JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim

Editorial

Joint Virtual Special Issue on Computational Toxicology

Cite This: *J. Chem. Inf. Model.* 2020, 60, 1069−1071

Read Online

ACCESS | Metrics & More | Article Recommendations

Computational models have earned broad acceptance for assessing chemical toxicity during early stages of drug discovery or environmental safety assessment.[1] Some academic groups and many companies have developed platforms for chemical toxicity prediction. For instance, MultiCASE Inc.[2] was a pioneer in developing a platform for Ames mutagenicity, which was later expanded to many other end points. Many academic and industrial groups have been developing methods and models to predict chemical toxicity for products used in agriculture, cosmetics, the chemical industry, and, of course, the pharmaceutical industry. A recent review summarized major types of toxicity end points and *in silico* tools to predict different types of toxicity available commercially or via open access.[3]

Prediction of chemical toxicity can be generally viewed as a subset of general bioactivity prediction, but there is an important caveat that places the challenge of toxicity prediction into an important separate category. The issue is that chemical safety assessment is a matter of regulatory concern by both the FDA and EPA in the US and by respective regulatory organizations worldwide (such as the ECHA in Europe). Historically, animal testing has been the standard approach to safety assessment. However, in October 2016, the European Union's REACH passed a regulation to make nonanimal testing for skin sensitization the default requirement for regulatory assessment of chemical product safety.[4] In September 2019, the U.S. EPA released a memo to promote research to *reduce* animal testing, with a goal to eliminate all mammal study requests and support by 2035.[5] Given the federal mandate to identify alternative approaches to animal testing of chemical and medical agents, the *Interagency Coordinating Committee on the Validation of Alternative Methods* (ICCVAM) created a U.S. Strategic Roadmap for establishing new methods to evaluate the chemical safety of medical products[6] and to expedite the development of NAMs, i.e., chemical hazard/risk assessment technologies that combine a series of *in vitro* toxicity assays and computational approaches, which do not involve animal use. These recent developments in the regulatory space placed a new emphasis on the importance of accurate *in silico* tools for chemical safety assessment. Not surprisingly, recent market trend analysis[7] suggested that the "*in-vitro* toxicology testing market is expected to reach USD 12.7 billion by 2024 from an estimated USD 8.1 billion in 2019". The same report went on to state that "the increasing focus of the pharmaceutical and cosmetics industries on using *in-vitro* methods for product testing along with the improvement *in-silico* methods for predictive toxicology studies are expected to offer significant growth opportunities for players in this market".

It should not then be surprising that both the *Journal of Chemical Information and Modeling* (JCIM) and *Chemical Research in Toxicology* (CRT) have experienced a growth in the submission of studies in the area of chemical toxicity prediction. In fact, this trend led CRT to recruit a new Associate Editor specializing in computational toxicology. Reflecting these trends, both journal editors felt that a joint special virtual issue highlighting most interesting and innovative contributions published in both JCIM and CRT in the last three years would be in order.

The papers selected for this special virtual issue reflect recent and developing trends in employing increasingly more sophisticated data modeling methods such as deep learning to toxicity databases that continue to grow in both size and complexity. For instance, Xu et al.[8] have pursued acute oral toxicity, one of the well-known and important end points. However, by using a novel molecular graph encoding and convolutional neural network they have achieved an improved accuracy as compared to models built with more traditional approaches. Jiménez et al.[9] have explored a large collection of compounds selected from ChEMBL and used deep self-normalizing neural network model for the prediction of molecular pathway association of these chemicals to provide mechanistic prediction of compound effects. Their models reached accuracy as high as 81−83% for external industrial data sets. In fact, many recent publications have reported the use of deep neural network (DNN) methods to the task of toxicity prediction. Perhaps the most uncommon was the study by Fernandez et al.[10] who used unconventional representation of molecules by 2D drawings (i.e., molecular images) rather than by any of conventional chemical descriptors. Despite the simplicity of such representation, their models achieved accuracy comparable to state-of the-art conventional chem-informatics approaches and a standard Tox21 benchmark data set. Wenzel et al.[11] have explored both single-task and multitask DNN models to predict several properties including microsomal lability in different species for a large data set of 50 000 both publicly available and proprietary compounds. Interestingly, they found that multitask DNN models provided higher prediction accuracy for human metabolic liability data.

Published: February 26, 2020

https://dx.doi.org/10.1021/acs.jcim.0c00140
J. Chem. Inf. Model. 2020, 60, 1069−1071

Wu et al.[12] also used DNN along with more traditional machine learning approaches in their study of aquatic and acute oral toxicity. Unique to their study was the use of a novel approach to compound characterization, element specific persistent homology (ESPH), which employs special topological abstraction of geometric complexity of molecules. The authors found that using their novel descriptors, multitask deep learning models had higher accuracy than ensemble methods such as random forest and gradient boosting decisions trees. Another example demonstrating the power of multitask learning over single-task approaches was presented by Sosnin et al.[13] These authors also studied acute toxicity for a large compound data set extracted from the Registry of Toxic Effects of Chemical Substances (RTECS). Importantly, they made their models publicly available via their OCHEM platform (ochem.eu/multitox) in an effort to promote regulatory acceptance of multitask modeling approaches.

Finally, two interesting studies in the difficult area of predicting drug induced liver injury (DILI) have been published by Wu et al.[14] and Kotsampasakou and Ecker.[15] Reflecting the complexity of DILI, Wu et al. have introduced a novel quantitative structure activity relationship (QSAR) approach incorporating the drug's mode of action (MOA) as part of the modeling process. The authors used the MOA knowledge to divide compounds in their data sets into respective groups, developed QSAR models for each group, and then combined individual models to obtain an integrated MOA-DILI model. They showed that this approach achieves appreciable statistical accuracy while allowing to discriminate DILI-causing drugs by the MOA. Kotsampasakou and Ecker approached one of the major forms of DILI, namely cholestasis prediction, using somewhat similar hybrid modeling scheme except besides chemical descriptors, they also used predictions of selected hepatic transporters' inhibition (BSEP, BCRP, P-gp, OATP1B1, and OATP1B3) as additional "biological" descriptors of drugs. They found that the inclusion of these descriptors makes models more mechanistically relevant while having appreciable prediction accuracy.

The computational articles in CRT covered a wide variety of topics including classical toxicity modeling, the development of new descriptors and generation of hypothesis that can be confirmed experimentally in new studies. The first two selected works deal with toxicity estimation for single end points. Prediction of DILI is important to improve hepatotoxicity risk assessment in drug discovery. The Bayesian model based on in vitro assays as well as lipophilicity and exposure variables was developed[16] to classify compounds and providing a probability of the compound to belong to the target category. One of the most interesting contributions of this work is the developed visualization tools to interpret the importance of variables for each prediction. Compounds that absorb light in UV/vis spectra can demonstrate phototoxicity. A very nice and comprehensive photosafety assessment protocol was described by Schmidt et al.[17] It incorporates a multitude of computational approaches, such as identification of phototoxophores, and, importantly, calculation of physically relevant descriptors, e.g., quantum chemical, UV/vis spectral absorption descriptors, which were used as pat of machine learning modeling efforts to eliminate the problematic compounds. The importance of descriptors for appropriate modeling of biochemical processes on the surface of nanoparticles of nanoparticles was a topic of the Boukhvalov and Yoon study.[18] The authors considered ion extraction from the surface of a specimen to aqueous media and water dissociation on the surface to come with new descriptors to model toxicity of nanomaterials. The analysis of whether bioinformatics knowledge can help to predict synergetic toxicity of binary mixtures was investigated by Kim et al.[19] The author used protein−chemical and protein−protein interaction networks to characterize shared biological targets and their neighborhoods for analyzed mixtures, which were the most important features for classifying chemicals into synergistic and nonsynergistic groups.

The methodology on how to use results of in vitro assays for in vivo predictions is amid of the intensively developing area of studies of the journal. The application of Tox21 HTS assays within computational models to identify compounds with genotoxicity potential was reported by Hsieh et al.[20] The authors noticed that, despite current quantitative high-throughput screening (qHTS) alone have low sensitivity and cannot replace traditional genotoxicity assays, they can be used to prioritize chemicals for further studies. In another study the results of HTS screening were found to contribute the more accurate classifiers for prediction of 35 in vivo target organ toxicity outcomes[21] than the studied chemical descriptors alone. The combination of both types of descriptors further improved the models. These preliminary results indicate perspectives to predict organ toxicities of untested chemicals without a need to perform animal testing. In the last selected study Poussin et al.[22] reported crowd-source validation of computational results of Philip Morris, which suggest that mice exposed to MRTP (Modified Risk Tobacco Product) have gene expression profiles more close to nonsmoke exposed groups. While this study does not directly show that consumption of MRTP has a decreased health impact, it does suggest such working hypothesis to be confirmed experimentally.

In summary, this joint JCIM/CRT virtual issue highlights current important developments in the area of computational toxicology that continue to be important components of modern drug discovery and chemical safety assessment research. Both journals continue to be committed to publishing innovative and impactful papers in this important research discipline. We encourage scientists working in this field to submit their new manuscripts to either journal; more computationally inclined methodological papers may perhaps go to JCIM whereas manuscripts emphasizing specific applications including regulatory science may be more suitable for CRT.

Igor V. Tetko
Alexander Tropsha

AUTHOR INFORMATION

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.0c00140

Notes

Views expressed in this editorial are those of the authors and not necessarily the views of the ACS.

REFERENCES

(1) Rusyn, I.; Daston, G. P. Computational toxicology: realizing the promise of the toxicity testing in the 21st century. *Environ. Health Perspect.* **2010**, *118*, 1047−1050.
(2) MultiCASE. *Powerful and reliable software for in silico safety assessment*, http://www.multicase.com/products (6 February 2020).

https://dx.doi.org/10.1021/acs.jcim.0c00140
J. Chem. Inf. Model. 2020, 60, 1069−1071

(3) Yang, H.; Sun, L.; Li, W.; Liu, G.; Tang, Y. In Silico Prediction of Chemical Toxicity for Drug Design Using Machine Learning Methods and Structural Alerts. *Front. Chem.* **2018**, *6*, 30.
(4) ECHA. *Non-animal methods now a default for skin sensitisation—submit correct information*, https://echa.europa.eu/-/non-animal-methods-now-a-default-for-skin-sensitisation-submit-correct-information (accessed 6 February 2020).
(5) EPA. *EPA Administrator Wheeler Signs Memo to Reduce Animal Testing, Awards $4.25 Million to Advance Research on Alternative Methods to Animal Testing*, https://www.epa.gov/newsreleases/administrator-wheeler-signs-memo-reduce-animal-testing-awards-425-million-advance. (6 February 2020).
(6) ICCVAM. *A Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medical Products in the United States*, https://ntp.niehs.nih.gov/ntp/about_ntp/bsc/2017/december/meetingmaterials/iccvam_roadmap_508.pdf (6 February 2020).
(7) *In-Vitro Toxicology/Toxicity Testing Market by Product & Services, Toxicity Endpoint & Test, Technology, Method, Industry, Region-Global Forecast to 2024*, https://www.reportlinker.com/p05129363/In-Vitro-Toxicology-Testing-Market-by-Products-Services-Toxicity-Endpoints-Tests-Technology-Method-Industry-Forecast-to.html. (6 February 2020),.
(8) Xu, Y.; Pei, J.; Lai, L. Deep Learning Based Regression and Multiclass Models for Acute Oral Toxicity Prediction with Automatic Chemical Feature Extraction. *J. Chem. Inf. Model.* **2017**, *57*, 2672–2685.
(9) Jimenez, J.; Sabbadin, D.; Cuzzolin, A.; Martinez-Rosell, G.; Gora, J.; Manchester, J.; Duca, J.; De Fabritiis, G. PathwayMap: Molecular Pathway Association with Self-Normalizing Neural Networks. *J. Chem. Inf. Model.* **2019**, *59*, 1172–1181.
(10) Fernandez, M.; Ban, F.; Woo, G.; Hsing, M.; Yamazaki, T.; LeBlanc, E.; Rennie, P. S.; Welch, W. J.; Cherkasov, A. Toxic Colors: The Use of Deep Learning for Predicting Toxicity of Compounds Merely from Their Graphic Images. *J. Chem. Inf. Model.* **2018**, *58*, 1533–1543.
(11) Wenzel, J.; Matter, H.; Schmidt, F. Predictive Multitask Deep Neural Network Models for ADME-Tox Properties: Learning from Large Data Sets. *J. Chem. Inf. Model.* **2019**, *59*, 1253–1268.
(12) Wu, K.; Wei, G. W. Quantitative Toxicity Prediction Using Topology Based Multitask Deep Neural Networks. *J. Chem. Inf. Model.* **2018**, *58*, 520–531.
(13) Sosnin, S.; Karlov, D.; Tetko, I. V.; Fedorov, M. V. Comparative Study of Multitask Toxicity Modeling on a Broad Chemical Space. *J. Chem. Inf. Model.* **2019**, *59*, 1062–1072.
(14) Wu, L.; Liu, Z.; Auerbach, S.; Huang, R.; Chen, M.; McEuen, K.; Xu, J.; Fang, H.; Tong, W. Integrating Drug's Mode of Action into Quantitative Structure-Activity Relationships for Improved Prediction of Drug-Induced Liver Injury. *J. Chem. Inf. Model.* **2017**, *57*, 1000–1006.
(15) Kotsampasakou, E.; Ecker, G. F. Predicting Drug-Induced Cholestasis with the Help of Hepatic Transporters-An in Silico Modeling Approach. *J. Chem. Inf. Model.* **2017**, *57*, 608–615.
(16) Williams, D. P.; Lazic, S. E.; Foster, A. J.; Semenova, E.; Morgan, P. Predicting Drug-Induced Liver Injury with Bayesian Machine Learning. *Chem. Res. Toxicol.* **2020**, *33*, 239.
(17) Schmidt, F.; Wenzel, J.; Halland, N.; Gussregen, S.; Delafoy, L.; Czich, A. Computational Investigation of Drug Phototoxicity: Photosafety Assessment, Photo-Toxophore Identification, and Machine Learning. *Chem. Res. Toxicol.* **2019**, *32*, 2338–2352.
(18) Boukhvalov, D. W.; Yoon, T. H. Development of Theoretical Descriptors for Cytotoxicity Evaluation of Metallic Nanoparticles. *Chem. Res. Toxicol.* **2017**, *30*, 1549–1555.
(19) Kim, J.; Fischer, M.; Helms, V. Prediction of Synergistic Toxicity of Binary Mixtures to Vibrio fischeri Based on Biomolecular Interaction Networks. *Chem. Res. Toxicol.* **2018**, *31*, 1138–1150.
(20) Hsieh, J. H.; Smith-Roe, S. L.; Huang, R.; Sedykh, A.; Shockley, K. R.; Auerbach, S. S.; Merrick, B. A.; Xia, M.; Tice, R. R.; Witt, K. L. Identifying Compounds with Genotoxicity Potential Using Tox21 High-Throughput Screening Assays. *Chem. Res. Toxicol.* **2019**, *32*, 1384–1401.
(21) Liu, J.; Patlewicz, G.; Williams, A. J.; Thomas, R. S.; Shah, I. Predicting Organ Toxicity Using in Vitro Bioactivity Data and Chemical Structure. *Chem. Res. Toxicol.* **2017**, *30*, 2046–2059.
(22) Poussin, C.; Belcastro, V.; Martin, F.; Boue, S.; Peitsch, M. C.; Hoeng, J. Crowd-Sourced Verification of Computational Methods and Data in Systems Toxicology: A Case Study with a Heat-Not-Burn Candidate Modified Risk Tobacco Product. *Chem. Res. Toxicol.* **2017**, *30*, 934–945.

https://dx.doi.org/10.1021/acs.jcim.0c00140
J. Chem. Inf. Model. 2020, 60, 1069–1071