Analyzing Study-Specific Estimates of Exposures Associated with a Defined Relative Risk vs U.S. Background Exposure (RRBs) for Inorganic Arsenic (iAs) Health Outcomes (Poster 4)

Kevin Hobbs¹, J. Allen Davis², Kan Shao³, Cara Henning⁴, William Mendez Jr.⁴, Janice S. Lee⁴, Ila Cote⁵, Ingrid Druwe⁶, Jeff Gift⁴


Purpose and Scope

➢ National Research Council (NRC) recommended that EPA derive risk estimates for iAs for health effects with adequate epidemiologic evidence (NRC, 2013).
➢ EPA developed an approach to provide an efficient, yet also effective, means of focusing dose-response analysis efforts given the extent of the epidemiologic evidence base, and the variance in data quality across health outcomes.

Relative Risk Exposure vs Background Exposure (RRB)

EPA developed an approach that allows for comparison of relative risk estimates across studies that use various exposure metrics. Dose-response modeling is used to estimate exposures associated with a given increase in relative risk (RR). The RRB is divided by an estimate of the U.S. background level for that exposure metric. This approach involves:

➢ Selection of datasets: starting from health outcomes with robust/moderate databases, and considering author-performed trend tests.
➢ Exposure-response modeling: case-control and cohort studies were modeled to predict exposures where relative risk (RR) changed by 20% (regardless of endpoint severity or prevalence) compared to the RR estimated at U.S. background (Table 2) (RRUB).

Deriving relative risk (RRE) values from estimates of U.S. background (RRUB). Exposure units for U.S. background estimates differ to match RRB units, but are based on similar water and dietary intake assumptions (see Table 2).

Selection of Datasets

➢ Hazard Identification – Focused on epidemiological studies of iAs health outcomes having robust/moderate databases (see Poster 1).
➢ Initial screen – Focused on datasets from cohort and case-control studies. Ecological, cross-sectional and continuous (e.g., neurocognitive response measures) datasets not considered for purposes of RRUB derivation for purposes of the RR analysis.
➢ Secondary screen – Each dataset received a score of 0, 1, or 2 for each rating element (Table 1). Datasets for which the sum of scores was > 5 were excluded.
➢ Final screen – Studies with inadequate or conflicting dose-response data were removed if issue(s) could not be resolved through communications with authors.

Data Preprocessing

➢ Estimating Group-Level Mean Exposures – Exposure ranges were fit to lognormal distributions using maximum likelihood (ML) methods. Group mean estimates were derived by drawing large Monte Carlo samples (10 million) from fitted distributions, and sampling randomly in each exposure range for appropriate numbers of "subjects." 
➢ Adjusting Incidence to Account for Covariates – "Effective exposure" values derived from reported ORs that were adjusted for covariates (see Poster 1).
➢ Identifying Background Exposure for the U.S. Population – For RRB and RRUB derivations, relative risk for central tendency background exposures (Table 2) set to 1.0. Thus, the RREi is exposure or dose that produces the calculated relative risk is 1.2. This allows for comparison of U.S.-specific risk results across studies.
➢ Categorizing Outcomes – To facilitate comparing across RRBs, outcomes categorized by types (clinical-fatal, clinical-non fatal, preclinical, subclinical, and other outcomes) (e.g., fetal loss, infant mortality and stillbirth for pregnancy outcomes).

Exposure Response Modeling

➢ Case-control studies – adjusted case and control numbers were fit by a logistic model: \( \text{f(dose)} = 1 / (1 + \exp(-a - b \times \text{dose})) \). Use of a logistic model allows for analysis of case-control studies with prospective studies, both having the same binomial-based likelihood contributions from their exposure groups (Prentice and Pyke, 1979).
➢ Cohort studies – counts of cases in each exposure group follow a Poisson distribution: \( \text{r}_i = \text{Poison}(\lambda_i, \mu) \), where \( \text{r}_i \) and \( \mu_i \) are observed cases and expected case number in the ith exposure group, respectively. Seven continuous dose-response models used for \( f(c) \), including the linear model, power model, 2nd-degree polynomial model, Michaelis-Menten model, and the Exponential 2, 3, and 4 models.
➢ Model Fit Assessment and Model Selection – for each dataset, the modeling generated estimates of log-logistic, AIC and x² p-value, estimates of model parameters, and predicted risks (ORs for case-control; RRbs for cohort) at each exposure level, with confidence limits. EPA (2012) BMD modeling methods were used to select a best fitting model from the multiple models used to fit cohort study data.
➢ Selection of a Benchmark Relative Risk - for this comparative analysis, a 20% relative risk dose, or RRBs, is estimated. The 20% effect level was chosen to avoid extrapolating far outside the range of data and because, for the bulk of the epidemiologic data sets, an increase in odds ratio or relative risks of about 20% was near the smallest increase that could be resolved based on the data.

Results

➢ Final screening of studies led to the identification of 262 datasets within 68 studies.
➢ The figure shows individual and median preclinical/subclinical, clinical nonfatal and clinical fatal RRB results organized by most to least number of datasets.
➢ Table 3 presents RRB ranges, means and medians for each health outcome.

Conclusions

As indicated in Poster 1, all of the outcomes in this RRB analysis, as well as neurocognitive effects for which RRb values could not be derived, were identified as having Robust or Moderate evidence overall and will therefore be considered for dose-response analysis. However, NRC (2013) identified priority health outcomes for EPA to focus on and recommended that EPA further prioritize. EPXs RRB analysis approach supports this prioritization effort by providing a method for comparing the results of diverse studies of health outcomes, and identifying key endpoints and datasets that are suitable for use in more detailed dose-response analyses (see Posters 5, 6, and 7).

Consistent with key outcomes identified by the NRC (NRC, 2013), DCS, bladder cancer and lung cancer were identified as having the largest databases of adequate dose-response datasets, increasing confidence in the RRB summary statistics (e.g., median estimates), as well as low RRB values relative to most outcomes. RRB values for diabetes and liver cancer data are also low, but are associated with smaller databases and a lower degree of certainty in the RRB summary statistics.

References

