

Summary

A method for deriving transcriptomic BMDs using BMRs based on relative effect size rather than standard deviation of the data was compared to the method specified in the US EPA ETAP document. The proposed method performs slightly better based on multiple concordance metrics and has higher replicability of results.

Associated files:

GO_replicates_EPA_DTT.xlsx

TranscriptomicMethodComparison_Tables.pdf

TranscriptomicMethodComparison_Figures.pdf

DTT_tBMD_pipeline.bm2

DTT_tBMD_pipeline_filtered.txt

EPA_tBMD_pipeline.bm2

EPA_tBMD_pipeline_filtered.txt

SyntheticKidney.bm2

SyntheticLiver.bm2

Synthetic_FDR_Results.xlsx

Data

Data were transcriptomic dose-responses from five-day gavage studies of twenty test substances. [Full details elsewhere]. Four of the chemicals were tested in replicate studies: BDCA, furan, PFOA, and TBBPA. All four were tested in 3 studies. However, the results of the first study for TBBPA were discarded due to poor quality.

Methods

DTT method

The proposed DTT method for deriving tBMDs is as follows.

Curve fitting

Dose-response data were log₂-transformed. Probe data that had values of 0 for any response after the log₂ transformation were omitted.

Data were prefiltered in BMDEExpress3 version 3.20.0141 BETA (<https://github.com/auerbachs/BMDEExpress-3/wiki>) using a Williams trend test (requiring a p value of less than 0.05), and then a curve fit prefilter requiring a response of 1.581 standard

deviations. Probe BMDs were derived by curve fitting using a BMR of 25% relative deviation, using the ToxicR model averaging method in BMDEExpress3.

Gene BMDs

Gene BMDs were derived from probe BMDs using the Category Analysis/Individual Gene Analysis tool in BMDEExpress3. The following criteria for inclusion of probes were used:

1. R^2 for the curve fit ≥ 0.6
2. BMD/BMDL < 10
3. No step function > 0.5 of total response if BMD $<$ the lowest non-control dose
4. Promiscuous probes were NOT removed

GO biological process (GO BP) BMDs

Gene BMDs were derived from probe BMDs using the Category Analysis/Gene Ontology Analyses tool in BMDEExpress3. Probe-level criteria given above for the gene BMDs were used in this calculation. GO BPs used in the analysis were ones with from 40-500 genes on the platform, with at least 3 genes with BMDs in the dose range, and at least 5% of the genes in the BP having BMDs in the dose range.

The gene BMD at the 5th percentile of all genes in a GO BP was used as the BMD for that GO BP. The lowest GO BP BMD across all GO BPs for both kidney and liver data was used as the tBMD for the chemical tested. The tBMDL was calculated using the same method but using the gene BMDL values.

The BMDEExpress file containing the analysis is DTT_tBMD_pipeline.bm2. Output of the GO BP analysis is in DTT_tBMD_pipeline_filtered.txt. Each row of data contains results for one chemical replicate, one tissue, and one GO BP. The data columns used to generate the tBMDs and tBMDLs are 'BMD at 5th Percentile of Total Genes' and 'BMDL at 5th Percentile of Total Genes'. GO BPs in this file are only those that satisfied the criteria (above) for number and percentage of genes in the GO BP having BMDs in the dose range.

EPA method

The U. S. EPA has developed a method for deriving transcriptional points of departure (tPODs) (U. S. EPA 2024) for use in EPA transcriptomic assessment products (ETAPs). Their procedure chooses a tPOD based on the BMDs of genes in GO biological processes (GO BPs). The tPOD is the BMDL associated with a transcriptomic BMD, namely, the BMDL for the GO BP with the lowest BMD. For purposes of this analysis, we will compare the tBMDs (BMDs for the GO BP with the lowest BMD) with the tBMDs from the proposed DTT method.

BMDEExpress version 2 was used to carry out the calculations for the EPA method. The BMDEExpress2 file is EPA_tMBD_pipeline.bm2. GO BP analyses from this file are saved in EPA_tMBD_pipeline_filtered.txt. This includes only those GO BPs with at least 3 genes passing the filters (as specified in U. S. EPA 2024).

Apical BMDs

Gwinn et al. (2020) derived neoplastic and non-neoplastic apical endpoint BMDs from histopathological endpoints in chronic and subchronic studies for 18 of the 20 studied chemicals. BMDs were based on 10% extra risk. To supplement these results, additional sources of dose-response for apical endpoints were consulted and (in some cases) analyzed. In some cases, the available data required analysis using methods for continuous data. In those cases, the BMD was based on 10% change in response. In one case (for triclosan) data could not be obtained from the literature but a BMD for 20% change in response was reported. Transcriptomic and apical BMDs and BMDLs are shown in Table 1. Lowest apical BMDs for both neoplastic and non-neoplastic endpoints are shown. The apical BMD used for comparison with transcriptomic BMDs was the lower of the neoplastic and non-neoplastic values, shown in boldface in Table 1.

Comparisons between replicates

Results of the two tBMD methods were compared across replicate chemical studies. There were 2 replicate studies used for TBBPA and 3 for BDCA, furan, and PFOA.

Comparison 1: The range of tBMD values across replicates was examined.

Comparison 2: GO BPs satisfying the criteria for inclusion in the tBMD calculation were compared across replicates, with replicability of kidney and liver data considered separately. GO BPs were ranked by the value of their BMDs. Ranks were examined up to rank 150 (with rank 1 being the lowest BMD). Agreement between the replicates at ranks $r = 5, 10, \dots, 150$ or less was measured as follows.

1. The GO BPs with rank $\leq r$ were listed for each replicate.
2. For each of those GO BPs, the number of times it occurred in the lowest- r list across replicates was calculated, giving a replicate count of 1-3 (maximum value of 2 for TBBPA and 3 for the other chemicals).
3. A concordance score was calculated as $(1/r) \times (\text{the number of GO terms that had a replicate count} = \text{the maximum possible value of 2 or 3})$.

Comparison of transcriptomic and apical BMDs

Transcriptomic and apical metrics were compared using concordance metrics from Weitekamp et al. 2025. Values were calculated both for the full set of 20 chemicals and for the subset of chemicals used in the data analysis performed during the development of the ETAP. For the latter, the chemicals omitted were BPAF, fenofibrate, ginseng, milk thistle extract, and triclosan. The chemicals not omitted are underlined in Table 1. These metrics were calculated using only the first of the three replicate studies for the replicated chemicals BDCA, furan, and PFOA and the second study for TBBPA.

The concordance metrics from Weitekamp et al. 2025 can be divided into three groups

1. Mean absolute difference, RMSD, and residual RMSE: a lower value indicates a better fit to the data.
2. Correlation, rank correlation, R^2 : higher value indicates a better fit to the data.
3. Mean signed difference: a value closer to zero indicates a better fit to the data.

Simulated null data

False discovery rates (FDRs) for the DTT method were calculated by applying the method to data generated by sampling from the estimated distribution of the control data, omitting data that had responses (on the log₂ scale) of zero. Data and analysis results (filtering, BMD calculation, GO BPs) are in BMDEpress3 files SyntheticKidney.bm2 and SyntheticLiver.bm2.

Results

Replicability

Figure 1 shows the tBMDs for all replicates for each replicated chemical. For PFOA, the chemical with the lowest tBMD, the fold change between the lowest and highest replicate BMDs was slightly larger for the DTT method than for the EPA method. For furan, with the second highest tBMD, the fold change between lowest and highest tBMD was much lower for the DTT method. For BDCA and TBBPA, the fold change was higher for the DTT method.

Table 2 shows the number of GO terms with rank ≤ 150 for each chemical/tissue/replicate method. The ranking algorithm assigns averaged values to tied numbers, so not all ranks are integers. Therefore, the number of GO BPs ranked ≤ 150 is not always 150 even when there are 150 or more GO BPs with BMDs. Also, the EPA criteria for inclusion of GO BPs are less strict than those for the DTT method. There are many cases with the DTT method where there are not 150 GO BPs.

Figure 2 shows the scores for concordance across replicates for PFOA in both kidney and liver, and for furan and TBBPA in liver. Concordance scores are higher for the DTT method.

The file GO_replicates_EPA_DTT.xlsx has tables of low-ranking (i.e. low-BMD) GO BPs for all chemical-tissue pairs except TBBPA in kidney. Results from both the DTT and EPA method are shown. There are tables of GO BPs at rank 25 or less. For the more strongly responding chemical-tissue pairs (furan and TBBPA in liver, and PFOA in both kidney and liver), there are also tables of GO BPs at rank 50 or less. There are generally many more replicated GO BPs with the DTT method than with the EPA method.

Transcriptomic and apical BMDs

Metrics from the comparison of transcriptomic to apical BMDs are shown in Table 3. All metrics were calculated using the log₁₀ of the apical and transcriptomic BMDs. Figures 3-5 are scatter plots of (3) DTT vs apical BMDs; (4) ETAP vs apical BMDs; (5) DTT vs ETAP transcriptomic BMDs.

The DTT method gives lower values for the mean absolute difference, the RMSD, and the residual RMSE than the EPA method does.

Correlations are slightly higher for the DTT method, while rank correlations are slightly higher for the EPA method. R^2 is higher for the DTT method.

Mean signed difference is closer to zero for the DTT method.

False discovery rate

Results of the false discovery rate calculations are shown in Synthetic_FDR_Results.xlsx. The fraction of simulated probe dose-response curves remaining after each step in the analysis (prefiltering and curve fitting) is given.

About 1/216 probes in the kidney data set and 1/236 in the liver data set had BMDs in the dose range after filters were applied. These values are calculated relative to the number of simulated probes after probes with zeros at any point are omitted. About 1/283 genes in the kidney data set and 1/315 in the liver data set had BMDs in the dose range. These values are calculated relative to the number of possible genes represented in the probes in the data set.

Out of 40 data sets (20 each for kidney and liver), 37 had no GO BPs that satisfied our criterion of 3 or more active genes and 5 percent or more of genes in the GO BP active. The other 3 data sets each had 2 GO BPs satisfying the criteria. For comparison, all the chemicals in the 5-day study had at least 5 such GO BPs for at least one tissue, and all but two of the chemicals had at least 11 such GO BPs.

Discussion

Replicability

As shown in Figure 1, the range of tBMDs across replicates is similar for all chemicals when using the EPA method. The DTT method gives a smaller range of tBMD values for the most potent chemicals, PFOA and furan, but a larger range for the less potent chemicals. Comparison of replicated GO BPs shows much higher replicability of low-BMD GO BPs for the DTT method than for the EPA method (Figure 2 and tables in GO_replicates_EPA_DTT.xlsx). At low numbers of ranks compared, PFOA in liver shows the best replicability of GO BPs.

Comparison with apical BMDs

Concordance metrics for the two methods are mostly similar, with the relative performance of the methods depending on the chemical set used for comparison.

The proposed DTT method performs slightly better based on multiple concordance metrics and has higher replicability of results.

False discovery rate

The false discovery rate is low. No simulated null data set had more than 2 GO BPs with the required number of genes to be included in the tBMD calculation. However, the method for simulating data does not take into account correlation between probes, which may affect the results.

References

U. S. EPA 2024, Standard Methods for Development of EPA Transcriptomic Assessment Products (ETAPs).

Weitekamp et al., 2025. Quantitative and qualitative concordance between clinical and nonclinical toxicity data. *Toxicological Sciences*, 2025, 206(2), 253-272.

Software used

BMD analysis was performed using BMDEExpress 2 (<https://github.com/auerbachs/BMDEExpress-2/wiki>) and BMDEExpress3 (<https://github.com/auerbachs/BMDEExpress-3/wiki>).