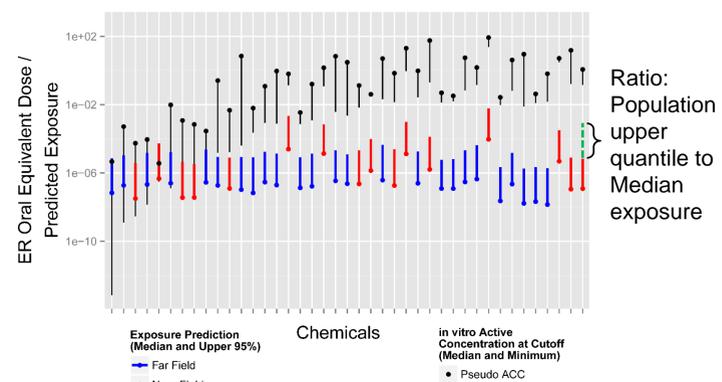


Introduction

- The ExpoCast models (Wambaugh et al, 2013; Wambaugh et al, 2014) predict population *median* exposures with confidence intervals.
- High-throughput approaches are being considered for roughly estimating hazard (Judson et al, 2011)
- A high-throughput, margin-of-exposure-like approach could be developed, if we could put bounds on the likely population variability, so the lower bound of predicted effective doses could be compared to an upper percentile of exposure variability.



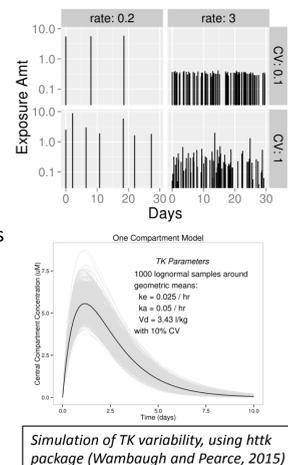
Exposure forecasts for median exposure (median estimate and upper 95% confidence limit on median exposure) for some ToxCast chemicals (Wambaugh et al, 2013), and predictions of human equivalent exposures resulting in estrogen receptor activity based on ToxCast assays (in prep).

Proposition

- Use population variability of urinary analytes from biomonitoring from, e.g., the National Health and Examination Survey (NHANES) for 45 chemicals to get a plausible distribution of the ratio of population median (P50) to upper quantiles (e.g. 90, 95, 99 %-iles: P90, P95, P99) of exposure.
- Use the distribution (and estimation uncertainty) to quantify the ratio between the population median exposure and various upper bounds of population exposure, for 90%, 95%, 99% of chemicals.

Contributors to Urinary Biomarker Variation

- Exposure variability:
 - Magnitude of exposure events
 - Interval between exposure events
 - Among individuals
- Toxicokinetic variability among individuals
 - Absorption
 - Clearance
 - Volume of distribution
 - Individual variation in renal function parameters
- Sampling
 - Timing of sample relative to previous exposures



Ratio Urine Conc. P95/P50 to Exposure P95/P50

		Elimination half-life / exposure interval				
GSD(dose)	CV (%)	0.083	0.25	0.5	1	2
Dose rate only varies among people						
1	0	8.2	2.3	1.5	1.2	1.1
2	78.5	4.1	1.4	1.1	1	1
3	153.1	3.4	1.3	1.1	1	1
4	241.5	3	1.2	1.1	1	1
Dose rate varies among and within people						
1	0	8.2	2.3	1.5	1.2	1.1
2	78.5	3.9	1.3	0.9	0.7	0.6
3	153.1	3.2	1.1	0.7	0.5	0.5
4	241.5	2.9	1	0.6	0.5	0.4

From Aylward et al, (2012). Overestimate of exposure variation by variation in urine concentrations decreases as:

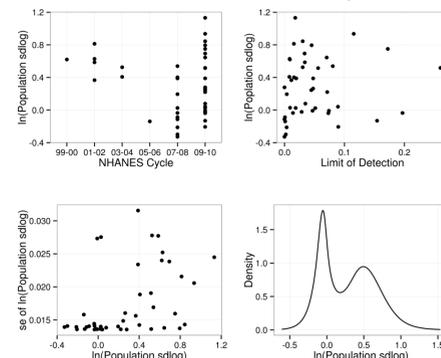
- elimination half-life increases
- exposure variance increases, and
- correlation of measurements within people decreases.

Going Forward:

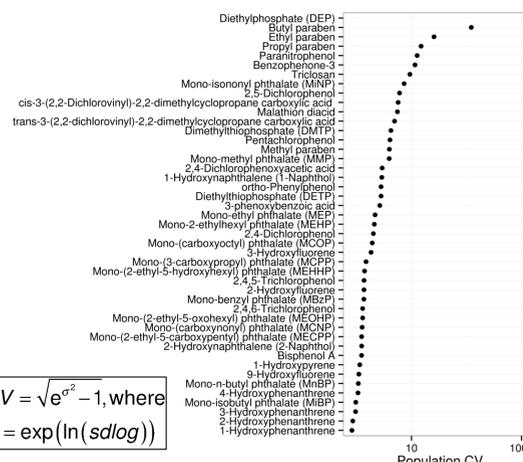
- Variance among urinary samples is likely to be greater than actual exposure variance, once sources other than sampling are considered.
- Using urinary sample variance to extrapolate from predicted medians should overestimate the upper quantiles of the exposure distribution.
- Better understanding of the kinetics for individual chemicals may allow reduction of this overestimate.

Methods and Results

- Select urinary analytes from NHANES datasets (Wambaugh, et al, 2014, and references therein)
- Using appropriate weighting procedures (to account for NHANES cluster sampling: Lumley 2004, 2014) and censored data techniques (to account for observations < detection limit), fit lognormal distribution to individual data for each analyte (more detail in Wambaugh et al., 2014), estimating *meanlog* and *ln(sdlog)* (the natural logarithm of the sd of *ln(observations)*)
- Urinary analytes are often metabolites of the parent compound to which exposure occurred. CV estimates are matched to the corresponding parent compounds



Several NHANES cycles (2 years each) are used for the estimates. Estimates of sdlog vary somewhat among cycles, for unexplained reasons. There seems to be little relationship between sdlog estimates and detection limits. Standard errors of the estimates are small, but are larger for larger values. The estimates of the *ln(sdlog)* values over all the urinary analytes considered appears to be bimodal.



$$CV = \sqrt{e^{\sigma^2} - 1}, \text{ where } \sigma = \exp(\ln(sdlog))$$

Population CV, computed from *ln(sdlog)*, estimated for each urine analyte

Population CV of urine analytes, plotted against associated parent compounds. Note that, multiple analytes mapping to the same parent compound have similar estimated CVs.

- For chemicals with multiple estimates of *ln(sdlog)*, estimate a single value as a weighted mean of the estimates, using reciprocal squared standard errors as weights.
- Use the logspline smoother to estimate the distribution of values of *ln(sdlog)* among chemicals (Kooperberg, 2013).
- Estimate confidence limits for population variability ratios using the bootstrap.
- If Q_p is the $p \times 100$ -ile of the standard normal distribution, the ratio of the $100p^{th}$ percentile to the median of the exposure distribution is $\exp(Q_p \times sdlog)$.

Population %-ile	Chemical %-ile		
	90%	95%	99%
90%	21.4 (11.2 - 41)	35.4 (14.2 - 88.1)	115 (22.1 - 600)
95%	51.1 (22.2 - 117)	97.2 (30.1 - 314.0)	442 (53.2 - 3680)
99%	261.0 (80.4 - 845)	647.0 (124.0 - 3390)	5520 (276.0 - 110,000)

The distribution of sdlog among chemicals, as measured by the ratios of exposure percentiles to the median exposure. For example, the 95th percentile of sdlog among chemicals corresponds to a population 95th percentile to median ratio of 97.2, which has 95% confidence limits of 30.1 and 314.0

Conclusions

- Protection of 95% of the population against 95% of chemicals requires multiplying estimates of population medians by 100 – 300.
- It is likely that, with more information about chemicals, including even crude estimates of TK parameters, the multiplier may be decreased.

Caveats and Future Work

- This is all based on the idea that population exposures follow (approximately) lognormal distributions. It is likely that real distributions are mixtures of approximately lognormal distributions.
- Refine treatment of parent chemicals with multiple urinary analytes, and chemicals that share urinary analytes.
- Expand datasets by looking at multiple NHANES cycles, expanding to analytes in blood as well as urine.
- It may be possible to improve the upper bound on exposure variability by modeling the likely exposure process and subtracting out an estimate of the variance that comes from using a single sample per individual. (e.g., Aylward, et al, 2012).

References

Aylward L, Kirman C, Adgate J, McKenzie L, Hays S. 2012. Interpreting variability in population biomonitoring data: Role of elimination kinetics. *JEEF* 22(4): 398-408.

Aylward L, Hays S, Smolders R, Koch H, Cocker J, Jones K, Warren N, Levy L, Bevan R. 2014. Sources of variability in biomarker concentrations. *J Toxicol. Environ. Health. Part B, Critical Reviews* 17(1): 45-61.

Judson R, Kavlock R, Setzer R, Cohen Hubal E, Martin M, Knudsen T, Houck K, Thomas R, Wetmore B, Dix D. 2011. Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chem. Research Toxicol.* 24(4): 451-462.

Kooperberg C. 2013. logspline: Logspline density estimation routines. R package version 2.1.5. <http://CRAN.R-project.org/package=logspline>

Lumley T. 2004. Analysis of complex survey samples. *Journal of Statistical Software* 9(1): 1-19.

Lumley T. 2014. "survey: analysis of complex survey samples". R package version 3.30. <http://CRAN.R-project.org/package=survey>

Wambaugh J, Setzer R, Reif D, Gangwal S, Mitchell-Blackwood J, Arnot J, Jollet O, Frame A, Rabinowitz J, Knudsen T, Judson R, Eggeghy P, Vallero D, Cohen Hubal E. 2013. High-throughput models for exposure-based chemical prioritization in the ExpoCast project. *ES&T* 47(15): 8479-8488

Wambaugh J, Wang A, Dionisio K, Frame A, Eggeghy P, Judson R, Setzer R. 2014. High throughput heuristics for prioritizing human exposure to environmental chemicals. *ES&T* 48(21): 12760-12767.

Wambaugh J, Pearce R. 2015. *httk: High-Throughput Toxicokinetics*. R package version 1.1. <http://CRAN.R-project.org/package=httk>

The views expressed in this poster are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.