Variability in organ-level effects in repeat dose animal studies

Miran (MJ) Foster[1,2*], Richard S. Judson[1], R. Woodrow Setzer[1], Katie Paul Friedman[1],

[1]Center for Computational Toxicology and Exposure, US EPA, Research Triangle Park, NC USA ; [2]Oak Ridge Associated Universities

*Miran (MJ) Foster | foster.miran@epa.gov

Abstract 4028/Poster P726
March 29, 2022
Society of Toxicology Virtual Annual Meeting

www.epa.gov

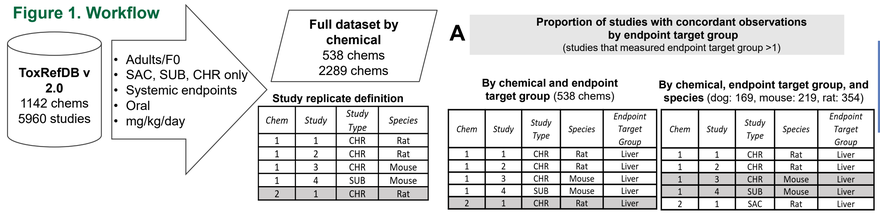
This poster does not necessarily reflect EPA policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Building scientific confidence in the use of new approach methodologies (NAMs) in safety assessment may include performance comparison to *in vivo* study outcomes. This work defines the variability in organ-level effects and suggests qualitative and quantitative benchmarks for maximum NAM performance for prediction of organ-level effects in repeat dose studies of adult animals. Previous work suggests that the root mean square error (RMSE) for study-level lowest effect level (LEL) values (on a $\log_{10}$-mg/kg/day basis) approaches 0.5 $\log_{10}$-mg/kg/day†.

Observations of liver, kidney, stomach, spleen, thyroid and adrenal gland from the Toxicity Reference database (v2.0) were included in this analysis to understand the reproducibility of organ-level effects.

Abbreviations Key
CHR=Chronic; SUB=Subchronic; LEL=Lowest effect level; AED=Administered equivalent dose

Figure 1. Workflow



A **Proportion of studies with concordant observations by endpoint target group**
(studies that measured endpoint target group >1)

By chemical and endpoint target group (538 chems)

Chem	Study	Study Type	Species	Endpoint Target Group
1	1	CHR	Rat	Liver
1	2	CHR	Rat	Liver
1	3	CHR	Mouse	Liver
1	4	SUB	Mouse	Liver
2	1	CHR	Rat	Liver

By chemical, endpoint target group, and species (dog: 169, mouse: 219, rat: 354)

Chem	Study	Study Type	Species	Endpoint Target Group
1	1	CHR	Rat	Liver
1	2	CHR	Rat	Liver
1	3	CHR	Mouse	Liver
1	4	SUB	Mouse	Liver
2	1	SAC	Rat	Liver

By chemical, endpoint target group, and study type (dog: 169, mouse: 219, rat: 354)

Chem	Study	Study Type	Species	Endpoint Target Group
1	1	CHR	Rat	Liver
1	2	CHR	Rat	Liver
1	3	CHR	Mouse	Liver
1	4	SUB	Mouse	Liver
2	1	SAC	Rat	Liver

B **Variance analysis on subsets by endpoint target group**
(studies that measured endpoint target group >1)

Method: Multilinear regression (MLR)

Descriptors used for LEL data by organ:
- Study type
- Species
- Administration method
- Dose number
- Dose spacing
- Substance purity
- Study year

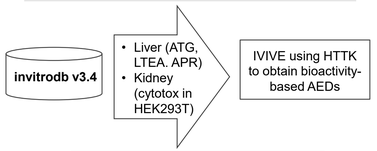
Used to calculate total variance =
Unexplained variance (MSE)
+
Explained variance

C **Analysis of differences of SUB and CHR findings by endpoint target group, paired by chemical**

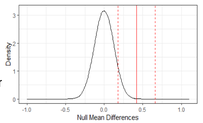
Method 1: Odds Ratios
- For each of the 6 endpoint target groups and species, filter by chemicals that have both study types present.
- Calculate the odds ratio for CHR and SUB.
- Convert odds ratio to probability of a positive in CHR given a positive in SUB.

Method 2: Paired Randomization Test
- For each of the 6 endpoint target groups, filter by chemicals that have both study types present.
- Calculate log10 differences of LELs.
- Perform a paired randomization test to check for significant differences in the distributions of SUB/CHR LELs.

D **Comparison of PODs for liver and kidney for bioactivity-based AEDs**

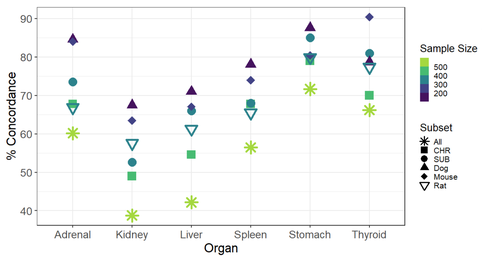


Method: Paired Randomization Test
- For liver and kidney, gather chemicals that have both LEL and AED data.
- Calculate log10 differences between LELs/AEDs.
- Perform a paired randomization test to check for significant differences in the distributions of SUB/CHR LELs.



A: Qualitative reproducibility of organ-level effect observations in repeat dose studies of adult animals was 33-88%, depending on grouping.

Figure 2. Repeated concordance of organ-level findings



$$\%\ Concordance = \frac{chemical\ with\ positive\ finding\ in\ all\ studies + chemicals\ with\ negative\ finding\ in\ all\ studies}{total\ chemicals\ tested}$$

- Organs associated with more negative chemicals (stomach, thyroid, adrenal) had higher rates of concordance in this range.
- Within-species concordance tended to be greater than within-study concordance

B: Variance in organ-level effects in repeat dose studies was smaller than study level variance, but the organ-level effect RMSE was similar to study-level RMSE and approaches ~0.5 $\log_{10}$-mg/kg/day.

Table 1. Results of MLR to estimate unexplained and explained variance in organ LELs

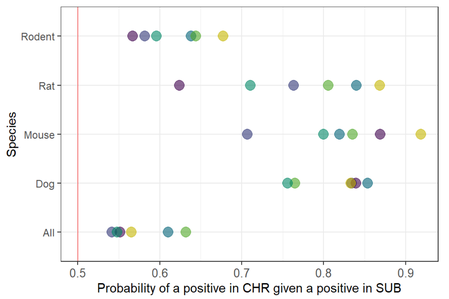
$$organLEL \sim b_0 + chemical * b_1 + species * b_2 + study\ type * b_3 + administration\ method * b_4 + dose\ spacing * b_5 + number\ of\ dose\ levels * b_6 + study\ year * b_7 + \%\ substance\ purity * b_8$$

Endpoint Target Group	Chem	N	Var	MSE	RMSE	% var explained
adrenal	81	208	0.756	0.349	**0.591**	53.8
kidney	263	790	0.765	0.316	**0.562**	58.7
liver	359	1318	0.745	0.355	**0.596**	52.3
spleen	127	336	0.671	0.318	**0.564**	52.6
stomach	55	146	0.553	0.173	**0.416**	68.7
thyroid	73	198	0.721	0.378	**0.615**	47.6

Chems = # chemicals; N = number of studies; Var = total variance; MSE = mean square error on the model; RMSE = root residual mean square error; % var explained = % of total variance explained.by study descriptors

C. Organ-level findings in SUB appear qualitatively predictive of organ-level findings in CHR studies (ignoring adversity).

Figure 3. Probability of a positive CHR outcome by organ/species subgroups



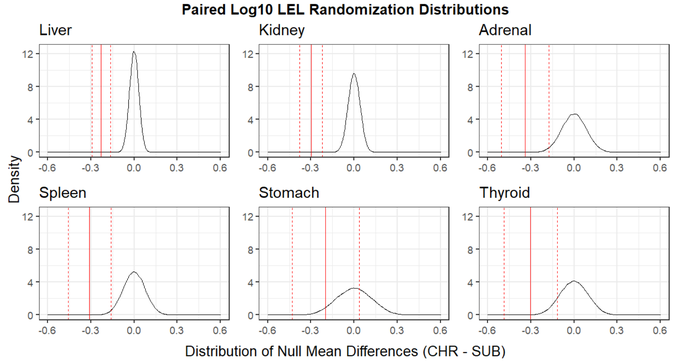
- Odds ratios are converted to probabilities using the formula:

$$P = \frac{OR}{1 + OR}$$

- The red line represents 0.5 probability, which would indicate having information about SUB would inform CHR no better than random.
- Note that "All" and "Rodent" sample sizes were significantly reduced because of the matching procedure.
- An "All" chemical must have been present in all species groups and both study types.

C. SUB organ-level LEL values are typically within 0.5 log10-mg/kg/day of CHR organ-level LEL values; for some organs, available data suggests that SUB and CHR studies produce similar LEL values.

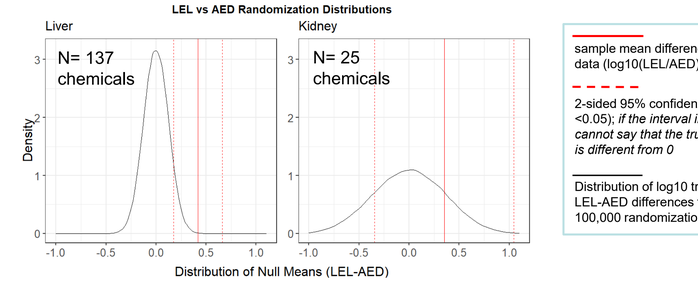
Figure 4. Matched randomization distributions for CHR/SUB LELs by organ



- Liver, kidney, spleen, and thyroid have significantly different log10 LEL difference from 0 at the α = 0.05 level [95% confidence interval (red dotted lines) does not include 0].
- Differences between CHR and SUB organ-level LELs may be within estimates of replicate study variance (0.5 log10-mg/kg/day).

D. *In vivo* LELs are higher than bioactivity-based AEDs on average for liver, and they may be within 0.5 $\log_{10}$-mg/kg/day for liver and kidney predictions.

Figure 5. Matched randomization distributions for LEL/AED differences



- Liver LEL and AED values are significantly different at the α = 0.05 level [95% confidence interval (red dotted lines) does not include 0], but the difference is well within estimates of variance in organ-level LELs (~0.5 log10-mg/kg/day).
- No significant difference for LEL-AED in kidney, the interval overlaps both 0 and 0.5, but N is small.
- Chemical diversity is limited for this preliminary result. The primary ToxPrint chemotypes represented are benzene ring and other aromatic bonds and chains.

U.S. Environmental Protection Agency
Office of Research and Development

†Based on Pham LL, Watford S, Pradeep P, Martin MT, Thomas RS, Judson RS, Setzer RW, Paul Friedman K. Accepted. "Variability in in vivo studies: Defining the upper limit of performance for predictions of systemic effect levels." *Computational Toxicology*. https://doi.org/10.1016/j.comtox.2020.100126