EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Forming Chemical Categories..

Grace Patlewicz

European Chemicals Bureau
Institute for Health & Consumer Protection (IHCP)
Joint Research Centre (JRC), European Commission
21020 Ispra (Va), Italy

E-mail: grace.patlewicz@jrc.it http://ecb.jrc.it/QSAR

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Overview

- What is a Category?
- What do we mean by Chemical Categories?
- Regulatory definitions e.g. OECD
- So what?
- How do we form Categories?
  - What guidance is out there, their strengths and limitations
- Possible approaches & examples
  - A couple of ECB activities
- Challenges?

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

What is a Category?

Some potential definitions..

- A collection of objects sharing a common attribute
- A division of objects according to appearance, quality, functionality..
- A group of objects that have some features that are the same

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

What do we mean by Chemical Categories?

- A group of chemicals that have some features that are common
  - Structurally similar e.g. common substructure
  - Property e.g. similar physicochemical, topological, geometrical, or surface properties
  - Behaviour e.g. (eco)toxicological response underpinned by a common Mechanism of Action
  - Functionality e.g. preservatives, flavourings, detergents, fragrances

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

OECD Definition of Category

- A chemical category is a group of chemicals whose physicochemical and toxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity
- These structural similarities may create a predictable pattern in any or all of the following parameters: physicochemical properties, environmental fate and environmental effects, and human health effects

OECD Manual for Investigation of High Production Volume (HPV) Chemicals. Chapter 3, Section 2.

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Annex IX of REACH

Substances whose physicochemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity may be considered as a group, or "category" of substances.

Application of the group concept requires that physicochemical properties, human health effects and environmental effects or environmental fate may be predicted from data for a reference substance within the group by interpolation to other substances in the group (read-across approach). This avoids the need to test every substance for every endpoint.

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

So what?

Why might we want to develop categories?

- If we can characterize chemicals sufficiently in order to group similar chemicals together with respect to a given property/endpoint… then we can use that relationship to make predictions, rationalizations about the behaviour for new chemicals likely to belong to the same group.
- In the context of safety assessments – grouping has the potential to save resources and reduce hazard testing.
- BUT the challenge is to characterize chemicals appropriately in order to group them in the first place.

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

How do we form a Category?

What Guidance is out there?

OECD HPV Manual

- This addresses:
  - Definitions and explanations of the chemical category concept.
  - General approach for developing categories.
  - Differences in grouping for different endpoints.
  - Use of (Q)SARs for the development of a category
  - Guidance on different types of categories
  - Provides a number of examples of categories that have been adopted within the OECD HPV Chemicals Programme. NB: Examples appear largely limited to chemical classes or structurally similar chemicals.
- **but** it lacks detail on the practical steps on how
  - to formulate a category,
  - justify it
  - document it

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

How do we form a Category

What Guidance is out there?

- **In development for REACH (under the RIP 3.3 Task 3):**
  - To develop *non-prescriptive* guidance that explains and illustrates:
    - the commonalities and differences between SARs, read-across and categories (including sub-categories)
    - how to justify and report qualitative and quantitative read-across
    - how to build a category (practical details), including examples of qualitative and quantitative read-across
    - how to evaluate the robustness and applicability domain of a category
    - how to justify and report a category proposal

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Possible and Practical approaches

- Grouping could be nominally categorised into one of four classes:
  - knowledge-based
  - analogue-based
  - unsupervised
  - supervised

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Knowledge-based approach

- Encompasses "human expert rules". Human experience linked to specific endpoints is encoded in the form of structural alerts or chemical reactivity mechanisms.
- Examples include Cramer rules, Ashby Tennant alerts, Derek alerts..

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Example for Skin sensitization

- Skin sensitization results from a T-lymphocyte mediated immune response to a chemical allergen that comes into contact with the skin.
- A chemical need to overcome a number of hurdles in order to induce skin sensitization.
- These comprise:
  - Penetration into the viable epidermis across the *stratum corneum*.
  - **Formation of a stable association with protein** to create an immunogenic complex. This requires that a **chemical is** inherently **protein-reactive**, or can be transformed chemically or metabolically to a protein-reactive species. Typically the stable association is thought to be a **covalent** one
  - Deliver dermal trauma sufficient to induce and up-regulate those epidermal cytokines that are necessary for the mobilisation, migration and maturation of LC.
  - Be inherently immunogenic such that a T lymphocyte response of sufficient magnitude is stimulated.
- If these hurdles are not successful then skin sensitization will either not occur or will be sub-optimal.

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Example for Skin sensitization

Joint Research Centre

- The **KEY** step is the formation of a stable association (usually covalent) with a protein.
- Skin sensitization is underpinned by an understanding of chemical reactivity and chemical mechanisms.
- The chemical behaves as an electrophile whereas the protein behaves as a nucleophile.
- There are various types of electrophile-nucleophile reactions encountered in skin sensitization, perhaps the most common are: Michael-type reactions; SN2 reactions; SNAr reactions; acylation reactions and Schiff-base formation.
- It is possible to group chemicals and rationalise their skin sensitization potential on the basis of these chemical reaction mechanisms.

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

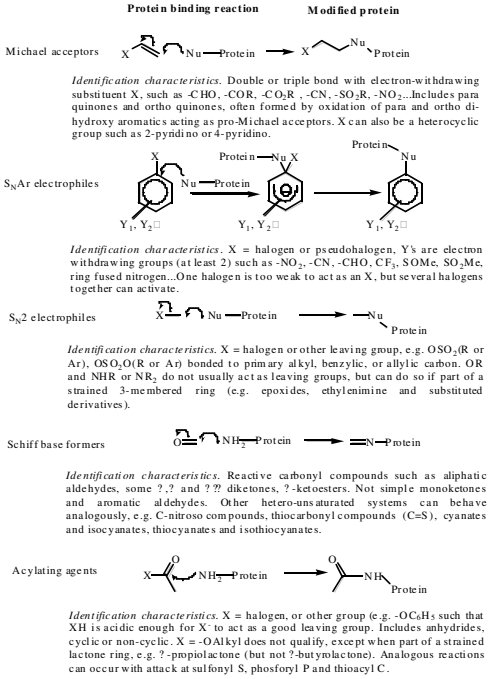


Figure taken from Aptula AO et al, 2006 *Tox in Vitro*, 20(2):239-47

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Analogue-based approaches

- Search and retrieval of structurally similar chemicals either on the basis of a substructure, fingerprint/fragment or on the basis of a similarity index measure such as Tanimoto
- The Tanimoto similarity index compares one compound with another on the basis of fingerprints to arrive at a number between zero and one. A index of 0 signifies that the two compounds have nothing in common whereas a value of 1 reveals them to be identical.

Tanimoto index =
(number of bits in common between A and B)/
(no of bits in molecule A + no of bits in molecule B - no of bits in common between A and B)

- There are a number of tools both freely available and commercially available to facilitate this type of searching.

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Analogue-based approaches: Search engines

- **AIM:** US EPA's Analog Identification Methodology identifies publicly available, experimental toxicity data on closely related chemical structures. Contains **31,031** records.
- **AMBIT (IDEA Consult Ltd):** Tool for defining applicability domain of QSAR models. Contains **463,426** records.
- **ChemFinder:** Portal of free and subscription scientific databases.
- **ChemID Plus Advanced:** Free database from the National Library of Medicine. Contains over **379,000** records.
- **Danish (Q)SAR Database:** Free internet-accessible version of a QSAR database originally developed by DK EPA. Contains **166,000** records.
- **Leadscope:** Commercial tool – databases + (Q)SAR functionalities

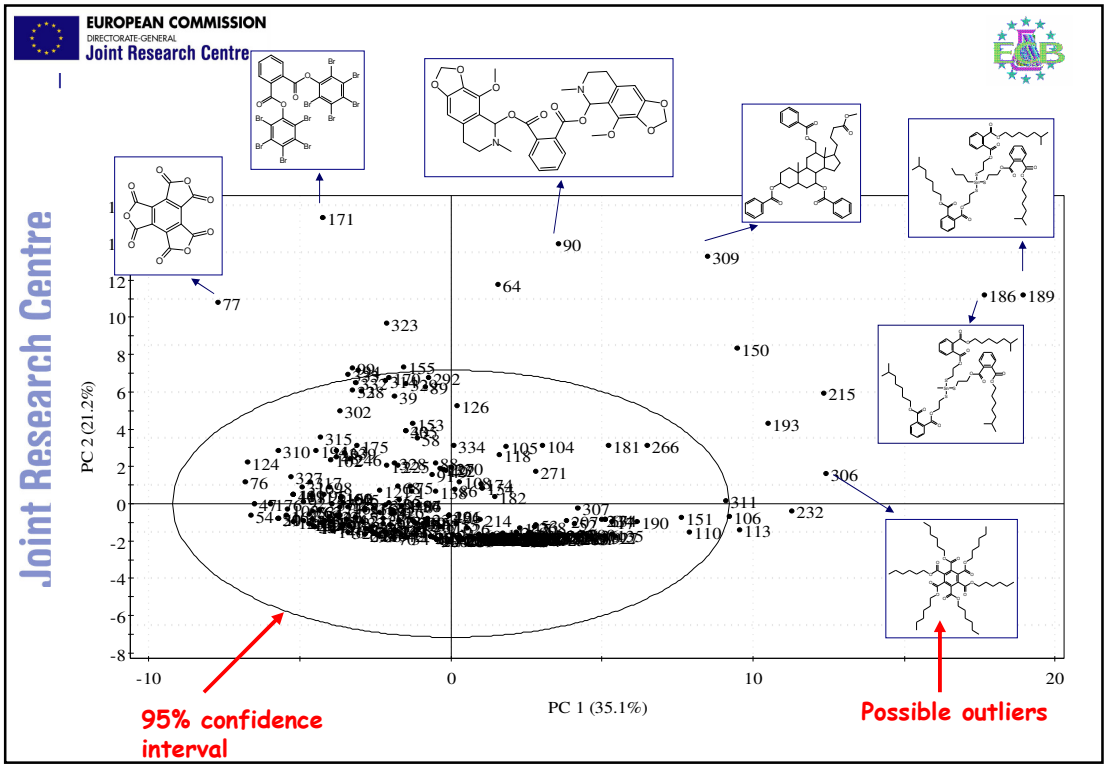
ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Unsupervised

Joint Research Centre

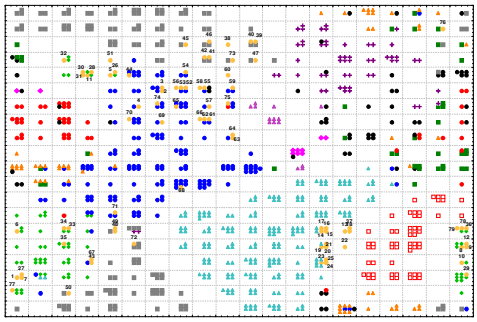
- Unsupervised approaches
- Based on independent variables (descriptors) only
- Methods involve the use of statistical techniques e.g. principal components analysis (PCA), clustering, self organising maps
- The approach relies on a starting dataset/inventory of chemicals and computing different numerical parameters (such as geometrical, topological, structural, physicochemical, electronic descriptors) for those chemicals or characterising them through the use of fingerprints or structural features.



EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

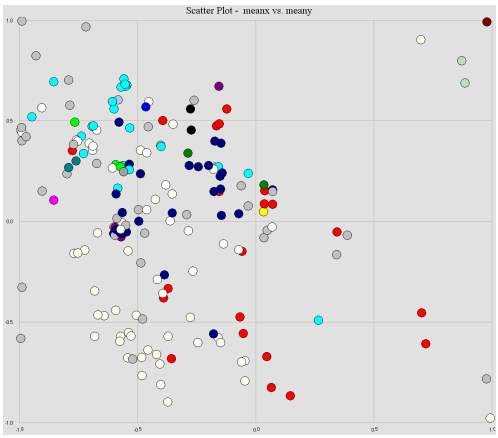
Examples of SOMs



- Aromatics
- Diox+Furans
- PAHs
- PCBs
- Pest. various
- Aliphatics
- List6
- OrgCl
- OrgPh
- Phenylureas
- Triazines
- Carbammates
- Dinitrobenzenes

With permission from M. Pavan

Grouping using descriptors thought to be relevant for the endpoint of concern



EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Supervised learning

- Many of the same statistical techniques that might be applied.
- The key difference is that information about the activity/toxicity of chemicals is taken into account in addition to the structural/descriptor information
- For example
  - Clustering techniques employed with the criteria that clusters are extracted to discriminate for the toxicity present.
  - Recursive partitioning and simulated annealing which aims to find active or statistically correlated subsets based on the presence or absence of a particular combination of substructural features.

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

ECB contributions: REACH work on Categories

- Application of QSAR and ranking methods to organic chemicals
- Application of ranking methods to organic chemicals
- Investigation of similarity-based grouping methods
- Review of examples discussed in EU Working groups

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

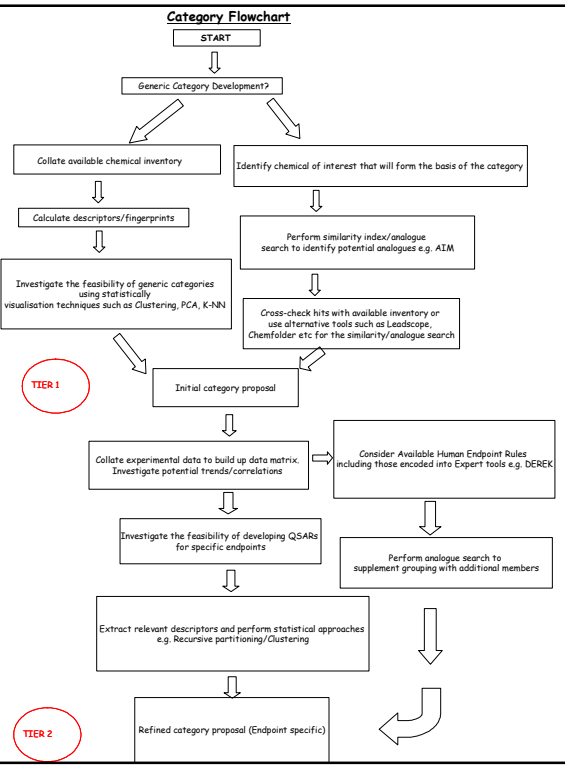
ECB contributions: Other

- ECB Tool for Chemical Similarity
  - ECB contract with IdeaConsult Ltd, Sofia, Bulgaria
    - to facilitate the development of generic (atom environments) and endpoint-specific categories
    - inclusion of functionality for read-across
    - prototype (due end year) is being developed using several endpoints as examples including skin sensitization, skin irritation, aquatic toxicity and bioaccumulation

ihcp

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre



EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Joint Research Centre

Challenges?

- Guidance is in development for regulatory use to provide more practical guidance on how to formulate categories
- BUT More work needs to be done to characterize chemicals appropriately for the formation of groupings i.e. what are the drivers for a particular endpoint
  - Gaining the mechanistic understanding to underpin each endpoint
    - Being able to characterize reactivity
      - Through extracting insights from properly structured toxicity databases
      - Through appropriate experimentation where necessary

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Acknowledgements

Joint Research Centre

Dave Roberts
Nora Aptula
Terry Schultz
Gil Veith
Chihae Yang
Glenn Myatt
Manuela Pavan
Nina Jeliazkova
Joanna Jaworska
Ana Gallegos
Aldo Benigni
David Basketter
Ovanes Mekenyan