An evaluation of selected in silico models for the assessment of skin sensitization potential – performance and practical utility considerations

Jeremy Fitzpatrick¹ and Grace Patlewicz¹

¹National Center for Computational Toxicology (NCCT), US Environmental Protection Agency, Research Triangle Park, NC, 27711, USA

Skin sensitization remains an important endpoint for consumers, manufacturers and regulators. Although the development of alternative approaches to assess skin sensitization potential has been extremely active over many years, the implication of regulations such as REACH and the Cosmetics Directive in EU has provided a much stronger impetus to actualize this research into practical tools for decision making. Thus there has been considerable focus on the development, evaluation, and integration of alternative approaches for skin sensitization hazard and risk assessment. This includes in silico approaches such as (Q)SARs and expert systems. This study aimed to evaluate the predictive performance of a selection of in silico models and then to explore whether combining those models led to an improvement in accuracy. A dataset of 473 substances that had been tested in the local lymph node assay (LLNA) was compiled. This comprised 295 sensitizers and 178 non-sensitizers. Four freely available models were identified - 2 statistical models VEGA and MultiCASE model A33 for skin sensitization (MCASE A33) from the Danish National Food Institute and two mechanistic models Toxtree’s Skin sensitization Reaction domains (Toxtree SS Rxn domains) and the OASIS v1.3 protein binding alerts for skin sensitization from the OECD Toolbox (OASIS). VEGA and MCASE A33 aim to predict sensitization as a
binary score whereas the mechanistic models identified reaction domains or structural alerts which may lead to skin sensitization. VEGA had an accuracy of 62% for the 310 substances which were not associated with experimental data. The MCASE A33 model was only able to make predictions for 212 substances, the remainder were outside of the applicability domain. It had an accuracy of 51%. The utility of the reaction domains from Toxtree and the alerts from OASIS were explored. 73% of substances firing a domain in Toxtree were sensitizers, whereas 59% of substances without a domain were non-sensitizers. 85% of the 184 substances with OASIS alerts were found to be sensitizing, for those with no alerts, 46% were found to be non-sensitizing. The VEGA, Toxtree, and OASIS predictions were then combined. Substances for which OASIS gave no prediction or VEGA contained experimental information were excluded. The combination model had an accuracy of 85% for the resulting set of 245 substances. Combining predictions from several models together results in a better overall performance than any one model on its own.

Disclaimer: The views expressed are those of the authors and do not necessarily reflect the views or policies of the US Environmental Protection Agency.