A novel method for combining chemically induced outcomes with different gene-level classifications

Main author: Salomon Sand (Swedish Food Agency)

Co-authors:

INTRODUCTION

Short-term studies in rodents that include high-throughput transcriptomics as an alternative data stream have been considered to serve as a bridge between in vitro approaches and traditional animal toxicity studies. The lowest gene-set benchmark doses (BMDs) derived from such data have been shown to correlate reasonably well to corresponding BMDs from resource-intensive guideline studies in the same species. It has therefore been suggested that in vivo gene-set BMDs may be used for MOE-based prioritisations when guideline studies are not available. The National Toxicology Program defines gene level potency in terms of median BMDs, which provide a central estimate for each gene-set within an analysis. Along the lines of traditional risk assessment the lowest median BMD has been used as a form of overall reference point (also called point of departure) for comparative purposes. The present work advances this by combining all gene level BMDs to support quantitative hazard and risk characterisation.

METHODOLOGY

A reference point profile (RPP) rather than a single reference point is herein derived that describes variability within and across gene-sets, making use of all individual BMDs. The RPP is estimated by an iterative approach. In each iteration a value is simulated from every unique BMDs uncertainty distribution. The gene-set allocation for simulated BMDs is randomised across iterations, since a gene may be part of several gene-sets, and combined with rank values between 0 and 1 that describe the fraction of BMDs exceeded within associated gene-sets. A single iteration provides a snapshot along BMD sequences across all gene-sets, and fitting the RPP model, gives one set of estimates for the location (potency), shape (within gene-set variability), and standard deviation (between gene-set variability). At a given exposure the probability for exceeding the entire RPP can be assessed, describing the overall gene-level disturbance, and this response can also be summarised by integrating contributions across ranks. Conversely, dose equivalents corresponding to specified responses can be estimated. This method includes a weight function that can modulate the results of rank integration.

RESULTS

The observed variation in RPPs across chemicals indicate that joint consideration of multiple effects can modify conclusions on potency. Evaluation of the integrated response at doses corresponding to traditional potency estimates does not result in the same impacts since the RPP shape and standard deviation differ across chemicals. Also, the uncertainty in dose estimates from the RPP model was clearly lower than that associated with the mean/median log BMD for specific gene-sets. The increase in dose-response information that follows when moving away from apical response enhances the opportunity to derive less uncertain exposure guidelines by the combination of data. As noted earlier the RPP describes the sequence of BMDs within and across gene-sets. Individual BMDs are in this context ranked in the same manner within the different gene-sets. However, results from comparison of semantic similarity between gene-sets (i.e. gene ontology categories) indicated that gene-sets that are more distant in dose location are less similar, suggesting that such information may be used to refine the ranking of gene-sets, and possibly also guide parameterisation of the attached weight function.

DISCUSSION

Determination of exposures below in which no significant perturbation of toxicity pathways occur, in line with National Research Council's vision, is conceptually similar to the definition of the critical effect as the first adverse effect or known precursor. Moving away from apical response may, however, to a higher extent require probability to be part of the consideration of what is a significant alteration e.g. within the network of toxicity pathways. The probability of exceeding the RPP provides a quantitative description of the extent of gene-level disturbance, and evaluating the gradual increase of this effect allows for impact assessment above traditional guidance values that are often exceeded. Overall, the increasing amount of effect data that may result when focus is shifted upstream from adverse apical response, more strongly than before, promotes or requires methods for combining data and evaluation of the total impact. In this process the greater context described by e.g. gene ontology allows for quantitative comparison of effects using bioinformatics tools, instead of value-based comparison of apical/traditional outcomes, which may better facilitate data integration.