrosivity.