United States Environmental Protection Agency

# Considerations for the Dose-Response Analyses of Inorganic Arsenic Health Outcomes

**Jeff Gift, PhD**
**EPA National Center for Environmental Assessment**
**October 3, 2017**

**Office of Research and Development**
National Center for Environmental Assessment, IRIS

## Disclaimer

The views expressed in this presentation are those of the author(s) and do not necessarily reflect the views or policies of the US EPA.

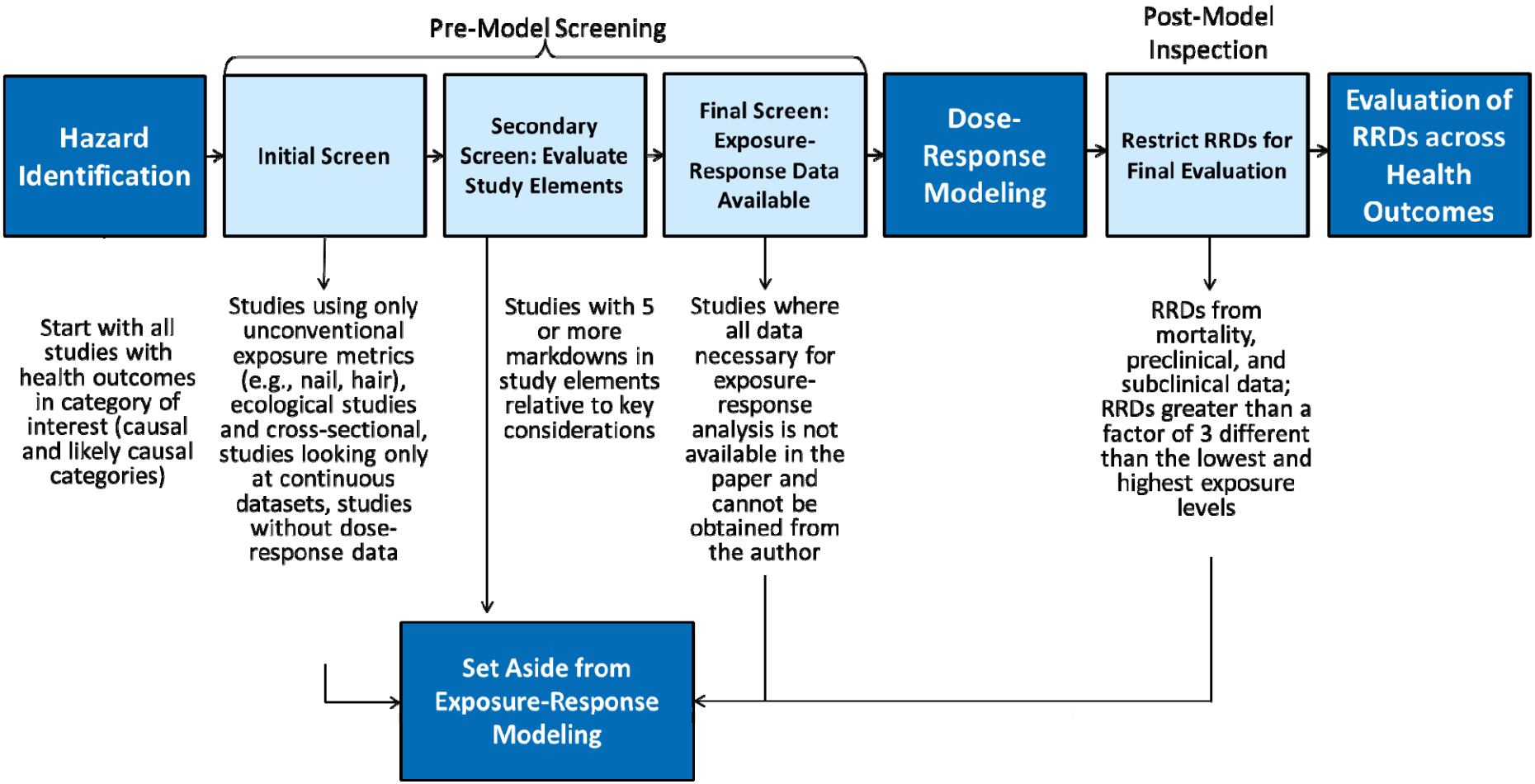
## Presentation Outline

- Overview of Potential Tiered Dose-response Approach
  - Margin of Exposure (MOE) analysis
  - Model averaging to assess within study model uncertainty
  - Bayesian meta-regression multiple study analysis
  - Extrapolation to target, U.S. population using lifetable approach
- Potential Science Issues
- Acknowledgments

# D-R Methods – Margin of Exposure (MOE) Screening Level Analysis

- **Screening Level Margin of Exposure (MOE) Analysis:**
  - **Purpose – verify NRC health outcome tiers, prioritize studies and endpoints for more complex analyses and inform cost-benefit analyses for a broad set of endpoints**
  - **Approach:**
    - Develop Relative Risk Dose (RRD; like BMD) for causal & likely causal health outcomes
    - Identify exposure-response datasets from adequate studies
    - Model unaltered exposures and adjust responses for selected datasets
    - Focus on single study, single best fitting standard model

# MOE Analysis – Identification of Adequate Dose-response Datasets



# "Mark-Down" Elements for Pre-model Study Screening

| "Mark-Down" Element | Explanation |
|---|---|
| **Endpoint Selection** | Incidence preferred to mortality; odds ratio or relative risk preferred to SMR |
| **Number of subjects and cases reported** | Numbers of cases & controls/subjects highly preferred for dose-response; reporting only summary measures (OR, RR, SMR) is serious shortcoming |
| **Exposure ascertainment** | Prefer individual measurements, then small-groups, then large groups |
| **Exposure uncertainty** | Mean or median with variance preferred; ranges of exposures less desirable |
| **Exposure/Dose metric** | Cumulative intake or exposure preferred over point-in-time measurements. Urinary markers (adequately characterized) is considered a reliable indicator of intake; other biomarkers (hair, nails, blood) less desirable |
| **Exposure timing, duration** | Evidence supporting long-term exposure patterns highly desirable |
| **Measurement of, adjustment for, covariates** | Estimates control for at least smoking, gender, age, as well as other covariates (nutrition, genetic polymorphisms, etc.) as appropriate |
| **Number of exposure grps** | Referent plus two or more exposure groups preferred; having only on exposed group is not a categorical disqualification for meta-regression |
| **Representativeness of referent group** | Referent group characteristics (case-control) should be well-documented; referent group should be similar to exposed, with regard to key covariates |
| **Numbers of cases, controls, subjects** | Having sufficient numbers of cases, controls, subjects to support reliable statistical analyses is highly desirable; most serious issue is small number of referents in case-control studies |
| **Sensitive populations** | Where sensitive populations are identified, it is desirable to have separate exposure/dose-response evaluations for them |

# D-R Methods – Multiple Study Bayesian Meta-Regression

- **Bayesian Meta-Regression Analysis of Select NRC (2013) "Tier 1" health outcomes**:
  - **Purpose - Combine multiple studies to benefit from more (e.g., low dose) data**
  - **Approach:**
    - Multiple studies, single model, Bayesian meta-regression, and lifetable analysis
    - Model select NRC (2013) "Tier 1" health outcomes (e.g., bladder cancer, lung cancer, diseases of the circulatory system)

# EPA PBPK Model

**EPA PBPK model used for evaluation of Tier 1 health outcomes (El-Masri and Kenyon, 2008)**

- validate with data from 11,500 people in Taiwan and Nevada
- Allow the incorporation of studies that reported urinary concentrations of iAs

**Dose-Response variabilities and uncertainties evaluated**

- Study and model choice
- Exposure variability
- Population characteristics: e.g. age, BMI, smoking, sex, background incidence of disease, nutrition, arsenic metabolism and genetic polymorphisms

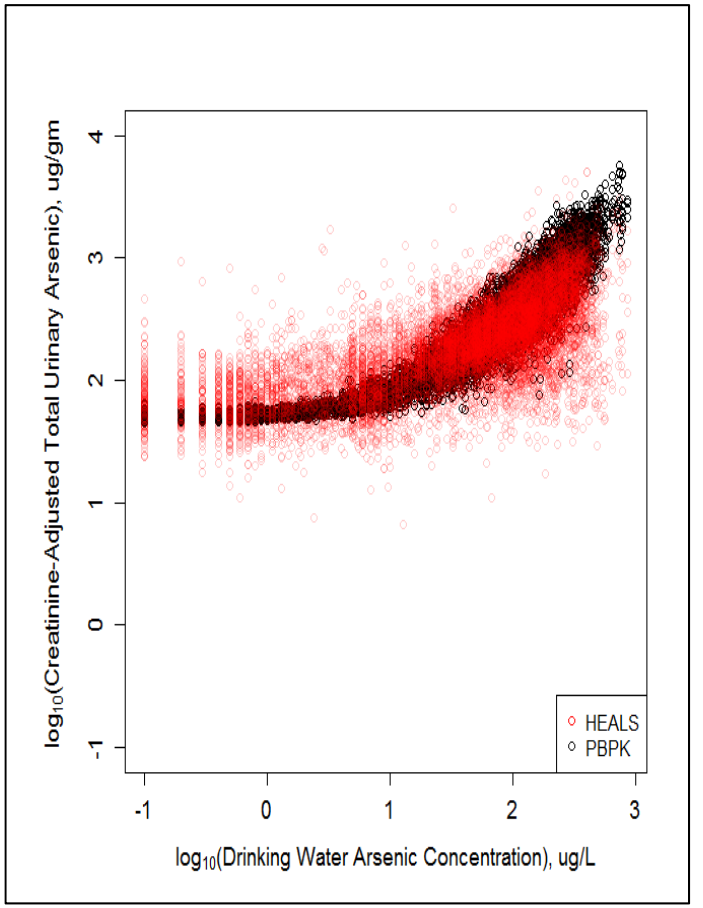
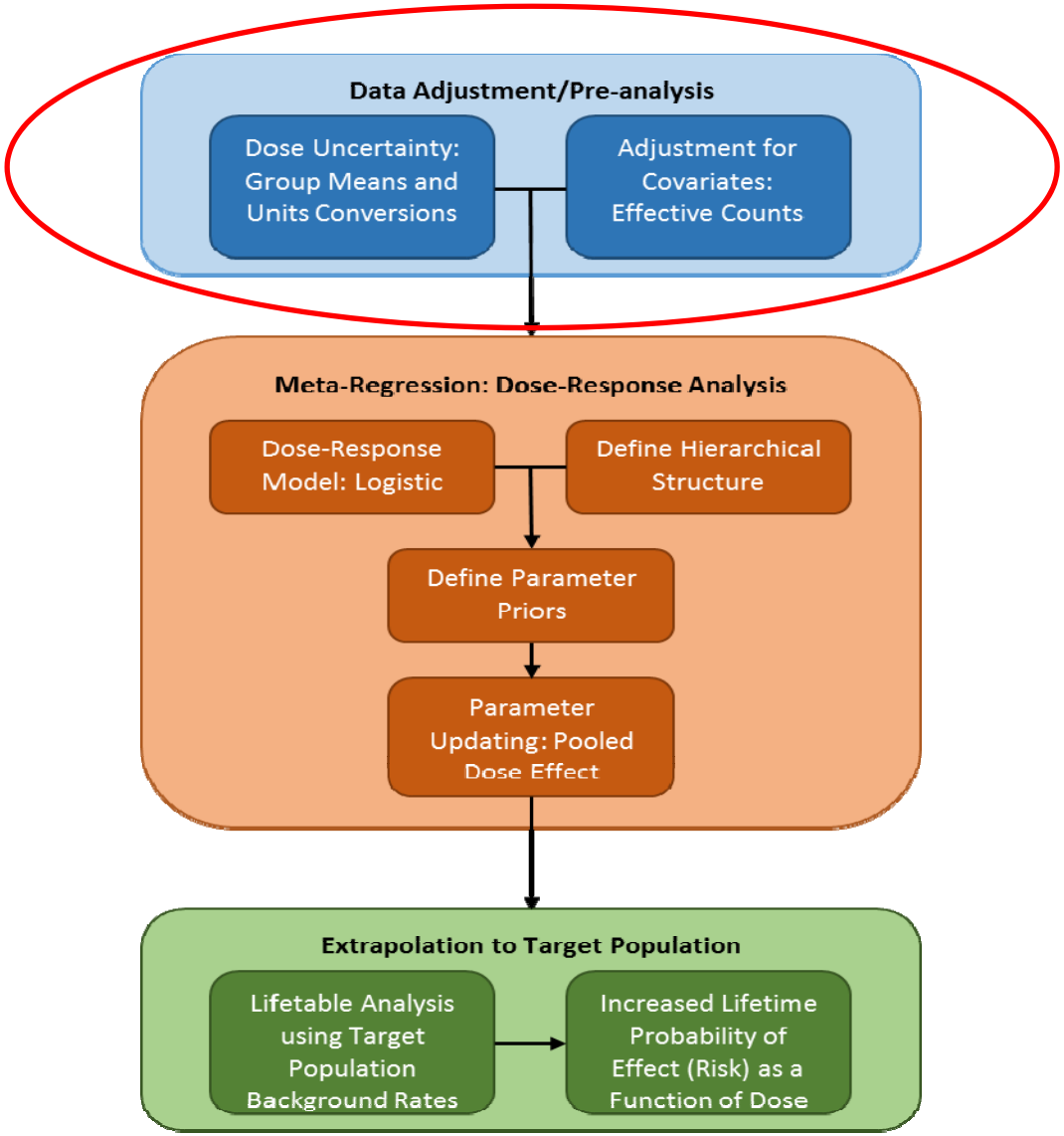



Figure 2. El-Masri and Kenyon (2008) PBPK model calibration against measured iAs total urinary concentrations and drinking water concentrations.

# D-R Methods - Meta-Regression Analysis Flow Chart



# Meta-Regression - Dose Uncertainty Pre-analysis

- **Most published reports provide ranges of exposures. Proposed approach:**
  1. **Compute means** (in units given in the published papers) – assume lognormal distribution; estimate mean for maximized profile likelihood (best estimate); then find MLE while minimizing (low estimate) and maximizing (high estimate) value of the high dose group (most uncertain and most influential dose group)
  2. **Convert to daily intake estimates.** If intake data not provided in study convert exposures to intake using population-specific estimates of key factors such as average duration of well exposure (yrs), average age at diagnosis (yrs), low water exposure (µg/L), water consumption rate (ml/kg-d) and dietary intake (µg/kg-d)
  3. **Estimate uncertainty about the mean daily intake estimates** via a Monte Carlo analysis that takes into account confidence intervals for key factors
     a. Each iteration consists of drawing values from all the "key factor" distributions necessary for conversion to dose, then averaging according to the number of individuals in the exposure group
     b. This will be repeated 1,000 times to derive a distribution and median, $2.5^{th}$ and $97.5^{th}$ percentiles obtained to characterize the best, low-end and high-end dose values

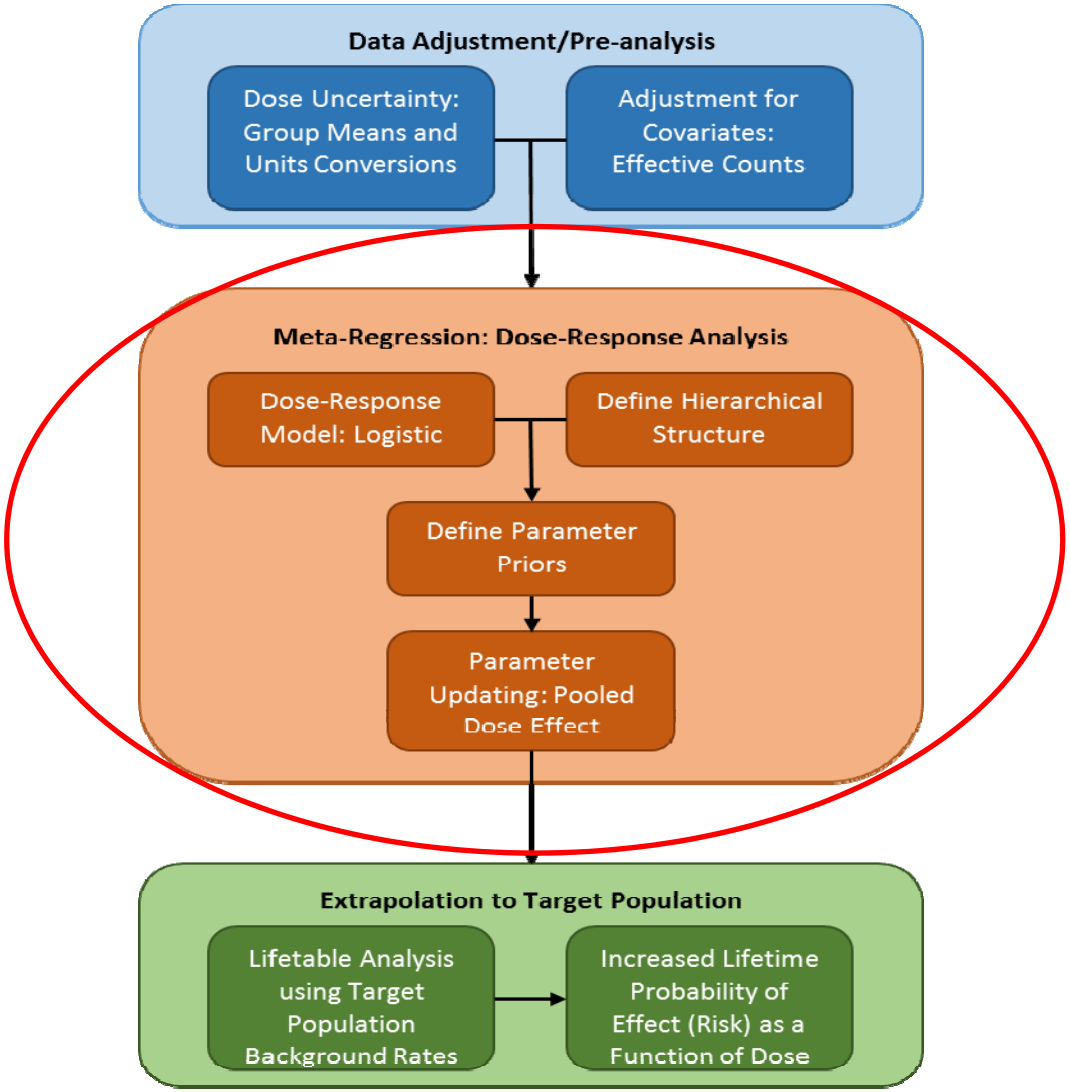
# Meta-Regression – "Effective Counts" Pre-analysis

- **The proposed Bayesian approach is based on likelihoods of observing a particular number of cases** (e.g., the number of observed cases in a cohort study) where the expected number given is described by a Poisson distribution

- **"Effective counts"** are adjusted counts of cases and controls that reflect only the effect of arsenic

- **For all groups**, adjustments are made so as to mimic data that might have been collected had the covariate levels remained the same as in the referent group.

# iAs D-R Methods Overview- Meta-Regression Analysis Flow Chart



# Meta-Regression – Dose-response Modeling

- **The Bayesian hierarchical under consideration** involves simultaneous modeling of individual studies to yield study-specific and pooled logistic model "β" parameter estimates that describe the relationship between iAs exposure and a health outcome.
- **Use of hierarchical (mixed effects) modeling** recognizes that there are differences among studies that lead to variation in study-specific estimates.
- **Prior information about specific model parameter values** are incorporated.
- **Variation** is expressed in terms of Bayesian posterior distributions rather than single, constant-across-studies point-estimates.
- **The degree of heterogeneity across study-specific estimates** is estimated to indicate whether more complex hierarchical methods are necessary to account for study-specific differences in effects, or whether a fixed-effects model is more appropriate.

# Meta-Regression – Bayesian Priors

- **Study-specific βs –** normal distribution assumed; allows positive and negative β values.

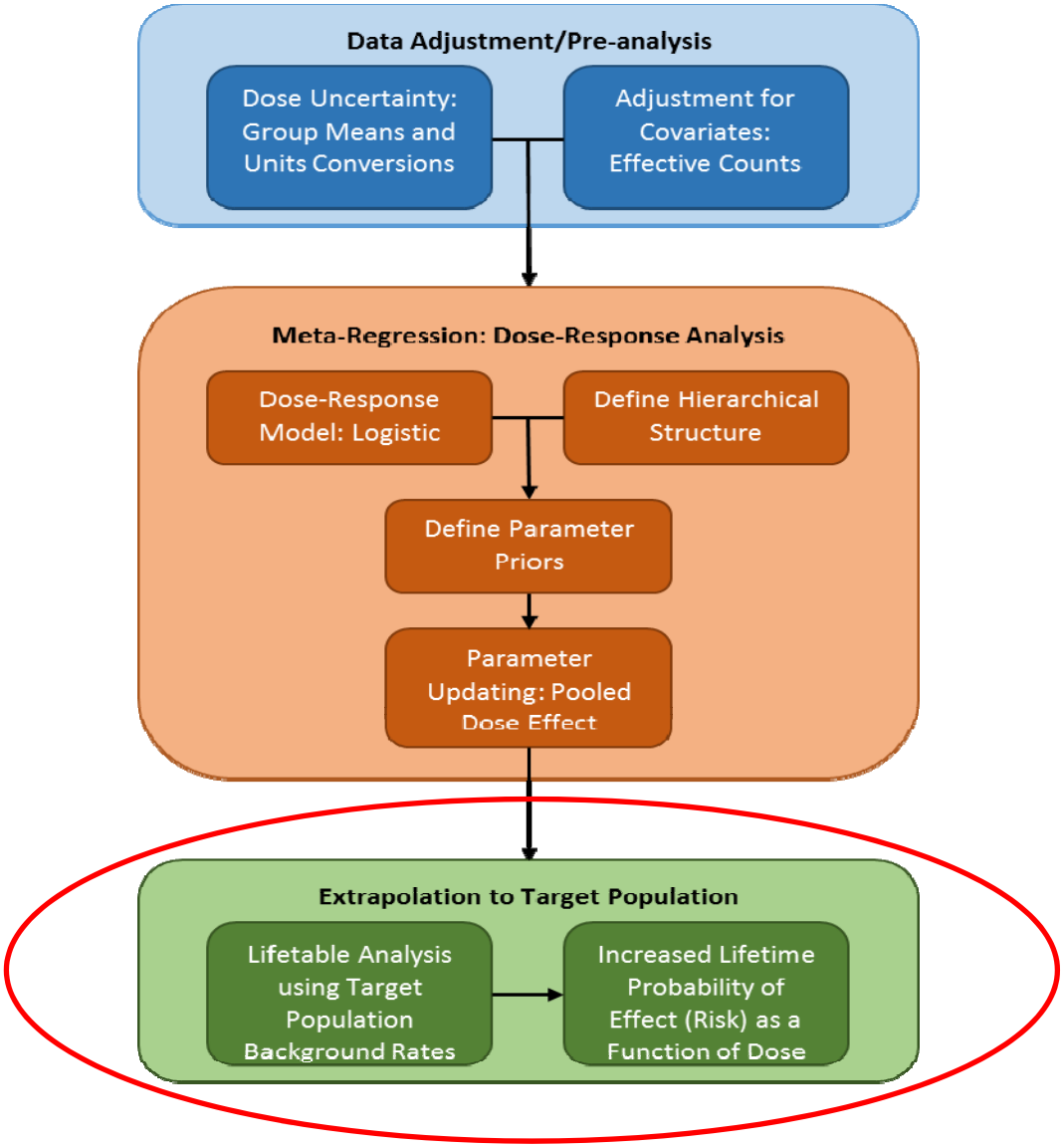
- **Pooled βs -** Gamma distribution assumed; allows only positive β values; reflects NRC (2013) determination that arsenic is causal for bladder cancer.

- **The α and *b* parameters of the standard Gamma distribution (below)** can be set to reflect expected upper and lower risk boundaries at a given iAs dose.

$$f(x) = \alpha e^{-\alpha x}(\alpha x)^{b-1} / \Gamma(b) \qquad x \geq 0$$

- **Gamma distributions** give greatest weight to β_mean values closest to zero.

# iAs D-R Methods Overview- Meta-Regression Analysis Flow Chart



# Meta-Regression – Extrapolation to Target Population

- **Life-table type analyses for extrapolating extra risk to target populations** typically involve summing the probability of cancer at age intervals adjusted by the probability of dying from other causes at each age interval

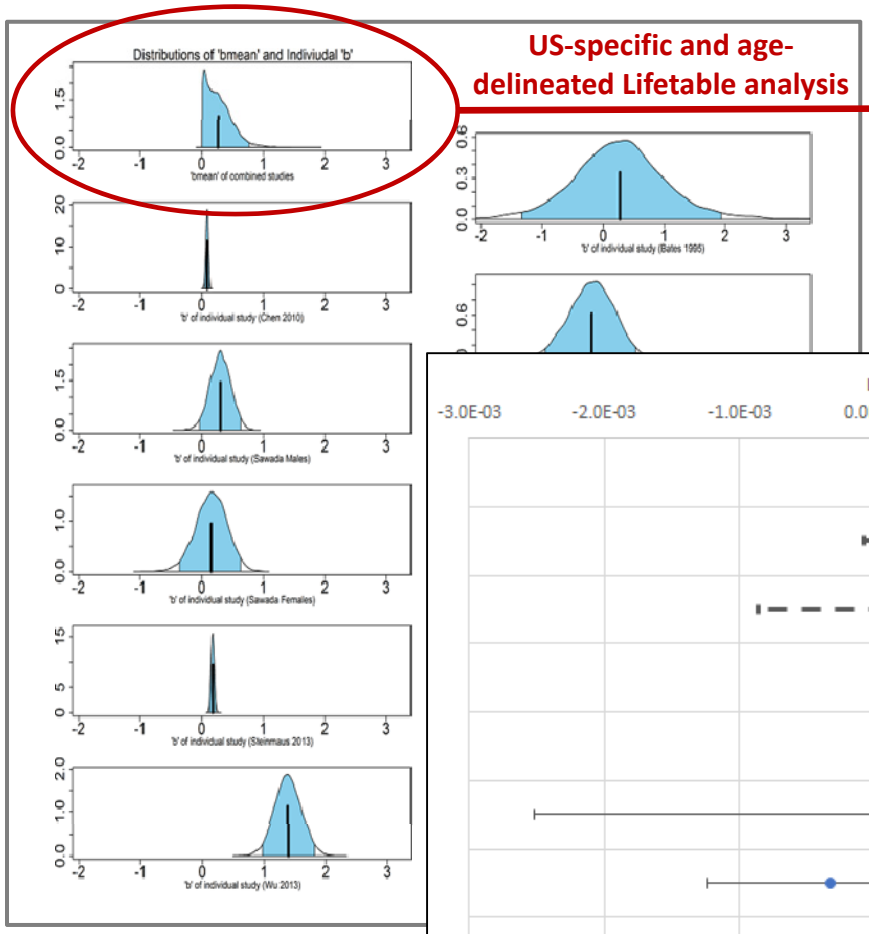
- **Posterior distribution of a pooled "slope" (i.e., logistic β parameter)** can be used to characterize uncertainty about risk estimates at any level of daily intake.

- **Background rates reported for U.S. health outcomes can be expressed as zero extra risk from iAs at a mean U.S. background dose**

- **A benefit of this approach** is that one can consider impact of both positive changes (additional sources of iAs above background) and negative changes (reduction in background associated with, for example, clean-up of background sources) on lifetime cancer risks.

# Bayesian Meta-Regression (example; not an actual assessment)

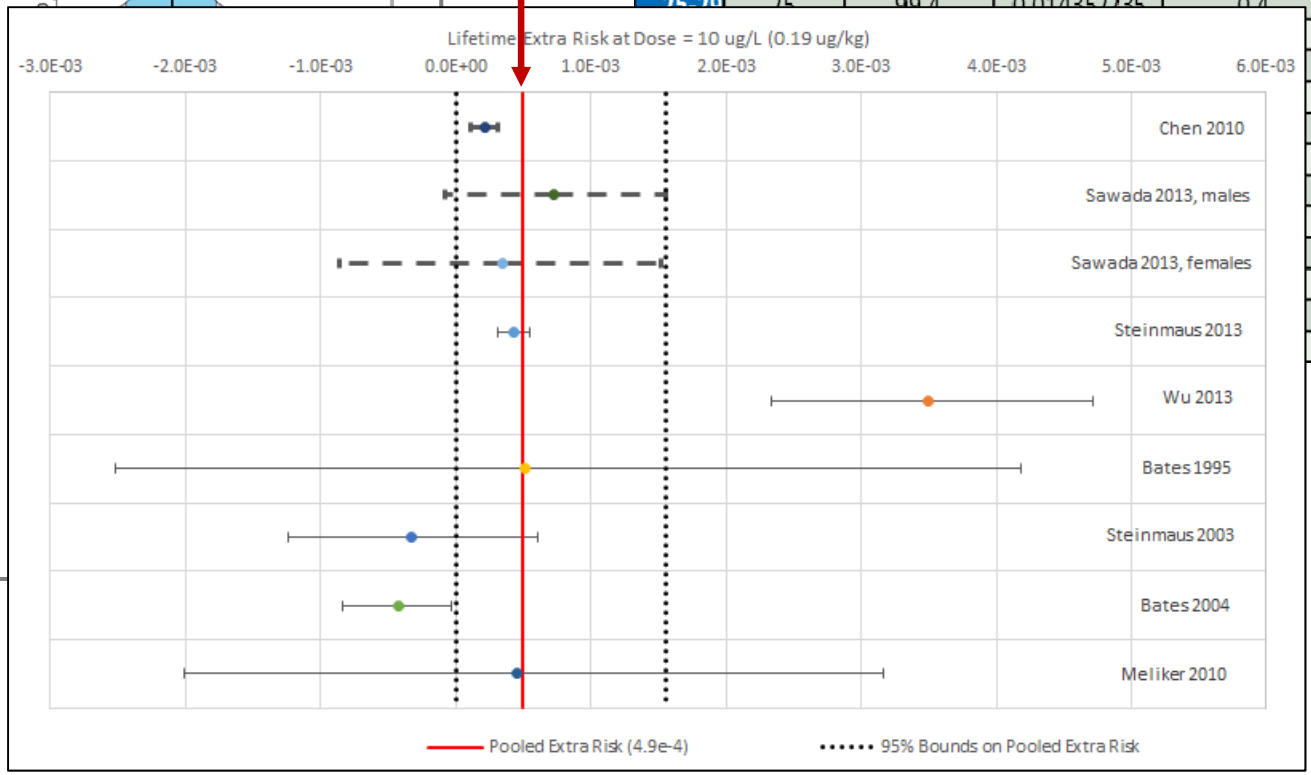
## *Combining 10 Studies, Logistic Modeling, and Lifetable Analysis*




**US-specific and age-delineated Lifetable analysis**




| Age Range | Starting Age | Annual all-Cause Deaths | Annual Bladder Cancer Deaths per 100,000 | annual Bladder Cancer Incidence per 100,000 |
|---|---|---|---|---|
| 0-1 | 0 | 678.9 | 0 | 0 |
| 1-4 | 1 | 29.3 | 0 | 0 |
| 5-9 | 5 | 12.6 | 0 | 0 |
| 10-14 | 10 | 15.2 | 0 | 0 |
| 15-19 | 15 | 55.9 | 0.004502298 | 0.1 |
| 20-24 | 20 | 93.4 | 0.009442598 | 0.1 |
| 25-29 | 25 | 99.4 | 0.014352235 | 0.4 |



# Meta-regression – Additional Sensitivity Analyses

**Additional sensitivity analyses could include:**

- Expansion of dose-uncertainty Monte Carlo pre-analysis (slide 10),
- Choice of datasets (“leave-one-out” type analysis)
- Assumption of zero inhalation background exposure
- Consideration of alternative gamma prior distribution values

## Potential Science Issues

- "Risk at a dose" (with confidence intervals) tables and equations versus traditional cancer CSF and noncancer RfD?
- Prior assumptions for meta-regression
  - Normal distribution for individual study slope estimates
  - gamma distribution of pooled slope estimate
  - α and b parameters of the standard Gamma distribution
- Model choice - possible future incorporation of model averaging with multiple fractional forms of the logistic model (raises issues for implementation of lifetable analysis)
- Nationwide background estimate and CDC cancer statistics - more targeted lifetable analyses could be done given population-specific background exposure estimate and age-interval cancer mortality and incidence data

# Acknowledgments for Developement of Dose-Response Methods


- iAs Dose-Response Team – Jeff Gift, Allen Davis, Janice Lee, Ila Cote, Ingrid Druwe, Tom Luben, Ellen Kirrane, Hisham El-Masri, David Thomas

- Contract Support - William Mendez, Natalie Blanton, Kevin Hobbie, Audrey Turley (ICF), Kan Shao (Indiana University), Bruce Allen (Allen Consulting, LLC)