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

**Sub-Committee of Experts on the Globally Harmonized
System of Classification and Labelling of Chemicals**   **27 June 2018**

Thirty-fifth session
Geneva, 4-6 July 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**

Use of non-animal testing methods for classification of health hazards: Status report and proposed continuance of work

Transmitted by the experts from the United Kingdom and the Netherlands on behalf of the informal working group

Introduction

1.   This informal paper provides an update on the work performed by the Working group on "Use of non-animal testing methods for classification of health hazards" since the thirty-fourth session of the Sub-Committee.

Background

2.   The Sub-Committee agreed to keep the work on the use of non-animal testing methods for classification of health hazards on its programme of work for the 2017-2018 biennium (see ST/SG/AC.10/C.4/64). Information on the mandate/terms of reference of the correspondence group is in INF.27/Rev.2 (thirty-first session) and the report of the Sub-Committee on its thirty-first session (ST/SG/AC.10/C.4/62 paragraph 26).

3.   The Working Group presently has approximately 50 members, reflecting the importance of, and interest in, its work. Its membership includes experts with specialised knowledge of test methods and their application to classification and experts on National legislation and their consequences. Discussions are often lively and detailed, but overall are propelled by a strong desire to make progress on the Working Group's mandate and ensure that non-animal test methods are consistently incorporated in the GHS in a way that reflects their growing importance and scientific relevance, whilst recognising their limitations.

Status report

4.   Since the last meeting of the Sub-committee in December 2017, the Working Group convened 4 times by webex (on 26 February, 9 March, 8 May and 15 June 2018). After each meeting the Netherlands and the United Kingdom, joint leads, revised the draft text of Chapter 3.2 and prepared papers on specific topics to take forward the discussions taking into account written comments and information on specific topics provided by the participants.

5.   The working draft considered in part by the Working Group in June 2018 is in Annex I. Comments received from members during the June 2018 meeting have not yet been incorporated into this draft. New text relative to GHS rev 7 is shown in red; for clarity deleted text is not shown. To some degree, this is still work in progress, and the wording of some sections has not yet been finally discussed by the Working Group. This working draft is presented so the Sub-Committee can see what has been achieved so far, and steer the Working Group as it considers appropriate, in particular with a view to discussing whether and, if so, how the revised section 3.2 should be processed further to achieve adoption by the Sub-Committee in the present biennium.

6.   To provide the Sub-committee with an indication of the nature of the work that has been undertaken, the issues that have been identified, and the progress that has been made, paragraphs 7 to 18 below summarise the main issues that have arisen so far in the Working Groups consideration of Chapter 3.2.

In vitro/ex vivo and non-test methods

7.   New sub-sections have been added setting how to classify for skin corrosion and skin irritation based on in vitro/ex vivo data (3.2.2.3) and on non-test methods such as computer models predicting structure-activity relationships, computer expert systems and read-across using analogue and category approaches (3.2.2.6).

8.   A negative result in an internationally accepted and validated in vitro/ex vivo test for skin irritation, can now be used to conclude as not classified for skin irritation, but only where category 3 is not adopted by the competent authority. Acceptance of a negative result of non-test methods for conclusion as not classified is also included, but not yet discussed.

9.   For classification based on in vitro/ex vivo data it was considered important by some members to set out in the GHS the classification criteria where data is available from the relevant validated test methods, in the same way as is already done for data based on, for example, the standard animal test method using OECD TG 404. Other members felt that specific classification criteria should not be included due to possible additions to the available methods and changes of their criteria. A compromise was reached to provide criteria for in vitro/ex vivo data as background guidance in Table 3.2.6 (for skin corrosion) and in Table 3.2.7 (for skin irritation). These Tables provide the core content of a new sub-section in the background guidance (3.2.5.3.2, Classification based on in vitro/ex vivo data).

10.   The Working Group noted that for non-test methods the text to be included in the GHS may be similar for all health hazard classes, and in future may be better located as general text in Section 1.3. However, for now a new sub-section on non-test methods is retained in Chapter 3.2.

11.   In preparing Table 3.2.6 a problem was encountered because within OECD TG 431 a number of methods are available and the classification criteria for each are not the same. Presently the different methods are identified at OECD level by trademarked names such as EpiSkin™ and EpiDerm™. However, trademarks are not allowed in UN publications so another way had to be found to link the specific OECD methods to the criteria in Table 3.2.6. With the help of the OECD secretariat, this problem is being addressed by identifying the various methods within TG 431 as methods 1, 2, 3 … etc. The Working Group is very grateful to the OECD secretariat for their assistance in resolving this administrative problem.

Tiered vs Integrated approach

12.   The Working Group had extensive discussions comparing the current GHS tiered plus an overall weight of evidence approach to classification for skin corrosion and skin irritation with the integrated approach now adopted at OECD level. The

overall conclusion was that the two approaches are not contradictory, and the tiered approach can be considered as one possible formulation of the integrated approach. Some members of the Working Group considered the integrated approach (which considers all data from the start in a weight of evidence approach) to be preferable; however, others wanted to retain the tiered approach. The compromise is to retain the tiered approach with some amendments and to apply a (total) weight-of-evidence assessment in certain data combinations, especially in case of inconsistent data and to review the relative positions of some of the current tiers (see paragraph 13 below).

13. The detailed consideration of the tiered and integrated approaches, and the need to review the appropriate positions within the tiered approach of the existing validated ex vivo/in vitro tests and Structure Activity Relationship (SAR) tiers, also led to discussion of what conclusions can be drawn from human data and from 'other existing skin data in animals', as well as exactly what these cover. Specific papers on theses aspects were prepared and discussed, and a first revision of Figure 3.2.1 illustrating the tiered approach, has been developed but is still under discussion. There have also been discussions on what is meant by "sufficient" data quality which have not yet been finalised.

14. As well as an early revision of Figure 3.2.1, the outcomes of these discussions are two new sub-sections in the background guidance to Chapter 3.2:

(a)   Guidance on the use of human data for classification as skin corrosion or skin irritation; and

(b)   Guidance on the use of other existing skin data in animals for classification as skin corrosion or skin irritation.

These sub-sections also provide guidance on the limitations of these methods for classification.

15. The latest working draft of Chapter 3.2 pulls together in one sub-section existing text on classification using human data and provides a sub-section on classification based on each type of data in Figure 3.2.1, together with a sub-section on the tiered approach. The order of these sub-sections is aligned as much as possible with the order in the tiered approach. The relevant sub-sections under 'Classification criteria for substances' are:

(a)   3.2.2.1 – Classification based on human data, with a cross reference to the guidance in 3.2.5.3.3

(b)   3.2.2.2 – Classification based on standard animal test data

(c)   3.2.2.3 – Classification based on in vitro/ex vivo data

(d)   3.2.2.4 – Classification based on other, existing skin data in animals, with a cross reference to the guidance in 3.2.5.3.4

(e)   3.2.2.5 - Classification based on chemical properties

(f)   3.2.2.6 – Classification based on non-test methods

Test method neutrality

16. For health and environmental hazards, paragraph 1.3.2.4.3 of the GHS sets out the principle that "tests that determine hazardous properties, which are conducted according to internationally recognized scientific principles, can be used for purposes of a hazard determination […]. The GHS criteria […] are test method neutral, allowing different approaches as long as they are scientifically sound and validated according to international procedures and criteria already referred to in existing systems for the hazard of concern and produce mutually acceptable data." Nevertheless, to be helpful to users of the GHS specific examples of such methods are often given. The Working Group will maintain this principle in its work, whilst

continuing to provide practical information to GHS users. In this, as in all other aspects, the latest working draft in Annex 1 remains a work in progress, and discussions in the Working Group continue.

Other changes

17.   The strong interconnection between section 3.2.2 of the classification criteria for substances and the sub-section on the standard animal test has been removed to make clearer that different types of information can be used for classification for this hazard class.

18.   The decision logics 3.2.1 in paragraph 3.2.5.1 and 3.2.2 in paragraph 3.2.5.2 need to be updated based on any final changes to the order of the tiers and the paragraph numbers.

4 July 2018 meeting

19.   The Working Group on non-animal test methods will continue its work in its meeting on 4 July (for the agenda see Annex II) followed, if needed, by teleconferences. There is still the tentative hope that it will be possible to finalise the revision of section 3.2 in time for adoption by the Sub-Committee in this biennium.

20.   Further, activities planned for the next biennium include updates to the chapters on the hazard classes eye irritation/severe eye damage and skin sensitisation in line with the Working Group's mandate.

Annex I

Changes considered within the working group on non-animal methods for chapter 3.2 of GHS on the criteria for classification of substances for skin corrosion / irritation.

Changes considered are indicated in red (based on the clean version as partly discussed on 15 June 2018). Text in black is the current text in the 7th revision of GHS. Text that was removed, or moved was not indicated to increase readability. Text in blue was used for a statement whether or not a change is applied or needed for certain paragraphs.

It is noted that discussions are ongoing and there remains known significant differences of opinion in the Annex below. This includes but is not limited to issues including when to apply a weight-of-evidence evaluation, Figure 3.2.1 on the tiered approach, guidance on the use of human data and guidance on the use of other existing skin data in animals and issues indicated in the text below using comment fields. Additional detailed discussions regarding most texts are warranted, therefore text indicated in red should not be considered as final draft at this point and likewise it is possible that text in black may also change.

NL/UK June 2018

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 assessed. Wherever possible classification should be based on data generated using internationally validated and accepted methods, such as OECD Test Guidelines or equivalent methods. Sections 3.2.2.1 to 3.2.2.6 below provide classification criteria for the different types of information that may be available.

3.2.1.3 A *tiered approach* (see section 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 satisfy the criteria. However, where the available information gives inconsistent and/or conflicting results within or between tiers, or where data individually are insufficient to conclude the classification a total weight of evidence approach is used (see section 1.3.2.4.9).

3.2.1.4 Further 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 total weight of evidence approach, is provided in OECD Guidance Document 203, An Integrated Approach on Testing and Assessment (IATA) for Skin Corrosion and Irritation.

3.2.2 **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***

3.2.2.1.1 Existing reliable and good quality human data on skin corrosion/irritation should be given high weight where relevant for classification (see section 3.2.5.3.3). Information from human exposure 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, paragraph 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***

The current version of OECD TG 404 using up to 3 animals is the standard animal test (See Tables 3.2.1 and 3.2.2 for classification as skin corrosive or irritant respectively). Other existing animal studies using more than 3 animals may have been conducted under earlier versions of OECD TG 404, and are also considered standard animal test when interpreted in accordance with Section 3.2.5.3.1.

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. 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. A separate irritant criterion accommodates cases when there is a significant irritant response but less than the mean score criterion for a positive test.

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 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 from gradings in at least 2 of 3 tested animals from grades at 24, 48 and 72 hours 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 3.2.5.3.1*

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* data may require data from one or more tests. For authorities implementing category 3 it is important to note that the currently available *in vitro/ex vivo* test methods do not allow identification of substances classified as category 3.

3.2.2.3.2   *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 published literature should also be taken into consideration.

3.2.2.3.4.   *Skin corrosion*

3.2.2.3.4.1   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. Where tests have been undertaken in accordance with OECD Test Guidelines (TGs) 430, 431, or 435 a substance is classified as skin corrosion in category 1 (and, where possible into sub-categories 1A, 1B or 1C) based on the criteria in paragraph 3.2.5.3.2.1 and in Table 3.2.6.

3.2.2.3.4.2   Some *in vitro/ex vivo* methods do not allow differentiation between sub-categories 1B and 1C (See Table 3.2.6). Where available, additional information has to be taken into account in a total weight of evidence evaluation to differentiate these two sub-categories. Where no or insufficient additional information is available, category 1 is applied.

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

3.2.2.3.5   *Skin irritation*

3.2.2.3.5.1   The classification criteria in internationally validated and accepted *in vitro/ex vivo* test methods should be applied. Where sufficient data (e.g. from OECD TG 430, TG 431 or TG 435) show that a substance is not corrosive to the skin, classification as a skin irritant in category 2 is required where the criteria from OECD TG 439 in paragraph 3.2.5.3.2.1 and Table 3.2.7 are fulfilled.

3.2.2.3.5.2   Where information from an *in vitro/ex vivo* study for skin irritation indicates that the substance can be classified in category 1 or 2, additional information is needed to specify the correct classification of the substance., in either category 1, and in sub categories 1A, 1B or 1C as appropriate, or in category 2.

3.2.2.3.5.3   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 where category 3 is not adopted.

3.2.2.3.5.4   The 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. Where this category is adopted, and the classification criteria for either category 1 or 2 are not fulfilled, additional information is required to differentiate between category 3 and no classification.

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

3.2.2.4.1   Other existing skin data in animals may be used for classification, but there may be limitations on the conclusions that can be made (see Section 3.2.5.3.4). If a substance is highly toxic by 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 are made of skin corrosion/irritation in acute toxicity studies these data may be used for classification, provided that the dilutions used and species tested are relevant. Guidance on the use of other existing skin data in animals including acute dermal toxicity tests, dermal repeated dose toxicity tests

and other tests is provided in Section 3.2.5.3.4. Solid substances (powders) may become corrosive or irritant when moistened or in contact with moist skin or mucous membranes.

3.2.2.5 *Classification based on chemical properties*

3.2.2.5.1   Likewise, pH extremes like ≤ 2 and ≥ 11.5 may indicate skin effects, 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 by data 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 *Classification based on non-test methods*

3.2.2.6.1   Classification, including non-classification can be based on non-test methods, with due consideration of differing reliability and applicability, on a case-by-case basis. Such methods include computer models predicting structure-activity relationships (SAR); quantitative structure-activity relationships (QSARs); computer expert systems; and read across using analogue and category approaches. General information on the use and documentation of non-test methods is provided in 1.3.2.4.8.

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   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 (Q)SAR is not sufficient evidence of 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.

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, existing animal data and then other sources of information. Where information from data within the same tier is inconsistent and/or conflicting, the overall conclusion from that tier is determined by a weight of evidence approach.

3.2.2.7.3   Although information might be gained from the evaluation of single parameters within a tier, consideration should be given to the totality of existing information and making an overall weight of evidence determination.

3.2.2.7.4   Where information from several tiers is inconsistent and/or conflicting, 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 (more potent category or sub-category) than information from a higher tier, the overall classification is determined by a total weight of evidence approach.

Figure 3.2.1: Tiered evaluation for skin corrosion and irritation

Step	Parameter		Finding		Conclusion
1a:	Existing human [a] or standard animal skin corrosion/irritation data ↓ Not corrosive/Insufficient/ Inconclusive/No data ↓	→	Skin corrosive	→	Classify as **skin corrosive** [b]
1b:	Existing human [a] or standard animal skin corrosion/irritation data ↓ Not irritant/Insufficient/ Inconclusive/No data ↓	→	Skin irritant	→	Classify as **skin irritant** [b]
1c:	Existing human [a] or standard animal skin corrosion/irritation data ↓ Insufficient/Inconclusive/No data ↓	→	Not a skin corrosive or skin irritant	→	**Not classified**
2a:	Existing *ex vivo/in vitro* skin corrosion data ↓ Insufficient/Inconclusive/ No data/Negative response ↓	→	Skin corrosive	→	Classify as **skin corrosive** [b]
2b:	Existing *ex vivo/in vitro* data ↓ Insufficient/Inconclusive/ No data/Negative response for skin irritation ↓	→	Negative: Skin corrosive Positive: Skin irritant 2	→	Classify as **skin irritant 2**
2c: (only for authorities not applying Category 3)	Existing *ex vivo/in vitro* skin irritation data [j] ↓ Insufficient/Inconclusive/No data /Negative response ↓	→	Negative: Skin irritant 2 [j]	→	**Not classified**
3:	Other, existing skin data in animals [c] ↓ Negative/Insufficient/Inconclusive/ No data ↓	→	Yes; other existing data showing that substance may cause skin corrosion or skin irritation	→	May be deemed to be a **skin corrosive** [b] or a **skin irritant** [b g]
4:	pH-based assessment (with consideration of acid/alkaline reserve of the chemical) [e] ↓	→	pH ≤ 2 or ≥ 11.5 [i] with high acid/alkaline reserve or no data for acid/alkaline reserve	→	Classify as **skin corrosive** [g]

	Not pH extreme, no pH data or extreme pH with data showing low/no acid/alkaline reserve [h] ↓				
5:	Non-test methods → ↓		Skin corrosive →	→	Deemed to be **skin corrosive** [b]
			Skin irritant ↘	→	Deemed to be **skin irritant** [b]
			Negative ↘	→	Deemed to be not classified
	Insufficient/Inconclusive/No data ↓				
6:	Consideration of the total weight of evidence [f] → ↓		Skin corrosive →	→	Deemed to be **skin corrosive** [b]
			Skin irritant ↘	→	Deemed to be **skin irritant** [b]
			Negative ↘	→	Deemed to be not classified
	Insufficient/Inconclusive/No data ↓				
7:	**Not classified**				

(a)   *See Section 3.2.5.3.3 for guidance on the use of human data.*

(b)   *Classify in the appropriate category/sub-category as applicable, depending on the availability and quality of data;*

(c)   *See the guidance provided in background guidance Section 3.2.5.3.4.*

(d)

(e)   *Measurement of pH alone may be adequate, but assessment of acid or alkali reserve (buffering capacity) would be preferable.*

(f)   *All information that is available should be considered and an overall determination made on the total weight of evidence. This is especially true when there is conflict in information available on some parameters. Expert judgment should be exercised prior to making such a determination.*

(g)   *In case there is a conflict in available data, e.g. negative/irritation human data but positive/corrosive in vitro/ex vivo data, a weight of evidence assessment should be performed, see footnote f.*

(h)   *Non corrosivity needs to be confirmed by other data and preferably by data from an appropriate validated in vitro/ex vivo test.*

(i)   *For the case of mixtures with no human or animal data on skin corrosion/irritation but with extreme pH see Figure 3.3 in 3.2.3.2.1.1.*

(j)   *The method should be validated such that a determination for non-classification can be reached (e.g. OECD TG 439).*

3.2.3   Classification criteria for mixtures

No changes proposed.

3.2.4   Hazard communication

No changes proposed

3.2.5   Decision logics and guidance

Changes needed to 3.2.5.1 and 3.2.5.2 when finalised because reference is made to individual paragraphs and the order of the information types has changes.

3.2.5.3 Background guidance

3.2.5.3.1   Classification based on tests with more than 3 animals

No change to paras 3.2.5.3.1 to 3.2.5.3.5, though renumbered as 3.2.5.3.1.1 to 3.2.5.3.1.5.

3.2.5.3.2   Classification criteria based on in-vitro/ex vivo data

3.2.5.3.2.1   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 into sub-categories 1A, 1B or 1C) as skin corrosion and in category 2 as skin irritant 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

		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 value TER value in kΩ and sometimes 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) the mean TER value ≤ 5 kΩ and the skin discs are obviously damaged (e.g. perforated), or (b) the 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 min exposure; or < 35% after 60 min exposure; or < 35% after 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							> 3 to 60 min.	> 3 to 30 min

1C		≥ 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	> 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 UN 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 UN 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.

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

3.2.5.3.3.1   Relevant, reliable and good quality human data is generally given high weight for classification. However, human data also have some limitations depending on the type of data. For two specific types of data, human patch test (HPT) and case studies, the limitations on the use of such data for classification are described below.

3.2.5.3.3.2   HPTs are performed to discriminate between irritant and non-irritant substances. Application of a corrosive substance to human skin is generally avoided. Therefore, the HPT alone does not discriminate between irritant and corrosive substances. However, the combination of an HPT and sufficient other information on skin corrosion can be used for classification within a weight of evidence assessment.

3.2.5.3.3.3   Specific criteria for HPT results leading to classification as category 2 (skin irritation), category 3 (mild irritation) or not classified have not been established. Therefore, the results of an HPT are generally used within a weight of evidence assessment. However, clearly negative results in an HPT with sufficient volunteers after exposure to the undiluted substance for 4 hours can justify no classification.

3.2.5.3.3.4   Human case study reports may be used for classification as corrosive if irreversible damage to the skin was observed. However, these case studies are typically short exposure duration and lack longer-term follow-up information. This, together with the absence of accepted classification criteria, means that human case studies resulting only in irritation or no effects are generally not conclusive, but can be used in a weight of evidence assessment.

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

3.2.5.3.4.1   General approach

3.2.5.3.4.1.1   All existing other animal data should be carefully reviewed to determine if the data are conclusive for classification, and if so, it should be used. In evaluating all 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 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.

3.2.5.3.4.2   Other data limitations and consequences for classification

3.2.5.3.4.2.1   Acute dermal toxicity tests, repeated dose animal studies, skin sensitisation studies and skin absorption studies may all differ from the *in vivo* acute dermal irritation/corrosion test (e.g. OECD TG 404) with regard to exposure duration, dose per cm^2, the use of dissolved substances, scoring and follow-up of the skin lesions and the test species.

3.2.5.3.4.2.2   Other specific data considerations and consequences for classification are set out in paragraphs 3.2.5.3.4.2.3 to 3.2.5.3.4.2.7 below.

3.2.5.3.4.2.3   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. In general skin thickness decreases with body weight. 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.

3.2.5.3.4.2.4   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 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.

3.2.5.3.4.2.5   In skin sensitisation studies (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 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.

3.2.5.3.4.2.6   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. 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.

3.2.5.3.4.2.7   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 allometric scaling of skin thickness, the mouse deviates the most from rabbits and humans.

Annex II Draft agenda

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

**Sub-Committee of Experts on the Globally Harmonized
System of Classification and Labelling of Chemicals**   **4 July 2018, 10:00 – 12:00 CET**

Informal working group on use of non-animal testing methods for the classification of health hazards

Room S4

Suggested draft agenda: Working Group on Non-animal test methods

1)   Welcome and introduction

2)   Summary notes of the meeting on 15 June (document 1)

3)   Discussion of possible revision of chapter 3.2 (document 2, and 3)
            The discussion will be based on the line numbers of the clean version (document 3)

        3a   3.2.2.3 In vitro methods including the tables in 3.2.5.3.2
        3b   Figure 3.2.1
        3c   3.2.1.2-4 and 3.2.2.7 (tiered approach and WoE)
        3d   3.2.2.1 (Human data) and 3.2.5.3.3 (guidance human data)
        3e   3.2.2.5 (other animal data) and 3.2.5.3.4 (guidance other animal data)
        3f   Remaining paragraphs

4)   Update for eye irritation/corrosion and skin sensitisation

5)   Next steps

        Documents

        For discussion:
        1   Non-animal working group - draft note of webinar - 15 June 2018 document 1
        2   GHS Chapter 3-2 non-animal testing - for discussion July 2018 - TC document 2
        3   GHS Chapter 3-2 non-animal testing - for discussion July 2018 - clean document 3

        For information:
        4   Minutes of meeting 8 May 2018 V3 document 4

———————