Mixture-based QSAR Models of Ocular Toxicity for Regulatory Hazard Categories

A Sedykh[1], N Choksi[2] , D Allen[2] , N Kleinstreuer[3] , W Casey[4], R Shah[1] | [1]Sciome LLC, [2]ILS , [3]NIH/NIEHS/DNTP/NICEATM, [4]NIH/NIEHS/DNTP, Research Triangle Park, NC,USA

Abstract ID 479:

For further information please contact **nchoksi@ils-inc.com**

Computational modeling can be used to design effective nonanimal approaches, if grounded in reliable experimental data. We have developed of a set of computational models to predict eye irritation and corrosion. The models were developed using a curated database of in vivo eye irritation studies from the scientific literature and stakeholder-provided data. The database contains over 500 unique substances, including many mixtures, tested at different concentrations. Substances were categorized according to Globally Harmonized System (GHS) and U.S. Environmental Protection Agency (EPA) hazard classifications. Two modeling approaches were used to predict classification of mixtures. A conventional approach generated predictions based on the chemical structure of the most prominent component of the mixture. A mixture-based approach used weighted feature averaging to consider all known components in the mixture. Ranking accuracy rates (calculated based on the area under the receiver operating curve) for EPA hazard classification of undiluted test substances were 74-81% and 75-80% for the conventional and mixture-based models, respectively. Ranking accuracy rates for EPA hazard classification of substances diluted to 10% in the conventional and mixture-based models were 90-95% and 92-96%, respectively. Ranking accuracy rates for GHS hazard classifications for undiluted test substances were 79-82% and 80-91% for the conventional and mixture-based models, respectively. Rates ranged from 89-95% for the diluted GHS classification predictions for both approaches. We observed a strong correlation between a substance's pH and activity. Our results suggest that these models are useful for screening compounds for eye irritation potential. Future efforts to increase the models' utility will focus on expanding their applicability domains and using them in conjunction with other input variables (e.g., in vitro data) to establish defined approaches for eye irritation testing.

Eye Toxicity Hazard Classifications

In vivo effect on eye tissues	EPA	GHS
Corrosive or not reversible in 21 days	Category I	Category 1
Irritation, reversible in 8-21 days	Category II	Category 2A
Irritation, reversible in 1 – 7 days	Category III	Category 2B
Minimal effects, disappearing in 24h	Category IV	No category

Concordance of EPA vs GHS calls across data records

	EPA categories				
	I	II	III	IV	No call
GHS Cat.1	135	3			56
GHS Cat.2A	3	29	10		36
GHS Cat.2B		3	37		8
GHS No Cat.		6	114	201	72
GHS No call	2	2	10	1	62

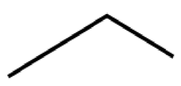
NICEATM Ocular Toxicity Data ("OCUTOXDB")

Contains **810** curated data records with *in vivo* ocular toxicity for **604** unique substances: **36**% of test substances have multiple data records either from different sources or test doses), **23%** of test substances are either salts or mixtures.

Data record examples:

Ethanol

CAS RN#: **64-17-5**



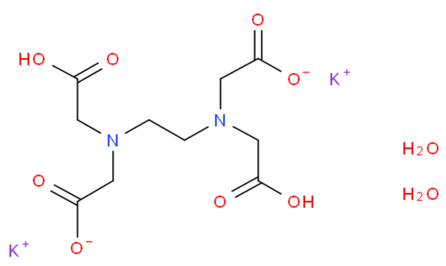
at 10% dose:
GHS: No Category
EPA: Category IV

at 79% dose:
GHS: Category 2B
EPA: Category III

at 100% dose:
GHS: Category 2A
EPA: Category I
EPA: Category II
EPA: Category III

EDTA, dipotassium

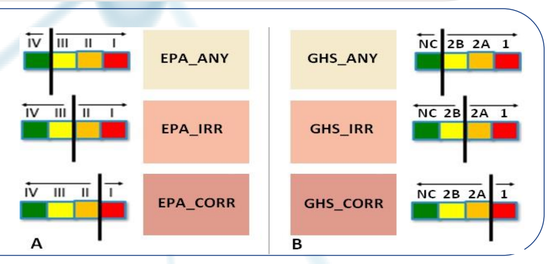
CAS RN#: **25102-12-9**



at 20% dose:
GHS: No category
EPA: Category III

Constructing OcuTox Datasets

For each classification system (EPA and GHS), 3 types of binary activity labels were formed (CORR, IRR, and ANY) at two dose levels (100% - "HI" and 10% - "LO").



Finalized ocular toxicity datasets and their composition

Dataset	Inactive conc. threshold	Active conc. threshold	Inactive	Active
EPA_CORR_HI	≥90%	≤100%	311	155
EPA_IRR_HI			258	184
EPA_ANY_HI			142	333
EPA_CORR_LO	>10% and <100%	≤10%	45	32
EPA_IRR_LO			39	35
EPA_ANY_LO			152	46
GHS_CORR_HI	≥90%	≤100%	330	152
GHS_IRR_HI			284	205
GHS_ANY_HI			261	230
GHS_CORR_LO	>10% and <100%	≤10%	49	32
GHS_IRR_LO			43	35
GHS_ANY_LO			282	38

Modeling details

814 chemical features (after removing redundant ones)

Mordred 2D descriptors, Chemotyper alerts, pH from ADMET Predictor

2 representation approaches of substances

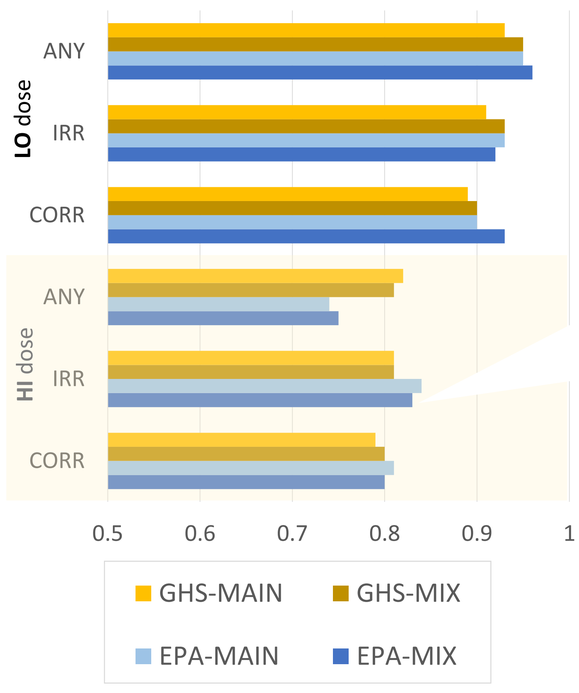
MAIN – largest chemical component (conventional way)

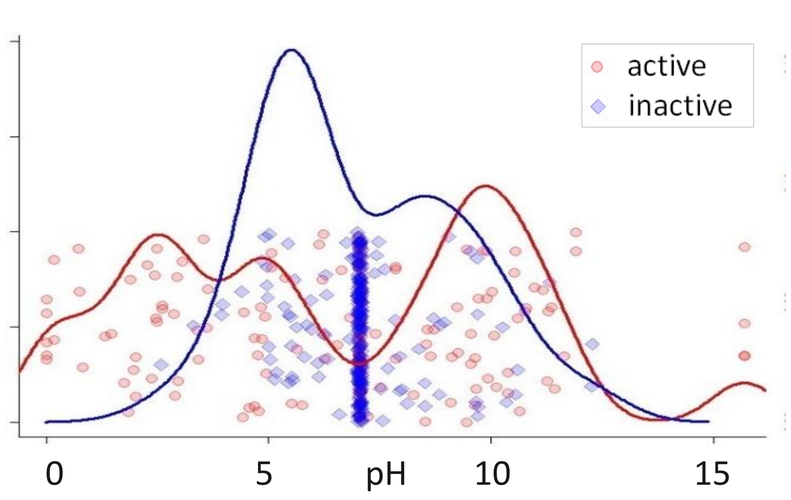
MIX – fraction-weighted average of features for all components

Machine learning methods (caret R package)

Random Forest, Support Vector Machines with radial kernel, Generalized Linear Models, Extremely Randomized Trees –applied in consensus.

Cross-validation accuracy (as ROC AUC) of ocular toxicity models





Distribution of calculated pH values among active (red) and inactive (blue) substances in EPA_IRR_HI (MIX) model.
Most of the inactive substances cluster in the middle, neutral pH area, while many actives are well spread to extreme pH values (strong acidic or alkali areas)

Discussion and Conclusions

- Mixture-based models slightly outperform conventional QSAR versions for low-dose datasets. For high-dose datasets such difference was insignificant.
- High-dose model performance (*i.e.*, EPA_IRR > EPA_CORR > EPA_ANY) may indicate stronger mechanistic relationship for the EPA activity binning schemes that are based on combining corrosives and severe irritants (US EPA Categories I-II). For the GHS datasets, validation results for all three endpoints were quite close.
- Comparing high-dose models between hazard classifications, they performed similarly for corrosives endpoint ("CORR"). For irritants ("IRR"), the EPA models were more accurate by 2-6%, and for all irritants ("ANY"), the GHS models were more accurate by 6-8%.
- On external ECHA data, models showed sensitivity similar to that in cross-validation, but specificity was 5-10% lower, even for GHS models. This overprediction (false positives) could be due to outdated toxicity records and/or missing test dose information.

External evaluation on ~700 ECHA substances (HI dose models)

To further evaluate our models, we gathered additional 700 substances with ocular toxicity data from ECHA dossiers. These are based on **GHS categories** and have no test dose information. Best performing models are highlighted red.

These ocular toxicity data for this ECHA test set is currently being re-examined by NICEATM to update hazard classifications.

	Model	Sensitivity	Specificity	PPR	NPR	Coverage
ANY	GHS_MAIN	0.69	0.68	0.47	0.84	0.82
	GHS_MIX	0.59	0.69	0.45	0.80	0.83
	EPA_MAIN	0.80	0.51	0.43	0.84	0.84
	EPA_MIX	0.70	0.63	0.44	0.83	0.78
IRR	GHS_MAIN	0.70	0.66	0.40	0.88	0.89
	GHS_MIX	0.75	0.67	0.42	0.90	0.79
	EPA_MAIN	0.73	0.60	0.38	0.87	0.86
	EPA_MIX	0.66	0.68	0.39	0.87	0.88
CORR	GHS_MAIN	0.55	0.75	0.23	0.92	0.86
	GHS_MIX	0.73	0.73	0.27	0.95	0.85
	EPA_MAIN	0.74	0.64	0.22	0.95	0.89
	EPA_MIX	0.79	0.65	0.24	0.96	0.90

Acknowledgements

This project was funded with federal funds from the National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN273201500010C.

Enabling Science via Analytical Informatics

Advancing Science, Improving Lives

NTP National Toxicology Program U.S. Department of Health and Human Services