Machine Learning Approaches for Predicting Human Skin Sensitization Hazard

Judy Strickland[1], Qingda Zang[1], Michael Paris[1], Nicole Kleinstreuer[1], David M Lehmann[2], David Allen[1], Neepa Choksi[1], Joanna Matheson[3], Abigail Jacobs[4], Anna Lowit[5], Warren Casey[6]

[1]ILS/NICEATM, RTP, NC, USA; [2]EPA/ORD/NHEERL, RTP, NC, USA; [3]CPSC, Rockville, MD, USA; [4]FDA/CDER, Silver Spring, MD, USA; [5]EPA/OCSPP/OPP, Washington, DC, USA; [6]NIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

One of ICCVAM's top priorities is the development and evaluation of non-animal approaches to identify potential skin sensitizers. The complexity of biological events necessary for a substance to elicit a skin sensitization reaction suggests that no single alternative method will replace animal tests. We used data from the direct peptide reactivity assay (DPRA), human cell line activation test (h-CLAT), KeratinoSens assay, six physicochemical properties, and an *in silico* read-across prediction of skin sensitization hazard as inputs to two machine learning approaches to predict human skin sensitization hazard. Support vector machine and logistic regression were each applied with 12 feature set combinations. Models were trained on a set of 72 substances and tested on an external set of 24 substances. The feature set containing DPRA, h-CLAT, KeratinoSens, read-across, and log P performed the best for both approaches: accuracy = 99% (71/72), sensitivity = 98% (50/51), and specificity = 100% (21/21) for the training set; and accuracy = 96% (23/24), sensitivity = 93% (14/15), and specificity = 100% (9/9) for the test set. The performance of this integrated approach was better than the local lymph node assay or any of the *in chemico*, *in vitro*, or *in silico* methods alone. These data suggest that computational methods are promising tools to effectively identify potential skin sensitizers without animal testing. *This abstract does not represent EPA policy.* *This project was funded in whole or in part with Federal funds from the NIEHS, NIH under Contract No. HHSN273201500010C.*

Keywords:

1. Predictive toxicology
2. Computational toxicology
3. Adverse (toxicological) responses