Testing Strategies for UN GHS Classification for Serious Eye Damage/Eye Irritation of Chemicals: Cosmetics Europe Analysis.

Nathalie Alépée[1], Els Adriaens[2], Daniel Bagley[3], Bertrand Desprez[4], Jalila Hibatallah[5], Karsten R. Mewes[6], Uwe Pfannenbecker[7], Àlvar Sala[8], Pauline McNamee[9]

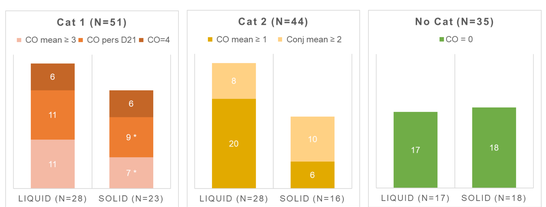
[1]L'Oréal Research & Innovation, Aulnay Sous Bois, France, [2]Adriaens Consulting, Aalter, Belgium, [3]Colgate-Palmolive, Piscataway, New Jersey, United States, [4]Cosmetics Europe - The Personal Care Association, Brussels, Belgium, [5]Chanel Parfums Beauté, Neuilly sur Seine, France, [6]Henkel AG & Co. KGaA, Düsseldorf, Germany, [7]Beiersdorf AG, Hamburg, Germany; [8]Kao Corporation, S.A., Barcelona, Spain, [9]The Procter & Gamble Company, Egham, United Kingdom

Introduction

A core part of the Cosmetics Europe (CE) eye programme focuses on data integration/evaluation of testing strategies/approaches for identification of serious eye damage/eye irritation of chemicals that can be advocated for external/regulatory acceptance. To enable this, CE curated an initial database of chemicals for which *in vivo* and partial *in vitro* data exist. This database was used for selection of 80 chemicals tested in *in vitro* test methods in the CEFIC CON4EI project. After integration of all *in vitro* data on an industry platform level, remaining data gaps were identified. CE completed *in vitro* testing to fill these data gaps resulting in a comprehensive *in vivo/in vitro* database of more than 110 chemicals to date. Building on proposed CON4EI testing strategies, CE has analysed the comprehensive database to determine the robustness of such testing strategies and to identify where opportunities exist for refinement.

Materials and methods

A set of 110 up to 130 chemicals was tested with seven test methods. Distribution of the 130 chemicals (73 liquids & 57 solids) according to the Drivers of Classification as defined by Barroso et al. (2017; definition see poster abstract #516) is shown in the barplot.



* Two Draize eye studies for this Cat 1 solid (CO mean ≥ 3; CO pers D21)

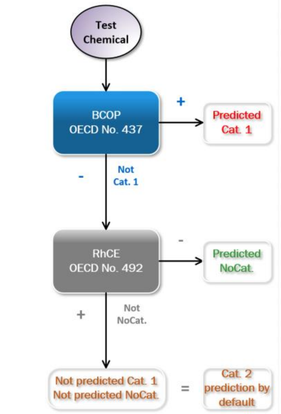
In order to identify which test method (original or optimized) was most suitable as a fist step in a Top-Down (identification of Cat 1) or Bottom-Up (identification of No Cat) approach (Scott et al., 2010), the performance of each individual test method was assessed by comparing the prediction results with the existing proposed UN GHS classification (data not shown). Next, the most promising test methods were combined into a testing strategy. The performance of the testing strategy was evaluated in terms of correct predictions and the predictive capacity (predictive value).

Results & Discussion

Performance of two step Top-Down testing strategy (TS)

OECD TG test methods

Fig 1



Endpoints BCOP: Opacity + 15 x Permeability (OD) = IVIS

IVIS cut-off values 3 / 55

Endpoint RhCE: Viability (%)

Performance (n=130) (bootstrap predictions)

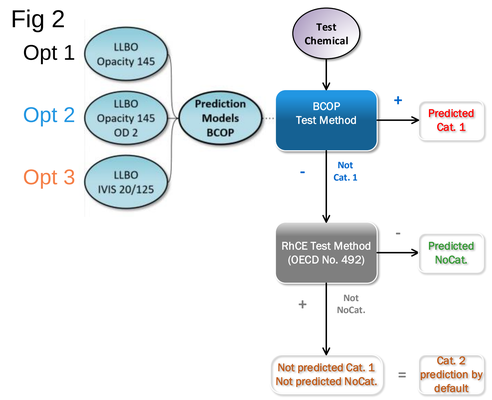
UN GHS (Draize) / UN GHS (In-vitro TS)	Cat 1	Cat 2	No Cat
Cat 1 (N = 51)	≈ 62%	≈ 38%	< 1%
Cat 2 (N = 44)	≈ 34%	≈ 60%	≈ 6%
No Cat (N = 35)	0%	≈ 28%	≈ 72%

Accuracy: 63 - 64%

- None of the No Cat predicted Cat 1
- Cat 2 over-prediction (OP) rate ± 34%
- Cat 1 under-prediction (UP) rate ± 38%
- Low false negative rate: < 1% Cat 1; ± 6% for Cat 2

BCOP LLBO/RhCE test methods

Fig 2



Performance (n=111) (bootstrap predictions)

UN GHS (Draize) / UN GHS (In-vitro TS)	Cat 1	Cat 2	No Cat
Cat 1 (N = 46)	**65%** **77%** **79%**	35% 23% 19%	0% 0% 2%
Cat 2 (N = 35)	26% 36% 50%	**64%** **56%** **40%**	10% 8% 10%
No Cat (N = 30)	0% 0% 0%	30% 30% 58%	**71%** **71%** **42%**

Accuracy for 2 strategies and the BCOP LLBO stand-alone:

BCOP LLBO (Op: 145) / RhCE: 65.3% - 66.4%

BCOP LLBO (Op: 145 & OD: 2) / RhCE: 67.4% - 69.3%

BCOP LLBO stand-alone (IVIS: 20/125): 56.8%

- None of the No Cat predicted as Cat 1
- Cat 2 OP rate 36% and Cat 2 FNR = 8%
- Cat 1 FNR = 0% and Cat 1 UP only 23% (added value over OECD TG TS)

Predictive capacity of the TS

The predictive value (PV) of a TS is influenced by (1) the prevalence (true distribution of eye effects) in a specific population of chemicals and (2) the accuracy of the TS. The values presented in the Table are based on the accuracy of TS Fig 2, Opt 2.

Top 1st test method
BCOP LLBO (Opacity/Permeability: 145/2)
= 77% sensitivity

Bottom 2nd test method
OECD TG 492 RhCE
= 71% specificity

Assuming a random selection of 100 chemicals from a population with prevalence of outcomes distributed according to substances tested with the OECD TG 405 in REACH registrations (2008-2014) (Luechtefeld et al., 2017) and applying TS Fig 2 Opt 2, the following predictive values are obtained.

UN GHS	Draize Eye test Prevalence (REACH registrations 2008-2014)			Testing Strategy
	Cat 1 = 10%	Cat 2 = 16%	No Cat = 74%	
Cat 1	**57%** **(N=8)**	43% (N=6)	0%	**14%**
Cat 2	7% (N=2)	**28%** **(N=9)**	65% (N=21)	**32%**
No Cat	0% (N=0)	2% (N=1)	**98%** **(N=53)**	**54%**

- Negative PV: 98% of the test outcomes are correct (correspond with an *in vivo* No Cat)
- Cat 1 and Cat 2 PVs: 57% and 28% of test outcomes are correct, respectively
- Overall, under predictions will be very low
- The TS tends to result in over-classification, *in vivo* Cat 2 predicted as Cat 1 and *in vivo* No Cat predicted as Cat 2

Conclusion

TS performs better than a stand-alone method. Furthermore, under-predictions were often related to low water solubility and over-predictions were more often related to the "Cat 2 CO mean ≥ 1" Driver of *in vivo* classification.

TS tends to over-predict (regarding predictive value), a possible solution could be to take into account physico-chemical properties (e.g. increase specificity of test method that can identify No Cat).

Draize rabbit eye test is known to over-predict versus humans and in turn test methods/TS tend to over-predict compared to Draize - what does this mean for prevalence in the long term?

Adriaens et al., 2014. Retrospective analysis of the Draize test for serious eye damage/eye irritation: importance of understanding the in vivo endpoints under UN GHS/EU CLP for the development and evaluation of in vitro test methods. Arch. Toxicol. 88,701-723.

Barroso et al., 2017. Cosmetics Europe compilation of historical serious eye damage/eye irritation *in vivo* data analysed by drivers of classification to support the selection of chemicals for development and evaluation of alternative methods/strategies: the Draize eye test Reference Database (DRD). Arch. Toxicol. 91, 521-547;

Luechtefeld et al., 2017. Analysis of Draize eye irritation testing and its prediction by mining publicly available 2008-2014 REACH data. ALTEX. 2016;33(2):123-34. doi: 10.14573/altex.1510053.

Scott et al., 2010. A proposed eye irritation testing strategy to reduce and replace in vivo studies using Bottom-Up and Top-Down approaches. Toxicol. In Vitro 24, 1-9.

Acknowledgements

The CON4EI project was funded by the European Chemical Industry Council (CEFIC)-Long range Research Initiative (LRI) – AIMT6 Eye irritation testing in vitro in practice – Database and testing strategy. The data of this project were kindly provided by Dr. An Van Rompay – VITO N.V., Mol, Belgium.