Secretariat

Distr.: General
25 September 2018

Original: English

**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**

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 GHS 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 working document summarises the work of the informal working group on "Use of non-animal testing methods for classification of health hazards" on Chapter 3.2, Skin Corrosion/Irritation, and, with informal document INF.6, presents for the agreement of the Sub-Committee a revision of this Chapter to better reflect the increased capability, availability and utility for classification of *in vitro/ex vivo* test methods and of non-test methods such as computer models and read-across.

Background

2.   The terms of reference the Sub-Committee gave to the informal working group (informal document INF.27/Rev.2, (thirty-first session)) set out four main activities:

---

[*] In accordance with the programme of work of the Sub-Committee for 2017–2018 approved by the Committee at its eighth session (see ST/SG/AC.10/C.3/100, paragraph 98 and ST/SG/AC.10/44, paragraph 14).

(a)   To identify and evaluate the available *in vitro* and *in chemico* test methods, validated at the international level, and the existing guidance on *in silico* methods, including grouping approaches, quantitative structure activity relationship (QSARs) and read-across, that could be useful for GHS hazard classification for health hazard and environmental hazard classes, using a step-wise approach, starting with a hazard class to be determined by Group;

(b)   To assess for each hazard class whether an integrated or tiered evaluation approach should be developed, taking into account all relevant scientific information and combination of methods for hazard classification, and where substances and mixtures may be classified using non-animal methods, whether new or amended GHS classification criteria are needed;

(c)   To prepare draft amendments and additions to the GHS to facilitate hazard classification using non-animal methods where appropriate, taking into account relevant limitations and uncertainties. The amendments and additions should include as appropriate classification criteria, notes, decision logics, tiered evaluation and guidance, and should take into account the needs of all sectors;

(d)   To report progress to the GHS Sub-Committee as appropriate. The latest status update is informal document INF.25 in the thirty-fifth session.

3.   The informal working group has around 50 members, reflecting the importance of, and interest in, its work.

4.   The informal working group identified at an early stage that skin corrosion/irritation would be a good starting point for this work, as internationally accepted and validated *in vitro/ex vivo* test methods are well established.

5.   Discussions in the informal working group are lively and detailed, and the group has worked intensively to address its mandate and to resolve issues and challenges. For example, in the period from January to early September 2018 six webinars have been held (26 February, 9 March, 8 May, 15 June, 24 July and 7 September), with a focus on completing the group's work on Chapter 3.2 in this biennium. After each meeting the Netherlands and the United Kingdom, as 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 members of the group.

6.   Paragraphs 7 to 22 below 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 solutions that have been adopted. The changes made to Chapter 3.2 are summarised in the Annex to this paper. For clarity the full text of the Chapter 3.2 as revised is set out in informal document INF.6 with indication of where the text has changed relative to the seventh revised edition of GHS.

Issues and Outcomes

Tiered vs integrated approach

7.   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 the Organisation for Economic Co-operation and Development (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.

8.   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 weight of evidence assessment where the available information gives inconsistent and/or conflicting results within a tier, and an overall weight of evidence assessment where there are inconsistent and/or conflicting results between tiers. Within the overall weight of evidence evaluation classification is made on higher tier data where this indicate greater concern, and on the basis of a more detailed overall weight of assessment where lower tier data indicate a greater concern than higher tier data. In addition, the relative order of some of the current tiers has been reviewed (see paragraphs 16 and 17 below).

9.   The detailed consideration of the tiered and integrated approaches, and the need to review the appropriate order within the tiered approach of the existing validated *in vitro/ex vivo* test and non-test 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 and their limitations for classification in practice.

10.   Specific papers on these aspects were prepared and discussed. The outcomes of these discussions are three new sub-sections in the background guidance to Chapter 3.2:

      (a)   A pointer to helpful information on the strengths and weaknesses of the different test and non-test methods, and to useful guidance on how to apply a weight of evidence approach (3.2.5.3.1);

      (b)   Guidance on the use of human data for classification for skin corrosion or skin irritation (3.2.5.3.2); and

      (c)   Guidance on the use of other existing skin data in animals for classification as skin corrosion or skin irritation (3.2.5.3.5).

In vitro/ex vivo and non-test methods

11.   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). The broader term 'non-test methods' now replaces 'validated Structure Activity Relationship (SAR) methods' in the seventh revised edition of GHS.

12.   For classification based on *in vitro/ex vivo* data it was considered important by some group 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 provided now for data based on, for example, the standard animal test method using OECD Test Guidelines (TG) 404. Other members felt that specific classification criteria should not be included due to possible additions to the available methods over time and changes to the 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 Table 3.2.7 (for skin irritation). These Tables provide the core content of a new sub-section in the background guidance (3.2.5.3.4, Classification based on *in vitro/ex vivo* data).

13.   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 United Nations 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 TG will be updated accordingly and published before the end of 2018. The informal working group is very grateful to the OECD secretariat for their assistance in resolving this administrative problem.

14.   Clarity is now provided that where category 3 is not adopted by the competent authority a negative result in an internationally accepted and validated *in vitro/ex vivo* test for skin irritation can be used to conclude as not classified for skin irritation. Similarly, there is now explicit provision that, with due consideration of reliability and applicability case-by-case, a negative result in non-test methods can lead to the conclusion not classified.

15.   The informal working group noted that as further work is done on other health hazard classes it may become apparent that the same or similar text on classification using non-test methods applies to all GHS health hazard classes. If so, such text may be better located in Section 1.3. However, for now a new sub-section on non-test methods (3.2.2.6) is retained.

Presentation of classification criteria and order of tiers

16.   As well as adding new sub-sections on 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 (3.2.2.6), the opportunity has been taken to pull together in one sub-section (3.2.2.1) existing text on classification using human data, with a cross reference to the related background guidance.

17.   As noted in paragraph 8 above, the introduction of more detailed classification criteria using *in vitro/ex vivo* and non-test methods prompted the informal working group to review the order of the tiers. An outcome is that in Figure 3.2.1 the tiered evaluation now follows the order below, and the text in section 3.2.2, Classification criteria for substances, also reflects this order. In particular, *in vitro/ex vivo* data is now tier 2, whereas in the seventh revised edition of GHS it is tier 3.

Tier 1 - existing human or standard animal skin corrosion/irritation data

Tier 2 – existing *in vitro/ex vivo* skin corrosion/irritation data

Tier 3 – other existing skin data in animals

Tier 4 – pH-based assessment

Tier 5 – non-test methods

Tier 6 – consideration of the overall weight of evidence

18.   The informal working group developed an amended version of Figure 3.2.1 reflecting the revised order of the tiers and other detailed changes. However, the working group also experimented with a new flow diagram to illustrate more clearly the process associated with the tiered approach. The new flow diagram is more conceptual and has less detail than Figure 3.2.1 in GHS, seventh revised edition. However, it brings out clearly that where a classification decision can be made within a tier, but other data from a lower tier indicate a stricter classification, an overall weight of evidence approach is needed. In discussion the informal working group considered the flow diagram to be a significant improvement.

19.   The new flow diagram has been included in the proposed revised text of Chapter 3.2 (see Annex and informal document INF.6). The informal working group is satisfied that all significant details in the original Figure 3.2.1 have been incorporated in the revised Chapter 3.2, either in the text or in the new flow diagram and associated footnotes. However, the working group is also aware that if the Sub-Committee agrees the revised Chapter 3.2 there will be an interim period where Figure 3.2.1 takes a different form than the corresponding

Figure 3.3.1 for serious eye damage/eye irritation. As noted in paragraph 23 below, if the Sub-Committee extends the mandate of the informal working group to the next biennium, amendment of Chapter 3.3 to better reflect non-animal test methods will be a high priority.

20.   Corresponding changes have been made in the decision logics 3.2.1 and 3.2.2.

21.   In checking the application of the tiered approach when applying the classification criteria using pH measurements (tier 4), the informal working group noted that in certain circumstances there was a lack of consistency in GHS, seventh revised edition, between the text in 3.2.2.2.5 (and for mixtures 3.2.3.1.2), Figure 3.2.1, and Decision logic 3.2.1. Different classifications could result where there is extreme pH, low buffer capacity and no other data to confirm the substance or mixture is not corrosive. In discussion, the working group considered that in these circumstances the outcome should be classification as inconclusive, and the corresponding adjustments have been made in the proposed revision of Chapter 3.2 to ensure consistency in these circumstances.

Test method neutrality

22.   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. In revising Chapter 3.2 the informal working group has tried hard to maintain this principle, whilst continuing to provide practical information to GHS users.

Action and next steps

23.   The Sub-Committee is invited to agree the revised Chapter 3.2 as set out in the Annex to this paper and in informal document INF.6.

24.   Looking ahead, the informal working group recognises the longer-term nature of this work to 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. With the agreement of the Sub-Committee, 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 informal working group's mandate.

Annex

Proposed amendments of chapter 3.2

Amend chapter 3.2 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 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."

Insert two new paragraphs to read as follows:

"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 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   Replace the text of 3.2.2 with the following:

"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   Amend the heading of 3.2.2.1 to read as follows: "Classification based on human data"

Replace the text of 3.2.2.1 and its paragraphs 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). 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, 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".

3.2.2.1   Renumber the heading "Classification based on standard animal test data" as 3.2.2.2.

Insert new text in 3.2.2.2, below the heading "Classification based on standard animal test data", as follows:

 "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.

Renumber the heading "Skin corrosion" with current number 3.2.2.1.1 as 3.2.2.2.1.

Renumber paragraph 3.2.2.1.1.2 to 3.2.2.2.1.1 and paragraph 3.2.2.1.1.4 to 3.2.2.2.1.2 and delete paragraphs 3.2.2.1.1.2 and 3.2.2.1.1.3.

Delete footnote a from Table 3.2.1.

Renumber section 3.2.2.1.2 and paragraphs 3.2.2.1.2.1 to 3.2.2.1.2.5 as section 3.2.2.2.2 and paragraphs3.2.2.2.2.1 to 3.2.2.2.5 respectively.

Delete footnote c (Evaluation of a 4, 5 or 6-animal study should follow the criteria given in 3.2.5.3) from Table 3.2.2.

3.2.2.3   Insert the following new sub-heading: "3.2.2.3 Classification based on *in vitro/ex vivo* data"

3.2.2.3.1-3.2.2.3.4.3 Insert the following new paragraphs:

"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 *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 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 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 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   Insert the following new sub-heading: "3.2.2.4 Classification based on other, existing skin data in animals"

Insert the following new paragraph:

"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   Insert the following new sub-heading: "3.2.2.5   Classification   based   on chemical properties"

Insert the following new paragraph:

"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. 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   Insert the following new sub-heading: "3.2.2.6   Classification   based on non-test methods"

Insert the following new paragraphs:

"3.2.2.6.1 Classification, including 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 require 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   Renumber the heading "Classification in a tiered approach" with current number 3.2.2.2 as 3.2.2.7.

Renumber paragraph 3.2.2.2.1 and amend the text as follows:

"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".
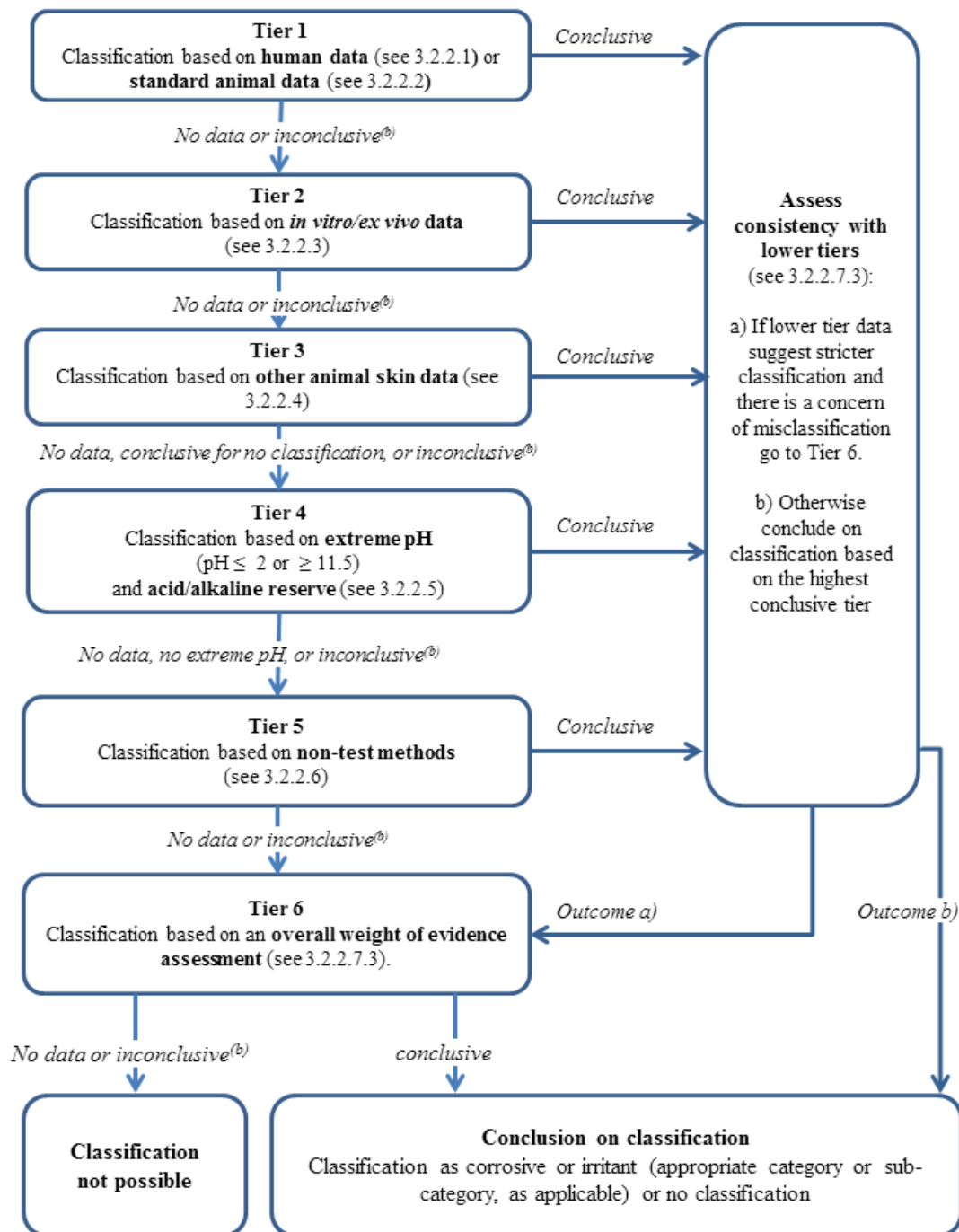
Delete paragraph 3.2.2.2.2 to 3.2.2.2.7 and insert the following new paragraphs with indicated numbering:

"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."

Amend the heading of Figure 3.2.1 to read as follows: "Figure 3.2.1: Application of the tiered approach for skin corrosion and irritation [(a)]"

Insert the new Figure 3.2.1 and notes, to read as follows:

Tier 1
Classification based on **human data** (see 3.2.2.1) or **standard animal data** (see 3.2.2.2)

Conclusive

No data or inconclusive[(b)]

Tier 2
Classification based on *in vitro/ex vivo* **data** (see 3.2.2.3)

Conclusive

No data or inconclusive[(b)]

Tier 3
Classification based on **other animal skin data** (see 3.2.2.4)

Conclusive

No data, conclusive for no classification, or inconclusive[(b)]

Tier 4
Classification based on **extreme pH** (pH ≤ 2 or ≥ 11.5) and **acid/alkaline reserve** (see 3.2.2.5)

Conclusive

No data, no extreme pH, or inconclusive[(b)]

Tier 5
Classification based on **non-test methods** (see 3.2.2.6)

Conclusive

No data or inconclusive[(b)]

Tier 6
Classification based on an **overall weight of evidence assessment** (see 3.2.2.7.3).

Outcome a)

Assess consistency with lower tiers (see 3.2.2.7.3):

a) If lower tier data suggest stricter classification and there is a concern of misclassification go to Tier 6.

b) Otherwise conclude on classification based on the highest conclusive tier

Outcome b)

No data or inconclusive[(b)]   *conclusive*

Classification not possible

Conclusion on classification
Classification as corrosive or irritant (appropriate category or sub-category, as applicable) or no classification

Replace footnotes (a) and (b) with the following:

"(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."

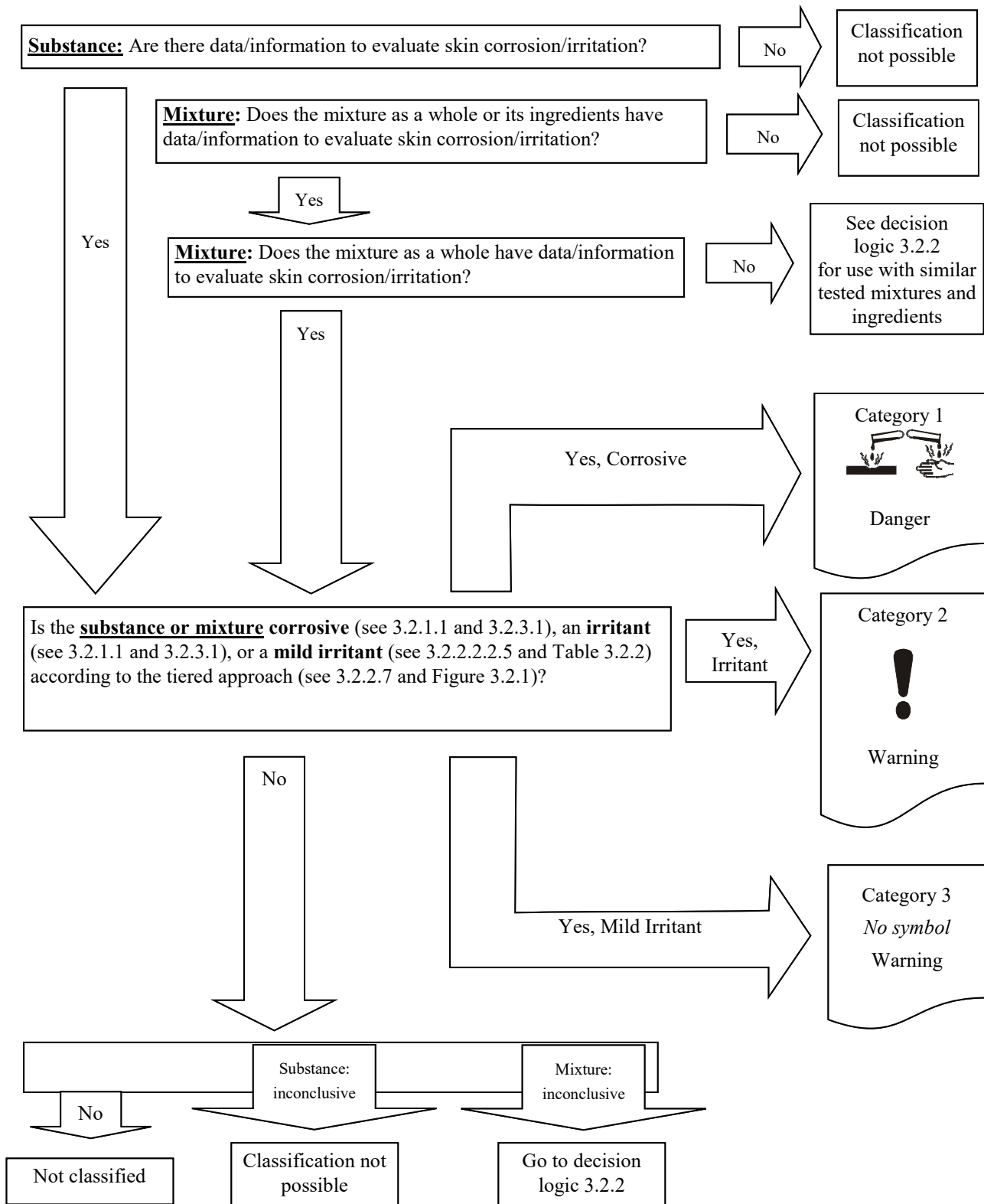Remove footnotes (c), (d), (e) and (f).

3.2.3.1 Replace paragraph 3.2.3.1.1 with the following text:

"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 3.2.3.2 or 3.2.3.3 as appropriate."

Replace paragraph 3.2.3.1.2 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 they 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 are reasons 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."

3.2.5.1　　　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 ↓

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, Corrosive → Category 1 / Danger

Is the **substance or mixture corrosive** (see 3.2.1.1 and 3.2.3.1), an **irritant** (see 3.2.1.1 and 3.2.3.1), or a **mild irritant** (see 3.2.2.2.2.5 and Table 3.2.2) according to the tiered approach (see 3.2.2.7 and Figure 3.2.1)?

Yes, Irritant → Category 2 / Warning

Yes, Mild Irritant → Category 3 / *No symbol* / Warning

No

Substance: inconclusive

Mixture: inconclusive

No → Not classified

Classification not possible

Go to decision logic 3.2.2

3.2.5.3.1 Insert a new sub-heading "Relevant guidance documents" to be numbered 3.2.5.3.1 and insert the following new text:

"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 Insert a new sub-heading "Guidance on the use of human data for classification as skin corrosion or skin irritation" to be numbered 3.2.5.3.2.

Insert the following new paragraphs:

"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 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 sufficient 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, expert judgement may be required to evaluate the sufficiency of the exposure duration and the availability of sufficient long-term follow-up information and to conclude 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 Insert a new sub-heading "Classification based on standard animal tests with more than 3 animals"

Renumber existing paragraphs 3.2.5.3.1 to 3.2.5.3.5 as 3.2.5.3.3.1 to 3.2.5.3.3.5.

3.2.5.3.4 Insert a new sub-heading "Classification criteria based on *in-vitro/ex vivo data*" to be numbered 3.2.5.3.4 and insert the following new text:

"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) for skin corrosion and in category 2 for skin irritation are set out in Tables 3.2.6 and 3.2.7."

Insert the following new heading of a new Table:

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

Insert the following new Table:

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

Insert the following new heading of a new Table:

"Table 3.2.7 Skin irritation criteria for *in vitro* methods"

Insert the following new Table:

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 ($\leq$) 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 $\leq$ 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 Insert a new sub-heading "Guidance on the use of other existing skin data in animals for classification as skin corrosion or skin irritation" to be numbered 3.2.5.3.5 and insert the following new paragraphs:

"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 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 deviates the most from 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".

―――――――――