Leadscope predictive models for acute oral systemic toxicity

Glenn J. Myatt,
Leadscope Inc., 1393 Dublin Road, Columbus, Ohio 43215
Email: gmyatt@leadscope.com

Leadscope®

Agenda

- How the models were built
- Model results
- *In silico* toxicology protocol project
  - Acute toxicity *in silico* protocol

Leadscope®

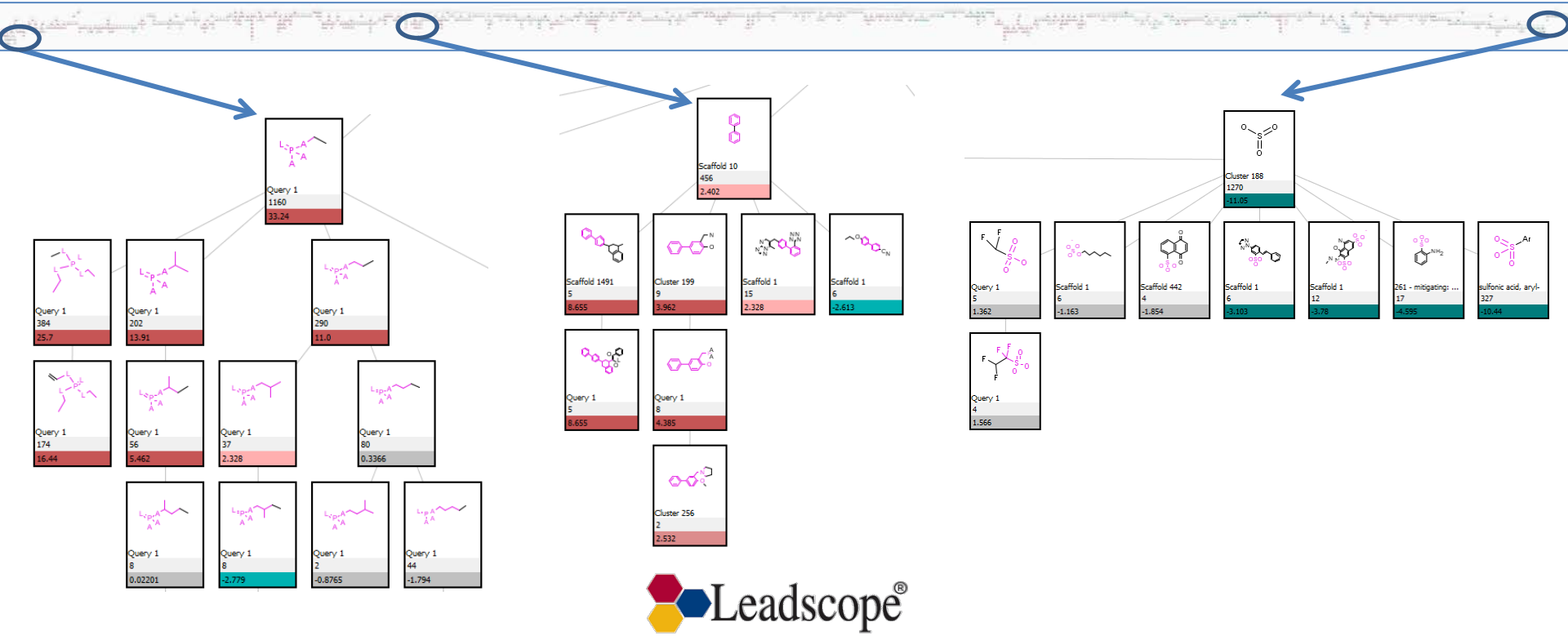
HOW THE MODELS WERE BUILT

Approach to modelling data

- Structure-based classification
- Developed a series of QSAR models and alerts
- Integrated results to generate final endpoints

Leadscope®

Structure classification-
830 structural classes organized hierarchically



Leadscope®

QSAR and alert models Built

- **-log(LD50) QSAR**: Negative log of the LD50 point estimate
- **EPA**
  - **EPA 1 QSAR**: 1 for all EPA hazard category 1, 0 for all other categories
  - **EPA 1 2 QSAR**: 1 for all EPA hazard category 1 or 2, 0 for all other categories
  - **EPA 1 2 3 QSAR**: 1 for all EPA hazard category 1 or 2 or 3, 0 for all other categories
- **GHS**
  - **GHS 1 QSAR**: 1 for all GHS hazard category 1, 0 for all other categories
  - **GHS 1 2 QSAR**: 1 for all GHS hazard category 1 or 2, 0 for all other categories
  - **GHS 1 Alerts**: 1 for all GHS hazard category 1, 0 for all other categories
  - **GHS 1 2 Alerts**: 1 for all GHS hazard category 1 or 2, 0 for all other categories
  - **GHS 1 2 3 QSAR**: 1 for all GHS hazard category 1 or 2 or 3, 0 for all other categories
  - **GHS 1 2 3 4 QSAR**: 1 for all GHS hazard category 1 or 2 or 3, 0 for all other categories
- **Very_toxic QSAR**: 1 for TRUE and 0 for FALSE
- **Nontoxic QSAR**: 1 for TRUE and 0 for FALSE

QSAR descriptors

- Leadscope pre-defined features
- 830 structural classes
- Calculated phys-chem properties

Leadscope®

QSAR modelling approaches

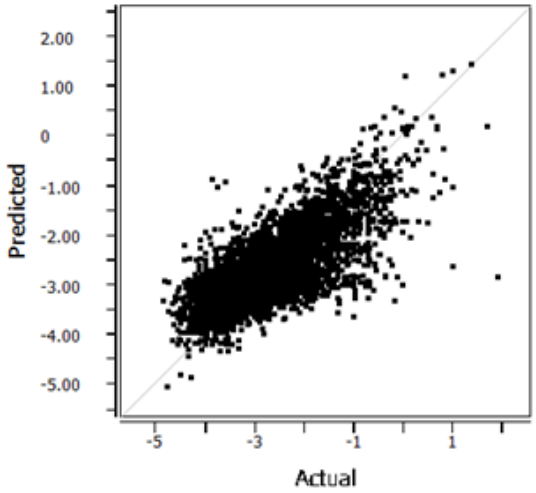
- Features selection
- Continuous prediction - partial least squares (PLS) regression
- Close to balanced classification - partial logistic regression
- Imbalanced classification - average of a series of balanced partial logistic regression (PLR) models
- No optimization of the models

Leadscope®

MODEL RESULTS

-log(LD50) QSAR

Training Set



R-Square: 0.5404

Press: 2268.0

Leadscope®

Model results

Models/Endpoint	Cross-validated Performance
EPA 1 QSAR	Concordance: 88.3 Sensitivity: 75.0 Specificity: 89.6 Positive Predictivity: 39.9 Negative Predictivity: 97.5
EPA 1 2 QSAR	Concordance: 79.0 Sensitivity: 76.1 Specificity: 80.2 Positive Predictivity: 62.3 Negative Predictivity: 88.7
EPA 1 2 3 QSAR	Concordance: 78.7 Sensitivity: 79.4 Specificity: 75.8 Positive Predictivity: 93.2 Negative Predictivity: 46.9

Models/Endpoint	Cross-validated Performance
GHS 1 QSAR	Concordance: 96.7 Sensitivity: 70.3 Specificity: 97.2 Positive Predictivity: 31.1 Negative Predictivity: 99.5
GHS 1 2 QSAR	Concordance: 88.6 Sensitivity: 73.0 Specificity: 90.1 Positive Predictivity: 40.2 Negative Predictivity: 97.3
GHS 1 2 3 QSAR	Concordance: 81.9 Sensitivity: 75.3 Specificity: 83.7 Positive Predictivity: 56.5 Negative Predictivity: 92.4
GHS 1 2 3 4 QSAR	Concordance: 74.1 Sensitivity: 78.6 Specificity: 67.9 Positive Predictivity: 77.0 Negative Predictivity: 69.9

Models/Endpoint	Cross-validated Performance
GHS 1 Alerts	Concordance: 97.6 Sensitivity: 53.8 Specificity: 98.4 Positive Predictivity: 39.6 Negative Predictivity: 99.1 (Not cross-validated)
GHS 1 2 Alerts	Concordance: 93.7 Sensitivity: 42.5 Specificity: 98.4 Positive Predictivity: 61.4 Negative Predictivity: 95.5 (Not cross-validated)

Models/Endpoint	Cross-validated Performance
Very_toxic QSAR	Concordance: 88.8 Sensitivity: 73.4 Specificity: 90.2 Positive Predictivity: 40.6 Negative Predictivity: 97.4
Nontoxic QSAR	Concordance: 76.9 Sensitivity: 72.3 Specificity: 79.7 Positive Predictivity: 68.6 Negative Predictivity: 82.4

Endpoint Calculation

Endpoint	Prediction process
Very toxic	If the Very_Toxic QSAR result is within the applicability domain and the probability is greater than 0.6, a TRUE value is returned, if the probability is less than 0.4 a FALSE value is returned, if the probability is between 0.4 and 0.6, then no result is returned.
Nontoxic	If the Nontoxic QSAR result is within the applicability domain and the probability is greater than 0.5, a TRUE value is returned, if the probability is less than 0.3 a FALSE value is returned, if the probability is between 0.3 and 0.5, then no result is returned.
LD50	If the result is within the applicability domain of the model, the -log(LD50) QSAR model prediction is used to calculate the LD50, using the inverse log of the prediction multiplied by -1
EPA hazard category	A prediction will only be made as long as the all the EPA QSAR models are in domain. If the EPA 1 QSAR probability is greater than 0.6 then an EPA category of 1 is given. If the EPA 1 2 QSAR probability is greater than 0.6 then an EPA category of 2 is given. IF the EPA 1 2 3 QSAR probability is greater than 0.6 then a EPA category of 3 is given, otherwise an EPA category of 4 is given.
GHS hazard category	A prediction will only be made as long as the all the GHS QSAR models are in domain. If the GHS 1 QSAR probability is greater than 0.6 or the GHS 1 Alerts are positive or indeterminate then a GHS category of 1 is given. If the GHS 1 2 QSAR probability is greater than 0.6 or the GHS 1 2 Alerts are positive or indeterminate then a GHS category of 2 is given. IF the GHS 1 2 3 QSAR probability is greater than 0.6 then a GHS category of 3 is given. IF the GHS 1 2 3 4 QSAR probability is greater than 0.6 then a GHS category of 4 is given, otherwise a GHS category of 5 is given.

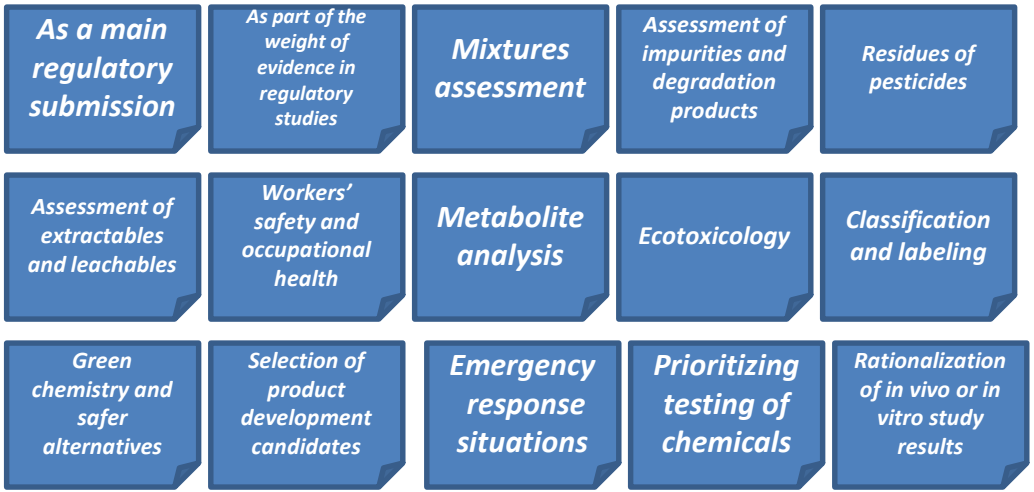
GHS Model Results

		Predicted				
		1	**2**	**3**	**4**	**5**
	1	35	13	2	4	3
	2	53	72	19	20	10
Actual	**3**	21	99	56	139	36
	4	8	73	90	560	200
	5	7	66	48	387	654

Leadscope®

IN SILICO TOXICOLOGY PROTOCOLS

Applications that currently can benefit from *in silico* methods



Leadscope®

In silico toxicology project

- The standardization of *in silico* tool use and interpretation
- Reduce the burden on both industry and regulators to provide justification for the use of these methods
- Results can be generated, recorded, communicated, and archived in a uniform, consistent and reproducible manner
- Incorporating these principles routinely into the use of *in silico* methods will support a more transparent analysis of the results and mitigate “black box” concerns
- Provides an important step towards a quality-driven science for *in silico* toxicology

Leadscope®

Status of *in silico* toxicology project

- Completed the overall strategy
- Initial drafts for the genetic toxicity protocol completed
- Forming subgroups to develop the protocols for:
  - Skin/respiratory/oral sensitization, Carcinogenicity, Reproductive/developmental toxicity, **Acute toxicity/lethality**, Endocrine disruption, Liver toxicity, Cardiac toxicity, Neurotoxicity, Repeated dose, Bone marrow toxicity, Renal toxicity, Brain/CNS toxicity, Gastrointestinal toxicity, Respiratory system toxicity, Skin/eye irritation/corrosion, Physical hazards, Ecotoxicity, Photosensitization/phototoxicity, Physical chemical parameters, Immunotoxicity

Leadscope®

Conclusion

- *In silico* toxicology
  - Fast and inexpensive approach to support toxicological assessments
  - Support the principles of the 3Rs
  - Accepted as part of regulatory submissions
- Protocols provide support for implementation of *in silico* toxicology across many applications

Leadscope®

Acknowledgements

Organizers of Workshop: Predictive Models for Acute Oral Systemic Toxicity

All the coauthors of the *in silico* toxicology paper as well as collaborators in the *in silico* toxicology protocol project

Leadscope scientists: Dave Bower, Kevin Cross, Candice Johnson, Scott Miller, Pat Quigley

Research reported in this presentation was supported by the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number R44ES026909. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Leadscope®

Thank you!

Glenn J. Myatt
gmyatt@leadscope.com
Leadscope, Inc.