A Case Study to Leverage Public and Commercial Resources to Improve In Silico Chemical Safety Assessment

Mostrag A,[1] Arvidson K,[2] Cronin M,[3] Ebbrell D,[3] Enoch S,[3] Firman J,[3] Madden J,[3] Marusczyk J,[1] Volarath P,[2] Rathman J,[1] Schwab C,[1] Yang C[1]*

[1] MN-AM, Nurnberg Germany and Columbus OH USA [2] US FDA, Center for Food Safety and Applied Nutrition, College Park, MD USA [3] LJMU, Liverpool John Moores University, Liverpool, UK

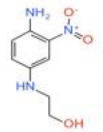
http://www.ng.cosmosdb.eu/

Motivation

- **COSMOS Project (2011-2015):** part of SEURAT-1 cluster
- **COSMOS DB:** >3000 registered users (>70% industry)
  - Cosmetics inventory, regulatory inventories
  - Assessment and in vitro/in vivo toxicity data
- **COSMOS Next Generation (NG) (https://ng.cosmosdb.eu/)**
  - Developed from COSMOS DB as 1st step towards **knowledge hub** for chemical safety evaluation leveraging public data systems
  - Database and in silico tools
  - Gateway for the commercial ChemTunes.ToxGPS (https://www.mn-am.com/) to publicly share data and new methods with industry and regulatory agencies
  - Assessment and toxicity data from EFSA, ECHA, US FDA covering cosmetics ingredients, food additives, agrochemicals, their metabolites, and impurities
- **Cosmetics Europe (CE)** joined the commercial-public resource sharing effort - COSMOS NG captures workflows based on CE case studies for TTC and read-across utilizing in silico, biokinetics, metabolism results, and high throughput and high content mode-of-action and toxicity studies
- We present an **ab initio workflow based on case study experiences**

Case Study

TARGET:
- **HC RED NO. 7**
- **Systemic toxicity**

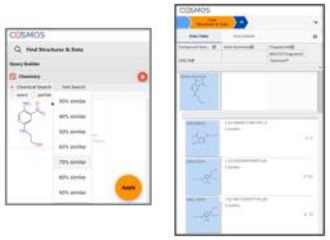


- Hair dying agent
- CMS-23938
- CAS 24905-87-1

Download **Guided Read-Across Workflow Using COSMOS NG @** https://ng.cosmosdb.eu/

- *Yang et al. 2021. Comp Tox, Submitted*

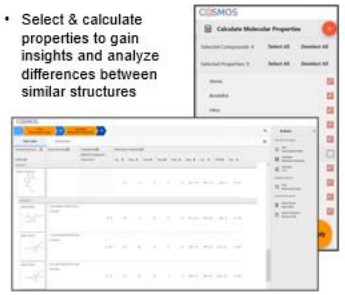
1 Step 1: Similarity Searching & Selection of Similar Structures



- Structure search for similarity >70% in COSMOS NG (RDKit Topological fingrtprints & Tanimoto coefficients)
- **Selected 3 potential analog candidates with systemic toxicity data: CMS- 60520, CMS-43204, CMS-72054**

2 Step 2: Calculating Properties Within COSMOS NG Data Table

- Select & calculate properties to gain insights and analyze differences between similar structures



- **Export results into XLS table**

3 Step 3: Analyzing Structure- & Property-Based Similarities & Analogue Quality

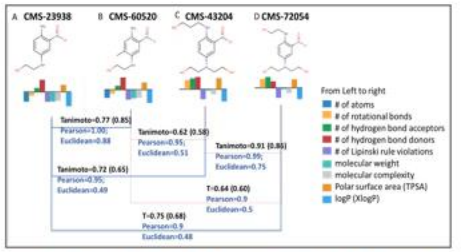
Structure-based similarity
- RDKit topological fingerprints, ToxPrints, public MACCS Keys
- Tanimoto coeff. calculated for each analog candidate against a target structure

Property-based similarity
- Pearson Correlation
- Euclidean Distance

Analog Quality (AQ)
- Combination of both structure and property-based similarities

$$AQ = \sqrt[N]{\prod_{i=1}^{N}(SimilarityMeasure)_i}$$

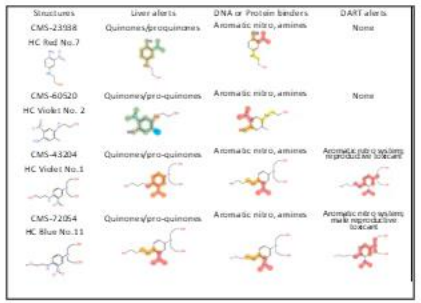
Tanimoto coeff based on RDKit Topological fingerprints (https://www.rdkit.org/) and (in parentheses) on ToxPrints (www.toxprint.org)



- **CMS-43204 & CMS-72054 are less similar to the target by all similarity measures**

4 Step 4: Chemical Profiling by Categories & Pathways

- Chemotype Matching by Chemotype Profilers in COSMOS NG



- CMS-43204 & CMS-72054: aromatic nitro system matching chemotype alerts for male reproductive toxicants – not present in the target and the closest structural analogue (this argues against considering these two compounds as analogues of the target)
- **CMS-60520 selected as a final analog**

5 Step 5: Attaching Appropriate Toxicity Endpoint Data & Study Reliability

- Find toxicity data and COSMOS MINIS Grade for analogs in COSMOS NG **(=> 90-d rat gavage study from SCCS; MINIS Grade = 4)**
- Consult expert toxicologist for Opinion Score **(=> Opinion Score = 4)**
- Calculate the Study Reliability Likelihood based on the above **(=> 0.9)**

Study Quality
- **COSMOS MINIS Grade (1-5)**
  - *Does study design give enough parameters to be considered "usable" in safety assessment?*
  - *Are experimental results supported by study design?*
- **Opinion Score (1-5)**
  - *Are conclusions deemed "reliable/ believable/ interpretable" by regulators or domain experts?*

Study Reliability
- **Study Reliability Likelihood**:
  - Composite reliability score based on MINIS Grade and Opinion Score
  - Defined as a likelihood
  - Amenable for probabilistic treatment when used for weight of evidence combinations

- *Yang et al. 2021. Comp Tox, Submitted*
- *Yang et al. 2017. Food Chem Toxicol 109:170-193, http://dx.doi.org/10.1016/j.fct.2017.08.043*

6 Step 6: Setting-Up An Evidence Table for Read-Across Assessment

- Compile each piece of evidence in Weight of Evidence (WoE) table including quantitative measures to derive final read-across (RA) reliability
- In COSMOS NG – simple joint probability between AQ (83% based on Euclidean distance) and Study Reliability Likelihood (90%) **(=> 0.75)**
- In commercial system – Dempster Shafer Theory (DST) combination available **(=> 0.79-0.96 with uncertainty of 17%)**
- For the case of one analogue/one study, simple joint probability may be a quick indicator, although it yields the most conservative estimation

- *Rathman et al. 2018. Comp Toxicol 6:16-31, https://doi.org/10.1016/j.comtox.2018.03.001*

7 Step 7: NOAEL Estimation With Data from COSMOS NG

- Nearest Neighbor analysis (structure-based similarity) using a dataset of 1538 compounds with qualified NO(A)EL/LO(A)EL values provided within COSMOS NG 17 similar structures with aromatic amine/nitro system for hair dyes
  - 11 when removing the structures hitting the reproductive toxicant profile (see Step 4).
- The NOAEL bounds at 95% were estimated (by assuming normal distribution) to be **14-33 mg/kg-bw/day**

- *Yang et al. 2021. Comp Toxicol, Submitted*
- *Yang et al. 2021. Chem Res Toxicol 34(2): 616-633, https://doi.org/10.1021/acs.chemrestox.0c00429*
- *Yang et al. 2020. Food Chem Toxicol 143:111561, https://doi.org/10.1016/j.fct.2020.111561*

Evidence	CMS-23938	CMS-60520
Compound Role	Target	Analogue
Structure-Based Similarity		
RDKit Tanimoto Coeff.	1	0.77
ToxPrint Tanimoto Coeff.	1	0.85
Property-Based Similarity		
Euclidean similarity	1	0.88
Pearson similarity	1	0.999
Analogue Quality (AQ)		
Euclidean similarity		0.83
Pearson similarity		0.87
Chemical Category Profile		
Liver	YES	YES
DNA binders	YES	YES
Developmental/Reproductive	NO	NO
Toxicity Data		
Study Design	Assumed no data	90-day SD rat oral-gavage/intubation 50-500 mg/kg-bw/day
Study Results		NOAEL=17 mkd / LOAEL=50 mkd (clin signs; clin chem; organ weight dec; liver pathology at higher dose)
Study Source	N/A	SCCS
Study Quality: COSMOS MINIS	N/A	4
Study Quality: Opinion Score	N/A	4
Study Reliability Likelihood	N/A	0.9
Read-Across Reliability		
Simple Estimation		0.75
Dempster-Shafer Rules (only in commercial system)		*0.79-0.96*
Read-Across Endpoint Value		
NOAEL bounds (at 95% confidence interval)	14-33 mkd	

Conclusions

- **COSMOS NG**
  - Public web-based searching/retrieval system and forum for sharing resources, models and supporting workflow developments
  - Database (regulatory inventories, assessment and toxicity data)
  - COSMOS MINIS Grade and expert Opinion Score to evaluate the reliability of tox data
  - Fingerprinting schemes for structure-based similarity and molecular properties for property-based similarity
  - Action tools for compiling the data for the read-across evidence table
- **Case study** illustrates utilitarian value of COSMOS DB and NG as part of the read-across workflow

Perspectives

- **Future data updates**: HESS database from Japanese NITE, EFSA OpenFoodTox, US EPA ToxRefDB, Cosmetics Safety data from KCII, and CE
- **Continued collaboration** between public and commercial entities - expanding the knowledge domain, increasing reliability, possibility of sharing workflows with analysis of safety/risk assessment across multiple entities

- **Exhibitor Hosted Session "Novel Chemoinformatics platform ChemTunes.ToxGPS and its public version COSMOS Next Generation"**
  Tuesday, March 23rd, 4-5 pm Eastern time
- **ToxExpo Booth 2052**

ACKNOWLEDGEMENT: This work is funded through the Long Range Science Strategy (LRSS), a program of Cosmetics Europe, with participation of industrial members

Contact: chihae.yang@mn-am.com