Multivariate analysis of toxicity experimental results of environmental endpoints.

Mansouri K., Martin M., Judson R.,

National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, RTP, NC, USA

The toxicity of hundreds of chemicals have been assessed in laboratory animal studies through EPA chemical regulation and toxicological research. Currently, over 5000 laboratory animal toxicity studies have been collected in the Toxicity Reference Database (ToxRefDB). In addition, the ToxCast research program released over 800 high-throughput screening (HTS) bioassays for over 1800 chemicals across 3 multi-year phases.

In order to optimize the use of these $ billions worth *in-vivo* and $ millions worth *in-vitro* studies in the toxicity assessment of chemicals, a more understanding of the information encoded in these databases is needed. In this work, different multivariate analysis techniques were employed in the learning process about the relationships between the different layers of information represented by the chemicals, the bioassays and the biological targets.

The results provided by these methods; Principal Component Analysis (PCA) Kohonen-Networks Self Organising Maps (SOM) and Spanning trees were interpreted and used to map the linkage between chemical structures and biological activity. This analysis aims to better understand the relationship between chemical structure, bioactivity and *in vivo* response.

This abstract does not necessarily reflect U.S. EPA policy.