Estimating EC3 (Effective Concentration for a Stimulation Index of Three) Bounds and Uncertainties for Skin Sensitization Based on Similar Structures and Assay Profiles

J Rathman,[1,2] C Yang,[1,2] M Cronin,[3] S Enoch,[3]

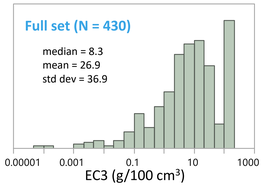
[1] MN-AM, Nurnberg Germany and Columbus OH USA, [2]Ohio State University, Columbus OH USA, [3]Liverpool John Moores University, Liverpool, UK

1 OBJECTIVES

- When reliable toxicity data for a target substance are not available to address its skin sensitization potential, QSARs and Adverse Outcome Pathway (AOP) based assays are known to reasonably identify the hazard and potency categories. However, predicting the continuous-valued effective concentration for a stimulation index of three (EC3) is more challenging.
- To estimate confidence bounds for EC3, we apply here a method similar to one recently published to estimate NOAEL bounds and uncertainties based on qualified analogues in a read-across workflow from a well-curated NOAEL database.[1]
- The chemical categories in skin sensitization are clearly defined due to their well-understood chemical reactivity with proteins and the roles in the induction of the immune response. EC3 profiles for structures from a given reactivity class (e.g., SN2, SNAr, acylating agents, Michael acceptors, and Schiff base formers) were evaluated using the ToxPrint chemotypes and then the $EC3_{Diff}$ was calculated between the pairs of nearest neighbors.
- The analogue quality was estimated by taking structural similarity, physico-chemical properties, and AOP assays (e.g., KeratinoSense™ and h-CLAT) profiles into account. This analysis allowed us to estimate bounds for EC3 values at a confidence level acceptable to users, where the width of the interval reflects the uncertainty.

2 EC3 DATASET

- EC3 is the concentration (g per 100 cm^3 vehicle) observed to give a three-fold stimulation index in the murine local lymph node assay (LLNA) for skin sensitization potency.[2]
- Dataset of 430 structures compiled from
  - SCCS/SCCP/SCCNFP[3]
  - OECD QSAR ToolBox
  - ICCVAM/NICEATM[4]
  - ECHA Registered Substance Database
  - Open literature publications[5,6]



Literature cited

[1] Yang, et al. *Chem Res Toxicol* 2021, *34*, 616-633.

[2] Basketter, et al. *Contact Derm* 2007, *57,* 70–75.

[3] https://ec.europa.eu/health/scientific_committees/consumer_safety_en

[4] https://ntp.niehs.nih.gov/whatwestudy/niceatm/test-method-evaluations/skin-sens/index.html

[5] Jaworska, et al. *Arch Toxicol* 2015, *89,* 2355–2383.

[6] Urbisch, et al. Reg Tox Pharm 2015, *71*, 337-351.

[7] C. Yang et al. *J Chem Inf Model* 2015, 55, 510–52.

3 HYPOTHESIS

- The EC3 of a target molecule of interest can be estimated by identifying appropriate analogues for which experimental EC3 values are available.
- The ***analogue quality*** score, which considers both structure- and property-based similarity as well as mechanistically-relevant structure categories, can be used to identify appropriate, robust and justifiable analogues for read-across.
- The uncertainty in estimated EC3 value for the target varies inversely with analogue quality.

4 DESCRIPTORS

- **ToxPrint Fingerprints (ToxPrints)**[7]
  - https://chemotyper.org/
  - Molecular fragments encoding connectivity and topology, and properties of atoms, bonds and electronic systems.
  - Mechanistically designed for chemical reactivity and toxicity.
  - Features designed to reflect categories.
- **Properties**
  - ChemTunes.ToxGPS
  - ***Physicochemical:*** # atoms, # rotatable bonds, # H-bond donors and acceptors, # Lipinski Rule of 5 violations, complexity, topological polar surface area, log P
  - ***Quantum mechanical:*** heat of formation, HOMO energy, LUMO energy, HOMO-LUMO gap

5 PAIRWISE $EC3_{Diff}$ and Analogue Quality

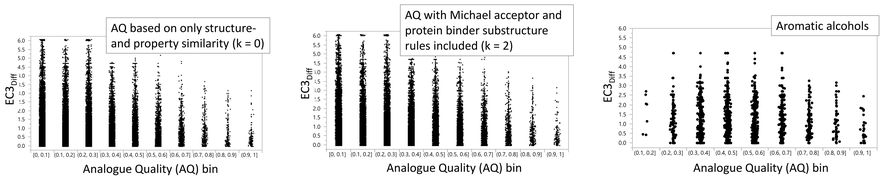
- The absolute difference between the EC3 values for structures in each pair is quantified by

$$EC3_{Diff} = \left|\log_{10}(EC3_2) - \log_{10}(EC3_1)\right| = \left|\log_{10}\left(\frac{EC3_2}{EC3_1}\right)\right|$$

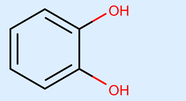
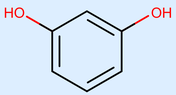
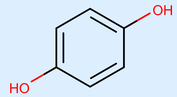
- Analogue quality (AQ) is defined as the geometric mean of the pairwise structure-based similarity (Tanimoto), property-based similarity (scaled Pearson correlation coefficient), and the number (k) of common chemotype profile skin sensitization rules.

$$AQ = \left((Tanimoto)(Pearson+1)/2\right)^{1/(2+k)}$$

- EC3Diff distributions support the hypothesis that structure pairs having a high AQ tend to have similar EC3 values. Including substructure rules increases number of high-AQ pairs.
- $EC3_{Diff}$ can also be reduced by identifying high-AQ pairs within specific chemical category subsets (e.g., aromatic alcohols).



6 CASE STUDY

Target	Analogue 1	Analogue 2
 Catechol		
Analogue Quality (AQ)	0.999	0.995
EC3 (g/100 cm^3)	6.3	0.1

EC3 BOUNDS ESTIMATION

- ***Method 1:*** EC3 bounds are estimated by analyzing the quantiles in the [0.9, 1] AQ bin of the $EC3_{Diff}$ vs. AQ plot for aromatic alcohols. A confidence interval for EC3 is calculated from the 95% quantile.

	Analogue 1	Analogue 2
EC3Diff @ 95% confidence:	$\left\lvert\log_{10}\left(\frac{EC3_{Target}}{EC3_{Analogue}}\right)\right\rvert \le 0.72$	
95% confidence bound for Target EC3	1.2 to 33 g/100 cm^3	0.02 to 0.52 g/100 cm^3

- ***Method 2:*** Distribution of $\log_{10}$(EC3) values for nearest neighbors based on high AQ scores against target is fit to a normal distribution and confidence interval is estimated. There were 13 neighbors in this group and the resulting CI is:

95% confidence bound for Target EC3: 0.60 to 2.9 g/100 cm^3

7 CONCLUSIONS

- EC3 bounds for a target compound are estimated at a desired confidence level based on EC3 values from one or more qualified analogues.
- Analogues are selected based on AQ scores, which take structure and property similarities into account, and also by extracting subsets from the EC3 dataset based on chemical categories relevant to the target and analogues.

ACKNOWLEDGEMENT: Petra Kern, Els Deconinck, Greg Dameron, Cindy Ryan at Procter & Gamble

Contacts: james.rathman@mn-am.com, chihae.yang@mn-am.com