A. Tropsha[1]

CURRENT TRENDS IN QSAR MODELING

[1] CB # 7568 Beard Hall, UNC Eshelman School of Pharmacy, UNC Chapel Hill, NC 27599

More than fifty years of continuous improvements, interdisciplinary breakthroughs, and community-driven developments were needed to make QSAR modeling one of the commonly employed approaches to modeling the physical and biological properties of chemicals in use today. In fact, the analysis of published literature indicates that the continuing growth of chemical data and databases especially in the public domain has stimulated the concurrent growth in QSAR publications. However, throughout its entire history the QSAR approach has drawn both praise and criticism concerning its reliability, limitations, successes, and failures. In this presentation, we will discuss: (i) the development and evolution of QSAR; (ii) the current trends, unsolved problems, and pressing challenges; and (iii) several novel and emerging applications of QSAR modeling [1]. Throughout the discussion, we will provide guidelines for QSAR development, validation, and application, which are summarized in best practices for building rigorously validated and externally predictive QSAR models. We emphasize the importance of communications between computational and experimental chemists towards collaborative development and use of QSAR models. We also address the issue of data accuracy and reproducibility that are particularly important for computational scientists such as bio- and cheminformaticians whose success inherently depends on the quality of experimental data used as inputs for their models. We stress that the exploitation of modern chemogenomics repositories containing huge sets of heterogeneous data requires the use of powerful, transparent, and robust data curation workflows. Supplementing and enriching our previous chemical curation protocols [2], we will describe the enhanced chemical and biological data curation workflow. This global data curation workflow can be utilized to improve the quality of data analysis and interpretation as well as boost the prediction performances of computational models built with available chemical genomics data. We will also discuss guidelines concerning the use of stringent scientific standards to manuscripts reporting new QSAR studies that should enable easier manuscript acceptance by journal reviewers and editors.

1. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R, Consonni V, Kuz'min VE, Cramer R, Benigni R, Yang C, Rathman J, Terfloth L, Gasteiger J, Richard A, Tropsha A. *J Med Chem*. 2014, **57**: 4977-5010.
2. Fourches D, Muratov E, Tropsha A. *J Chem Inf Model*. **2010**, 50: 1189–1204.

The support from NIH grant GM066940 is acknowledged.