

Characterisation of the Chemical Space of the COSMOS Cosmetics Inventory and the COSMOS Non-Cancer TTC Dataset



M Pavan¹, MTD Cronin², SJ Enoch², M Fuart-Gatnik³, J Gasteiger⁴, P Alov⁵, I Pajeva⁵,
E Fioravanzo¹, CH Schwab⁴, L Terfloth⁴, I Tsakovska⁵, AP Worth³, C Yang⁶

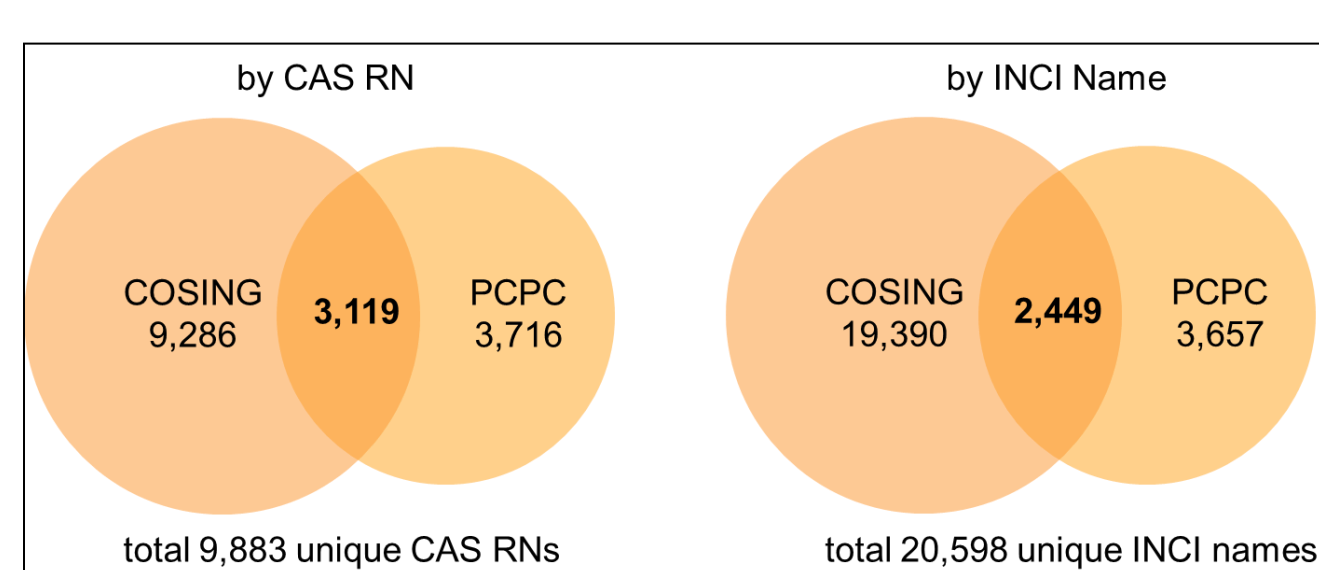
manuela.pavan@s-in.it

¹ S-IN Soluzioni Informatiche SRL, Vicenza, Italy, ² Liverpool John Moores University, England, ³ European Commission – Joint Research Centre, IHCP, Systems Toxicology Unit, Ispra, Italy, ⁴ Molecular Networks GmbH, Erlangen, Germany, ⁵ Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Sofia, Bulgaria, ⁶ Altamira LLC, Columbus, OH, USA

Introduction and Aims

COSMOS datasets:

- COSMOS non-cancer Threshold of Toxicological Concern (TTC) dataset** (version 1.0), containing 558 repeated-dose toxicity data for cosmetic ingredients; a complete subset of the **COSMOS Inventory**
- COSMOS Cosmetics Inventory** (version 1.0), compiled from the EU CosIng and US Personal Care Products Council (PCPC) lists and made of 4460 structures.



Objectives:

- Exploring the **applicability of the TTC approach to cosmetics**, by comparing the chemical space of the two COSMOS datasets with the one of the **Munro non-cancer dataset**¹, i.e. the standard non cancer TTC dataset.
- Prove the capability of the **COSMOS TTC dataset to represent the chemical space of cosmetics in general**, i.e. **Cosmetics Inventory**

Methods

Descriptors employed:

1. Structure and subgraph features

- SMARTS and RDK searching in KNIME²
- MOSES³ subgraph features

grouped by types of atom, bond, ring, functions and connectivity. US FDA CFSAN features coded in the CSRML format⁴

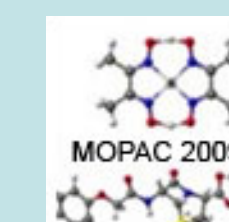
2. Physical/chemical property descriptors

size: molecular weight
shape: diameter, number of rot. bonds
partitioning behaviour: logP
solubility: logS
general characteristics: H acceptors, H donors



ADRIANA.Code⁵

reactivity: dipole, HOMO, LUMO energies, electronegativity, hardness, softness and electrophilicity

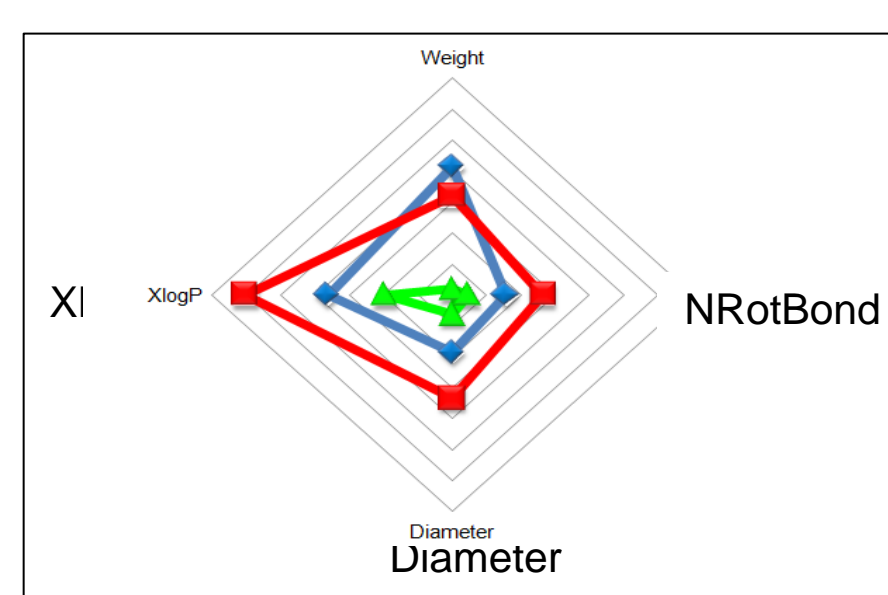


MOPAC⁶

This analysis is considered preliminary and further analysis is planned to better highlight the dissimilarities among the datasets. The outcome of the datasets comparison based on structural categories is illustrated in the **WP2 poster (C Yang et al)**.

Results

1. Applicability of the TTC approach to cosmetics :



BLUE = Munro; GREEN = COSMOS;
RED = Cosmetics Inventory

MW:

Munro, Cosmetic Inv. > COSMOS TTC

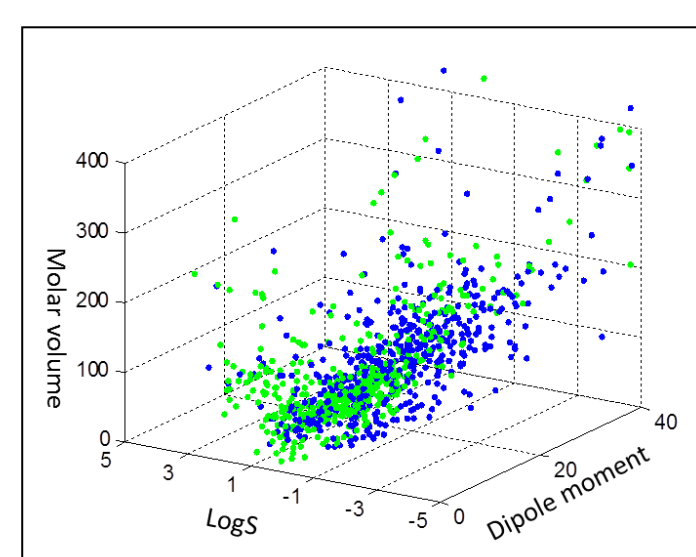
Long linear chains structures:

Cosmetics Inv. > Munro, COSMOS TTC

Lipophilicity:

Cosmetics Inv. > Munro, COSMOS TTC

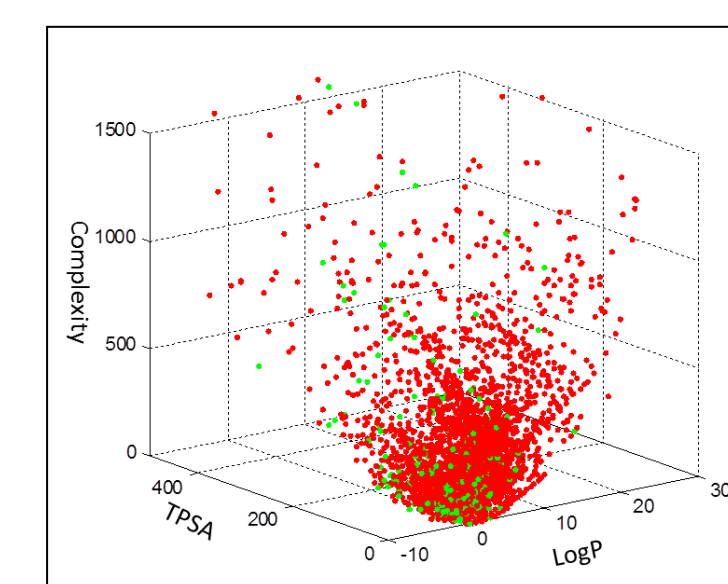
Overlap of the **Munro** and **COSMOS** datasets according to molar volume, solubility and dipole moment.



Next step: molecular descriptors and fingerprints analysis

Results

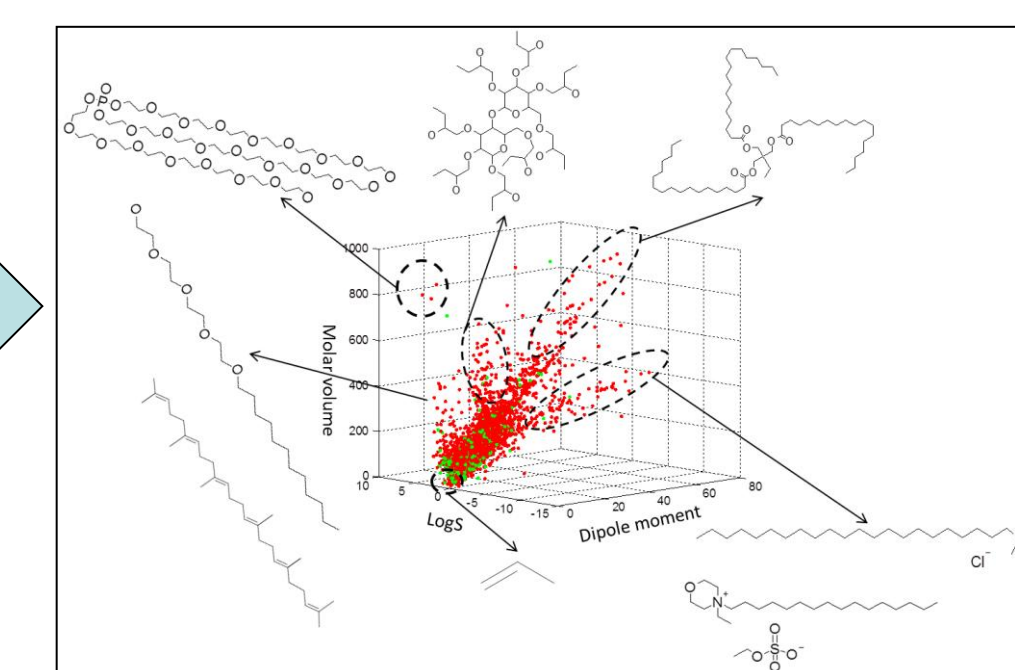
2. COSMOS TTC representativeness of the Cosmetics Inventory:



The **COSMOS TTC dataset** is representative of the chemical space of cosmetics in general, i.e. **Cosmetics Inventory**.

Clusters of cosmetics ingredients can be found based on logS, dipole moment, and molar volume.

Next step: characterisation of the clusters



Conclusions

- The **COSMOS TTC dataset** showed a good representation in terms of physicochemical property ranges of the **Cosmetics Inventory**.
- The **COSMOS TTC dataset** was considered to be suitable for investigating the applicability of the TTC approach to cosmetics.
- The analysis will be implemented in **KNIME workflows**.
- Additional descriptors such as molecular descriptors and fingerprints as well as other statistical approaches will be employed.
- Cluster analysis and cluster characterisation will be performed.

References

- Munro IC, Ford RA, Kennepohl E, Sprenger JG 1(996) *Food and Chem Toxicol.* 34: 829 – 867
- <http://www.knime.org/>
- MOSES software, Molecular Networks GmbH
- Chemical Substructure Representation Mark-up Language (CSRML) <http://bulletin.acscinf.org/node/224#W7>
- ADRIANA.Code software, Molecular Networks GmbH, version 2.2.4
- MOPAC2009, Stewart JJP, Stewart Computational Chemistry, Colorado Springs, CO, USA 2009.

Acknowledgements

The authors are grateful to the US Food and Drug Administration (FDA) for the oral repeated dose toxicity data provided for the cosmetic ingredients.

The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement n° 266835 and from Colipa.

