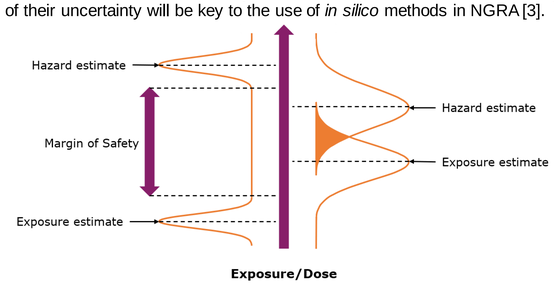
Quantitative Molecular Activity Calculation Using Bayesian Learning Neural Networks

Timothy E H Allen[1], Alistair M Middleton[2], Jonathan M Goodman[3], Steve Gutsell[2], Predrag Kukic[2], Paul J Russell.[2]

1. **MRC Toxicology Unit,** University of Cambridge, Gleeson Building, Tennis Court Road, Cambridge, United Kingdom, CB2 1QR
2. **Unilever Safety and Environmental Assurance Centre**, Colworth Science Park, Sharnbrook, Bedfordshire, United Kingdom, MK44 1LQ
3. **Centre for Molecular Informatics**, Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, United Kingdom CB2 1EW

Introduction

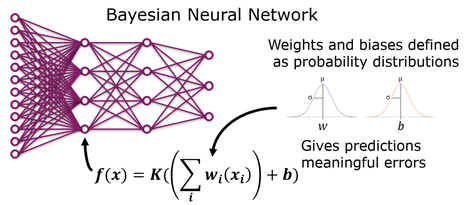
Next generation risk assessment (NGRA) [1] involves making a comparison between the exposure of an individual to a chemical and the hazard associated with that chemical to determine a margin of safety. Exposure and hazard are typically treated as probability distributions, due to factors such as uncertainties in experimental measurements and population variance. Any overlap in these distributions represents a potential risk. Because of this, an understanding of the quantitative activity of a molecule and the uncertainty in such an estimate is required for NGRA [2]. Many current *in silico* approaches are based on the classification of molecules, for example as "active" or "inactive", and as such cannot assist in answering these questions. The ability of computational approaches to produce robust numerical activity estimates coupled with an understanding of their uncertainty will be key to the use of *in silico* methods in NGRA [3].



The molecular initiating event (MIE) [4,5,6] is the initial chemical-biological interaction that can be thought of as a gateway to the adverse outcome pathway (AOP) [7]. Modelling of MIEs allows for mechanistic understanding of the kinds of interactions molecules can make, and hence the kinds of adverse outcomes they might cause. This makes them ideal targets for *in silico* modelling aiming to contribute to a mechanistic risk assessment decision – one of the goals of NGRA.

Bayesian Neural Networks

In this study we have built machine learning models for the prediction of quantitative activity and uncertainty estimation at human MIEs. Pharmacologically important human MIE targets from the Bowes list [8] have been investigate, using open source data from the database ChEMBL [9]. Initially, neural networks were trained on quantitative experimental data (p(IC50), p(EC50), p(Ki), p(Kd)) to provide regression models using linear outputs and mean absolute error (MAE) as the loss function and primary evaluation statistic. Coefficient of determination values (R^2) against the x=y diagonal have also calculated for completeness. These were improved using Bayesian learning in TensorFlow Probability [10] with Dense Variational Layers to replace point value weights and biases throughout the network with probability distributions. This process allows the quantitative output to be produced, using a Monte Carlo simulation of 500 iterations, as a mean and standard deviation representing the activity estimate and its uncertainty. A training set was used in model training, with a randomly removed validation set (for hyperparameter optimization) and test set (for final model evaluation). An external validation set was obtained by extracting new compounds from a more recent version of ChEMBL (version 25), for additional validation on chemicals less similar to the training data.



All models constructed, model build codes developed and datasets extracted in this work are freely available online through GitHub;

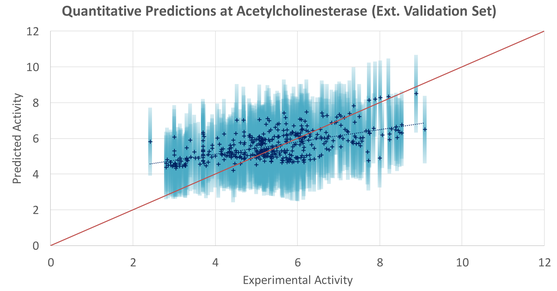
https://github.com/teha2/chemical_toxicology

Quantitative Predictions

Average MAE and R^2 values are shown in the table below for the held out Test and External Validation data sets.

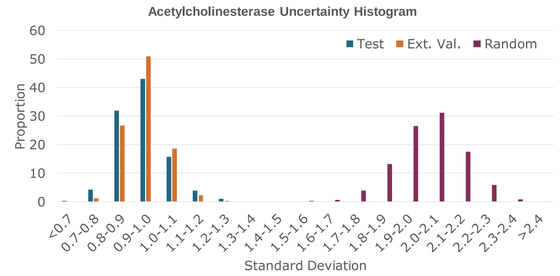
	Test Set		Ext. Validation Set	
	MAE	**R^2**	**MAE**	**R^2**
Mean	0.621	0.572	0.943	0.128
St. Dev.	0.050	0.094	0.223	0.437

Overall, the best performing models for each MIE have produced MAEs within one log unit on both the test and external validation data. A decrease in model performance on moving from the test to external validation data is expected, and this is observed. Of the 21 cases studied, eight show external validation R^2 values greater than 0.4, suggesting similar performance to the test set. Seven of these cases though show negative R^2, meaning these models are not predictive on these datasets when compared to the mean activity of that dataset. This can be caused by small external validation datasets containing molecules with a similar experimental activity, and additional data may be able to assist with this in future work. A graph of predicted vs experimental external validation values is shown below for acetylcholinesterase, with 95% confidence intervals.



Understanding Uncertainty

As well as quantitative activity values, the Bayesian neural networks also produce standard deviation values which can be treated as uncertainties in their estimates. To evaluate their use, the standard deviations were compared between the test set, external validation set and a random set, of randomly shuffled molecular fingerprints. A histogram of these values is shown below for the acetylcholinesterase data.



Each set shows a slightly different distribution of standard deviations, with the test set being the lowest, external validation set compounds shifted slightly to the right and random set the highest. This suggests the model understands which fingerprints are closest to the test data, and which are dramatically different – resulting in increased uncertainties.

Quantification of uncertainty is considered one of the biggest challenges required for NGRA to succeed. The Bayesian learning models presented here provide the desired uncertainty values that can be fed into an NGRA procedure or a quantitative adverse outcome pathway for the safety evaluation of a novel chemical.

Conclusions

- NGRA requires a quantitative understanding of molecular activity and uncertainty in that value.

- Bayesian learning neural networks can provide quantitative predictions and uncertainties suitable for risk assessment.

- Leveraging tools such as these will allow computational toxicology to contribute to more areas of risk assessment decision making.

Acknowledgements

Unilever
MRC Toxicology Unit
Centre for Molecular Informatics
St. John's College, Cambridge

References

1. Baltazar M.T. *et. al., Toxicol. Sci.* (2020) 176, 236-252.
2. Dong Z. *et. al., Environ. Int.* (2015) 85, 120-132.
3. Thomas R.S. *et. al., Toxicol. Sci.* (2019) 169, 317-332.
4. Allen T. E. H. *et. al., Chem. Res. Toxicol.* (2014) 27, 2100-2112.
5. Gutsell S, Russel P. J., *Toxicol. Res.* (2013) 2, 299-307.
6. Allen T. E. H. *et. al., Chem. Res. Toxicol.* (2016) 29, 2060-2070.
7. Ankley G. T. *et al., Environ. Toxicol. Chem.* (2010) 29, 730-741.
8. J. Bowes. *et al, Nature Reviews* (2012) 11, 909-922.
9. ChEMBL Database (Version 23): *www.ebi.ac.uk/chembl*
10. TensorFlow Probability: *www.tensorflow.org/probability*