DSSTox and Chemical Information Technologies in Support of Predictive Toxicology

Ann M. Richard, National Center for Computational Toxicology, USEPA, RTP, NC

The EPA NCCT Distributed Structure-Searchable Toxicity (DSSTox) Database project initially focused on the curation and publication of high-quality, standardized, chemical structure-annotated toxicity databases for use in structure-activity relationship (SAR) modeling. In recent years, the project has expanded to include: creation of DSSTox files for high-interest EPA chemical inventories; strengthening structure-based linkages among public resources; tailoring chemical and bioassay DSSTox content for incorporation into NIH's PubChem; creating local structure-browsing capabilities of DSSTox content and inventories; and expanding comparability of and linkages to gene expression data. Within the NCCT, the DSSTox project framework is applying strict quality standards for chemical information, pertaining to both generic (ACToR, ToxRefDB) and actual test substances (ToxCast^{TM}, Tox21). Within these projects, we are working to expand comparability and linkages of summarized toxicity data in the context of a standardized cheminformatics environment. Future research will build on these cheminformatics data foundations and enriched data resources to develop data mining strategies to explore new and flexible ways to relate chemical structure to biological endpoints (e.g., reactivity groupings, biofunctional or toxicity-informed similarity, chemical feature space), and new representations of biological endpoints in relation to structure-based modeling (e.g., HTS clusters, bioassay profiles, summarized or grouped effects, qualitative active and inactive classes). Incorporating traditional SAR concepts into this new HTS data-rich world poses conceptual and practical challenges, but also holds great promise for improving toxicity prediction capabilities. *This work was reviewed by EPA and approved for publication but does not necessarily reflect official Agency policy.*