Bayesian Hierarchical Meta-Regression of Epidemiologic Studies: Dose-Response Modeling and Target Population Predictions (Poster 7)

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, [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 second of two (see also Poster 6) 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 dose-response and target population prediction steps described here employ methods to:

- Apply a flexible logistic model to cohort and case-control epidemiological studies of inorganic arsenic (iAs) in a hierarchical Bayesian framework to estimate study-specific and pooled slopes
- Extrapolate predictions of risk to a target population of interest using lifetable methods
- This method explicitly uses as inputs the results of the pre-analysis steps described in Poster 6.

Dose-Response Modeling and Lifetable Analysis

- The purpose the dose-response analysis described herein is to perform a meta-regression to combine multiple studies for two kinds of epidemiological studies: case-control and cohort studies
- We assume that the *prospective likelihood* is given by a logistic equation applied to a vector of p explanatory variables $X = (X_1, \ldots, X_p)$:

$$logit\{\Pr(D = 1|X)\} = \alpha^* + \beta^T s(X)$$

- Due to the differing designs of case-control and cohort studies, methods were developed for each study type independently in order to predict the *prospective likelihood* of each study
- For the Bayesian implementation of the meta-regression:
  - All analyses were conducted in the Stan programming language
  - Defined necessary parameters for modeling and set priors:
    - Case-control studies: β (slope parameter) and λ (true proportion of doses in a dose-interval)
    - Cohort studies: μ(δ) (expected number of cases in the referent group)
  - Calculated the parameter α or α*
  - Defined the log-likelihood contribution for each dose group
- Typical lifetable analysis methods, including consideration of background exposure to iAs, were used to estimate extra risk of disease in the target population:
  - Background rates of disease assumed to represent zero extra risk from iAs
  - A mean background iAs dose of 0.071 μg/kg-day was assumed (0.05 μg/kg-day from dietary sources, 0.021 μg/kg-day from drinking water, and 0 μg/kg-day from inhalation) (Xue et al., 2010; Mendez et al., 2017).

Dose-Response Modeling and Lifetable Analysis cont.

- Table 1 summarizes the data used in the case study of iAs and bladder cancer, including the estimated intake values and effective counts calculated as described in the Poster 6

Table 1. Maximum Likelihood Dose Estimates, adjusted relative risks (RR) or odds ratios (OR) and Effective Counts for Bladder Cancer Studies Used in Meta-Regression

Data Set Name (Reported dose units)	Exposure Ranges (in reported dose units)	"Best" Dose Values for Analysis (avg daily μg/kg)	Adjusted RR or OR (LCL – UCL)	Raw Counts: Cases	Raw Counts: Expected or Controls[2]	Effective Counts: Cases	Effective Counts: Expected or Controls[2]
Cohort Studies							
Chen et al. (2010) (cumulative water exposure, μg/L-years)	< 400	0.80	1	6	6.00	6.00	6.00
	400-1000	1.05	1.11 (0.27-4.54)	3	2.70	2.84	2.56
	1000-5000	1.89	2.33 (0.86-6.36)	12	5.15	10.65	4.57
	5000-10000	4.24	3.77 (1.13-12.6)	5	1.33	4.72	1.25
	>10000	19.56	7.49 (2.7-20.8)	11	1.47	9.56	1.28
Sawada et al. (2013) Males[1] (water concentration, μg/L)	40.5	0.85	1	28	28.00	28.00	28.00
	54.7	1.15	1.45 (0.89-2.37)	41	28.28	37.44	25.82
	63.5	1.33	0.89 (0.51-1.55)	26	29.21	22.37	25.14
	99.1	2.08	1.56 (0.95-2.55)	46	29.49	36.06	23.12
Sawada et al. (2013) Females[1] (water concentration, μg/L)	37.1	1.07	1	6	6.00	6.00	6.00
	51.2	1.42	1.96 (0.7-5.53)	10	5.10	8.98	4.58
	64.2	1.65	2.06 (0.72-5.87)	10	4.85	8.34	4.05
	107.6	2.79	1.54 (0.5-4.73)	7	4.55	6.18	4.01
Case-Control Studies							
Steinmaus et al. (2013) (μg/d from water)	<41	0.92	1	32	197	32.00	83.32
	41-136	1.64	1.08 (0.62-1.87)	39	194	39.23	94.59
	137-307	3.20	3.06 (1.75-5.35)	64	154	57.22	48.69
	>307	10.40	5.85 (3.41-10.05)	97	95	99.06	44.09
Wu et al. (2013) (μg/gm Creatinine)	≤11.74	0.28	1	44	196	44.00	108.33
	11.74-20.94	0.55	1.42 (0.9-2.25)	63	196	69.52	120.54
	>20.94	1.23	4.13 (2.69-6.35)	192	202	166.80	99.44
Bates et al. (1995) (cumulative water iAs intake, mg)	<19	0.091	1	14	47	14.00	40.24
	19-33	0.097	1.56 (0.8-3.2)	21	36	18.98	34.98
	33-53	0.105	0.95 (0.4-2)	17	39	9.30	28.14
	>53	0.121	1.41 (0.7-2.9)	19	38	16.49	33.62
Steinmaus et al. (2003) (cumulative Intake, mg)	<6.4	0.09	1	66	101	66.00	73.03
	6.4-82.8	0.10	0.77 (0.48-1.24)	57	111	56.96	81.85
	> 82.8	1.09	0.73 (0.45-1.17)	58	116	54.29	82.29
Bates et al. (2004) (water concentration, μg/L)	0-50	0.39	1	87	80	87.00	51.35
	51-100	1.17	1.11 (0.3-3.7)	8	8	7.58	4.03
	101-200	2.12	0.81 (0.3-2)	13	13	11.67	8.51
	>200	7.98	0.28 (0.1-1.4)	3	10	3.49	7.36
Meliker et al. (2010) (water intake, μg/d)	<1	0.11	1	189	252	189.00	210.37
	1-10	0.13	0.83 (0.62-1.11)	162	234	145.13	194.62
	>10	0.36	1.01 (0.62-1.64)	43	48	37.01	40.79

[1] Sawada et al. (2013) reported medians for the exposure groups. Thus, we used those estimates (reported here) rather than working from the ranges.
[2] Expected raw and effective counts for cohort studies or control raw and effective counts for case-control studies.

- For the purpose of dose-response modeling, the α* parameter was assumed to be independent for each dataset
- Methods also assume study-specific β values that are normally distributed around a mean = β_mean, with standard deviation = β_sigma. Both β_mean and β_sigma were assigned priors and updated (Table 2)

Table 2. Prior parameter distributions used in the Bayesian Meta-Regression

Parameter	Prior Distribution
β(i)[1]	Normal (β_mean, β_sigma)
β_mean	Gamma (a=0.52, b=0.89)
β_sigma	Half-Cauchy (scale=5)

[1] β(i) is the dose coefficient for data set i.

- The gamma distribution for β_mean reflects determination that iAs is causally associated with the development of bladder cancer
  - prior judgement that exposure to 1 μg/kg-day iAs (~14-fold average background exposure) is highly likely to result in 1.0001 < OR < 20.
  - 1st and 99th percentiles of gamma distribution ($f(x) = \alpha e^{-\alpha x}(\alpha x)^{b-1} / \Gamma(b)$) set equal to ln(1.0001) and ln(20), results in parameters listed in Table 2
  - Important to note that gamma distribution gives greatest weight to values of x closest to zero (hence, prior assumption is weaker association with iAs unless data are sufficient to override prior)
- Estimates of pooled and study-specific β values derived from the hierarchical model and estimated lifetime extra risks in the target population are summarized in Tables 3 and 4 and Figures 1-3.

Table 3. Summary of Bayesian Meta-Regression Outputs, Including Parameters Important for Risk Estimation in the Target Population

β Parameter	Mean	SE of the Mean	SD	Percentiles: 2.50%	25%	50%	75%	97.50%	Effective Sample Size	Rhat
β_mean	0.2018	0.002	0.1775	0.0008	0.0572	0.1636	0.3018	0.6274	8219	1
β_sigma	0.6232	0.0026	0.2315	0.3198	0.4679	0.5767	0.724	1.2205	7788	1.0001
Chen et al. (2010)	0.0879	0.0002	0.0212	0.0434	0.0742	0.0885	0.1022	0.1284	9221	0.9998
Sawada et al. (2013) - males	0.2968	0.0017	0.1686	-0.0322	0.1823	0.2972	0.4107	0.6255	9294	1.0005
Sawada et al. (2013) - females	0.1455	0.0026	0.2497	-0.3659	-0.0196	0.1545	0.3166	0.6147	9346	0.9999
Steinmaus et al. (2013)	0.1774	0.0003	0.0246	0.1303	0.1607	0.1771	0.1936	0.2272	8530	0.9998
Wu et al. (2013)	1.349	0.0023	0.2164	0.9246	1.2018	1.3461	1.4953	1.7746	8934	0.9999
Bates et al. (1995)	0.2135	0.0075	0.6863	-1.1259	-0.1985	0.1969	0.6078	1.5945	8484	1.0004
Steinmaus et al. (2003)	-0.1389	0.0021	0.2004	-0.5335	-0.2717	-0.1366	-0.006	0.2499	8747	0.9997
Bates et al. (2004)	-0.1787	0.0009	0.0875	-0.3562	-0.2359	-0.1764	-0.1192	-0.0144	9141	1.0001
Meliker et al. (2010)	0.1869	0.0058	0.5348	-0.8819	-0.1539	0.181	0.5292	1.2308	8523	1.0001

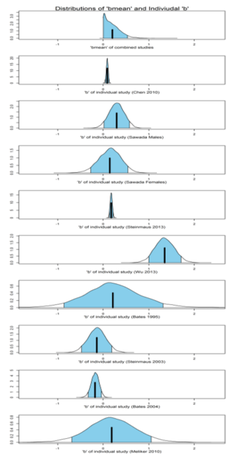
Summaries of Stan model runs: 4 chains were run, each with 10000 iterations; "warm-up" = 5000; remaining 5000 iterations per chain were thinned by 2 (every other iteration was dropped). Therefore, the total number of post-warmup draws = 10,000.
Effective samples = effective sample size used to estimate the parameters and Rhat is a measure of convergence (at convergence Rhat = 1). Effective sample sizes are large enough and the convergence criterion is satisfied for all parameters.

Table 4. Pooled Meta-Regression Estimates of Extra Lifetime Bladder Cancer Incidence Risk at Various Doses (per 10,000) and Drinking Water Exposures using MLE Dose Estimates

Extra Lifetime Risk[b,c]	Average Daily Arsenic Dose (ug/kg-day)						
	0.071[a]	0.12	0.19	0.26	0.33	0.75	1.45
	Average Daily Arsenic Drinking Water Concentration (ug/L)						
	1.5	5	10	15	20	50	100
0	2.0 (0.01 - 6.3)	4.8 (0.02 - 16)	7.7 (0.03 - 25)	11 (0.04 - 35)	29 (0.1 - 106)	64 (0.2 - 271)	

[a] U.S. daily background dose is estimated at 0.071 ug/kg, with 0.05 ug/kg from diet (Xue et al. 2010), 0.021 ug/kg from water, 1.5 ug/L median U.S. water level (Mendez et al. 2017) x 0.014 L/day mean U.S. water consumption rate (U.S. EPA, 2011, Table 3-1, "All Ages") and 0 ug/kg from air. Thus, 1.5 ug/L in water is associated with a background dose of 0.071 ug/kg and an extra risk of 0.
[b] These extra risk estimates assume a mean U.S. background rate for bladder cancer of 2% (NCI, 2017). Predicted additional cases in a cohort of size 10,000 for extra risk, x, when the background rate is b, would be 10,000*(1-b)*x. Thus, additional cases of bladder cancer at an extra risk of 2/10,000 (0.02%) would be 10,000*(1−2%)*0.02% = 1.96.
[c] Mean, 2.5% and 97.5% of Bayesian posterior slope distributions were used with US lifetables to estimate mean and credible intervals for extra risk above average background risks.

Figure 1. Posterior Distributions for Pooled and Study-specific Logistic Slope Parameters Using the MLE Dose Estimates.



Black vertical lines indicate means of posterior distributions. 95% credible intervals for the logistic slope parameters are highlighted in blue

Dose-Response Modeling and Lifetable Analysis cont.

Figure 2. Extra Lifetime Bladder Cancer Risk due to Oral iAs Exposure, using MLE Dose Estimates

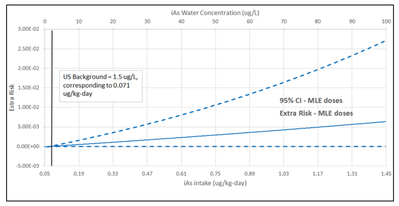
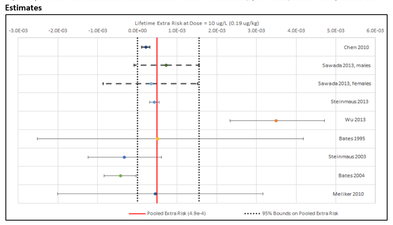


Figure 3. Forest Plot of Extra Lifetime Bladder Cancer Risk at 10 μg/L iAs Exposure, using MLE Dose Estimates



- The sensitivity of the hierarchical model and its outputs were examined regarding four sources of uncertainty:
  - Characterization of exposure levels used in the modeling: this was addressed using the "high" and "low" dose estimates discussed in Poster 6; using different estimates of dose did not result in pooled β_mean that differed greatly (0.19, 0.20, or 0.21)
  - Choice of datasets: a leave-one-out analysis was performed which showed that no one study had a disproportionately large influence on the final pooled β_mean value (Table 5)
  - Zero background inhalation assumption: assuming background inhalation exposures of 0.2 to 0.6 μg/day decreased mean extra risk estimates from 4.88×10^{-4} μg/kg-day (Table 5, no data set excluded) to 4.68 or 4.51×10^{-4} μg/kg-day

Table 5. Impact of Leave-One-Out Cross Validation (Dataset Exclusion) on Lifetime Extra Risk of Bladder Cancer

Excluded Data Set	Extra Lifetime Risk at 10 μg/L (0.19 μg/kg): 2.5 %tile	Mean	97.5 %tile
None	1.91E-06	4.88E-04	1.56E-03
Chen et al. (2010)	1.91E-06	5.48E-04	1.85E-03
Sawada et al. (2013), males	1.67E-06	4.88E-04	1.72E-03
Sawada et al. (2013), females	1.67E-06	5.17E-04	1.75E-03
Steinmaus et al. (2013)	1.44E-06	5.27E-04	1.76E-03
Wu et al. (2013)	4.78E-07	1.62E-04	5.77E-04
Bates et al. (1995)	1.91E-06	4.85E-04	1.54E-03
Steinmaus et al. (2003)	2.39E-06	5.86E-04	1.85E-03
Bates et al. (2004)	2.63E-06	6.11E-04	1.86E-03
Meliker et al. (2010)	1.44E-06	4.95E-04	1.64E-03

- The consideration of alternative gamma prior distributions for β_mean: alternative distributions that considered different 1st or 99th percentile values did not overly influence final risk estimates (Table 6)

Table 6. Posterior β_mean distribution values resulting from various prior Gamma distributions

Alternative Prior	2.5th percentile	Mean	97.5th percentile	% Mean Difference
1.0001 – 10	0.0012	0.2108	0.6512	4.46
1.0001 – 30	0.0005	0.1966	0.6311	-2.58
1.00001 -20	0.0001	0.1707	0.5922	-15.41
1.001 - 20	0.0045	0.237	0.6673	17.44
Original Prior (1.0001 – 20)	**0.0008**	**0.2018**	**0.6274**	--

Conclusions

- These Bayesian meta-regression methods (Posters 6 and 7) allow for inclusion of more studies than other meta-regression methods by reconciling different study designs and exposure metrics, and could potentially be applied to any endpoint for which multiple studies and incidence/mortality/morbidity lifetables are available
- The logistic dose-response model used could be extended to consider fractional-polynomial forms of the logistic model, $logit(p(x)) = \alpha^* + \beta\_1(x^{p1}) + \beta\_2(x^{p2})$, to allow more flexibility in fitting datasets for the investigation of whether the data suggest a J-shaped dose-response (e.g., negative slopes in the low dose region)

References

- Bates, M.N.; Rey, O.A.; Biggs, M.L.; Hopenhayn, C.; Moore, L.E.; Kalman, D.; Steinmaus, C.; Smith, A.H. Case-control study of bladder cancer and exposure to arsenic in Argentina. Am J Epidemiol 2004;159:381-389
- Bates, M.N.; Smith, A.H.; Cantor, K.P. Case-control study of bladder cancer and arsenic in drinking water. 1995;141:523-530
- 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
- Meliker, J. R.; Slotnick, M.J.; Avruskin, G.A.; Schottenfeld, D.; Jacquez, G.M.; Wilson, M.L.; Goovaerts, P.; Franzblau, A.; Nriagu, J.O. Lifetime exposure to arsenic in drinking water and bladder cancer: A population-based case-control study in Michigan, USA. Cancer Causes Control 2010;21:745-757
- Mendez, W.M.; Eftim, S.; Cohen, J.; Warren, I.; Cowden, J.; Lee, J.S.; Sams, R. Relationships between arsenic concentrations in drinking water and lung and bladder cancer incidence in U.S. counties. J Expo Sci Environ Epidemiol 2017;27:235-243
- NCI (National Cancer Institute). (2017). The cancer genome atlas: Urothelial bladder carcinoma. Available online at https://cancergenome.nih.gov/cancersselected/UrothelialBladderCarcinoma.
- Sawada, N.; Iwasaki, M.; Inoue, M.; Takachi, R.; Sasazuki, S.; Yamaji, T.; Shimazu, T.; Tsugane, S. Dietary arsenic intake and subsequent risk of cancer: The Japan Public Health Center-based (JPHC) prospective study. Cancer Causes Control 2013;24:1403-1415
- Steinmaus, C.; Yuan, Y.; Bates, M.N.; Smith, A.H. Case-control study of bladder cancer and drinking water arsenic in the western United States. Am J Epidemiol 2003;158:1193-1201
- Steinmaus, C.M.; Ferreccio, C.; Romo, J.A.; Yuan, Y.; Cortes, S.; Marshall, G.; Moore, L.E.; Balmes, J.R.; Liaw, J.; Golden, T.; Smith, A.H. Drinking water arsenic in northern Chile: High cancer risks 40 years after exposure cessation. Cancer Epidemiol Biomarkers Prev 2013;22:623-630
- U.S. EPA. Exposure factors handbook: 2011 edition (final). Washington, DC: U.S. EPA, Office of Research and Development, National Center for Environmental Assessment; 2011
- Wu, C., -C; Huang, Y., -K; Chung, C., -J; Huang, C., -Y; Pu, Y., -S; Shiue, H., -S; Lai, L., -A; Lin, Y., -C; Su, C., -T; Hsueh, Y., -M. Polymorphism of inflammatory genes and arsenic methylation capacity are associated with urothelial carcinoma. Toxicol Appl Pharmacol 2013;272:30-36
- Xue, J.; Zartarian, V.; Wang, S.W.; Liu, S.V.; Georgopoulos, P. Probabilistic modeling of dietary arsenic exposure and dose and evaluation with 2003-2004 NHANES data. Environ Health Perspect 2010;118:345-350

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