Bayesian Hierarchical Meta-Regression of Epidemiologic Studies: Dose and Response Pre-Analysis (Poster 6)

Bruce Allen[1], Jeff Gift[2], Kan Shao[3], Kevin Hobbie[4], William Mendez Jr.[4], Janice S. Lee[2], Ila Cote[2], Ingrid Druwe[2], J. Allen Davis[5]

[1] Independent Consultant, [2] EPA, Office or Research and Development, National Center for Environmental Assessment – Research Triangle Park, [3] Indiana University, [4] ICF International, [5] EPA, Office or Research and Development, National Center for Environmental Assessment – Cincinnati

J. Allen Davis | davis.allen@epa.gov | 513-569-7024

Purpose and Scope

- National Research Council (NRC) has recommended the application of meta-analytical approaches, including Bayesian approaches, to well-studied health outcomes for the development of point estimates of risk and confidence intervals (NRC, 2013; NRC, 2014).
- NRC specifically recommended that EPA conduct dose-response meta-analysis for arsenic-related diseases in the IRIS assessment of inorganic arsenic (NRC, 2013).
- This poster is the first of two (see also Poster 7) that describe a case study highlighting an application of Bayesian hierarchical dose-response meta-regression to the analysis of arsenic exposure and human bladder cancer.

Case Study: Inorganic Arsenic (iAs) & Bladder Cancer

The pre-analysis steps described here employ methods to:

- address how doses are commonly reported in epidemiological studies
- calculate a common dose metric across all epidemiological studies
- calculate "effective counts" from reported effect measures in human studies to provide counts used in subsequent dose-response analyses to account for confounders. (see section on "Calculating Effective Counts")

Group Means and Uncertainty

- For dose-response analysis, a point estimate of dose is needed for each dose group, but epidemiologic data is often interval censored with an open ended reported for the high dose group (e.g., > 10,000 µg/L-yrs, Table 1)

Table 1: Calculated Effective Cases from Selected Arsenic Epidemiology Results - Information Presented in Tables 1 and 3 of Cohort Study by Chen et al. (2010)[a]

Cumulative water exposure, µg/L: years	Cases	Adjusted RR (95% CI)	Effective Cases	Effective Expected Number
< 400	6	--	6.00	6
400–1,000	3	1.11 (0.27 – 4.54)	2.84	2.56
1,000–5,000	12	2.33 (0.86 – 6.36)	10.65	4.57
5,000–10,000	5	3.77 (1.13 – 12.6)	4.72	1.25
>10,000	11	7.49 (2.70 – 20.8)	9.56	1.28

[a] The information in the first three columns is directly from Chen et al. (2010). The last two columns are computed as described subsequently in this poster.

- We assumed a log-normal distribution for exposures in the population of interest and calculated µ and σ as the log-scale mean and standard deviation using likelihood maximization.
- Given µ and σ, the mean within a exposure interval (c_g, c_{g+1}) is given by:

$$mean(g) = e^{\left(\mu+\frac{\sigma^2}{2}\right)} \times \frac{\theta(U_1(g)-\sigma)-\theta(U_0(g)-\sigma)}{\theta(U_1(g))-\theta(U_0(g))},$$

  - where $U_1(g) = \frac{(ln(c_{g+1})-\mu)}{\sigma}$, $U_0(g) = \frac{(ln(c_g)-\mu)}{\sigma}$, and $\theta()$ is the cumulative distribution function for the standard normal distribution
  - Group-specific means computed via this equation are used as the "MLE" doses
  - "High-end" and "low-end" doses were also estimated maximizing or minimizing the mean values for the highest exposure group
  - These "high-end" and "low-end" estimates correspond to a chi-squared-based 95% confidence interval around the maximum likelihood (MLE) estimate for the highest exposure group

Dose Conversions and Uncertainty

- For meta-analysis, it is imperative that all studies are expressed using a common dose metric, but iAs studies often report exposures in drinking water concentrations (µg/L), cumulative exposure (µg/kg-year), etc.
- For this analysis, we converted all reported studies into iAs daily intake values (µg/kg-day).
- For example, for a study that reports average iAs exposure (µg/L) or cumulative iAs exposure (µg/L-yr), daily intake (µg/kg-day) was calculated via:

$$dose = DI + f \times (WCR \times WE) + (1 - f) \times (WCR \times LE)$$

  - Where DI = dietary intake (µg/kg), f = fraction of lifetime exposed to the study reported iAs levels (WE), WCR = water consumption rate (L/kg), WE = arsenic exposure level (µg/L; if exposure is given in terms of cumulative exposure [CE], WE is estimated by dividing CE by the reported duration of exposure [RDWE]), and LE = low exposure value (µg/L).
- Parameters necessary for conversion determined on a study-by-study basis, according to study population.
- Factors for conversion were not treated as single values – a distribution of values was assumed over the individuals in the study to address interindividual variability and dose-group values were then averaged. Table 2 illustrates how this was done for one dose group.

Table 2: Example of Dose Calculations – Converting Reported cumulative exposure (µg/L-year) in Chen et al. (2010) to daily intakes (µg/kg-day) for one dose group

Variable	DI	f	WCR	RDWE	LE	WE – MLE	MLE Dose	WE - Low	Low Dose	WE - High	High Dose
Mean	0.65	0.630769	0.0345	42	5						
SD	0.333333	--	0.02319	3.333333	15						
Distribution	LogNormal	Beta	LogNormal	LogNormal	LogNormal						
Study Participant											
1	0.469875	0.739176	0.020187	47.73289	0.955895	3.29	0.524063	3.35	0.524879	3.24	0.523245
2	0.448139	0.548615	0.019016	37.84907	0.088638	4.15	0.492239	4.22	0.492959	4.09	0.491518
3	0.443911	0.525224	0.089145	37.46044	0.775256	4.20	0.673251	4.27	0.676512	4.13	0.669982
4	0.239939	0.716066	0.034483	39.79564	0.316072	3.95	0.340594	4.02	0.342213	3.89	0.338971
--	--	--	--	--	--	--	--	--	--	--	--
1000[1]	0.540283	0.868618	0.026601	45.62604	33.13435	3.45	0.735713	3.50	0.737035	3.39	0.734389
						Average	0.784552		0.785761		0.78334

[1] Assumed distributions with associated means and standard deviations are sampled a number of times equal to the dose-specific Ns, for doses with N > 1000, 1000 samples are taken to ease computational burden

- After averaging over all individuals within a dose-group, a Monte Carlo simulation was run with 1,000 iterations to derive a distribution of group-specific dose values.
- The median, 2.5th, and 97.5th percentiles from this distribution were used characterize the "best", "low-end", and "high-end" estimates of dose (Table 3).

Table 3: Results of the Monte Carlo Simulation Analysis to Calculate MLE, Low, and High Dose Estimates

Cumulative water exposure, µg/L: years	Dose Scenarios	Mean	Standard Deviation	0%	1%	5%	10%	25%	50%	75%	90%	95%	99%	100%
< 400	High	0.802	0.014	0.764	0.775	0.781	0.786	0.793	0.802	0.811	0.821	0.826	0.837	0.852
	Low	0.805	0.014	0.767	0.778	0.784	0.789	0.796	0.804	0.814	0.823	0.828	0.840	0.854
	MLE	0.804	0.014	0.765	0.776	0.783	0.787	0.794	0.803	0.813	0.822	0.827	0.838	0.853
400–1,000	High	1.052	0.016	1.008	1.017	1.027	1.033	1.042	1.051	1.062	1.072	1.078	1.091	1.111
	Low	1.052	0.016	1.008	1.017	1.027	1.033	1.042	1.051	1.062	1.073	1.078	1.091	1.111
	MLE	1.052	0.016	1.008	1.017	1.027	1.033	1.042	1.051	1.062	1.072	1.078	1.091	1.111
1,000–5,000	High	1.893	0.032	1.781	1.818	1.839	1.852	1.872	1.893	1.915	1.935	1.945	1.969	1.995
	Low	1.887	0.032	1.776	1.812	1.834	1.847	1.866	1.888	1.909	1.929	1.940	1.963	1.989
	MLE	1.890	0.032	1.779	1.815	1.837	1.849	1.869	1.891	1.912	1.932	1.943	1.966	1.992
5,000–10,000	High	4.245	0.123	3.880	3.961	4.032	4.086	4.164	4.245	4.329	4.397	4.443	4.526	4.677
	Low	4.238	0.123	3.874	3.955	4.026	4.080	4.157	4.239	4.322	4.390	4.436	4.519	4.669
	MLE	4.241	0.123	3.877	3.958	4.029	4.083	4.161	4.242	4.326	4.393	4.439	4.523	4.673
>10,000	High	20.487	0.597	18.322	19.152	19.481	19.720	20.083	20.478	20.867	21.292	21.477	21.858	22.285
	Low	18.687	0.543	16.718	17.475	17.774	17.992	18.320	18.679	19.032	19.420	19.588	19.935	20.325
	MLE	19.555	0.569	17.492	18.284	18.598	18.826	19.171	19.547	19.917	20.323	20.499	20.863	21.271

Calculation of Effective Counts

- For both cohort and case-control studies, published manuscripts almost always report relative risks (RR) or odds ratios (OR) that have been adjusted for some set of confounders
- The Bayesian dose-response meta-regression method described here is based on the likelihood of observing a particular number of cases
- The goal of computing "effective" counts of cases and controls is to construct of set of counts that reflect only the effect of exposure to iAs (Table 1)
- Essentially, the calculation results in counts of cases and controls that would have been calculated had all the covariates (other than dose) in all groups been the same as those observed in the referent group
- The methods employed to calculate these "effective counts" are based on those of Greenland and Longnecker (1992), Hamling et al. (2008), and Orsini et al. (2012)
- Studies included in the subsequent Bayesian dose-response meta-regression included incidence rate cohort, cumulative incidence cohort, and case-control studies

Conclusions

- The methods described herein were used to account for commonly encountered limitations in epidemiological studies in the context of dose-response analyses, including:
  - Reporting of interval-censored exposure groups
  - Use of divergent measures of iAs exposure across studies
  - And only reporting adjusted effect measures
- With respect to calculation of doses for use in a meta-regression, the current method calculates multiple exposure metrics and facilitates sensitivity analyses to investigate the degree of uncertainty in dose that exist across studies used in the analysis (Figure 1) (full set of sensitivity analyses discussed in Poster 7).

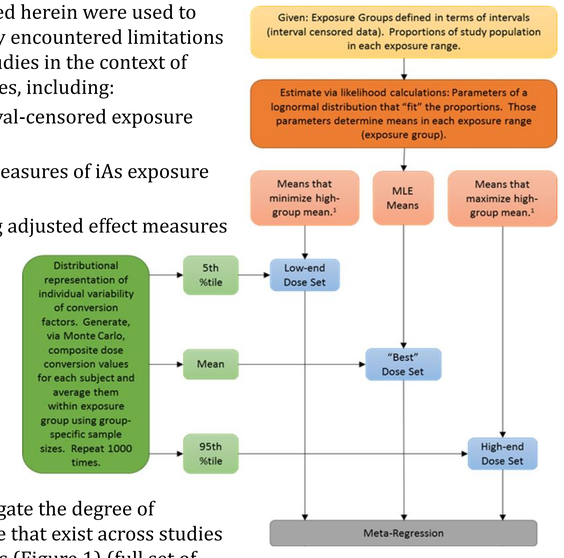

Figure 1: dose pre-analysis and uncertainty flowchart in relation to "best", "low-end", and "high-end" dose sets; [1] See Group Means and Uncertainty section

References

- Chen, C.L.; Chiou, H.Y.; Hsu, L.I.; Hsueh, Y.M.; Wu, M.M.; Wang, Y.H.; Chen, C.J. Arsenic in drinking water and risk of urinary tract cancer: A follow-up study from northeastern Taiwan. Cancer Epidemiol Biomarkers Prev 2010;19:101-110
- Greenland, S.; Longnecker, M.P. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. Am J Epidemiol 1992;135:1301-1309
- Hamling, J.; Lee, P.; Weitkunat, R.; Ambühl, M. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. Stat Med 2008;27:954-970
- NRC (National Research Council). (2013). Critical aspects of EPA's IRIS assessment of inorganic arsenic: Interim report. Washington, D.C: The National Academies Press.
- NRC (2014). Review of EPA's Integrated Risk Information System (IRIS) process. Washington, DC: The National Academies Press. http://www.nap.edu/catalog.php?record_id=18764
- Orsini, N.; Li, R.; Wolk, A.; Khudyakov, P.; Spiegelman, D. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. Am J Epidemiol 2012;175:66-73