Comparative analysis of predictive models used to identify drivers of nanomaterial toxicity

Bryan Harper[1], Rong Liu[2], Yoram Cohen[2] and Stacey L. Harper[1, 3]

[1]Dept. of Environmental and Molecular Toxicology, Oregon State University, Corvallis, Oregon, United States; [2]Center for Environmental Implications of Nanotechnology, University of California, Los Angeles, Los Angeles, California, United States; [3]School of Chemical, Biological and Environmental Engineering, Oregon State University, Corvallis, Oregon, United States

ABSTRACT

Understanding the inherent and conditional factors associated with nanomaterial toxicity is critical to the development nanotechnologies that pose minimal threats to humans and the environment over the life cycle of the nanomaterial. We compared the results of several different models built to predict toxicity from the open-source data on nanomaterial toxicity to embryonic zebrafish (*Danio rerio*) found in the Nanomaterial-Biological Interactions (NBI) knowledgebase at Oregon State University. Model comparisons included the ABMiner predictive models, MATLAB clustering analysis and the use of Self-Organizing Map (SOM) based consensus clustering conducted on the data in the NBI knowledgebase (nbi.oregonstate.edu). Overall results suggest that exposure concentration and outermost surface chemistry (and thus surface charge) both should be considered in conjunction with the core composition of nanomaterials when trying to develop predictive models for developing zebrafish. Core composition was found to be a significant contributor to the ABMiner predictive model. MATLAB clustering grouped materials into two clusters with outermost surface chemistry being the primary determinant. The SOM modeling identified five significant clusters (clustering index = 0.89); while no core materials occurred in all 5 clusters, 4 material types (based on core composition) occurred in 4/5 clusters and 15 material types occurred in 3/5 clusters. In addition, over half of the materials appeared in multiple clusters depending on the dose applied. Thus, classification of nanomaterials by simple descriptors such as core composition may not be sufficient for predicting nanomaterial toxicity or managing nanomaterial risks.

BACKGROUND

- Risk assessments for particulates are focused on determining risk based on material core composition.
- Engineered nanoparticles (NPs) can have a complex range of surface modifications that are dynamic in diverse biological environments.
- Rapid *in vivo* assays have provided a wealth of data for the development of predictive models of NP toxicity.
- Model development and refinement depends on identifying inherent and conditional NP features driving toxic responses to whole organisms.

OBJECTIVE

To provide the results of several independent analysis of NBI data and determine which predictive features are common across the various modeling efforts.

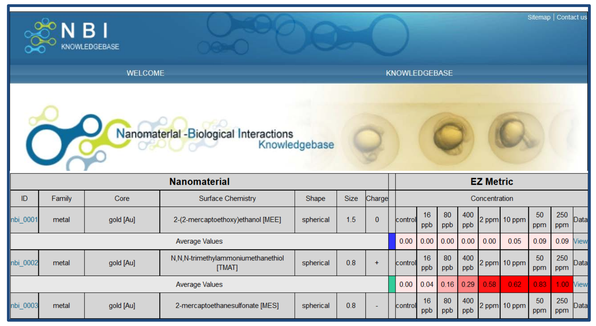
nanomaterial–biological interactions

NBI KNOWLEDGEBASE

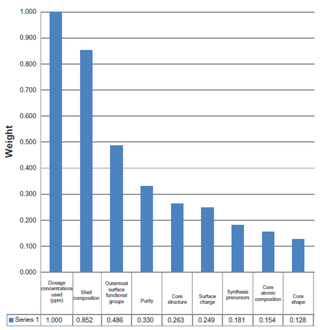
nbi.oregonstate.edu

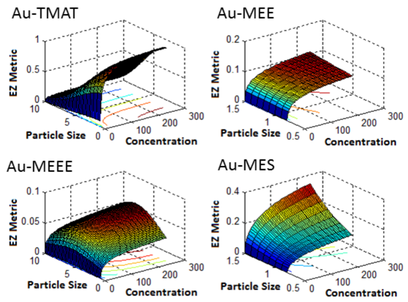
MODELING METHODS

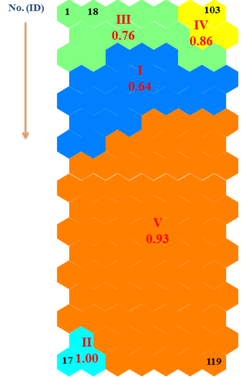
- Each analysis from the NBI database contains embryonic zebrafish responses to a wide-range of nanoparticles.
  - Data includes assessment of 21 endpoints encompassing mortality as well as behavioral and developmental endpoints as well as visual assessment of multiple morphological endpoints.
- Predictive features from previous models (Liu et al. 2013 and Harper et al. 2014, in review) were compared against the results of an additional clustering analysis using Self-Organizing Maps (SOMs).



	Nanomaterial						EZ Metric									
ID	Family	Core	Surface Chemistry	Shape	Size	Charge	Concentration									
nbi_0001	metal	gold [Au]	2-(2-mercaptoethoxy)ethanol [MEE]	spherical	1.5	0	control	16 ppb	80 ppb	400 ppb	2 ppm	10 ppm	50 ppm	250 ppm	Data
Average Values							0.00	0.00	0.00	0.00	0.00	0.05	0.09	0.09	View
nbi_0002	metal	gold [Au]	N,N,N-trimethylammoniumethanethiol [TMAT]	spherical	0.8	+	control	16 ppb	80 ppb	400 ppb	2 ppm	10 ppm	50 ppm	250 ppm	Data
Average Values							0.00	0.04	0.16	0.29	0.58	0.62	0.83	1.00	View
nbi_0003	metal	gold [Au]	2-mercaptoethanesulfonate [MES]	spherical	0.8	-	control	16 ppb	80 ppb	400 ppb	2 ppm	10 ppm	50 ppm	250 ppm	Data







NEI Miner Analysis

Liu, X., K. Tang, et al. (2013). "Predictive modeling of nanomaterial exposure effects in biological systems." International Journal of Nanomedicine **8**(1): 31-43.

- 82 nanomaterial records across 8 particle concentrations per material
- Evaluated against 20 measured end-points.
- Combined IBK and RELEIF approach.
- High correlation for modeling 24 hour mortality (0.837).
- Exposure concentration, outermost surface chemistry, core structure, and surface charge are significant attributes for predicting mortality.

MATLAB Hierarchical Clustering

Harper, B., D. Thomas, et al. (2014). "Comparative hazard analysis and predictive modeling of diverse nanomaterials using the embryonic zebrafish (EZ) metric of toxicity" *In review* , Nanotoxicology.

- Conducted on 68 nanomaterial records taking into account the summarized toxicity from multiple doses and endpoints (sumEZ).
- Based on Ward linkage rule with Euclidian distance.
- Measure gave well-separated clusters when data was assessed on the chemical constituents, primary particle size, and surface chemistry of the nanomaterials.

Surface Chemistry Based Model

Harper, B., D. Thomas, et al. (2014). "Comparative hazard analysis and predictive modeling of diverse nanomaterials using the embryonic zebrafish (EZ) metric of toxicity" *In review* , Nanotoxicology.

- Developed using 16 gold NP records on NBI.
- Particle concentration, surface chemistry and particle size were predictors of toxicity.
- High coefficient of determination (0.88) between actual and predicted scores for all 68 nanomaterials.

Self-Organizing Map Cluster Analysis

- Conducted on 74 nanomaterials across 8 concentrations
- Assessment of 20 lethal and sub-lethal endpoints.
- SOM, where each cell (i.e. the hexagons) contains a certain number of similar NPs.
- 5 clusters of similar SOM cells were identified with a clustering index of 0.89 indicating a significant clustering pattern.
- Particle concentration, surface chemistry and surface charge were found to relate to the clustering pattern.

RESULTS

	Exposure Concentration (ppm)	Outermost Surface Chemistry	Surface Charge	Core Composition	Partcile Size
NEI Miner RELEIF	X	X	X		
MATLAB Clustering		X		X	X
Surface Chem GLM	X	X			X
SOM Clustering	X	X	X		

- Out of 20 input variables considered in the analytical approaches, only 5 features were found to be predictive of nanomaterial toxicity.
- Outermost surface chemistry is the only determining feature common to all 4 modeling efforts.
  - Understanding changes in surface chemistry resulting from interactions with biological media should improve models.
- Interestingly, core composition was only found to be predictive in one case.
  - For single-composition nanoparticles like metal oxides core composition can be viewed as predictive.

CONCLUSIONS

- Data mining of large experimental databases comprised of heterogeneous nanoparticles are useful in developing predictive models of nanomaterial toxicity.
- Predictive model refinement can be achieved through consensus modeling of the same large datasets.
  - Data that includes thorough material characterization and multiple end-points provides the volume of information required for model development.
  - The surface chemistry-based predictive model was supported by other model findings.
- Risk assessments for nanomaterials based on core composition or size alone may not be sufficient for predicting *in vivo* hazard.

ACKNOWLEDGEMENTS

We thank the staff of the Northwest Alliance for Computational Science (NACSE), Pacific Northwest National Laboratory (PNNL) and Intelligent Automation, Inc. Studies were supported with funding from #ES017552-01A2; #P30ES03850; #ES0166896-01; #FA8650-05-1-15041; P30ES000210.