S03-02

Use of chemical informatics, quantum chemistry modelling and artificial intelligence algorithms to predict molecular initiating events

T. E. H. Allen[1]*, A. J. Wedlake[2], M. N. Grayson[3], A. M. Middleton[4], M. Folia[4], M. Baltazar[4], P. Piechota[4], E. Gelžinytė[2], J. M. Goodman[2], P. J. Russell[4], P. Kukic[4] and S. Gutsell[4]

1. University of Cambridge, MRC Toxicology Unit, Cambridge, UK
2. University of Cambridge, Centre for Molecular Informatics, Yusuf Hamied Department of Chemistry, Cambridge, UK
3. University of Bath, Department of Chemistry, Bath, UK
4. Unilever, Safety and Environmental Assurance Centre, Sharnbrook, UK

Molecular Initiating Events (MIEs) provide good targets for in silico modelling using a variety of approaches, as they are well defined chemical-biological interactions [1]. A wide variety of algorithms have been used to predict these important key events, and these models are seeing use in a variety of hazard identification and molecular screening safety decisions. However, using these algorithms in a chemical risk assessment remains a challenge. Next-generation risk assessment (NGRA) generally involves the comparison of quantitative exposure and hazard values considered as probability distributions [2]. Can computational approaches provide the mechanistic insight and uncertainty estimation required for such decisions?

We have deployed several different algorithms in this area. This includes transparent structural alert models constructed using a maximal common substructure search, random forest models developed using physicochemical descriptors and neural networks in Python 3 using TensorFlow [3,4]. All three computational approaches consistently provide models with over 90% accuracy against test set data and combining these models allows us to provide the best possible prediction – increasing model performance when compared to any individual model.

Quantum chemistry approaches have also been investigated for the prediction of chemical bond forming MIEs, specifically to predict the results of the Ames assay [5]. DNA can directly react with some electrophilic chemicals, modifying its structure and causing damage that can lead to genotoxic adverse outcomes. By computationally modelling the transition state of these reactions and calculating the activation energy we have been able to establish why some molecules can react with DNA, and return a positive Ames test, and why some chemically similar molecules cannot.

To further extend the use of some of our developed MIE models to NGRA, qualitative activity predictions and uncertainty estimates are required. Bayesian learning neural networks use probability distributions to produce an output prediction with a mean and standard deviation. This uncertainty accounts for both how close the new molecule is to the existing data and how much variation exists within the training set. These networks produce quantitative activity estimates with errors within one log unit, even on external validation

data. We have also started to apply our algorithms to the prediction of key events further down the adverse outcome pathway, allowing them to provide more mechanistic understanding of toxicity. Chemical risk assessment requires information on molecular potency, uncertainty and toxicity mechanism, and decision-makers and regulators need to have confidence in the tools being used. The models we have developed provide additional understanding and confidence, vital for their future use in risk assessment.

References

[1] Allen, T.E.H.; Goodman, J.M.; Gutsell, S.; Russell, P.J. 2014, Defining Molecular Initiating Events in the Adverse Outcome Pathway Framework for Risk Assessment, *Chem. Res. Toxicol*., 29, 1611

[2] Baltazar, M.T.; Cable, S.; Carmichael, P.L.; Cubberley, R.; Cull, T.; Delagrange, M.; Dent, M.P.; Hatherell, S.; Houghton, J.; Kukic, P.; Li, H.; Lee, M-Y.; Malcomber, S.; Middleton, A.M.; Moxon, T.E.; Nathanail, A.V.; Nicol, B.; Pendlington, R.; Reynolds, G.; Reynolds, J.; White, A.; Westmoreland C. 2020, A Next Generation Risk Assessment Case Study for Coumarin in Cosmetic Products, *Toxicol. Sci*, 176, 236

[3] Wedlake, A.J., Folia, M., Piechota, S., Allen, T.E.H.; Goodman, J.M.; Gutsell, S.; Russell, P.J. 2020, Structural Alerts and Random Forest Models in a Consensus Approach for Receptor Binding Molecular Initiating Events, *Chem. Res. Toxicol*., 33, 388

[4] Allen, T.E.H.; Wedlake, A.J.; Gelžinytė, E.; Gong, C.; Goodman, J.M.; Gutsell, S.; Russell, P.J. 2020, Neural Network Activation Similarity: A New Measure to Assist Decision Making in Chemical Toxicology, *Chem. Sci.*, 11, 7335

[5] Allen, T.E.H.; Grayson, M.N.; Goodman, J.M.; Gutsell, S.; Russell, P.J. 2018, Using Transition State Modeling to Predict Mutagenicity for Michael Acceptors, *J. Chem. Inf. Model.*, 58, 1266