The Application of Model Averaging Methods to Assess the Suitability of Taiwanese Studies for Predicting U.S. Bladder and Lung Cancer Risk from Inorganic Arsenic (iAs) Intake (Poster 5)

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

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

Jeff Gift | gift.jeff@epa.gov | 919-541-4828

Purpose and Scope

- National Research Council (NRC) recommended that EPA focus on high-quality epidemiologic studies that assess inorganic arsenic (iAs) exposures commonly experienced in the U.S., where mean background intake is estimated to be 0.071 µg iAs/kg-day (see Posters 6 and 7) and where intake levels above 1 µg iAs/kg-day are extremely rare (NRC, 2013).
- An analysis was performed to assess the suitability of two studies of bladder and lung cancer risk in a large Taiwanese population (Chen et al., 2010a,b) that:
  - meet EPA study quality criteria (see Poster 1),
  - form the basis of arsenic risk assessments performed by other international organizations (FDA, 2016; WHO, 2011), and
  - are associated with high iAs exposure levels relative to the U.S. (iAs intake for the reference group of these studies is ~0.9 µg/kg-day, more than 10 × higher than the estimated U.S. background intake level).

Modeling Approach - Overview

A model averaging approach was applied in an attempt to extrapolate lifetime bladder and lung cancer probabilities observed at µg/kg-day intake doses estimated for a large prospective cohort study of residents in northeast Taiwan (Chen et al., 2010a,b) to relevant U.S. doses. The approach is illustrated in Figure 1 and builds upon dose-response model averaging methods developed by the FDA. It involves:

- *Estimation of water and dietary intake variability for the Taiwanese population* to represent the variability in the input variables to the bootstrap model.
- *Bootstrap simulation* to incorporate uncertainty in the estimation of adjusted outcomes (cases of cancer) and daily arsenic intake dose.
- *Model Averaging* to extrapolate to U.S. relevant doses and assess model dependence.

Estimation of water and dietary intake variability

Multiple data sources and methods were used to derive inputs for the bootstrap estimation of arsenic intake. In summary:

- *iAs Drinking water intake* was estimated by fitting a mixed lognormal distribution to the drinking water concentration data from the Chen et al. cohort. Distributions of drinking water consumption were estimated based on age-specific survey data from the Taiwan Department of Health (TDOH, 2007).
- *iAs food intake* was estimated using food consumption from Taiwan Department of Health survey data (TDOH, 2007) and iAs concentration distributions (for rice and leafy vegetables) or central tendency estimates (tubers, pulses, meats and fish) estimated from multiple studies of Taiwanese and other Asian countries.

Bootstrap simulation

- A "bootstrap" methodology was applied to simulate the variability in arsenic intake and in outcome measures, and their impacts on risk estimates. As shown in Figure 1, the estimated arsenic intake doses from water and diet were summed for each subject in each bootstrap iteration, and average total daily intake doses were estimated across each exposure group. The 1,000 sets of group average arsenic intake dose served as inputs, along with the outcome data sets, to the dose-response estimation.

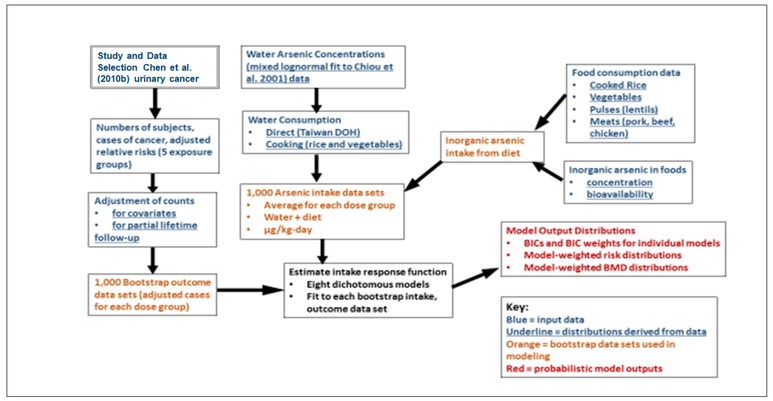

Figure 1. Summary of dose-response methodology for bladder and lung cancer. Note: BIC = Bayesian Information Criterion

Model Averaging

- Nine dose-response models available in EPA's Benchmark Dose Software (BMDS) were fit to each bootstrap data set (Table 1). A diverse set of models was chosen to cover "model space" and explore "model uncertainty" as fully as possible.
- Models were estimated by maximizing binomial likelihood with varying constraints.
- Outputs from the bootstrap analysis included 1,000 sets of maximum likelihood parameter estimates and model log likelihoods derived for each input data set.
- Bayesian Information Criteria (BIC) values were calculated as:

$$BIC = -2 \times \log(likelihood) + k \times \ln(n)$$

  where k = number of parameters estimated and n = number of observations.
- The weights employed in model averaging were based on the calculated average BIC values for each model. For each model (i), the Bayes weights were calculated as:

$$Weight_i = \frac{e^{(-0.5 \times BIC_i)}}{\sum_{i=1}^{9} e^{(-0.5 \times BIC_i)}}$$

- "Prior" model weights were assumed to be 1/9 (i.e., no *a priori* preferred model).
- Weibull, log logistic, log probit, Gamma, and dichotomous Hill models were run with power or slope terms both unconstrained and constrained to be >1.0 to better assess model dependence in the low dose region.
- Weighted estimates of lifetime bladder and lung cancer probabilities were calculated for a series of doses from 0 to 40 µg/kg-day, corresponding to the range of mean total arsenic intakes observed in the bootstrap data set.

Table 1. Models included in the dose-response assessment

Model	Form	Parameters
Quantal linear	r(dose) = a + (1−a) × (1−exp(−b × dose))	2
Logistic	r(dose) = 1/(1 + exp(−a−b × dose))	2
Probit	r(dose) = pnorm(a + b × dose)	2
Weibull	r(dose) = a + (1−a) × (1−exp(−c × dose^b))	3
Multistage 2	r(dose) = a + (1−a) × (1−exp(−b × dose−c × dose^2))	3
Log logistic	r(dose) = a + (1−a)/(1+exp(−c−b × log(dose)))	3
Log probit	r(dose) = a + (1−a) × pnorm(c + b × log(dose))	3
Gamma	r(dose) = a + (1−a) × pgamma(c × dose^b)	3
Dichotomous Hill	r(dose) = v × g + (v−v × g)/(1 + exp(−c−b × log(dose)))	4

Results

- In the range of the data, similar mean absolute risk, 2.5th and 97.5th percentiles are derived from unconstrained and constrained models (Figures 2 & 3; upper plots).
- At lower doses, absolute risks derived from the unconstrained models curve sharply downward compared to those from constrained models (Figures 2 & 3; lower plots).
- Differences in extra risk (i.e., the increase in risk relative to estimated "background risk") are more substantial, particularly in the low-dose range (see Figures 4 and 5).

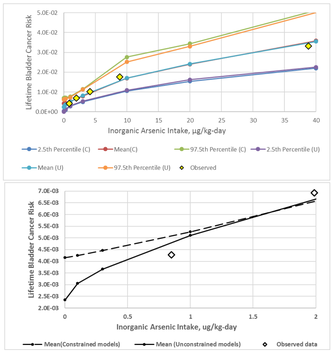

Figure 2. Predicted lifetime probability of bladder cancer versus all doses (upper plot) and low doses (lower plot) using constrained (C) and unconstrained (U) models compared to adjusted observed incidence from adjusted relative risks reported in Chen et al. (2010b).

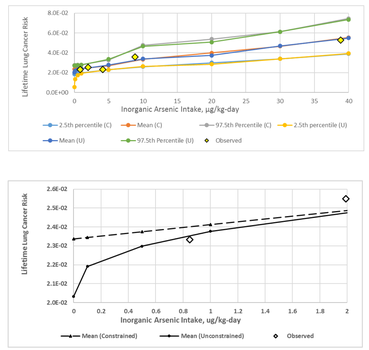

Figure 3. Predicted lifetime probability of lung cancer versus all doses (upper plot) and low doses (lower plot) using constrained (C) and unconstrained (U) models compared to adjusted observed incidence obtained from adjusted relative risks reported in (Chen et al., 2010a).

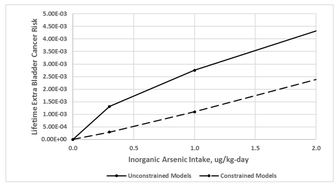

Figure 4. Predicted low dose extra risk of bladder cancer from Chen et al. (2010b) for constrained and unconstrained model averaging.

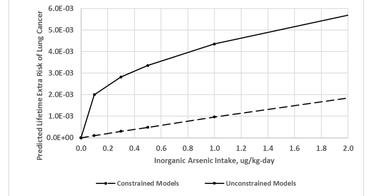

Figure 5. Predicted low dose extra risk of lung cancer from Chen et al. (2010a) using constrained and unconstrained model averaging.

Conclusions

As reflected in Figures 2 through 5, EPA's model averaging analysis shows substantial model uncertainty in extrapolating from the iAs doses estimated for the Taiwan cohort to the estimated U.S. background iAs dose of 0.071 µg/kg-day. This result, combined with the NRC (2013) recommendation to perform only "modest" (e.g., 1 order of magnitude) extrapolation from the lowest exposure group of a candidate study, suggests that the Chen et al. (2010a,b) studies should not serve as the sole basis for U.S.-specific cancer risk estimates. As a result, EPA has developed a multiple study Bayesian meta-regression approach that has the potential to better inform dose-response and provide more reliable risk estimates at U.S.-relevant arsenic dose levels (see Posters 6 and 7).

References

Chen, CL; Chiou, HY; Hsu, LI; Hsueh, YM; Wu, MM; Chen, CJ. (2010a). Ingested arsenic, characteristics of well water consumption and risk of different histological types of lung cancer in northeastern Taiwan. Environ Res 110: 455-462. http://dx.doi.org/10.1016/j.envres.2009.08.010

Chen, CL; Chiou, HY; Hsu, LI; Hsueh, YM; Wu, MM; Wang, YH; Chen, CJ. (2010b). Arsenic in drinking water and risk of urinary tract cancer: A follow-up study from northeastern Taiwan. Cancer Epidemiol Biomarkers Prev 19: 101-110. http://dx.doi.org/10.1158/1055-9965.EPI-09-0333

Chiou, H. -Y; Chiou, S. -T; Hsu, Y. -H; Chou, Y. -L; Tseng, C. -H; Wei, M. -L; Chen, C. -J. (2001). Incidence of transitional cell carcinoma and arsenic in drinking water: A follow-up study of 8,102 residents in an arseniasis-endemic area in northeastern Taiwan. Am J Epidemiol 153: 411-418. http://dx.doi.org/10.1093/aje/153.5.411

FDA, US. (2016). Arsenic in Rice and Rice Products. Risk Assessment Report. Center for Food Safety and Applied Nutrition, FDA, U.S.DHHS. https://www.fda.gov/downloads/Food/FoodScienceResearch/RiskSafetyAssessment/UCM486543.pdf

NRC (National Research Council). (2013). Critical aspects of EPA's IRIS assessment of inorganic arsenic: Interim report. Washington, D.C: The National Academies Press.

TDOH. (2007). Compilation of Exposure Factors. DOH96-HP-1801. Taipei City, Taiwan: Taiwan Department of Health.

WHO (World Health Organization). (2011). Safety evaluation of certain contaminants in food. (WHO Food Additives Series: 63. FAO JECFA Monographs 8). Geneva, Switzerland: prepared by the Seventy-second meeting of the Joint FAO/WHO Expert Committee on Food Additives (JECFA. http://whqlibdoc.who.int/publications/2011/9789241660631_eng.pdf

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

The views expressed in this poster are those of the author(s) and do not necessarily represent the views or the policies of the U.S. Environmental Protection Agency.