In this step, an internal assessment of the category is performed. The internal assessment consists of:

a) identification of the relational features that collectively generate the association between the category members. These relational features are proposed on the basis of *existing data*, which may be internal and/or external to the category.

b) use of the relational features to fill data gaps (empty cells in the category matrix) or fill in matrix cells containing data of uncertain quality.

In this context, the term "internal" is borrowed from the QSAR field, in which the internal assessment of a QSAR model refers to an assessment of the model performance by using the same data that were used to develop the model.

Evaluate the category approach to determine whether there is a correlation among category members and each SIDS endpoint by looking for patterns in the matrix. The same category members do not have to be used for each evaluation, i.e., the members selected for environmental fate may be different from those used to evaluate toxicology effects.

- If there are adequate data for a given SIDS endpoint, but no apparent pattern, the proposed category may not be appropriate and so testing may be required for all remaining category members for that SIDS endpoint. However, an alternative category proposal may be developed e.g. the analysis might suggest that the category should be divided. (go back to Step 1).
- If there are adequate data that correlate well, the category may be appropriate and a category test plan proposal should be prepared (Step 6).
- If adequate data do not exist, but the structure-based category is reliable for one or more SIDS endpoints, then a category approach may still be proposed (go to Step 6).

When establishing trends in data, laboratory and experimental variations should be considered. Similar species/strains, endpoints and test protocols should be compared. Deviations from a trend should be clearly identified and possible reasons for the deviations laid out in the category analysis.

The category approach is most robust when a quantitative trend between the category members can be established. A lack of observed effects for a chemical substance in a study of a specific end-point (especially if no dose-relationship can be established because no effects are observed at the highest dose tested) is of limited value to establish the robustness of the category.

- **Step 6: Prepare category test plan.**

Category test plans (Step 6 of Figure) should include a category definition, rationale, and matrix of data availability (see example category test plans in Annex 1.) and be accompanied by SIDS Dossiers for each category member.

The rationale supporting a category definition should be as simple and transparent as possible, and should explain why the existing data and proposed testing data allow interpolation or extrapolation to other members of the category that have no data or proposed testing.

The test plan needs to summarise the adequacy of the existing data, and how the proposed testing will adequately characterise the category.

The matrix of data is an essential part of the test plan and provides a useful tool for consideration and presentation of the available data (see Annex 1). Assuming the SIDS endpoints are rows in the matrix, each row must have data in at least one cell. If toxicity is expected to vary in a regular pattern from one end of the range of category members to the other end (e.g. high toxicity to low toxicity), samples chosen for testing should bracket both ends of toxicity. If the category is large, testing also needs to be performed and/or data should be available for one or more members in the middle of the range of toxicity. Any change in a tendency for a property should be accompanied by data in the adjacent cells in order to define the limits for the resulting subsets of the category or sub-categories. Assuming the columns are the category members, one or more columns may have all empty cells, i.e. no test data available. There are no rules for the number of columns and cells that must be filled nor the number that can be empty. Acceptability of the matrix will depend on the number of members in the category, the SIDS endpoint, and the confidence in the interpolation and extrapolation.

When selecting a sample to test, it should be representative of the substance marketed, including the presence of any manufacturing impurities (see section 2.3.3).

It should be noted that the category test plan is intended to provide information about the properties of the group as a whole rather than the properties of any specific, individual compound. This approach is very different from the approach widely used in the current evaluation of both new and existing chemicals, where the test plan is focussed on obtaining data on an individual compound of commercial interest. A category test plan may thus identify as key substances for testing substances of little or no commercial importance. Whilst in some cases this may even require the synthesis of chemicals specifically for this purpose, the approach may still prove more economical, both in terms of expense and numbers of animals used for testing, than a more conventional testing strategy based on individual commercially available chemicals.

At this point in the process, the sponsor country may consider to submit the test plan to the other OECD member countries for consultation.

- **Step 7: Conduct the necessary testing.**

- **Step 8: Perform an external assessment of the category and fill data gaps**

  In this step, some or all of the relational features are assessed by checking whether the predictions they make for data gaps (or data points of dubious quality) are accurate on the basis of *newly-generated* experimental data, obtained in Step 7.

  In this context, the term "external" is being borrowed from the QSAR field, in which the external assessment of a model refers to an assessment of the model performance by using independent data different from that used to develop the model.

  Add the new data to the SIDS Dossier for the relevant category member and evaluate whether the existing data and the new data support the proposed category.

  - If the results support the category, the testing phase is complete. A SIAR for the category of chemicals should then be prepared including a category analysis. The category analysis will include a summary of the one or more SIDS endpoints in which the category "holds", including the interpolation/extrapolation of test results to

the remaining, untested matrix cells (see below). The SIAR would receive Member country review at the SIAM meeting.

- If the results do not support the category return to Step 5. Further testing may be carried out, members of the category may be changed (e.g. dividing the category as appropriate), or the category proposal may be dropped altogether. The latter implies that testing will then be done to fill all appropriate SIDS endpoints for each HPV category member.

As indicated in section 3.2.2, data gaps are filled by read-across, extrapolation or interpolation. This is specific to each category. No definitive guidance can be provided for the moment. A few examples are provided in Annex 1.

Available options for filling data gaps include:

1. Qualitative: it is concluded that all the members of the (sub)category do or do not possess a particular property e.g. *in-vitro* mutagenicity.
2. Quantitative: it is concluded that all the members of the (sub)category possess a particular property with a similar potency or evolving according to a regular pattern. Data gaps can be filled e.g.
   - by using the value from the closest analogue in the (sub)category;
   - by using a worst case approach i.e. using the value from the most hazardous substance in the (sub)category, or in case of interpolation, the value from the most hazardous of the two closest analogues (see figure 1);
   - by estimating quantitatively the potency of the property from the potency of the two closest analogues or from the regular evolution of this potency over the different (sub)category members.

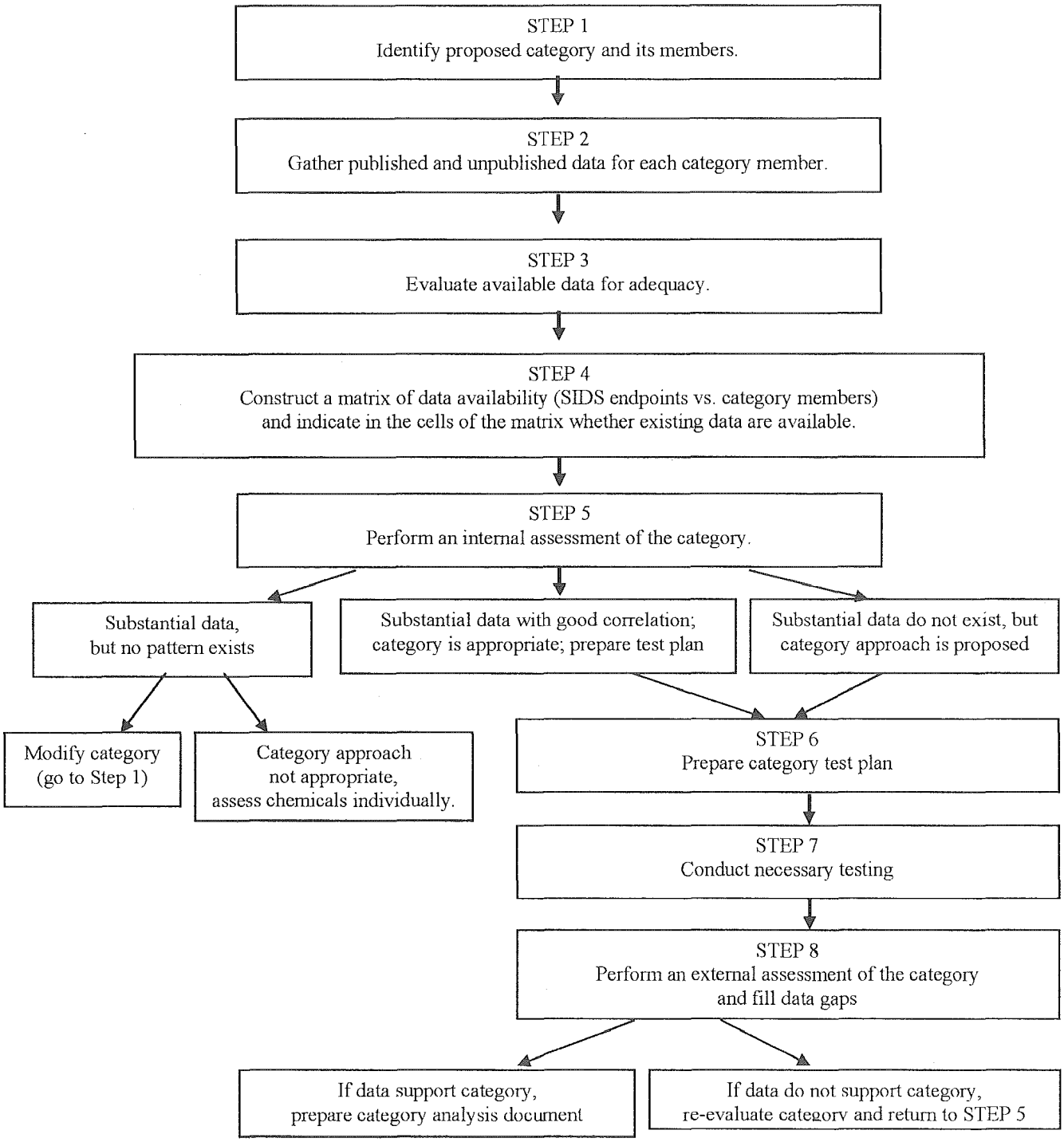
   There is currently only limited experience with quantitative data gap filling for toxicological endpoints. It should be applied with caution and the guidance will be revised as soon as more experience is available.

QSARs could be used to support proposals for filling data gaps by any of the mechanisms described above.

For categories composed of complex substances, approaches like the toxic equivalency factors or toxic units approach could be investigated to fill data gaps.

The mechanism by which the data gaps are filled in a given category should be any case be described transparently in the SIAR.

FIGURE 2: PROPOSED PROCESS FOR DEVELOPING CHEMICAL CATEGORIES



3.2.4 Use of QSARs for the Development of a Category

16. Greater confidence and further demonstration of the category approach may be gained through applying QSAR models on all category members for a given endpoint in case reliable QSAR models are available for the category members and the endpoint. QSARs can contribute at all stages of category development and consideration. Based on experienced assessment of the quality of output taking into account limitations and strengths of a range of models, QSARs may contribute not only for endpoints and compounds within categories where there are no relevant data but also in the interpretation of weight of evidence for mixed datasets and analysis of trends.

17. The output of QSAR modelling is particularly valuable in hypothesis generation and testing for step 1, "identifying the structure-based category and its members". Analysis of physico-chemical and (eco)toxicological data from the members of a category should demonstrate clear relationships between those members; outliers can then be investigated for their 'eligibility' for membership. It also permits hypothesis testing of several possible combinations and permutations for category definition. The more transparent, evolving analytical models provide access to detailed description of relevant data in the training sets. This can facilitate initial consideration of trends to establish the nature and bounds of the category.

18. QSAR modelling can also assist at this stage in defining the appropriate bounds of the proposed category, through consideration of measures of similarity for chemical descriptors in the models. In the more transparent, evolving analytical models, the bounds of this similarity can be specified and the category defined accordingly. In some cases, this can lead to the definition of a more extensive category than was originally envisaged.

19. In addition, QSAR can in some cases be used to assess similarities in metabolic pathways across the group, and this information can be helpful in assessing similarities and differences within the category.

20. Results of QSAR modelling are also relevant to step 2 ("Gathering published and unpublished data for each category member"). In addition to contributing to trends analysis for potential members of the category where no data have been identified, considered output of a battery of models can also add weight of evidence to increase confidence in trends analysis, where the pattern is not clear or consistent based on available data. For example, evaluated QSAR output may contribute where dose spacing or non-comparability of experimental protocols in available studies for different members of the category precludes meaningful analysis of quantitative trends of effect levels. In compiling this information, however, it is important to distinguish where the models contribute additionally to identified experimental data – i.e., that they are not simply duplicating the information, based on replication of its inclusion in their training set. The ease with which this information can be accessed for various models (if at all) varies, depending upon degree of transparency.

21. In relation to step 3 ("Evaluate available data for adequacy"), for QSAR modelling, this requires consideration of aspects related to the training sets and the models, themselves. Relevant aspects include criteria for inclusion of and nature of data in the training sets, the nature of the analysis for consideration of similarity, the criteria for weight of evidence for delineation of a positive/negative response and the nature of validation of the models and aspects thereof, including concordance, sensitivity and specificity for specific endpoints and subsets of chemicals. For characterization of hazard for related endpoints, critically evaluated QSAR output can be combined with weighting of the endpoints themselves (e.g., *in vivo* versus *in vitro* genotoxicity) as a basis for meaningful contribution to hazard characterization, particularly where data are lacking or mixed.

22. For step 4 ("Construct a matrix of data availability"), then, it will be important that results of QSAR modelling be clearly distinguished from those which are based on data. As indicated above, only evaluated results of QSAR modelling which contribute additionally to weight of evidence determinations or quantitative trends analysis should be included. This would include, then, only results for modelling, where evaluated output meaningfully contributes to weight of evidence or trend analysis (this could be for substances where there are no data or where datasets for category definition are uninformative or mixed).

23. For step 5 ("Perform an internal assessment of the category"), the output of QSAR modelling introduced and considered as outlined above can contribute to trend analysis for compounds in the series both for those for which there are data and those for which there are not. Through measures of similarity, it can also contribute to delineation of the bounds of the category.

24. For step 6 ("Prepare category test plan"), where critically evaluated output of QSAR contributes meaningfully to trend analysis, it may obviate the need for testing of certain members of the category. Rationales need be based on well documented critical evaluation of the output of batteries of models, with clear delineation of strengths and limitations and take into account availability for other members of the category and consistency overall of critically evaluated QSAR output and data.

25. For step 8 ("Perform an external assessment of the category and fill data gaps"), the principles outlined above for consideration of QSAR in development of the test plan are also relevant in considering their contribution to the initial assessment. This contribution must necessarily be based on critical evaluation of the output of a suite of models, based on an understanding of their relative limitations and strengths for the specified application.

3.2.5 Guidance on different types of categories

Chain length

26. These are defined as categories showing an incremental, and usually constant, increase in chain length across the category. There is an assumption that each category member exhibits the same toxic mode of action. Examples are the homologous series of alpha-olefins (see Example A in Annex 1) where each category member differs by a $-CH_2-$ unit and the ethylene glycols where there is an incremental increase in the number of CH_2CH_2O groups.

27. Categories defined by chain length generally show an incremental change in molecular weight and other physico-chemical properties such as water solubility or Log Kow. However, not all properties will necessarily exhibit a linear relationship with chain length and care must be taken in making assumptions about such trends. For example, the alpha-olefins show declining acute toxicity to fish with increasing chain length and decreasing water solubility. There is an apparent 'cut-off' point between the C8 and C10 chain length at which acute toxicity to fish is no longer observed due to the decreasing water solubility. For aquatic toxicity, the interplay between decreasing water solubility and increasing log Kow – a key indicator of uptake from water - with increasing carbon chain length is often important in determining this cut-off point. Similarly, a trend of increasing molecular weight may lead to decreasing systemic toxicity as absorption decreases and there may be a change of physical state of the category members as chain length increases.

28. Careful thought should be given to selecting the boundaries of a chain length category. The cut-off points described above may provide useful boundaries. The potential scope and size of a chain length category may be larger than that covered by a particular manufacturer or consortium. Where possible, well-characterised substances which are not HPV but which fit into the series should be included. There

may be cases when testing the end members of a chain length category is not appropriate. For example where the existing data indicates that the cut-off for toxicity occurs earlier in the series it may not be necessary to test the end member for that endpoint.

29. QSARs can be used to help justify the category and fill data gaps. In general, substances at either end of a chain length category should have all SIDS endpoints fulfilled, preferably with test data. This permits interpolation of data to the other category members rather than extrapolation and increases confidence in the read-across. For example, a linear regression has been used to predict acute aquatic toxicity of long chain alcohols. For categories where there is more than one variable, such as variation in chain length and degree of branching of the chains, more category members are likely to be required to bring confidence to the interpolations being made.

Metabolic pathways

30. The underlying hypothesis for a metabolic series is a sequential metabolism of a parent chemical to downstream blood metabolites that are chemicals of interest. Hazard identification studies with the parent compound could then be used to identify the hazards associated with systemic blood levels of the downstream primary and secondary metabolites and once quantified, can be used in place of studies using direct exposure to primary and secondary metabolites themselves. In certain instances, the metabolism of the parent compound within barrier tissue (e.g. lung or gut tissue) occurs so rapidly that the initial primary metabolite is the predominant chemical found within the blood. Under these circumstances data from hazard identification studies conducted with that primary metabolite itself can be used to identify hazards for the parent compound. PBPK or PBPD models may help to define categories. The metabolic pathway approach is usually reserved to some toxicological endpoints. For physico-chemical properties, environmental fate and ecotoxicity, information on the parent compound would need to be available.

31. The first technical issues faced when forming a metabolic series is to determine if the metabolism that is assumed to occur does occur. This is necessary before moving any further in developing a metabolic category and preferentially should be determined *in vivo*. In certain instances, *in vitro* metabolic studies can be used to help identify metabolic pathways, but the definitive evidence should be conducted in whole animals. The primary and secondary metabolites should be detected either in the blood or tissue. Primary and secondary metabolites that cannot be readily determined in blood or tissue should not be candidates for a metabolic series approach without some limitation placed upon the use of the information.

32. The second technical issue pertains to the level of evidence required to describe the metabolic processes. Direct measurement of the parent chemical and primary and secondary metabolites in the blood in an *in vivo* exposure is the recommended standard. The level of evidence required to presume that there will be blood-borne levels of primary and secondary metabolites following exposure to parent chemical, will have to be determined on a case by case basis. Certain metabolic processes are ubiquitous and well understood and these can be presumed to occur without performing *in vivo* experiments in every instance. Other metabolic processes are not part of normal metabolism or require enzyme induction. These metabolic processes may not be well characterized and should not be assumed without specific *in vivo* evidence of blood levels of primary and secondary metabolites.

33. The third technical issue provides a limitation for the metabolic approach to forming categories. The metabolic category reasoning is only useful for identifying hazards related to systemic blood levels of the parent compound and/or primary and secondary metabolites. Other endpoints of hazard identification studies that are dependent upon site of contact effects (e.g. eye, skin, respiratory tract irritation, irritation to gastric mucosa) cannot be addressed using the metabolic category logic. These sites of contact effects are often due to the physical chemical property of the chemical in question and therefore may differ considerably between the parent compound and primary and secondary metabolites. In addition, tests that

identify unique structural characteristics (e.g. skin or respiratory sensitization) or are dependant upon physical chemical properties (e.g. volatility and LC50 values) should not be considered as part of metabolic category because these properties may not be similar amongst the various members of the metabolic series.

34. An additional limitation of the metabolic categories approach is that metabolism and toxicokinetics experiments have to be conducted with the parent compound. Typically, these types of studies are not SIDS elements and therefore would require a sponsor of the chemical to do additional work beyond what is normally considered necessary. However, it should be recognized that the savings involved (numbers of animals used, testing costs) could be considerable compared with generating data for each metabolic category member for each endpoint of systemic toxicity. Since the OECD HPV Chemicals Programme is a screening level program that is interested in identifying hazards related to systemic blood levels, it should not become necessary to provide definitive toxicokinetic evidence or develop a toxicokinetic model for acceptance of hazard identification studies as relevant for the primary and secondary metabolites.

35. An additional advantage of using the metabolic category toxicity data is that in certain instances, higher systemic blood levels of a chemical can be achieved from metabolic pathways than if the primary or secondary metabolite was administered directly. For example, if a material is corrosive or has limited volatility, higher blood levels may be found following the administration of the parent compound than if the primary or secondary metabolite was administered directly to the animal.

36. The following specific issues should be taken into account when developing a metabolic pathway category, according to the stepwise procedure described in section 3.2.3

- *ad* step 1: Provide definitive information on the metabolism of the parent chemical to the primary and secondary metabolite. This information should also include, preferably, a time course data for either blood or tissue for both the parent chemical as well as the primary and secondary metabolites.

- *ad* step 2: The metabolism experiment should be examined to determine, if in fact, the primary and secondary metabolites are formed, if they achieve appreciable levels within the blood and/or tissues and determine basic toxicokinetic parameters for the parent material. For example, the $T_{1/2}$ for elimination for the parent chemical should be determined if possible. If the metabolism of the parent chemical to the primary metabolite is rapid and is thought to occur within barrier tissues, then it may be appropriate to use hazard identification studies from the primary metabolite to identify hazards associated with exposure to the parent chemical.

- *ad* step 3: If there are appropriate hazard identification studies that have been conducted with the parent chemical or primary or secondary metabolites for similar toxicity endpoints, then these studies should be examined to see if these materials have similar toxicity. If data is not available for the metabolic series in question and a study is to be designed and conducted, then the parent compound should be tested, so that blood levels of all category members will be present. The toxicokinetic and metabolic experiments that provide the basis for the metabolic category should have robust summaries prepared and be included in the SIDS Dossier for the parent chemical, primary and secondary metabolites. Within these robust summaries a table should be included detailing the relative blood levels of the parent chemical, primary and secondary metabolites.

– *ad* step 5: A quantitative analysis between exposures of the parent chemical and the primary and secondary metabolite is not necessary as the point of the OECD HPV Chemicals Programme is to provide hazard identification studies for these materials, not a quantitative analysis as would be done for risk assessment purposes. If the chemical becomes a chemical of concern, then additional toxicokinetic analysis (including preparing a model) may be appropriate, but for the purposes of the screening level OECD HPV Chemicals Programme it is not necessary.

37. The metabolic approach should not be used for environmental toxicity endpoints unless the metabolism of the parent compound to the primary or secondary metabolite can be demonstrated within the test species in question. Whereas it may be appropriate to extrapolate within mammals, it may not be appropriate to extrapolate between amphibia and fish or insects and other species due to the difference in the metabolic processes and enzymes present within those species.

38. On the other hand the same concept underlying the metabolic pathways can be used for environmental degradation processes. For example, for a substance which hydrolyses very rapidly in aquatic test systems (half-life < 1 hour), the aquatic toxicity endpoints can be covered by the test results with the degradation product(s) [see also the OECD Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures No. 23, ENV/JM/MONO(2000)6].

Chemical mixtures

39. Categories can sometimes apply to series of chemical reaction products or chemical mixtures[6] that are, again, related in some regular fashion. Analogous to the basic "discrete chemical" category model, in a mixture category some, but not all, of the individual mixtures may undergo testing. Annex 1 illustrates this using a category made up of linear alkylbenzene mixtures. This example was used to assess these chemicals in the OECD HPV Chemicals Programme. This is a relatively simple example of the type of approach that can work in the HPV Chemicals Programme. Additional guidance for other types of mixture categories is given below. In general, the use of the chemical category concept for chemical mixtures is not straightforward and further guidance will need to be developed as more experience is gained.

Isomers and their mixtures

40. Isomers are chemicals that have identical molecular formulas but different molecular arrangements. Although there are several types of isomers, the two that typically will be considered within the HPV Chemicals Programme are *structural* and *geometric*.

41. Structural isomers are molecules with differences in the arrangement of their atoms, such as butene-1 and isobutene. Structural isomers can include:

- chain isomers, for example hydrocarbon chains with identical or variable lengths and variable branching patterns

- position isomers, for example hydrocarbon chains with a functional group that varies in position along the chain

[6] The concept applies to reaction products or process streams and does not refer to intentional mixtures of substances (or preparations).

42. A third type of structural isomer is referred to as a functional group isomer. These isomers also have identical molecular formula, but contain different functional groups. Examples of two functional group isomers with $C_4H_{10}O$ as a molecular formula are 1-butanal and 2-butanone. Each of these isomers contain a carbonyl group (C=O), but are representative of two different chemical families, aldehydes and ketones. Although structural isomers, this type is less likely to be considered within a category for the Programme because functional isomers can have very different chemical and biological properties. Functional isomers are not included within the scope of this guidance.

43. Geometric, or stereo, isomers contain their molecules in the same arrangement, but a section or sections of each have different spatial arrangements. For example, *cis*-butene-2 and *trans*-butene-2 each have carbon groups on either side of a double bond, which cannot rotate, that are arranged on either the same side of the molecule (*cis*-) or opposite sides of the molecule (*trans*-).

44. Geometric and select structural isomers can have similar, somewhat different or very different chemical or toxicological properties. Even though they may behave identically in many chemical reactions, it is for example well known that the enzyme specificity in biological systems may be totally different and extreme caution is needed in case of such substances. An example of such specificity is select carbohydrates, which may be metabolised or not depending on the orientation of functional groups. Another example showing a profound difference in effects is the drug Thalidomide, which have one chiral atom and therefore exists as two enantiomers. The optical "R" isomer is an effective sedative and the optical "S"- isomer is a teratogen causing serious birth defects in children to mothers using the drug during pregnancy."

45. There are general rules for using read-across techniques as they apply to isomers:

- Relatedness - The substance(s) without data as well as the substance(s) with data are similar such that their physicochemical, biological, and toxicological properties would be expected to behave in a predictably similar manner or logically progress across a defined range.

- Structural Similarity - The substance(s) without data possesses a small incremental structural difference from the reference substance(s) or the difference between the two would not be expected to affect the property sufficiently such that it could not be accurately predicted.

46. There can be instances within a category of isomers, specifically as related to structural isomers, when read-across for an endpoint is not appropriate. An example is illustrated with two categories of isomers other than the butenes, the pentanes and hexanes. Though the pentanes may be broadly described as isomers, they actually represent three types of hydrocarbons, normal alkanes, branched alkanes, and cyclic alkanes. It is known that n-pentane, 2-methylbutane, 2,2-dimethylpentane, and cyclopentane exhibit distinct differences in potential biodegradability. n-Pentane and 2-methylbutane are readily biodegradable, whereas 2,2-dimethylpentane and cyclopentane are poorly biodegraded. Therefore, it would not have been possible to assess the biodegradability of the poorly biodegradable pentanes if they had no data using the results from the readily biodegradable pentanes even though the pentane isomers could still be considered a category for all other endpoints within the Programme. In such a case, the potential biodegradability of the two groups of pentanes would each have to be characterised separately within the context of the category. Likewise, the peripheral neurotoxicity in humans associated with n-hexane exposure has not been demonstrated to occur with exposure to other hexane isomers and a discussion of this effect within a hexane isomer category would have to isolate n-hexane from the other isomers.

47. An example of a category of isomers is provided in Annex 1 (Example D: Butenes and their mixtures). Based on this example, general principles of read-across/extrapolation and application of data within a category of isomers and their mixtures can include:

- Select properties of isomers may be read-across to another isomer(s) or to an isomeric mixture within a category if the data are similar and/or if the structure of the isomer(s) without data is similar to the isomers with data.

- Extrapolating properties to isomeric mixtures should take into account mode of action, potential additivity and synergy, as well as purity profiles, and mixture composition.

- For toxicological endpoints (e.g., LC_{50}, NOAEL) a range of toxicity or the lowest value in a range of toxicity may be used for read-across.

- Read-across from one isomer to another may not be straightforward. Metabolic data may be needed if existing knowledge of category members or related non category members suggests that differences may be expressed within a biological endpoint of interest.

Complex substances

48. Complex substances include a diverse range of materials which are frequently described as substances of *Unknown or Variable composition, Complex reaction products or Biological material* (UVCB Substances). There are many different types of complex substances, though generally they all have the following characteristics in common.

- They contain numerous chemicals (typically closely related isomers), and cannot be represented by a simple chemical structure or defined by a specific molecular formula. They are, however, assigned unique Chemical Abstract (CAS) numbers (see note[7] below about unique issues with CAS numbers for UVCB substances).

- They are not intentional mixtures of chemicals.

- Many are of natural origin (e.g., crude oil, plant extracts) and cannot be separated into their constituent chemical species.

- The concept of "impurities" typically does not apply to complex substances.

49. Category approaches for complex substances may vary, though generally the approach will be related to how the substances are manufactured, defined and used. For example, petroleum substances are generally defined by hydrocarbon chemistry (e.g., aliphatic hydrocarbons, aromatic hydrocarbons, etc.), physicochemical properties such as boiling range or carbon-number range, manufacturing and processing conditions, and common use categories. For hydrocarbon solvents [see example E provided in Annex 1], the categories are based on the typical chemistry and carbon-number range of hydrocarbon solvents and common uses. Under this approach, those hydrocarbon solvent substances with similar chemistry and

[7] CAS numbers are important for identifying substances; however, for UVCB substances they do not represent a unique chemical and the specificity of the CAS number definition may vary (some CAS number definitions are rather narrow, some are very broad). CAS numbers for complex petroleum substances are based on a hierarchy of considerations including hydrocarbon type, carbon number range, distillation range and the last processing step. Because of these numerous considerations, similar products sometimes have different CAS numbers. There are also historical and geographical reasons why similar substances may have been assigned different CAS numbers. Further, some CAS numbers have a broad definition that may fit different substances that would fall into different categories. Because of this, physical properties and chemical structure are the preferred way to construct categories of complex substances.

carbon-number range are grouped together in the same category and the category is defined by the composition of those substances. This approach is practical and has the benefit of making sure that similar commercial products are grouped together in the same category.

50. Based on the example described in Annex 1, some general guidance can be provided for developing chemical categories with complex mixtures:

- It is important to clearly characterise mixtures, details of the production process can be useful. It is necessary to identify the following attributes of a complex mixture:
  - Composition (what is present and in what proportion)
  - Impurities (substances present that are not wanted but need to be identified)
- Properties of the components of a complex mixture can be applied to the complex mixture if the properties of the single components are similar.
  - It is necessary to identify representative components of the mixture to cover the carbon range and structures of the mixture.
  - Components with outlying properties need to be identified (e.g. specific toxicity of hexane compared to other aliphatic hydrocarbons, higher water solubility of aromatic hydrocarbons compared to aliphatic hydrocarbons).
- Properties of a complex mixture can be read-across to another complex mixture if the composition of the two are similar.
- Quantitative read-across is more difficult (ranges can be used where applicable).It is necessary to carefully consider the dose for read across because of the nature of the mixtures and the amount of components of concern.
- It is necessary to carefully identify representative substances for testing purposes.

Metal and metal compounds

51. The concept of chemical categories has traditionally been widely used for inorganic substances. However, not much experience is available to date of a systematic use of this approach. The concept is being used for the assessment of nickel and nickel compounds (see example in Annex 1).

52. There are a number of assumptions underlying any grouping of metal compounds for estimating their biological properties. The main assumption is that it is the metal ion that is responsible for the effects to be assessed. This is considered to be a reasonable assumption for the majority of the inorganic and some organic anions. This implies that in the case of inorganic salts, the toxicity of the counter ion is assumed to be largely irrelevant in producing the effects to be assessed. If the counter ion influences significantly the effects of the compound to be assessed, it can not be part of the category. Where a metal can have different valence states (e.g. chromium), the toxicities of the different valence states may vary, and the different valence states considered separately.

53. The water solubility of the metal compounds is often used as the starting point for establishing a category, as this reflects the availability of the metal ion in the different compartments of interest. For inorganic nickel compounds, a number of sub-groups have been suggested, reflecting different ranges of aqueous solubility. In contrast to inorganic nickel compounds it is not obvious how to group organic nickel compounds based on solubilities alone.

54. Based on the example of nickel and nickel compounds, some tentative general guidance for metal and metal compounds can be proposed:

- The main assumption is that the metal ion (or ion complex) is responsible for the effects to be assessed (the toxicity of the counter-ion is assumed to be largely irrelevant in producing the effects to be assessed).
- One basis of grouping could therefore be water solubility (inorganic metal compounds), taking into account:
    - transformation/ dissolution of insoluble compounds
    - bioavailability of the metal ion in the environment
    - solubility in biological fluids
    - persistence in the body

  Other bases for grouping can be considered. For example if data is available, for systemic effects the solubility of the different compounds in the acidic environment of the stomach could be considered.
- The assumption that the metal ion (or ion complex) is mainly responsible for the effects rather than the counter-ion may not work for local mammalian toxic effects.
- Possible differences in the toxicity of different oxidation states of the metal ion (or ion complex) should be considered.
- Whilst the assumptions shown above can be expected to be valid for a wide range of inorganic compounds, these do not necessarily apply to organically based metal compounds. A different approach may be needed for grouping organic metal compounds. .

55. It should be noted that whilst this example considers groups of metallic cations, similar considerations would also apply to salts of anions where there are concerns for toxicity (e.g. cyanides, oxalates).

3.2.6 Experience in Developing Chemical Categories

56. OECD experience provides a framework for handling categories. However, since that experience is limited, lessons learned in the OECD, and other similar programmes such as the US HPV Challenge Programme, will provide a measure of feedback and review.

57. The largest categories applied in the OECD HPV Chemicals Programme to date contain eight to ten chemicals. This is not a formal maximum, but acceptable categories will tend to be self-limiting because endpoint trends are generally disturbed as structural variations become more complex. Practically the analysis of large categories can also become unwieldy tending to limit the size of categories proposed. In this regard, groups of related individual categories may be considered, each one contributing elements in the design and implementation of an overall category strategy. A larger category may be justifiable in certain cases, such as when toxicity of the category is generally low.

58. Annex 1 contains a number of examples of how a category approach has been used for the purpose of collecting, reporting, and assessing hazard information in the OECD HPV Chemicals Programme. Other examples of categorising chemicals for hazard assessment purposes include the CONCAWE (the European oil company organisation for environment, health and safety) approach of categorising chemicals in petroleum streams (CONCAWE, 1998), approaches to assess the ecotoxicity (Bowmer et al., 1998) and health effects (Clary, et al., 1998) of lactate esters, and a number of category/SAR analyses by the German authorities (Greim, et al., 1994, 1995, 1998; and Poelloth and Mangelsdorf, 1997).

REFERENCES

Bowmer, CT, RN Hooftman, AO Hanstveit, PWM Venderbosch and N van der Hoeven. 1998. *The ecotoxicity and the biodegradability of lactic acid, alkyl lactate esters and lactate salts.* Chemosphere (37)(7):1317-1333.

Clary, JJ, VJ Feron and JA van Velthuijsen. 1998. *Safety assessment of lactate esters.* Regul. Tox. and Pharm. (27):88-97.

CONCAWE. 1998. Heavy Fuel Oils Product Dossier, No. 98/109. *CONCAWE*, Brussels, May, 1998.

Greim, et al. 1994. Toxicity and ecotoxicity of sulfonic acids: structure-activity relationships. *Chemosphere* 28:2203-2236.

Greim, et al. 1995. Assessment of structurally related chemicals: Toxicity and ecotoxicity of acrylic acid and acrylic acid aklyl esters (acrylates), methacrylic acid and methacrylic acid esters (methacrylates). *Chemosphere* 31:2637-2659.

Greim, et al. 1998. Toxicity of aliphatic amines: structure-activity relationships. *Chemosphere* 36:271-295.

Poelloth, C. and I. Mangelsdorf. 1997. Commentary on the application of (Q)SAR to the toxicological Evaluation of existing compounds. *Chemosphere* 35:2525-2542.

ANNEX 1

EXAMPLES OF CATEGORY APPROACHES

1. Examples presented in this Annex are chemicals being investigated in the OECD HPV Chemicals Programme. They have been shortened for purposes of presentation in this document to illustrate the steps for identification and development of categories included in this guidance document. The examples are:

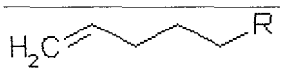
- A. Alpha-olefins - discrete chemicals with an incremental and constant change across the category;
- B. Linear alkyl benzenes - family of mixtures; and
- C. Brominated diphenyl ethers - family of congeners.
- D. Butenes – family of isomers and their mixtures
- E. Hydrocarbon solvents – family of complex mixtures
- F. Inorganic nickel compounds

Example A: Alpha Olefins Series

Step 1: Identification of structure-based category and its members:

2. The category was defined as olefins bearing a single medium-length, even-numbered, unbranched aliphatic chain with no other functional groups ("α-Olefins"). This category consists of discrete chemicals with an incremental and constant change across its members (dimethylene group). Because the double bond is terminal, possible metabolic reactions such as oxidation at the double bond or allyl position should not be unduly affected by the chain lengthening. The lower (C_6) and upper (C_{14}) boundaries were based on the available product lines of the sponsors involved in the OECD effort.

3. The chemical structure of the category is:



R = CH_3, n-Propyl, n-Pentyl, n-Heptyl, n-Nonyl

Step 2: Gather published and unpublished literature for each category member.

4. A literature search resulted in identifying a significant amount of available data for most category members in most of the major SIDS endpoints.

Step 3: Evaluate available data for adequacy.

5. Available data was evaluated at the individual study level and collected for each member of the category. Available data were compiled and included all SIDS endpoints and other relevant information; non-SIDS data were found and used in the hazard profile (e.g., aspiration hazard potential to humans).

Step 4: Construct a matrix of data availability.

6. Table A-1 is a matrix of SIDS endpoints and available/adequate data for each member of the alpha-olefin category. For simplicity, not all relevant data are presented.

Step 5: Perform an internal assessment of the category.

7. The information in Table A-1 identifies where data gaps exist (noted as "-" in the table). Adequate data (noted as "√" in the table) are available for most endpoints. Endpoint data were evaluated to determine whether they correlate with chemical structure to judge the acceptability of the category. Although not shown in Table A-1, the data suggested that water solubility decreased with increasing chain length and aquatic toxicity appeared to decrease with increasing chain length.

Table A-1
STEP 4: Matrix of Available and Adequate Data on Alpha-Olefin Category Members

Test	Hexene	Octene	Decene	Dodecene	Tetradecene
Physicochemical Properties					
Partition Coeff.	√	-	√	√	-
Water Solubility	-	-	-	√	√
Environmental Fate					
Biodegradation	√	-	√	√	√
Ecotoxicity					
Acute Fish	√	-	√	√	-
Acute Daphnid	√	-	√	√	-
Alga	√	-	√	√	-
Terrestrial	-	-	√	-	-
Human Health Effects					
Acute Oral	√	√	√	√	√
Acute Inhalation	√	√	√	√	√
Acute Dermal	√	√	√	√	√
Repeated Dose	√	√	-	-	-
Genotoxicity (in vitro - bacteria)	√	√	√	√	√
Genotoxicity (in vitro - non-bacterial)	√	√	-	√	√
Genotoxicity (in vivo)	√	-	-	-	-
Repro/Developmental	-	-	-	-	-

(√) = Data available and considered adequate; (-) = No data available, or available data considered inadequate.

Step 6: Prepare category test plan.

8. Table A-2 contains the proposed testing plan only for the endpoints for which new testing was recommended for the alpha-olefins. In this case it appears reasonable that if data gaps are filled by testing at the upper and lower ends of the homologous series (shaded regions in the table), and if the results suggest a pattern, then the remaining data gaps can be considered to fall within the ranges defined by the data.

Step 7: Conduct necessary testing.

9. The shaded cells in Table A-2 show where new testing was recommended for the category.

Table A-2 Alpha-Olefin Proposed SIDS Test Plan[1]					
Selected SIDS Endpoint	**Hexene**	**Octene**	**Decene**	**Dodecene**	**Tetradecene**
Water Solubility	√/-	-	-	√/+	√/+
Acute Fish	√/+	-	√/+	√/+	-
Acute Daphnid	√/+	-	√/+	√/+	-
Acute Algae	√/+	-	√/+	√/+	-
Repeated Dose	√/+	√/+	-	-	-[2]
Repro/Developmental	-	-	-	-	-[2]

[1] KEY: √/- = data available, but not adequate; √/+ = data available and considered adequate; - = no data available. Shaded cells represent those SIDS endpoints for which testing was recommended.
[2] A combined repeated dose and reproductive/developmental toxicity screen study design was recommended.

Step 8: Perform an external assessment of the category.

10. Table A-3 shows the results of the recommended testing and how it "fit" with available data for purposes of evaluating whether a pattern exists between some of the SIDS endpoints and the increase in 2-carbon increments from hexene to tetradecene. Note that there are four data points that exist in Table A-3 that were not present in Table A-2 (the octene water solubility and ecotoxicity results); these data were a late addition to the octene dossier and are included here to enhance the category analysis. This illustrates how all data should be considered in the evaluation of a category, even if it becomes available well after the literature search has been completed.

11. The new data show that patterns are clearly evident. For example, there is an apparent decrease in water solubility with increase in carbon chain length and a decrease in acute toxicity to fish and daphnids with an increase in carbon chain length. On the other hand, the mammalian toxicity data suggest a pattern of no difference between hexene and tetradecene for repeated dose (general) toxicity and developmental/ reproductive toxicity.

Table A-3
Results and Interpolation of Alpha-olefin SIDS Category Testing[1]

Selected SIDS Endpoint	Hexene	Octene	Decene	Dodecene	Tetradecene
Water Solubility	50 mg/L[2]	(4.1 mg/L)[3]	INSOLUBLE	"insoluble"	0.0004 mg/L
Acute Fish	5.6 mg/L (LC_{50})	(4.8 mg/L)[3] (LC_{50})	>Water solubility? (Reported value >10,000 mg/L (LC_{50})	>Water solubility? (Reported value >1000 mg/L (LC_{50})	>Water solubility (LC_{50})
Acute Daphnid	10 mg/L (NOEC)	($3 < EC_{50} > 10$)[3]	>Water solubility? (EC_{50})	>Water solubility? (EC_{50})	>Water solubility (LC_{50})
Acute Algae	>Water solubility (LC_{50})	(>Water solubility)[3] (LC_{50})	>Water solubility? (EC_{50})	>Water solubility? (EC_{50})	>Water solubility (LC_{50})
Repeated Dose	$NOEL_{oral}$ = 101 mg/kg (males) and >1000 mg/kg (females)	NOEL = 50 mg/kg (males)	SIMILARLY TOXIC		$NOEL_{oral}$ = 100 mg/kg (males) and >1000 mg/kg (females)
Repro/ Developmental	$NOEL_{repro}$ and $NOEL_{dev}$ = >1000 mg/kg	SIMILARLY TOXIC			$NOEL_{repro}$ and $NOEL_{dev}$ = >1000 mg/kg

[1] KEY: - = no data available; shaded cells represent those SIDS endpoints for which OECD recommended testing.
[2] Apparently this was the original value thought not adequate, but estimations of the water solubility were similar to this value, so a new study was not performed.
[3] These data were not identified as being available in the Testing Plan. However, because they were reported in the dossier, they are included here to enhance the category analysis.

Step 9: Fill the data gaps

Water solubility.

12. The 50 mg/L value for hexene and 0.0004 mg/L value for tetradecene suggest a wide range of solubility for the five members of the group. The octene value of 4.1 mg/L suggests that the pattern (decreasing water solubility with increasing chain length) holds. Therefore, water solubility tests were judged not necessary and computer estimates (consistent with the latter premise for decene and dodecene) were considered acceptable.

Acute aquatic toxicity

13. The data in Table A-3 suggests that hexene and octene may exhibit moderate acute toxicity to fish and daphnids based on measured values (NOEC, LC_{50}, EC_{50}). However, all other members of the category appear to show no effects on fish and daphnids at saturation. In the case of algae, all category members show no effects at saturation. From a category perspective, it appears that a declining pattern

exists for fish and daphnids (hexene and octene are more toxic than decene, dodecene, and tetradecene) but there was a flat pattern for algae (all members appeared equal). Based on this information, it was decided that no additional aquatic toxicity testing was necessary. The three literature values for octene noted in Table A-3 were considered acceptable. The aquatic acute toxicity for those endpoints correlate with water solubility, which in turn appear to determine (or limit) bioavailability of octene.

Repeated dose toxicity

14. The results presented in Table A-3 suggest that the general toxicity of hexene and tetradecene are similar, whereas octene appears more toxic than either hexene or tetradecene. In both cases, male rats were more sensitive than female rats. The effect observed in males, a male-rat specific kidney effect, does not appear to be relevant to humans. Also, both studies followed the OECD repeated dose/reproductive/developmental toxicity screening testing protocol. There were no data for either decene or dodecene. The octene data point suggests that any category pattern that might exist (equal toxicity across all members) given the hexene and tetradecene data might not exist for the middle members of the category. However, upon closer inspection of the octene data in the octene dossier, it is seen that the doses used in the repeated-dose study were 5, 50, and 500 mg/kg. Since the LOEL was 500 mg/kg, the "true" NOEL is anywhere from 50 to 500. Therefore, given these data, one could recommend that all members of the group likely have equal general toxicity under repeated dose conditions and testing of decene and dodecene is not required.

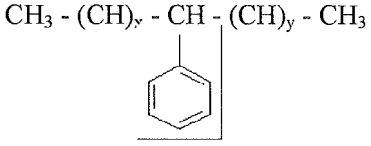
Reproductive/Developmental toxicity

15. The reproductive/developmental toxicity row in Table A-3 shows that data are available only for hexene and tetradecene. As with the repeated dose data, the results of the two studies were essentially the same. This suggests that it would not be necessary to test the middle three members of the category (octene, decene, and dodecene), especially given the results of the assessment of general toxicity (see above). The data suggest a consistent pattern across the category, or that all members are equally toxic for reproductive/ developmental effects under the conditions of the hexene/tetradecene studies (highest dose of 1000 mg/kg).

Example B: Linear Alkylbenzenes

Step 1: Identification of structure-based category and its members:

16. The linear alkylbenzene (LAB) category is comprised of nine different commercial formulations. Each formulation is a mixture containing various proportions of individual LABs with the following formulae:



Where x + y = 7-13 and x = 0-7, giving a linear carbon range of C_{10} to C_{16}.

17. Thus, this category would fall under "family of mixtures" in terms of category type. Table B-1 presents the nine commercial products evaluated. Note that the LAB category may be further subdivided into three subcategories based on the percentage of alkyl substituents with a low (C_{10}-C_{11}), mid (C_{11}-C_{13}), and high (C_{13}-C_{14}) proportion of carbon chain lengths.

Table B-1
Assignment of LAB SubCategories[1]

LAB Formulation	Carbon Chain Length for Substituted Alkyl Group (Numbers represent percent of total)				
	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}[(2)]
Nalkylene 500	21	39	31	7	<1
Nalkylene 500L	20	44	31	5	<1
Alkylate 215	16	43	40	1	<1
Nalkylene 550L	14	30	29	20	7
Alkylate 225	7	25	48	19	1
Nalkylene 575L	9	17	20	30	15
Nalkylene 600	<1	1	23	50	25
Nalkylene 600L	<1	1	23	50	25
Alkylate 230	1	2	16	50	30

1 The two shaded regions and the open area make three subcategories by presenting two ends of the spectrum in terms of a higher proportion (>50%) of shorter carbon chains (upper left) and a higher proportion (>50%) of longer carbon chains (lower right). Bolded formulations had available data in all SIDS categories.
2 The proportion of C_{15} and C_{16} is < 1% in all formulations except for an incidence of 1% C_{15} in Alkylate 230.