Weight of Evidence by Conditional Probabilities: A Bayesian Network Model for Predicting Fish Acute Toxicity Based on Fish Embryo Testing

Jannicke Moe[1], Anders Madsen[2], Scott Belanger[3], Thomas Braunbeck[4], Kristin Connors[3], Michelle Embry[5], Wayne Landis[6], Kristin Schirmer[7], Stefan Scholz[8], Raoul Wolf[1], and Adam Lillicrap[1]

[1]Norwegian Institute for Water Research (NIVA), Oslo, Norway
[2]HUGIN Expert A/S, Aalborg, Denmark
[3]Procter & Gamble (P&G), Mason, OH, USA
[4]Centre for Organismal Studies (COS), University of Heidelberg, Heidelberg, Germany
[5]Health and Environmental Sciences Institute (HESI), Washington, DC, USA
[6]Institute of Environmental Toxicology, Western Washington University, Bellingham, WA, USA
[7]Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Dubendorf, Switzerland
[8]Helmholtz Centre for Environmental Research (UFZ), Leipzig, Germany
E-mail contact: jmo@niva.no

1. Introduction

The fish embryo toxicity test (FET; OECD TG 236) has been proposed as an alternative to using juvenile fish in the acute fish toxicity test (AFT; OECD TG 203) to reduce the number of live animals required for hazard assessments of chemicals. However, FET data are not yet accepted as a replacement for AFT data for regulatory purposes such as REACH. The European Chemicals Agency has recommended the development of a Weight of Evidence (WoE) approach for strengthening ability of FET data to predict (juvenile) acute fish toxicity. To meet this challenge, we have developed a Bayesian network (BN) model (Figure 1) for using FET data in a probabilistic WoE approach [1, 2].

In the CEFIC LRI ECO51 project SWiFT (Strengthening Weight of evidence for FET data to replace Acute Fish Toxicity; http://swift.hugin.com), this BN model will be further developed and extended with additional lines of evidence (*e.g.*, fish gill cell-line cytotoxicity assay (ISO 21115), neurotoxicity and biotransformation), in order to provide more accurate predictions. More details on the lines of evidence and the underlying data will be presented in the session Alternative Approaches to Animal Testing.

This presentation will focus on the methodological aspects: how the main steps of a WoE approach (assembly, weighting, and integration of evidence) can be quantified and implemented in a BN model. We will address the opportunities and challenges related to topics such as: (1) quantification of lines of evidence - from associations to causal relationships; (2) accounting for uncertainty and variability in data, both within and between studies; (3) integration of the multiple lines of evidence; and (4) approaches to model evaluation.

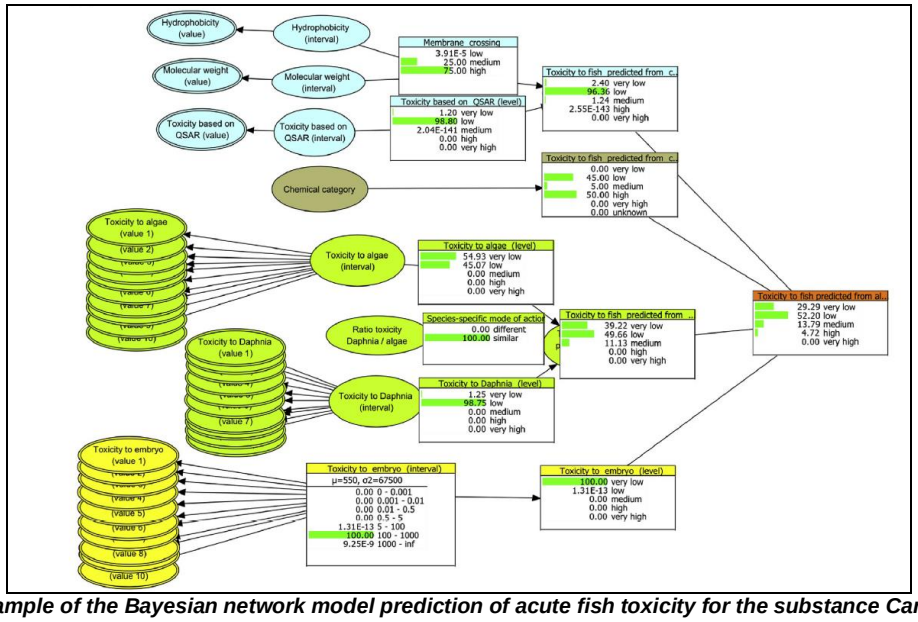

Figure 1: Example of the Bayesian network model prediction of acute fish toxicity for the substance Carbamazepine.

2. Materials and methods

A Bayesian network (BN) is is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph [3]. The purpose of the proposed BN model is to integrate information from large ecotoxicological and physico-chemical datasets, and apply it in a WoE context to predict fish acute toxicity of chemicals from data on fish embryo toxicity testing in combination with other types of information. This BN model was developed from data on fish embryo and juvenile acute toxicity in combination with other information for 237 chemicals, and aims to predict the toxicity interval of each chemical by combining information in four pathways (Figure 1): (i) fish embryo toxicity, (ii) physical and chemical properties, (iii) toxicity to fish of other chemicals in the same category, and (iv) toxicity to other species (algae and Daphnia). More details on this first version of the BN can be found in two papers [2, 3].

3. Results and discussion

3.1. Model performance under equal weighting of lines of evidence

Model performance was assessed by running the BN model with input data from a large number of substances and comparing the outcome, i.e. the predicted acute toxicity of selected chemical substances to juvenile fish, to the observed toxicity of the same substances to this endpoint. we applied four different criteria for selecting subsets of our dataset for validation in order to consider the robustness of the model from different perspectives. When the four lines of evidence were given equal weight, the model correctly predicted the most probable acute fish toxicity interval for 60-80% of the substances, depending on the criteria for data selection.

3.2. Calibration of the weights for lines of evidence

The weighting of the four (and later more) lines of evidence will be calibrated to further improve the model performance. Although the BN methodology allows for machine-learning methods to optimise such weighting (via the conditional probability tables), the data requirement would highly exceed the data availability. As a first step, we have tested a set of five alternative weight combinations: the Embryo line assigned the weights 0%, 25%, 50%, 75% and 100%, with the remaining weights equally distributed to the three other lines. Increasing weight to the Embryo line generally resulted in slightly higher precision (*i.e.*, percentage of correct predictions), as well as higher accuracy (*i.e.*, a better balance between over- and underestimation of toxicity).

4. Conclusions

We consider the model presented here as a good starting point for a probabilistic WoE approach, with high potential for further development and evaluation during the SWiFT project. Our planned next steps include: (1) revising the model structure to better represent causal relationships rather than associations; (2) incorporating more lines of evidence; (3) apply hierarchical Bayesian (statistical) modelling for better quantification of uncertainty within and among reported toxicity tests; (4) further develop the web-based user interface to the model. While the purpose of this model is predict acute fish toxicity from fish embryo toxicity, the approach can be relevant more generally for evaluating animal alternatives in regulatory toxicity testing, or even more generally for other types of environmental assessments.

5. References

[1] Moe SJ, Madsen AL, Connors KA, Rawlings JM, Belanger SE, Landis WG, Wolf R, Lillicrap AD. 2020. Development of a hybrid Bayesian network model for predicting acute fish toxicity using multiple lines of evidence. Environmental Modelling and Software 126:104655.

[2] Lillicrap A, Moe SJ, Wolf R, Connors KA, Rawlings JM, Landis WG, Madsen AL, Belanger SE. 2020. Evaluation of a Bayesian network for strengthening the weight of evidence to predict acute fish toxicity from fish embryo toxicity data. Integrated Environmental Assessment and Management 16:452–460.

[3] Kjærulff UB, Madsen AL. 2008. Bayesian networks and influence diagrams. Springer Science+ Business Media 200, 114.

Acknowledgement - This work is funded by the European Chemical Industry Council (CEFIC) as part of its Long-range Research Initiative (LRI) programme, project ECO51. The authors thank the project's Monitoring Team for useful feedback: Noemie Croze, Christopher Faßbender, Sylvia Gimeno, Marlies Halder, Sarah Hughes, Joop de Knecht, Mark Lampi, Wayne Landis, Teresa Norberg-King, Martin Paparella, Audrey Pearson, Eleonora Simonini, Marta Sobanska and Susanne Walter-Rohde. We also thank Ullrika Sahlin for constructive advice on improvement of the model structure.