

Assessing the Robustness of Chemical Prioritizations Based on ToxCast Chemical Profiling

Ander Wilson, Sumit Gangwal, Matthew Martin, Richard Judson, David Dix, David Reif

National Center for Computational Toxicology, Office of Research and Development, U.S. EPA, RTP, NC

A central goal of the U.S. EPA's ToxCast™ program is to provide empirical, scientific evidence to aid in prioritizing the toxicity testing of thousands of chemicals. The agency has developed a prioritization approach, the Toxicological Prioritization Index (ToxPi™), that calculates a comprehensive toxicity potential and a relative priority rank by incorporating information from ToxCast *in vitro* bioactivity data (high-throughput screening results from over 500 diverse assays), inferred toxicity pathways, *in vitro* to *in vivo* dosimetry estimates, chemical structural descriptors, and exposure considerations. Here, we explore the robustness of the prioritization assessing potential endocrine activity of 309 chemicals in the face of several sources of variation: 1) changes in the chemical makeup of the experiment, 2) missing data, and 3) spurious (false-positive) assay results. Bootstrap resampling was used to assess the effects of alternative chemical sets. Although missing data was not an issue in Phase I, it may be a concern in subsequent phases and in certain data domains (*e.g.* exposure data). To address this concern, we simulated both missing-at-random and missing-by-domain datasets for comparison with complete data. A similar approach was taken to assess the potential impact of false-positive assay results. Generally, the higher-scoring chemicals tended to be less sensitive to alternative chemical sets but were more sensitive to missing values and false positives than lower-scoring chemicals. However, initial results for all experiments showed 95% confidence intervals with mean width representing less than one decile, indicating that the multivariate endocrine rankings are relatively stable in the face of anticipated levels of common sources of data variation. This robustness, which is essential to a reliable prioritization scheme, arises out of the comprehensive nature of the scores, in that no single datum wields ultimate influence. *This abstract does not necessarily reflect Agency policy.*