

An evaluation of selected (Q)SARs/expert systems for the Prediction of Skin Sensitization Potential

Jeremy Fitzpatrick and Grace Patlewicz

Predictive testing to characterize substances for their skin sensitization potential has historically been based on animal models such as the Local Lymph Node Assay (LLNA) and the Guinea Pig Maximization Test (GPMT). In recent years, EU regulations have provided a strong incentive to develop non-animal alternatives – both in vitro and in silico. Here we selected three different types of expert systems: Derek Nexus (knowledge based), TIMES-SS (hybrid), and VEGA (statistical), and evaluated their performance using two large sets of animal data, one of 1249 substances from eChemportal (354 sensitizers and 895 non-sensitizers) and a second of 515 substances curated by NICEATM (329 sensitizers and 186 non-sensitizers). We considered a model to be successful at predicting skin sensitization if it had at least the same balanced accuracy as the LLNA and the GPMT had in predicting the outcomes of one another, which ranged from 79% to 86% depending on the dataset. We found that none of the expert systems evaluated was able to achieve such a high balanced accuracy in their global predictions, with balanced accuracies ranging from 56% to 65%. However, for substances within the domain of TIMES-SS, balanced accuracies were found to be 79% and 82% for the 2 datasets respectively, in line with the animal data. While no model performed as well as the animal skin sensitization tests globally, compounds within the domain of TIMES-SS were predicted with the same balanced accuracy as the animal results. This abstract may not reflect U.S. EPA policy.