Exponent®

Comments on the NRC and EPA Reviews of Epidemiologic Studies of Exposure to Arsenic in Drinking Water and Cancer

Prepared for

Wood Preservative Science Council
P.O. Box 293
Mt. Vernon, VA 22121

Prepared by

Exponent
1730 Rhode Island Ave. NW, Suite 1100
Washington, DC 20036

August 2005

Doc. no. BE02916.001 0101 0805 PM23

Contents

List of Figures

List of Tables

Acronyms and Abbreviations

CAD	cumulative arsenic dosage
CI	confidence interval
EPA	U.S. Environmental Protection Agency
Issue Paper	EPA's *Issue Paper: Inorganic Arsenic Cancer Slope Factor*
NRC	National Research Council
OR	odds ratio
PAC	peak arsenic concentration
QRA	quantitative risk analysis
SAB	Science Advisory Board
SMR	standardized mortality ratio
TCC	transitional cell carcinoma
Toxicological Review	EPA's *Toxicological Review of Ingested Inorganic Arsenic*

Comments on EPA's Review of Epidemiologic Studies of Exposure to Arsenic in Drinking Water and Cancer

Executive Summary

The Science Advisory Board (SAB) is asked in its current charge to consider whether the data set from southwestern Taiwan remains the "most appropriate choice" for estimating human cancer risks associated with inorganic arsenic in drinking water given the recent epidemiologic studies from U.S. and other populations with typically low-level exposures. In this document, we summarize and evaluate the reviews of epidemiologic studies from the NRC (2001) and U.S. EPA (2005a,b) reports on this issue. We describe the strengths and limitations of these studies and raise the question of whether the limitations, including potential biases, as noted by the National Research Council (NRC) and the U.S. Environmental Protection Agency (EPA) are likely to be of similar, lesser, or greater magnitude than the biases inherent in the studies from southwestern Taiwan. We also discuss issues related to the ability to generalize findings from the Taiwanese studies to the general U.S. population.

A major limitation of the review of the epidemiologic studies by NRC and EPA is that there is no systematic presentation of study strengths and limitations according to a uniform set of criteria. Although potential sources of bias are noted, the likely direction and magnitude of these biases are not evaluated formally. Thus, in some cases, the resulting impact of a given limitation could be minor. Finally, the recent epidemiologic studies are not reviewed in the context of a direct comparison with the southwestern Taiwanese studies in the NRC or EPA reports, and it appears that the criteria used to review the recent studies are not uniformly applied to the Taiwanese studies.

In general, the data from studies on exposure to low levels of arsenic in drinking water and risk of bladder cancer reviewed by either NRC or EPA (or both) do not suggest that low-level exposure is associated with increased risk. In addition, any potential biases in the studies reviewed do not appear to be likely to impact study results to a greater degree than biases

inherent in the studies from the Black Foot Disease endemic regions of southwestern Taiwan. Furthermore, the Taiwanese data are less likely to be generalizable to the U.S. general population than studies from the U.S. and other populations with low level exposures. With respect to the studies of bladder cancer reviewed, NRC and EPA have not provided justification as to why data from these studies should not be considered in the evaluation of human health risks associated with exposure to arsenic in drinking water or why the data from southwestern Taiwan should be considered more appropriate or valid.

Studies of lung cancer, skin cancer, and total cancers reviewed by NRC and/or EPA are also described and summarized. Although any observational epidemiologic study is subject to bias from a variety of sources, the studies reviewed are dismissed from further consideration without a "fatal flaw" being identified. Limitations are not discussed with respect to the southwestern Taiwanese studies, thus NRC and EPA do not make a case in support of the conclusion in EPA's *Issue Paper: Inorganic Arsenic Cancer Slope Factor* (U.S. EPA 2005a) that the southwestern Taiwan data set should be used in quantitative risk assessments for long-term exposure to arsenic in drinking water.

Finally, we review studies that were published in 2004 and were not reviewed by NRC or EPA. Notable among these studies was the finding by Michaud et al. (2004) that there was no association between toenail arsenic level and bladder cancer incidence, and that the relative risks observed in the study were below the range of relative risks predicted for exposure to arsenic levels of 50 μg/L based on dose-response curves derived from Taiwanese data. Similarly, in an ecologic study conducted in the United States, Lamm et al. (2004) reported results that were significantly below the risks predicted by NRC based on Taiwanese data.

In summary, NRC and EPA do not provide a systematic presentation of the epidemiologic studies, including the studies from southwestern Taiwan, that is based on consistent application of review criteria. As a result, there is no clear rationale or justification for the continued use of the southwestern Taiwan data set or the exclusion of potentially informative data from other studies, particularly those from the United States. The NRC (2001) report describes several sources of variability and uncertainty that apply to the application of the Taiwanese data to the general U.S. population. Many of the assumptions required to apply these data would not be

necessary if valid data from populations within the United States or similar to the United States were utilized in the risk assessment and estimation of dose-response curves.

Any dose-response curve that is derived should be validated with relevant data from the population of interest. That is, if data from southwestern Taiwan are used to estimate human cancer risks in populations with low exposure to arsenic in drinking water, then results from studies in these populations should be consistent with models derived from these sources. Based on results reported by Lamm et al. (2004) and Michaud et al. (2004), as well as data from other studies that generally have reported no consistent evidence of significantly increased risks at low exposure levels, this is clearly not the case. Similarly, the results of the meta-analysis being submitted by Exponent accompanying this report also found no increased risk of bladder cancer associated with low levels of exposure to arsenic in drinking water (Exponent, 2005). These results were not consistent with and most meta-relative risk estimates (mRRs) were below the range of relative risks predicted by the NRC report (2001). Thus, accurate models based on a more appropriate data set are needed to provide valid estimates of cancer risks to human populations exposed to low levels of arsenic in drinking water.

Introduction

In its evaluation of the long-term health effects of ingestion of inorganic arsenic, EPA has had to make decisions regarding which studies provide the best dose-response data to use in estimating human risks. The U.S. Environmental Protection Agency (EPA) and the National Research Council (NRC) (1999, 2001) have recommended that the primary source of data for this purpose should be the southwestern Taiwanese cancer mortality data from epidemiologic studies by Chen et al. (1985, 1988, 1992). In EPA's *Issue Paper: Inorganic Arsenic Cancer Slope Factor* (Issue Paper) (U.S. EPA 2005a), which has been made available to the Science Advisory Board (SAB) and the public, the Arsenic Cancer Slope Factor Workgroup agrees that the data from southwestern Taiwan should continue to be used as the critical studies in the dose-response assessment.

In EPA's *Charge to EPA Science Advisory Board Arsenic Review Panel* (U.S. EPA 2005c), it is noted in Section C2 that "since the NRC (2001) report on (inorganic arsenic), an additional body of literature has developed describing epidemiology data from populations in the United States exposed to (inorganic arsenic) in drinking water" (p. 5). The SAB is asked in this current charge to consider whether the Taiwanese data set remains the most appropriate choice for estimating cancer risk in humans in light of the additional epidemiologic studies.

The question of whether additional epidemiologic studies, particularly those evaluating risk of cancer in populations with exposure to low levels of arsenic in drinking water, is important because in the absence of such data and without information on specific modes of action, EPA uses linear extrapolation to estimate risks down to the origin. Whether this dose-response curve accurately predicts cancer risk among persons with low exposure has not been evaluated formally. Well-designed epidemiologic studies in such populations would prove useful and informative in this regard, in addition to providing data as to whether low-level exposures are likely to be causally associated with increased risk of bladder cancer and other cancers of interest.

In their reviews and critiques of epidemiologic studies, NRC (2001) and U.S. EPA (2005a,b) have considered potential sources of bias (i.e., systematic or non-random error) and issues of statistical power. NRC (2001) listed the criteria it used to evaluate the epidemiologic studies. These criteria were:

1. Accuracy of diagnoses or causes of death
2. Selection of an appropriate comparison population (or controls in case-control studies)
3. A clear definition of exposed and unexposed populations (for cohort and ecologic studies)
4. A high follow-up rate (cohort studies)
5. Adequate response rate among cases and controls (case-control studies)
6. Statistical power.

The NRC report (2001) states, "Findings from small studies, even those with excellent methodology, are of limited utility" (p. 38). This statement fails to take into consideration the magnitude of the association between the exposure and the disease. For example, small studies may be sufficient for large effect sizes. Furthermore, large studies with significant methodologic limitations are also of limited utility.

Although the series of studies conducted in southwestern Taiwan suffered from methodological limitations and were subject to potential sources of bias, these studies were selected as the primary data source, whereas other studies were dismissed due to potential bias and/or lack of precision. This is discussed in further detail below.

EPA's *Toxicological Review of Ingested Inorganic Arsenic* (Toxicological Review) (U.S. EPA 2005b) did not list criteria for reviewing the epidemiologic studies, but did note that few studies assessed the role of nutritional factors such as selenium or zinc deficiencies. Such deficiencies, however, are likely to be far less severe and less frequent in the United States and other

nutritionally sufficient populations compared to southwest Taiwan, Bangladesh, or West Bengal; thus, it seems unlikely that these variables would act as strong ***confounding*** factors in analyses of data from nutritionally sufficient populations. It has been suggested that these factors may act to ***modify*** or enhance arsenic toxicity (U.S. EPA 2005b), and that malnutrition in general may increase the susceptibility to adverse health effects associated with arsenic (NRC 2001). It seems plausible that accounting for these and other nutritional factors may be more important in studies conducted among nutritionally deficient populations, such as southwestern Taiwan, than in the studies of low arsenic exposure where the populations do not generally suffer from malnutrition.

The following discussion considers several of the studies reviewed by NRC and EPA, as well as some studies that were not reviewed, perhaps due to their recent publication date. We evaluate the potential biases in the studies and their likely impact on the study results. The best approach to address the question of whether bias could have caused a truly significant finding to be observed as a smaller, non-significant effect is a formal sensitivity analysis. This approach, however, is beyond the scope of these comments, although we do point out some specific examples of where a sensitivity analysis would have been particularly informative regarding the likely magnitude of any bias that could have been introduced. We consider the southwestern Taiwanese studies, the potential sources of bias and confounding in these studies, and whether the bias is likely to be less than, greater than, or equal to any bias in the studies that were dismissed by NRC and EPA. Throughout the discussion, we consider the NRC and EPA reviews of the epidemiologic studies and whether the relative strengths and limitations of the studies justify the current practice and recommendation by EPA to rely on the data from southwestern Taiwan to the exclusion of data from the United States and other populations with low exposure to arsenic in drinking water (U.S. EPA 2005a). To address this issue we evaluate whether the two bodies of literature are held to similar review criteria and standards.

Studies Reviewed in NRC (2001) and U.S. EPA (2005a,b)

Bladder Cancer

Chiou et al. (2001) evaluated the association between ingested arsenic and risk of transitional cell bladder cancer in a cohort of 8,102 residents from 18 villages in northeastern Taiwan. After adjustment for age, sex, smoking, and duration of well water drinking, relative risks of bladder cancer increased for each increasing category of arsenic concentration, although relative risks were not significant at arsenic water concentrations below 100 μg/L (1.0 [referent], 1.9 [95% confidence interval, or CI: 0.1–32.2]; 8.1 [95% CI: 0.7–98.2]; 15.1 [95% CI: 1.7–138.5]), respectively, for arsenic concentrations in well water of ≤ 10.0, 10.1–50.0, 50.1–100.0, and > 100.0 μg/L). NRC (2001) listed many strengths of this study, including the use of incident cancer cases, collection of information on residential history, water-use history, and cigarette smoking. Limitations noted included using only one measurement of well-water to represent long-term exposure and short duration of follow-up time. The latter resulted in a small number of cases and very wide confidence intervals. NRC (2001) concluded that the data from this study were "too imprecise" to be used in a quantitative risk assessment, but that they could serve as supplementary information with data from other studies. The EPA Toxicological Review characterized the findings of this study as follows: "A significant dose-response relationship was also observed after adjustment for age, sex, and cigarette smoking" (p. 27). The EPA Issue Paper states, "the study concluded that the increase in arsenic-induced transitional cell carcinoma (TCC) was more prominent for those individuals who were exposed to the contaminated drinking water for more than 40 years" (p. 5). The data from Chiou et al. (2001) do not support this statement, however, (relative risk = 1.0 for duration of well water drinking of 40 years or longer) and the authors state, "…there was no dose-response relation between the risks of urinary tract cancer and TCC and the duration of well water drinking" (p. 413). Thus, the EPA documents (U.S. EPA 2005a,b) seem to focus on the "positive" findings of this paper, yet dismiss it for its lack of precision. NRC (2001) acknowledged the imprecision of the study at this point in time, but recognized that it could provide useful information, particularly after additional follow-up, due to the strengths of the basic study design. The latter interpretation seems more reasonable.

Kurttio et al. (1999) assessed the levels of arsenic in drilled wells in Finland and examined the association between arsenic exposure and risk of bladder and kidney cancer in a case-cohort study. Information was collected on residential history, drinking water consumption, smoking, analgesic and diuretic use, education, and occupation. Analyses were stratified by "smoker" or "never or ex-smoker" status. Relative risks were below 1.0 (but not statistically significant) for never or ex-smokers in analyses based on arsenic in water concentration exposure and cumulative exposure categories. In contrast, relative risks were above 1.0 among smokers in each of the exposure categories. NRC (2001) notes that the findings from this study were not internally consistent. For example, the overall exposure level was low (the 95th percentile of arsenic exposure was 3.0 μg/L among the cases and 4.5 μg/L among the referent), yet the authors found a significantly increased risk of bladder cancer. This excess was limited to those in the "shorter" latency group rather than the "longer" latency group. This finding is not consistent with a causal explanation. The EPA Toxicological Review suggests the possibility of exposure misclassification. Unless misclassification is differential, this type of bias will tend to be toward the null. The review does not suggest reasons why exposure recall would likely differ for cases and controls in this study, which would be necessary for recall bias to be present. The EPA review also criticizes the study for not taking nutritional factors into account, yet there are no epidemiologic data suggesting that inclusion of such variables would significantly alter the results of the analysis, particularly in a population that presumably has a diet that is nutrient-sufficient. This criticism may be more applicable to the studies conducted in southwestern Taiwan, where it has been suggested that nutrient-poor diets may have contributed to susceptibility to cancer and other adverse health outcomes (Schoen et al. 2004; NRC 2001).

Lewis et al. (1999) examined the association between drinking water arsenic and mortality outcomes (including cancer) in a cohort of residents from Millard County, Utah. The study cohort was assembled from historical membership records of the Church of Jesus Christ of Latter-day Saints (Mormons). An arsenic exposure index score, derived from the number of years of residence in the community and the median arsenic concentration of community drinking water, was calculated for each person in the cohort, and divided into three categories for analysis. For comparison, mortality-specific expected death rates were generated from the white male and white female general population of Utah. Strengths of this study noted by NRC

were greater than 90 percent follow-up of the cohort for vital status, the large size of the cohort, and the large number of deaths upon which to base the analyses. The NRC report also discussed several limitations, including exposure measurement concerns and the use of an external comparison population. Although the exposure metric used was analogous to that commonly used in epidemiologic studies to quantify and characterize exposure to cigarette smoking (pack-years of smoking), the NRC report noted that use of this metric made comparison to other studies difficult because it was not commonly used and that individuals with exposure to average levels of arsenic over long periods of time would be grouped in the same exposure category as those with exposure to high levels over a shorter duration. Use of this type of metric has proved useful in studies of cigarette smoking and has not precluded the identification of causal associations with numerous adverse health outcomes.

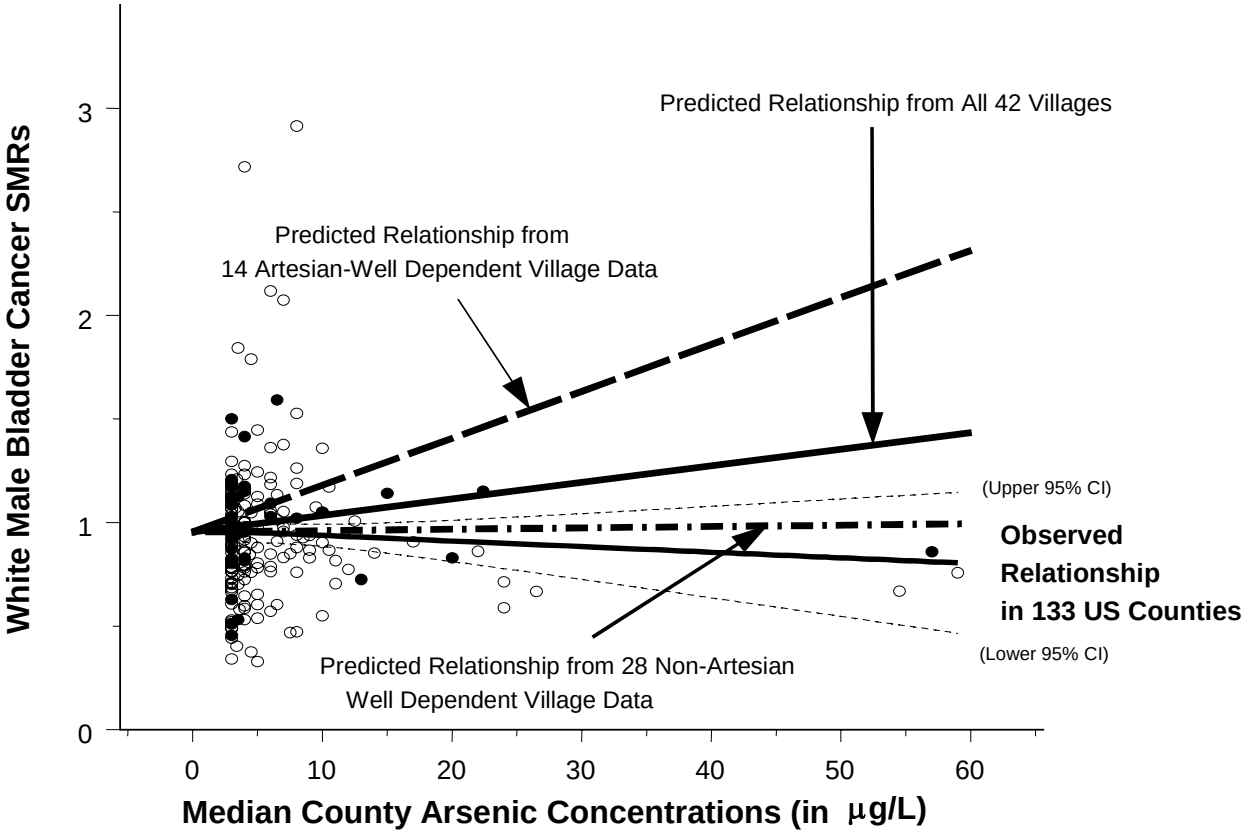
The NRC report states, "comparison rates used in the analysis of risk were for the state of Utah, and this comparison likely resulted in underestimates of risk for some causes of death and overestimates of others" (p. 55). The main concern was that the study population, who were members of the Church of Jesus Christ of Latter-day Saints were less likely to be smokers than the general population of Utah. By knowing the rate of smoking in Utah, the estimated association between smoking and a particular cancer (e.g., bladder cancer), and the rate of bladder cancer in the state of Utah, it would be possible to conduct a sensitivity analysis to determine how much of an observed deficit in bladder cancer is likely due to the difference in smoking rates in the study population versus the comparison population. It is possible that after taking this into account, the externally adjusted standardized mortality ratio (SMR) may still have been indistinguishable from unity. Even if the rates of smoking differed between the study population and the comparison population, it is worth noting that no dose-response association was observed based on the cumulative exposure analysis or by well water concentration. Finally, in the study by Lewis et al. (1999), a healthy population was compared to a population that was considered to be not as healthy. The converse is true in the southwest Taiwan studies, where a population that is considered to be nutritionally deficient has been compared to a population that is likely healthier (the general population of Taiwan or the United States). This caveat regarding comparison groups should be considered when interpreting the elevated risks observed in the Taiwanese studies, in addition to results from the U.S. studies. The paper by

Morales et al. (2000) illustrates the sensitivity of lifetime risk estimates to both the choice of comparison population (southwestern Taiwan, all of Taiwan) and whether or not a comparison population is used. Even though the ideal comparison population (including no comparison population) for the Taiwan studies is still being considered (U.S. EPA 2005a), alternative methods for analyzing the Lewis data (and/or for evaluating the potential bias associated with using the Utah population comparison) have not been given further consideration.

Lamm et al. (2003) obtained the southwest Taiwan data that were evaluated originally by Wu et al. (1989) and Chen et al. (1985, 1992), and reanalyzed by Morales et al. (2000), and created a continuous variable for arsenic exposure. The authors then classified villages as "shallow," "mixed," or "artesian," depending on the water source. For villages solely dependent on artesian well water, arsenic levels ranged from 350 to 934 μg/L, and for villages with non-artesian well water sources, arsenic levels ranged from 10 to 717 μg/L. Findings from this study are presented in Figure 1 together with the data from the Lamm et al. (2004) study. There was no association between arsenic level and risk of bladder cancer in the analyses of villages with non-artesian well sources. In contrast, an increasing dose-response relationship was observed for artesian well water sources. Lamm et al. (2003) propose, "that the best data from the SW Taiwan data set for QRAs [quantitative risk analyses] at low arsenic levels would be the data on the villages not dependent on artesian wells or not using artesian wells. Furthermore, QRAs that have used the SW Taiwan data to estimate risk in the USA have introduced further error by attempting to convert Taiwan exposures into US equivalent exposures with extrapolation to the US exposure range and by assuming that health care parameters in rural Taiwan are equivalent to those in the US." They suggest further that, "the best basis for the calculation of bladder cancer risk in the United States from ingestion of arsenic in drinking water is to use US data" (p. 367).

This interpretation by Lamm et al. (2003) is not in agreement with the recommendations made by U.S. EPA (2001) and NRC (1999, 2001) to use the southwestern Taiwan data set as the primary source for conducting QRAs for long-term exposure to arsenic in drinking water. The EPA documents submitted to the SAB (U.S. EPA 2005a,b) discuss several limitations of the Lamm et al. (2003) study. The EPA Issue Paper questions the validity of the well classifications

according to arsenic concentration. It states, "the authors artificially classified village well types into three categories (shallow, < 0.325 ppm; mixed, wells above and below 0.325 ppm; and artesian, > 0.325 ppm). … The validity of Lamm's reclassification is impossible to assess with the information provided" (p. 7). The EPA Toxicological Review also suggests that, "the classification into village well type was based on median arsenic concentrations such that well type and arsenic concentration are not independent variables" (U.S. EPA 2005b, p. 33). It may not have been possible to achieve this, however. Exposure misclassification for either artesian or non-artesian wells appears minimal based on the non-overlapping ranges in concentrations for these types of wells reported by Chen et al. (1985) as cited by Lamm et al. (2003). Lamm et al. (2003) make the point that all well types contained water with arsenic, yet the association was only apparent for the artesian wells, suggesting the possibility that something other than, or in addition to arsenic was contributing to the elevated risk estimates. Humic acids are found in artesian wells, for example, and have been proposed as a co-carcinogen or promoting factor in arsenic-related cancer (Lamm et al. 2003). Lamm and Kruse (2005) summarize other differences between artesian wells and shallow wells, including concentrations of specific compounds (e.g., ammonia, nitrogen, silicate), turbidity, and exposure of water to sunlight (and related algae growth). Regardless of water source, the classification of arsenic levels based on quantified cut-points will still yield results that are pertinent to the categorical range of exposure. This is evident by the finding of no association for low levels of arsenic exposure in the study population.



Source: Lamm and Kruse (2004)

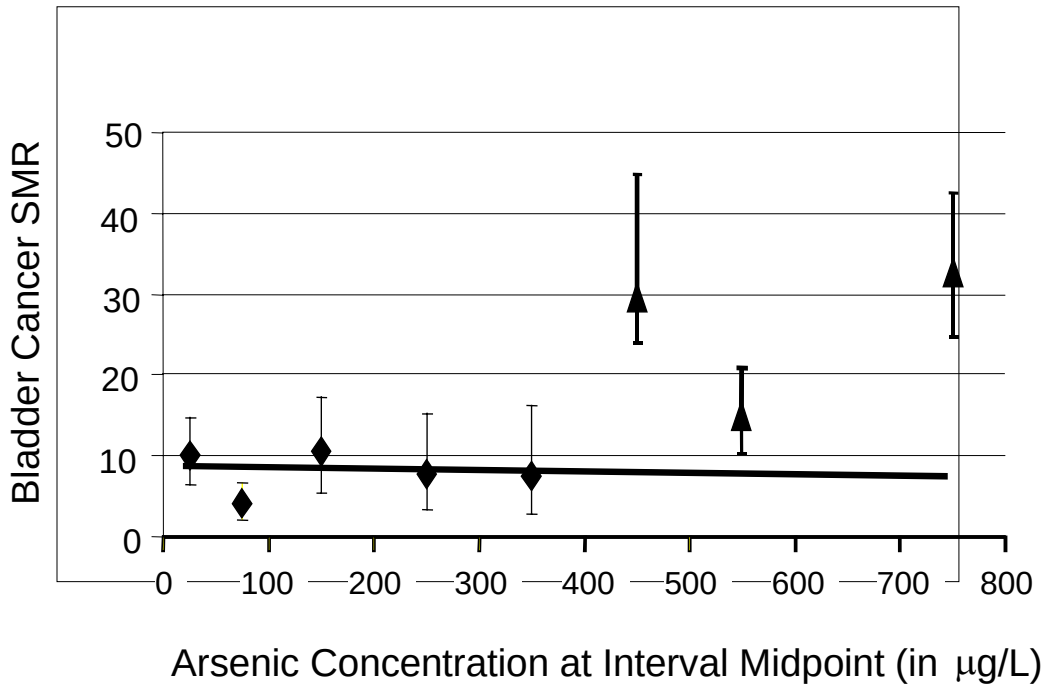
Figure 1. Observed vs. predicted bladder cancer SMRs in white males: a comparison of southwest Taiwan and US datasets

Information on smoking was not available for this southwest Taiwan data set analyzed by Lamm et al. (2003), and the EPA Issue Paper notes that, "it is evident from past literature that age and smoking can dramatically affect the onset of adverse health effects due to arsenic exposure" (U.S. EPA 2005a, p. 7). If smoking prevalence was greater in the comparison population, results may be biased downward. On the contrary, if smoking prevalence was greater in the study population, relative risk estimates may be biased upward. If the assumption can be made that age-specific smoking rates were the same in both populations, then the relative risk estimate may not be confounded by smoking status. Effect modification, however, may be present. If no comparison population is used, smoking may act as a positive confounder and bias relative risk estimates upward. The U.S. EPA (2005a,b) documents do not provide estimates for the magnitude or direction of potential confounding caused by the inability to take

smoking into account in the analyses. It is worth noting, however, that not all epidemiologic studies of bladder cancer have shown evidence of strong confounding or effect modification by smoking (see for example, Bates et al. 2004; Karagas et al. 2004).

In the EPA Toxicological Review, there is suggestion "that the use of linear regression may not be appropriate because bladder cancer rates are not normally distributed, and that a Poisson regression would have been more appropriate" (p. 33). Findings based on categorical data analysis revealed that there is no change in bladder cancer mortality for village arsenic exposure from 10 to 400 μg/L (Figure 2). Lamm et al. (2003) states that the discontinuity of data at approximately 400 μg/L may have been a "consequence of combining data from different populations of villages" (p. 362). In addition, two possible mechanistic explanations were suggested, "(1) that arsenic behaves like a high-dose phenomenon with respect to bladder cancer in that it only demonstrates its effect at high arsenic exposure level, or (2) that arsenic behaves like a co-carcinogen, for instance, with some constitutive factor that distinguishes artesian well water from other waters in SW Taiwan" (p. 362). U.S. EPA (2005a) dismisses the co-carcinogen hypothesis: "Little evidence exists for the co-carcinogen explanation" (p. 7). Nevertheless, as summarized by Lamm et al. (2003), several publications have reported on the presence of mutagenic substances in artesian wells in southwest Taiwan.

Lamm et al. (2004) conducted an ecologic study of bladder cancer and arsenic in drinking water based on data from 133 counties in the United States. Data on bladder cancer mortality in white males and arsenic contamination in groundwater were ascertained on the county level. The median exposure level ranged from 3 to 60 μg/L. The authors reported no arsenic-related increase in bladder cancer mortality in their study. Moreover, results from Lamm et al. (2004) excluded NRC's 2001 risk estimate that was based on data from southwest Taiwan. The upper 95% confidence limit for lifetime mortality due to bladder cancer for white males was 4.2×10^{-5} in Lamm et al. (2004) whereas the NRC risk estimate was 8.5×10^{-6}. EPA's 2001 estimates, however, were not excluded by the data from this study.



Source: Lamm and Kruse (2004)

Figure 2. Bladder cancer SMRs stratified by arsenic concentration, based on southwest Taiwanese data

Figure 1 shows the disparity between the findings from the United States reported by Lamm et al. (2004) and the southwest Taiwan data. There are diverging associations as arsenic levels are increasing. This study may be a more appropriate model for U.S. regulatory guidance and relative risk estimates, although the ecologic design of the study precludes causal interpretation of the results, and aggregate data such as those in this study cannot be assumed to apply at the level of the individual; these limitations also apply to the ecological data from southwest Taiwan.

The EPA Issue Paper discusses several limitations in this ecological study. As noted above, this study is ecological in design; therefore, individual exposure and outcome measurements were not ascertained. The EPA Issue Paper highlights possible sources of variability, including "age, duration of exposure, latency, occupation, and smoking" (U.S. EPA 2005a, p. 8). Although these factors were not individually accounted for in the southwest Taiwan studies, the NRC report stated that these studies "are the strongest sources of dose-response information for cancer endpoints" (p.181). It is not clear why the data from the Lamm et al. (2004) study,

which are clearly more generalizable to the U.S. population, should be disregarded in the EPA risk assessment. Furthermore, the EPA reports (U.S. EPA 2005a,b) do not provide rationale for weighing the data from the Taiwanese ecologic studies more heavily than the data from the Lamm et al. (2004) U.S. ecologic study.

The EPA Issue Paper claims that statistical power to observe associations below 1.7 is "severely limited in counties with arsenic greater than 50 μg/L" (U.S. EPA 2005a, p. 8). This may be true, but the emphasis should be on the increased power to detect an association among persons with low level arsenic exposure where most of the sample population occurred (these levels that are of particular relevance to exposure to arsenic in the United States). The EPA Issue Paper did not discuss statistical power for analyses of counties with arsenic less than 50 μg/L, and statistical power at these lower water arsenic levels presumably is not considered to be an issue of concern.

Finally, the EPA Issue Paper mentions that counties with zero cases were excluded from the analysis, and that, "this exclusion may have resulted in biases, especially if these counties were also areas with low arsenic exposure" (U.S. EPA 2005a, p. 8). Findings may be biased, but the bias would be towards an increased association. If the data for the other counties were added, the numerator would be the same (no more added cases) and the denominator would be larger (added person-years for the newly included counties). This provides further evidence that there is no risk for bladder cancer among low levels of arsenic in drinking water exposure.

Steinmaus et al. (2003) conducted a case-control study to evaluate the association between arsenic ingestion in drinking water and bladder cancer in six counties in western Nevada and Kings County in California. There were 181 incident bladder cancer cases and 328 controls who were frequency matched to cases by 5-year age group and gender. The EPA Toxicological Review lists as one of the limitations, "non-matching characteristics of the cases and controls involving income, education, and smoking" (U.S. EPA 2005b, p. 34). It is not clear whether EPA is suggesting that the study authors should have matched on these variables, or is merely pointing out that, for example, "controls were significantly more likely to be in a higher income bracket than cases" (U.S. EPA 2005a, p. 9). According to the study authors, "cases were more likely to be in the lower income bracket, less educated, and current smokers" (p. 1195). These

findings are not surprising given that smoking is considered to be the major cause of bladder cancer (Kogevinas and Trichopoulos 2002), and smoking behavior is more common among lower socioeconomic classes. Although adjustment for these factors changed the odds ratios only slightly, the analysis was done appropriately and it would not have been feasible to additionally match on these three variables. It is recommended to match only on strong confounders because efficiency may be lost otherwise. Income and education were not likely to be strong confounders given that their association with bladder cancer is not clear. Furthermore, it is not uncommon for participation rates among controls to be somewhat skewed toward higher education levels and income. Any potential bias due to the pattern of differences in the distribution of smoking, income, and education as reported by the authors, if incompletely adjusted for in the analysis, would likely bias results upward; that is, it would spuriously increase the magnitude of the observed odds ratios.

Each residence within the study area was linked to a water arsenic measurement for that residence. Daily arsenic intakes (μg/day) for a given year were estimated, as were cumulative exposure (mg) categories. Non-significant increased odds ratios were reported for persons with 6.4–82.8 mg (odds ratio, or OR = 1.63, 95% CI: 0.64–4.13) and > 82.8 mg (OR = 1.40, 95% CI: 0.73–2.70) of cumulative exposure (compared to those with < 6.4 mg of cumulative exposure), based on 40-year lag analysis. Findings that stratified by smoking status (never, ever) and lag time (5-year, 40-year) did not reveal consistent findings. In the 40-year lag analysis, increased associations that ranged between 2.25 and 4.01 among ever smokers were reported in the highest categories of year-average arsenic exposure (> 80 μg/day) and cumulative exposure (> 82.8 mg). Odds ratios were not significantly elevated for smokers with 40-year lags and with year-average arsenic exposure of 10–80 μg/day or cumulative exposure of 6.4–82.8).

The EPA Issue Paper states that the "most significant problem with this study may be the interpretation and conclusion" (U.S. EPA 2005a, p. 9). EPA suggests considering the elevated odds ratios and potential trends, as opposed to focusing on the lack of statistical significance. Any careful evaluation, however, would include examining the width of the 95% confidence intervals to get an indication of the precision of the relative risk estimates; this is different from

using the confidence intervals simply as a means of significance testing. The EPA Issue Paper does acknowledge a lack of trend significance, although the variability in smoking stratified analyses was not discussed. There were no "potential trends" in these analyses; all of the odds ratios for never smokers in the highest exposure categories with a 40-year lag were below 1.0.

Although the Steinmaus et al. (2003) was a case-control study, the EPA Toxicological Review discusses sample size requirements for ranges of relative risks (cohort studies). It is unclear as to why it does not address power and sample size considerations for a case control study, which is the design of Steinmaus et al. (2003), and would be more relevant to the discussion of this study. It is well known that a very large sample size is required for cohort studies of rare disease. A case-control study is an efficient design for rare diseases because cases are selected (i.e., from a registry, hospital, or other source), thus circumventing the need for "17,045,000" persons to detect a "very weak (1.15 relative risk)" as suggested by the EPA Toxicological Review (2005b, p. 35). Despite the fact that most funding agencies (e.g., the National Institutes of Health) would be unlikely to support a study with an aim of detecting a relative risk as small as 1.2, if one were to assume an alpha level of 0.05 (two-sided), a beta of 0.20, and an exposure prevalence among the controls of approximately 80% (that is, 80% of the controls would be expected to have arsenic concentrations in the "low" range of 0 to 19 μg/L, based on Table 2 in Steinmaus et al. 2003), a sample size of approximately 3,100 cases and 3,100 controls would be required to detect a significant odds ratio of 1.2 (Schlesselman 1982). For a one-sided test, 2,450 cases and 2,450 controls would be required. To detect a relative risk of 1.5, approximately 680 cases and 680 controls would be required, under the same parameters (535 cases and 535 controls for a one-sided test). To detect a relative risk of 2.0, the number of cases and controls required would be reduced to 260, respectively, for a two-sided test, and approximately 200 for a one-sided test. Given that the concern is that cancer risks will be increased, a one-sided test seems appropriate. For any of the above scenarios, these sample size estimates for an unmatched case-control study are significantly lower than the sample sizes estimated by the EPA Toxicological Review.

The EPA Issue Paper and Toxicological Review focus on the non-significant positive findings and suggest the study suffers from limited statistical power and potential residual confounding

and/or lack of external validity due to characteristics of the control group. This draws attention away from the finding that there was no suggestion of increased risk among nonsmokers with a 40-year lag, or that the odds ratios did not consistently increase with increasing exposure in all analyses. Furthermore, this study has sufficient power to detect a significant odds ratio of between approximately 1.6 and 2.0, depending on the frequency of the exposure in the population (controls).

Summary of Bladder Cancer Studies Reviewed by NRC (2001) and U.S. EPA (2005a,b)

Each of the above studies (Chiou et al. 2001; Kurttio et al. 1999; Lewis et al. 1999; Lamm et al. 2003; Lamm et al. 2004; and Steinmaus et al. 2003) was reviewed by NRC (2001), U.S. EPA (2005a and/or 2005b), or both. With the exception of the data from Chiou et al. (2001), none of the studies reviewed provided data that supports the hypothesis that exposure to low levels of arsenic in drinking water is associated with increased risk of (or death due to) bladder cancer. Chiou et al. (2001) reported elevated odds ratios, but the 95% confidence intervals were very wide, indicating imprecision in the estimates. Although limitations of these studies were discussed in the reviews (NRC 2001; U.S. EPA 2005a,b), none of the studies appeared to suffer from potential biases that would be any greater than those in the ecologic studies conducted in the Black Foot Disease endemic regions of southwestern Taiwan. In addition, the issue of "external validity," and which data would be most appropriate in terms of generalizability, was not given sufficient attention. Thus, neither NRC nor EPA have provided rationale or justification as to why the southwestern Taiwan data set should continue to be used, to the exclusion of data on bladder cancer from the United States and other areas with low exposure, in the quantitative risk assessment.

Lung Cancer

Ferreccio et al. (2000) conducted a case-control study in northern Chile based on 151 lung cancer cases diagnosed between 1994 and 1996 and concentration of arsenic in drinking water. Mean well water concentrations in cities in northern Chile were 860 μg/L during the period

1958–1970, and were subsequently reduced to 40 μg/L. Two frequency-matched hospital controls (total controls = 419) were selected for each case. The first control series included randomly selected patients admitted to a chosen hospital with cancer not known or suspected to be related to arsenic. The second control series was selected among patients admitted to "the next hospital on the list" with a diagnosis other than cancer, but excluding cardiovascular, skin, or neurologic diseases. The authors note that control selection is the main weakness of the study, as the use of hospital controls with matching by hospital could have resulted in matching by exposure. Furthermore, there was overselection of controls from the city of Antofagasta, which had higher exposure levels. Results indicated a pattern of increasing lung cancer odds ratios with increasing concentration of arsenic in water. Multivariate-adjusted odds ratios ranged from 1.0 (referent) to 7.7 for average water arsenic concentration categories ranging from 0–10 μg/L to 200–400 μg/L during the years 1930–1994. Results were generally similar, but somewhat attenuated, when peak exposure years of 1958–1970 were analyzed. In these analyses, multivariate-adjusted odds ratios for the exposure categories 10–29 μg/L and 30–59 μg/L were not significantly elevated when compared to exposure levels of 0–10 μg/L. In analyses stratified on smoking status, odds ratios were higher for ever smokers than never smokers.

The NRC report stated, "Strengths of this study include an acceptable response rate, unbiased ascertainment of exposure, individual estimates of exposure, exposure coverage of most of the life span for most study subjects, incorporation of individual data on other potentially confounding risk factors for lung cancer, appropriate analyses of study data, and an adequate study size" (p. 52). These strengths were also suggested in the EPA Toxicological Review. The NRC report and the EPA Toxicological Review consider the control selection methodology to be a "major" limitation of the study. The NRC report concludes that, "data from this study can be used in quantitative assessment of risk of arsenic in drinking water, along with data from other selected studies" (p. 52). The EPA Issue Paper concludes that the data from Ferreccio et al. (2000) are "not precise enough for quantified risk assessments" due to wide and overlapping confidence intervals (U.S. EPA 2005a, p. 11). For this study, issues surrounding the validity of odds ratios based on the control group should be the first concern in considering how much weight to give these data. The authors of the paper and the NRC report suggest that any bias

would be in the direction of the null (underestimation of relative risk estimates), particularly for the higher exposure categories. Ferreccio et al. (2000) and NRC (2001) do not mention that the analysis also overestimates risk at low exposures. This issue may warrant further formal evaluation.

Chen et al. (2004) conducted a prospective study of 2,503 residents in southwestern Taiwan and 8,088 residents in northeastern areas of Taiwan; both areas were arseniasis-endemic. One hundred thirty-nine incident lung cancer cases were identified over the study period (average = 8 years). Arsenic levels were measured in well water samples from each region and ranged from 350 to 1,140 μg/L in the southwestern cohort and from < 0.15 to 3,590 μg/L in the northeastern cohort. The authors reported a significant positive association and exposure-response pattern for increasing average arsenic levels in well water and lung cancer incidence. The relative risks were elevated further in smokers and were attenuated and no longer significant in analyses restricted to nonsmokers.

The EPA Issue Paper notes that, "after adjusting for cigarette smoking and other risk factors such as age, alcohol consumption, and years of schooling, a significant ($p < 0.001$) increasing trend in lung cancer was shown to result from increasing average levels of arsenic in well water. With levels < 10 μg/L as the referent, relative risks (with 95% confidence intervals) for those consuming drinking water with arsenic concentrations of 10–99, 100–299, 300–699, and ≥ 700 μg/L, were respectively, 1.09 (0.63–1.91), 2.28 (1.22–4.27), 3.03 (1.62–5.69), 3.29 (1.60–6.78)" (p. 11). Although these numbers are correct, the 95% confidence intervals largely overlap, and the trend statistic should be questioned. Arsenic levels in well water were relatively high in this study, however, there is no association for the second category of arsenic exposure (10–99 μg/L) (1.09 [0.63–1.91]). Of all exposure groups, this category is most relevant to addressing low level arsenic exposure and risk of lung cancer. Analyses that stratified by smoking intensity and average arsenic exposure were presented; however, the categories of arsenic exposure used are not conducive to estimating risk at low levels (categories of exposure: < 10, 10–699, ≥ 700).

Multiple Cancers

Tollestrup et al. (2003) conducted a community-based retrospective cohort study to evaluate the association between childhood exposure to ambient arsenic exposure and mortality. The cohort included 1,827 males and 1,305 females, born between 1895 and 1925, who had lived within 2.5 miles of a copper smelter and arsenic refinery in Ruston, Washington. The site and surrounding area are now part of a Superfund site primarily due to arsenic levels in soil. Exposure was computed as a function of duration and distance of residence from the smelter stack. Four intensity categories were created on the basis of number of years spent at a residence located less than 1.0 mile from the smelter stack. Follow-up status was determined through 1990. The U.S. EPA (2005a,b) documents note study limitations, including truncation of the study period to 1932, potential exposure misclassification, ambiguous exposure data, loss to follow-up, the use of crude mortality rates, and lack of smoking information. In addition to the above limitations, a direct measure of arsenic exposure was not used, which may have fostered a clearer interpretation of study findings than the surrogate measure of residential characteristics. The authors did use proportional hazards regression analysis to account for person-time during the study period. Although this study offers limited information to quantify arsenic exposure and risk of lung or bladder cancer, it provides additional evidence of low cancer risk in U.S. populations exposed to arsenic.

Skin Cancer

Karagas et al. (2001a) assessed arsenic levels in toenail clippings in a case-control study of 587 basal cell and 284 squamous cell skin cancer cases and 524 controls in New Hampshire. Toenail concentration ranges were 0.01–0.81 μg/g for controls, 0.01–2.03 μg/g for basal cell cases, and 0.01–2.57 μg/g for squamous cell cases. There was no evidence of increased risk of either type of skin cancer with arsenic levels, with the exception of persons with toenail arsenic concentrations above the 97th percentile. Among these individuals, the adjusted odds ratios were 2.07 (95% CI: 0.92–4.66) for squamous cell carcinoma and 1.44 (95% CI: 0.74–2.81) for basal cell carcinoma, as compared to those with toenail concentrations at or below the median value.

The EPA Toxicological Review states, "arsenic measured in toenail clippings may result from external exposure as well as internal exposure and typically relates exposure from a two week period occurring approximately a year prior to sampling. Therefore, the exposure measured in the toenail clippings did not occur during the critical period for development of skin cancer (NRC, 2001)" (p. 33). The NRC (2001) report estimates that the latency for skin cancer is more than a decade, but acknowledges that it has not been well-defined. There are examples where the latency period has been reported to be shorter (e.g., treatment of psoriasis with Fowler's solution). The NRC (2001) report goes on to note that more than 50% of the study population had been using the same water supply for over 15 years, suggesting that toenail clippings in this study might indicate "usual" or "longer term" exposure. Furthermore, Karagas et al. (2001a) use a biomarker to assess exposure, the lack of which is considered a limitation of many of the epidemiologic studies by NRC (2001). In addition, studies by Karagas et al. (2001b) and Garland et al. (1993) have reported data to support the notion that arsenic levels as measured in toenails may remain relatively constant for periods of time up to about 6 years.

The low participation rate in this study is troublesome, but without further evaluation comparing eligible non-participants with participants, it is difficult to estimate the impact this may have had on the study results.

Tucker et al. (2001) conducted a cross-sectional study to examine the association between prevalent skin cancer and two arsenic exposure measures: peak arsenic concentration (PAC) and cumulative arsenic dosage (CAD). PAC measures ranged from < 10 ppb (undetectable) to 2,000 ppb. CAD ranged from undetectable to 20,372 ppb-years. There were eight cases of skin cancer. The authors reported a statistically significant deficit of skin cancer among those with PAC levels between 50 and 150 ppb. The authors conclude that their results are consistent with a threshold-model analysis of the Taiwan data set, with the threshold at approximately 120 ppb: "The dose-response curve for skin cancer is best described with respect to the peak arsenic concentration (PAC) by a frequency-weighted model with a threshold at or near 150 ppb arsenic or by a most likely estimate hockey-stick model with a threshold at 122 ppb arsenic" (p. 32). The NRC report agrees that the data appear to be adequately described by the statistical models presented by the authors, but state that "they are also well-described by a nonthreshold linear

model" (p. 60). Tucker et al. (2001), however, clearly believe the non-linear models are more appropriate. The NRC (2001) report views the CAD and PAC exposure metrics as being limitations to this study. Issues regarding the CAD are similar to those raised for the Lewis et al. (1999) study; specifically that exposure intensity and temporal features (e.g., duration) cannot be separated. As discussed previously, this type of exposure measure is commonly used to assess smoking. The PAC metric is criticized for not referring to the time of peak exposures. Given the small number of skin cancers, the authors' recommendation that the study be replicated or expanded is reasonable.

Epidemiologic Studies Not Evaluated in the EPA or NRC Reports

The following recent epidemiologic studies (Karagas et al. 2004; Michaud et al. 2004; Bates et al. 2004) were not included in either the NRC (2001) report or the EPA Issue Paper or Toxicological Review (2005a,b), but are relevant to the issue of exposure to low levels of arsenic in drinking water and cancer.

Karagas et al. (2004) conducted a case-control study in New Hampshire, using incident cases (n = 383) of TCC of the bladder and 641 general population controls. Study participants submitted toenail clippings, and arsenic concentrations in toenails ranged from 0.014 to 2.484 μg/g in cases and 0.009 to 1.077 μg/g in controls. Interviews were conducted to collect sociodemographic, occupational, tobacco, medical, and household water supply information. Analyses were stratified by "ever" and "never" smoking status. Non-significant odds ratios ranging between 0.49 to 1.18 were reported for never smokers, whereas odds ratios ranging between 0.50 and 2.17 were reported for ever smokers. There was no evidence of an exposure-response pattern based on increasing categories of toenail arsenic concentrations.

The relationship between toenail arsenic concentrations and drinking water concentrations is presented in Karagas et al. (2004), and indicates that the majority of study participants had exposure levels below 100 μg/L. Findings in this study are consistent with results that were reported in other U.S. studies. Despite the novel exposure measurement technique used to estimate exposure to arsenic in drinking water in this study (i.e., toenail samples), a strength was the ascertainment of individual biomarker exposure information. Studies have found that arsenic levels as measured in toenails may remain relatively constant for periods of time up to about 6 years (Karagas et al. 2001b; Garland et al. 1993). In general, the results lacked statistical precision due to the analysis of many exposure categories, resulting in some sparse exposure-specific categories. For each analytical subgroup, however, the numbers of cases and controls are presented. This may allow for independent exposure category computations based on recategorized arsenic concentration exposure groups.

Michaud et al. (2004) evaluated the relationship between toenail arsenic levels and bladder cancer risk in a cohort of Finnish male smokers. The authors conducted a nested case-control study and ascertained 280 bladder cancer cases and 293 age, toenail collection date, smoking duration, and trial intervention group matched controls. All study participants were selected from the Alpha-Tocopherol, Beta-Carotene (ATBC) Cancer Prevention Study. Each study participant provided a toenail sample and information on food use. Arsenic levels in toenail samples were determined using neutron activation analysis. The median arsenic level was 0.110 μg/g among the cases and 0.105 μg/g among the controls. Quartiles of toenail arsenic concentration categories ranged from < 0.050 to > 0.161 μg/g, and these values are extrapolated to the approximate equivalent of < 0.01 to > 10 μg/L of drinking water exposure (Table 1). Toenail arsenic levels in this study were comparable to levels reported in previous U.S. studies (Karagas et al. 2001b; Garland et al. 1993).

Table 1. Estimation of drinking water arsenic concentrations based on correlations with toenail arsenic concentrations

Study	Toenail Arsenic Concentrations (μg/g)	Crude Extrapolations: Drinking Water Arsenic Concentrations (μg/L)
Michaud et al. 2004	< 0.050 (referent)	< 0.01
	0.050 to 0.105	0.01 to 2
	0.106 to 0.161	2 to 10
	> 0.161	> 10
Karagas et al. 2004	0.009 to 0.059 (referent)	< 0.01
	0.060 to 0.086	0.01 to 0.1
	0.087 to 0.126	0.1 to 3
	0.127 to 0.193	3 to 11
	0.194 to 0.277	11 to 36
	0.278 to 0.330	36 to 60
	0.331 to 2.484	> 60

Note: Estimations based on Karagas et al. (2001) and Michaud et al. (2004).

There were no associations between toenail arsenic levels and risk of bladder cancer (OR range: 1.09–1.13, *p*-value for trend = 0.65). The authors note that, "No statistically significant effect modification was observed for smoking dose, number of years smoking, place of residence, or beverage intake" (p. 856). Furthermore, Michaud et al. state the following:

> The US Environmental Protection Agency has used risk assessment models to estimate the maximum contamination level in drinking water, a level below which no known adverse health effects occur. For arsenic and bladder cancer, this agency has relied heavily on data from Taiwan. These risk assessment models make assumptions about dose-response curves because low-dose exposure data are not available or are not reliable. When these models are used, the relative risk of bladder cancer for being exposed to arsenic levels of 50 μg/L in drinking water has been estimated to be about 1.2–2.5. (p. 856)

Study strengths included the use of a nested case-control design, a biological marker used for quantifying low-level arsenic exposure, and an exposure metric that lends itself to reproducibility and comparability. Potential limitations include measurement error in the ascertainment of arsenic in levels in toenails and interpretation restrictions to male smokers (i.e., limited generalizability). Michaud et al. (2004) noted the limitations of using the Taiwanese data, which relied on external exposure information to make assumptions about dose-response curves at low-dose exposure, including differences in the environment, diet, and genetic susceptibility. An internal exposure biomarker, however, was used in this study (toenail concentrations) and results based on analysis of this metric are not supportive of an etiologic relationship between low-level arsenic exposure and risk of bladder cancer.

Bates et al. (2004) conducted a population-based case control study in two counties in Argentina. There were 114 incident transitional cell bladder cancer cases matched with 114 controls on age, sex, and county. Home interviews were conducted to collect information on residential history, water sources at each residence, beverage consumption, smoking, occupational history, and medical history. Water samples were collected from each study participant's current residence or from nearby "proxy wells." A fluid intake-adjusted arsenic exposure index was created. In the fluid intake-adjusted analysis, quartiles of exposure ranged from 0–1.0 μg/L to > 80 μg/L. Analyses were stratified by "ever" and "never" smoking status. There were no significant associations observed, based on analyses by average arsenic concentration or fluid-intake-adjusted arsenic exposure. In analyses that evaluated 10-year time windows of average arsenic exposure or consumption of well water, according to smoking status, there were two statistically significant results observed (out of 45 total odds ratios calculated). The significant finding was for 51–60 and 61–70 years of well water consumption among subjects who had ever smoked (OR for 51–60 years = 2.65, 95% CI: 1.2–5.8; OR for

61–70 years = 2.54; 95% CI: 1.0–6.4). This finding, however, should be interpreted cautiously as multiple comparisons may have produced a chance result. The completeness of case ascertainment was uncertain, because pathologists and urologists were instructed to identify cases in predominantly rural areas, where the wells are located. Other potential limitations included misclassification of arsenic exposure in water sampling measurements and water samples taken from the wrong well due to previous residential recall errors among the study participants. Strengths of this study include the extensive water sampling of each participant's current residence and "as many of [their] residential sources within the previous 40 years as practicable," (p. 382) use of 10-year window analyses that allow for relatively long exposure durations, and the evaluations of "ever" and "never" smokers combined (adjusted) and separately (stratified).

Thus, although NRC (2001) and the EPA Toxicological Review cite Ferreccio et al. (2000) as evidence that arsenic cancer risks in a nutritionally sufficient South American population are similar to those in southwestern Taiwan, the study by Bates et al. (2004), with lower exposures and a stronger study design, does not support the claim that the Taiwanese data can be used to accurately predict risks in a nutritionally sufficient South American population.

Additional Relevant Studies

The paper and analyses by **Morales et al. (2000)** are discussed throughout the EPA Issue Paper, yet the study is not reviewed or critiqued as are the epidemiologic studies. Morales et al. (2000) used 10 risk models based on the Taiwan data to estimate the risk of cancer mortality. EPA selected one of these models for deriving the Estimated Dose (ED_{01}) values for bladder and lung cancer. These models and the statistical techniques are explained in the U.S. EPA (2005a,b) documents, but potential limitations of the database that underlie these models are not addressed, nor are the models validated with data from studies outside of southwestern Taiwan (e.g., the United States). Morales et al. (2000) do, however, discuss potential limitations and biases: "Some factors that may affect risk could not be evaluated quantitatively: the ecologic nature of the data, the nutritional status of the study population, and the dietary intake of arsenic" (abstract). The authors continue to call these factors "sources of uncertainty." The authors used an ecological study design and data was ascertained at "village level," and there "appears to be variability in the exposure assessment, causing high variability in the risk estimates" (p. 660). There were no individual exposure measurements. The authors suggest that "poor nutritional status" could be a "contributing factor to the uncertainty" (p. 660). Moreover, they could not "account for dietary intake of inorganic arsenic in food" (p. 660). In addition, as discussed previously, information on smoking was not available in the data set from southwestern Taiwan and the EPA Issue Paper has noted the importance of the effect of smoking on arsenic-related health outcomes. The EPA Toxicological Review raised the above issues in the context of the low exposure epidemiologic studies, but a discussion of these issues is absent from the Morales et al. (2000) study description.

The EPA Toxicological Review quotes from Morales et al. (2000): "The concentrations are reported in U.S. equivalent concentrations of arsenic in drinking water, based on conversions that account for the average weight and average water intake for a male living in the United States compared to a male living in Taiwan" (Morales et al. 2000; p. 659). This statement corresponds to Tables 8, 9, and 10 in Morales et al. (2000) and the tables on page 26 of the EPA Toxicological Review. This may be a faulty extrapolation to U.S. risk, because this broad

generalization does not account for potentially important characteristics other than weight and average water intake that may vary between the two populations.

Lamm and Kruse (2005) conducted an analysis taking risk factors (slopes) from the ecologic study of 42 villages in southwestern Taiwan and applied them to data from an ecologic study of arsenic in groundwater and bladder cancer mortality in 133 U.S. counties (see description of Lamm et al. 2004). This was done to estimate the slope that was most predictive of the association between arsenic ingestion and bladder cancer mortality in the United States. The authors found that the U.S. data were not compatible with the southwestern Taiwanese data based solely on the subpopulation dependent on water from artesian wells. The U.S. data were consistent with the estimated slope derived from the subpopulation that used shallow aquifer water. These data are compared in Figure 1. There is an upward trend in predicted bladder cancer SMRs for concurrent increasing arsenic water concentrations (0 to 60 μg/L), based on southwestern Taiwan artesian well-dependent data. In contrast, there is no slope factor (i.e., no association) for non-artesian well-dependent data among low arsenic levels (0 to 60 μg/L), and there is a slight downward trend in the U.S. data. Furthermore, there appears to be a discontinuity of bladder cancer risk between relatively low levels (< 350 μg/L) of arsenic exposure compared to high levels (> 350 μg/L) (Figure 2), based on southwest Taiwanese data. Overall, there was no evidence of increased bladder cancer mortality risk from non-artesian well-dependent arsenic concentrations in drinking water in either population at relatively low arsenic levels (< 350 μg/L, southwestern Taiwan; < 60 μg/L, United States). The most accurate prediction of bladder cancer risk at low level arsenic exposure may be from the U.S. data and the non-artesian well-dependent data from southwest Taiwan. These data are similar with respect to water source, and therefore, more likely to be representative of low level arsenic exposure in the U.S. general population. EPA and NRC dismiss the notion of variable risks due to well water-source heterogeneity based on quantitative cut-points of arsenic concentration levels. The data presented in Lamm et al. (2003, 2004) and Lamm and Kruse (2005), however, provide evidence that there is not an increased risk of bladder cancer at low levels or arsenic exposure. This evidence is based on U.S. data and is supported by the southwest Taiwanese data for the type of water source that is most prominent in the United States (non-artesian sources).

The authors conclude that it is important to consider factors other than, or in addition, to arsenic as potential risk factors for bladder cancer in southwest Taiwan. They suggest further the possibility that there may be multiple cancer slope factors for arsenic and that the slope factor may be dose range-dependent, and point to studies suggesting that cigarette smoking may be a co-carcinogen with ***high*** arsenic exposures. Finally, the authors state that recent studies as well as reanalysis of the Taiwan data are consistent in their findings that arsenic concentrations in drinking water in the range of about 50 to 200 μg/L are not associated with increased bladder cancer risks.

Summary and Conclusion

To address the question of "which data set should be used in the quantitative risk assessment for long-term exposure to arsenic in drinking water" (U.S. EPA 2005a; Issue 3, p. 5), it is helpful to consider what would characterize the ideal study. A randomized controlled trial is generally considered to be the "gold standard" of study designs; however, such a study would not be possible in this case. Epidemiologic cohort studies or well-designed case-control studies are often considered to be the most informative of the observational study designs, and the least subject to bias. As discussed previously, ecologic studies have numerous limitations that are inherent to the study design. Even a well-designed ecologic study cannot be used to draw causal conclusions regarding cause and effect. Thus, it is surprising that EPA would choose to rely ***exclusively*** on data from an ecologic study for the purpose of quantitative risk assessment. Furthermore, it is not clear why ecologic data from a population in Taiwan, requiring numerous assumptions and adjustments in order to apply the data to the general U.S. population, are favored over ecologic data from a U.S. study that would be more easily generalizable and does not appear to suffer from bias to any greater degree than the Taiwanese studies. Given the potential for other substances found in artesian wells in Taiwan to influence the effect of arsenic (e.g., humic acid), studies conducted in the United States and other areas may be more likely to isolate an effect attributable to arsenic only. Finally, any observational study is subject to potential biases from many sources. The NRC and EPA reports have not provided persuasive arguments as to why the data from the case-control and cohort studies described above are more likely to suffer from bias than the Taiwanese data that have been used to calculate lifetime risk estimates (and excess risk) and slope factors. A thorough and careful review of the relevant studies under consideration requires that the same set of criteria be applied to each study and that the results of this evaluation be presented for each study systematically. The reviews presented in the NRC and EPA reports appear subjective in this regard, despite the listing of review criteria in the NRC report (2001). What is lacking is evidence of a review process with consistent criteria applied to all studies, and a direct comparison of the strengths and limitations of the recent epidemiologic studies with those of the southwestern Taiwan data set.

The EPA Toxicological Review provides a general summary of the limitations of the epidemiology studies. These limitations include "not specifying the form of arsenic, difficulties with the exposure assessment, and/or not examining possible confounding variables" (p. 17). They do acknowledge, however, that the pentavalent form of inorganic arsenic is the predominant form in water and soil. Furthermore, the EPA Toxicological Review states that many of the epidemiologic studies do attempt to examine the major confounders, including age and smoking. While few studies measure and assess the potential role of nutritional factors, such as selenium or zinc, the NRC report (2001) suggests that the findings from the Taiwanese studies are unlikely to be a result of confounding by diet. They do, however, recommend further research on the potential role of dietary and nutritional factors on the risk of arsenic-related adverse health outcomes. Given these comments by the NRC (2001), the omission of nutritional variables from a given epidemiologic analysis does not appear to be a "fatal flaw."

Careful and valid exposure assessment is critical to any environmental epidemiologic study. It is also a major challenge, particularly in studies of diseases with long latency periods, such as cancer. Evaluation of exposure assessment is critical to the review and interpretation of the epidemiologic studies of consumption of arsenic in drinking water and cancer and other health outcomes. What is notable is that studies of low-exposure populations have consistently observed a lack of evidence in support of an association for risk of lung or bladder cancer at low exposure levels, despite different methods for assessing and characterizing/modeling arsenic exposure.

Thus, the summary of limitations in the epidemiologic studies in the EPA Toxicological Review does not raise issues that would lead to the exclusion of a given study, nor does it raise issues that would apply to the studies conducted in the United States and other areas of lower exposure and not to the Taiwanese studies. Thus, a case for relying solely on the southwestern Taiwanese studies has not been supported.

The NRC (2001) report describes sources of variability and uncertainties in Chapter 4; many of these issues were raised on their 1999 report as well. Some of these sources of variability and uncertainty, many of which require assumptions and/or are associated with the potential for introducing bias, include the following:

1. The need for assumptions about the amount and source of water consumed due to lack of biomarkers of individual exposure, and resulting uncertainty in exposure estimates (and potential bias)
2. Lack of knowledge about intake of arsenic from food (both in U.S. and Taiwanese populations); intake could vary among populations, which would affect the appropriateness of extrapolation of risk from one population to another
3. Differences in per capita water consumption across populations
4. Variability in individual or population susceptibility due to age, arsenic metabolism (e.g., influences of genetic polymorphisms on metabolism; presence of other chemicals)
5. Dose metrics (e.g., cumulative dose, mean daily intake over lifetime, peak arsenic exposure)
6. Mode of action
7. Role of nutritional status/malnutrition.

Some of the above issues (e.g., dietary intake of arsenic in food; adjustment for water intake from drinking water and from cooking water) are included in the EPA Issue Paper, but the emphasis is on improving the extrapolation of the southwestern Taiwanese data to the U.S. population. The recommendation regarding "Issue 3: Choice of Study" is to continue to use the southwestern Taiwan data despite the recent data from the United States and other lower-exposed populations. Although the limitations of the recent studies are discussed (NRC 2001; U.S. EPA 2005a,b), a clear case outlining the features of the southwestern Taiwanese studies that make this data set less likely than the recent studies to be subject to bias, and more likely to be informative about health risks in the United States, has not been made.

As the SAB considers both the "Choice of Study/Studies" and the relative strengths and limitations of the entire epidemiologic literature relevant to the question of cancer risks associated with exposure to arsenic in drinking water, it is important to consider that for a dose-

response model to accurately predict health risks in humans, it must be able to accommodate observed data from valid epidemiologic studies at both ends of the exposure spectrum. For this reason, all of the relevant studies should be considered and weighed against a uniform set of objective criteria. Any observational study will be potentially biased to some degree and it is not difficult to identify sources of ***potential*** bias. To dismiss studies on these grounds, however, may result in omitting data that are potentially informative. It is preferable to take a more formal approach to the assessment of bias, and to conduct sensitivity analyses to ***identify and quantify*** the likely direction and magnitude of bias and the potential impact on relative risk estimates.

The recent epidemiologic studies offer data that will be informative to the SAB as its members consider how best to estimate potential human health effects associated with ingestion of inorganic arsenic through the consumption of drinking water. In response to C2 in the Charge to the SAB (U.S. EPA 2005c), we argue that there are limitations to the Taiwanese data set that are no less an issue than limitations to the majority of the U.S. studies (and other studies of low exposure populations). Furthermore, we argue that the Taiwanese data set is not the most appropriate choice for estimating cancer risk in humans because results from this population cannot be generalized to populations with dissimilar characteristics. It would be more appropriate to consider data from the populations to which the risk assessment will be applied (i.e., the United States). All of the relevant data should be used in the validation of the final dose-response model and risk assessment. Finally, as pointed out in the recent publication by Lamm and Kruse (2005), recent studies, together with their reanalysis of the Taiwanese data, indicate no consistent evidence of increased bladder cancer risks associated with arsenic concentrations in drinking water up to ranges of approximately 50–200 μg/L.

In conclusion, we recommend a careful and systematic review and summary of the relative strengths and limitations of the relevant studies. This would include evaluation of each study on its own merits with respect to whether it should be used to estimate cancer risk in humans, as well as a direct comparison of limitations of each study to the limitations in the southwestern Taiwanese data set. Our opinion is that the biases affecting the internal validity of the southwestern Taiwanese studies as well as issues related to the generalizability of data from this

unique population to the general U.S. population severely limits the utility of these data for the purposes of quantitative risk assessment and estimating cancer risks in humans. To address issues of precision of the relative risk estimates, options such as meta-analysis or pooled analysis could be considered. These approaches will improve the precision of the relative risk estimates and can be used to clarify associations across the body of scientific studies. There is evidence that models based on these data do not accurately predict the observed data in the U.S. and other populations with low exposure to arsenic in drinking water; the observed risks are lower than would be predicted by the models. Accompanying this report is a meta-analysis of low-level exposure to arsenic in drinking water and risk of bladder cancer (Exponent, 2005). These results combined with results of the individual studies indicate that more accurate models to predict risk in low exposure populations are needed. We suggest that using data from other populations would be more appropriate than continuing to rely solely on the data from southwestern Taiwan.

References

Bates, M.N., A.H. Smith, and K.P. Cantor. 1995. Case-control study of bladder cancer and arsenic in drinking water. Am. J. Epidemiol. 141(6):523–530.

Bates, M.N., O.A. Rey, M.L. Biggs, C. Hopenhayn, L.E. Moore, D. Kalman, C. Steinmaus, and A.H. Smith. 2004. Case-control study of bladder cancer and exposure to arsenic in Argentina. Am. J. Epidemiol. 159(4):381–389.

Chen, C.J., Y.C. Chuang, T.M. Lin, and H.Y. Wu. 1985. Malignant neoplasms among residents of a blackfoot disease-endemic area in Taiwan: High-arsenic artesian well water and cancers. Cancer Res. 45(11 Pt 2):5895–5899.

Chen, C.J., M.M. Wu, S.S. Lee, J.D. Wang, S.H. Cheng, and H.Y. Wu. 1988. Atherogenicity and carcinogenicity of high-aresenic artesian well water. Multiple risk factors and relate malignant neoplasms of blackfoot disease. Arterioscler. Thromb. Vasc. Biol. 8:452–460.

Chen, C.J., C.W. Chen, M.M. Wu, and T.L. Kuo. 1992. Cancer potential in liver, lung, bladder and kidney due to ingested inorganic arsenic in drinking water. Br. J. Cancer 66(5):888–892.

Chen, C.L., L.I. Hsu, H.Y. Chiou, Y.M. Hsueh, S.Y. Chen, M.M. Wu, and C.J. Chen. 2004. Blackfoot Disease Study Group. Ingested arsenic, cigarette smoking, and lung cancer risk: A follow-up study in arseniasis-endemic areas in Taiwan. JAMA. 292(24):2984–2990.

Chiou, H.Y., S.T. Chiou, Y.H. Hsu, Y.L. Chou, C.H. Tseng, M.L. Wei, and C.J. Chen. 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(5):411–418.

Exponent. 2005. Epidemiologic studies of low-level arsenic exposure in drinking water and bladder cancer: a review and meta-analysis (unpublished technical report). Exponent, Inc., Washington, DC.

Ferreccio, C., C. Gonzalez, V. Milosavjlevic, G. Marshall, A.M. Sancha, and A.H. Smith. 2000. Lung cancer and arsenic concentrations in drinking water in Chile. Epidemiology 11(6):673–679.

Garland, M., J.S. Morris, B.A. Rosner, M.J. Stampfer, V.L. Spate, C.J. Baskett, et al. 1993. Toenail trace element levels as biomarkers: Reproducibility over a 6-year period. Cancer Epidemiol. Biomarkers Prev. 2(5):493–497.

Karagas, M.R., T.A. Stukel, J.S. Morris, T.D. Tosteson, J.E. Weiss, S.K. Spencer, et al. 2001a. Skin cancer risk in relation to toenail arsenic concentrations in a US population-based case-control study. Am. J. Epidemiol. 153(6):559–565.

Karagas, M.R., C.X. Le, S. Morris, J. Blum, X. Lu, V. Spate, M. Carey, V. Stannard, B. Klaue, and T.D. Tosteson. 2001b. Markers of low level arsenic exposure for evaluating human cancer risks in a US population. Int. J. Occup. Med. Environ. Health 14(2):171–175.

Karagas, M.R., T.D. Tosteson, J.S. Morris, E. Demidenko, L.A. Mott, J. Heaney, and A. Schned. 2004. Incidence of transitional cell carcinoma of the bladder and arsenic exposure in New Hampshire. Cancer Causes Control 15(5):465–472.

Kogevinas, M., and D. Trichopoulos. 2002. Urinary bladder cancer. Chapter 20. In: Textbook of Cancer Epidemiology. H.O. Adami, D. Hunter, and D. Trichopoulos (eds). Oxford University Press, New York.

Kurttio, P., E. Pukkala, H. Kahelin, A. Auvinen, and J. Pekkanen. 1999. Arsenic concentrations in well water and risk of bladder and kidney cancer in Finland. Environ. Health Perspect. 107(9):705–710.

Lamm, S.H., and M.B. Kruse. 2004. Ingested arsenic and cancer, what's new? Available at: www.ruralwater.org/arsenic/lamm1.ppt. January 30, 2004. U.S. Senate Committee on the Environment and Public Works.

Lamm, S.H., and M.B. Kruse. 2005. Arsenic ingestion and bladder cancer mortality – what do the dose response relationships suggest about mechanism? Hum. Ecol. Risk Assess. 11:433–450.

Lamm, S.H., D.M. Byrd, M.B. Kruse, M. Feinleib, and S.H. Lai. 2003. Bladder cancer and arsenic exposure: Differences in the two populations enrolled in a study in southwest Taiwan. Biomed. Environ. Sci. 16(4):355–368.

Lamm, S.H., A. Engel, M.B. Kruse, M. Feinleib, D.M. Byrd, S. Lai, and R. Wilson. 2004. Arsenic in drinking water and bladder cancer mortality in the United States: An analysis based on 133 U.S. counties and 30 years of observation. J. Occup. Environ. Med. 46(3):298–306.

Lewis, D.R., J.W. Southwick, R. Ouellet-Hellstrom, J. Rench, and R.L. Calderon. 1999. Drinking water arsenic in Utah: A cohort mortality study. Environ. Health Perspect. 107(5):359–365.

Michaud, D.S., M.E. Wright, K.P. Cantor, P.R. Taylor, J. Virtamo, and D. Albanes. 2004. Arsenic concentrations in prediagnostic toenails and the risk of bladder cancer in a cohort study of male smokers. Am. J. Epidemiol. 160(9):853–859.

Morales, K.H., L. Ryan, T.L. Kuo, M.M. Wu, and C.J. Chen. 2000. Risk of internal cancers from arsenic in drinking water. Environ. Health Perspect. 108(7):655–661.

NRC. 1999. Arsenic in drinking water. National Research Council. National Academy Press, Washington, DC.

NRC. 2001. Arsenic in drinking water: 2001 update. National Research Council. National Academy Press, Washington DC.

Schoen, A., B. Beck, R. Sharma, and E. Dube. 2004. Arsenic toxicity at low doses: Epidemiological and mode of action considerations. Toxicol. Appl. Pharmacol. 198(3):253–367.

Schlesselman, J. 1982. Case-control studies. Design, conduct, analysis. Oxford University Press, New York, NY.

Steinmaus, C., Y. Yuan, M.N. Bates, and A.H. Smith. 2003. Case-control study of bladder cancer and drinking water arsenic in the western United States. Am. J. Epidemiol. 158(12):1193–1201.

Tollestrup, K., F.J. Frost, L.C. Harter, and G.P. McMillan. 2003. Mortality among children residing near the American Smelting and Refining Company (ASARCO) copper smelter in Ruston, Washington. Arch. Environ. Health 58(11):683–691.

Tucker, S.B., S.H. Lamm, F.X. Li, et al. 2001. Relationship between consumption of arsenic-contaminated well water and skin disorders in Huhhot, Inner Mongolia. Inner Mongolia Cooperative Arsenic Project (IMCAP) Study. Final Report. University of Texas, Department of Dermatology, Houston, TX, and Chinese Academy of Preventive Medicine, Institute of Environmental Engineering, Department of Environmental Epidemiology, Beijing, China.

U.S. EPA. 2001a. National primary drinking water regulations; arsenic and clarifications to compliance and new source contaminants monitoring. Final Rule. January 22, 2001 Federal Register (66 FR 6976). U.S. Environmental Protection Agency.

U.S. EPA. 2005a. Issue Paper: Inorganic arsenic cancer slope factor. July 23, 2005. Available at: www.epa.gov/waterscience/sab/ASIssues_SAB.pdf. U.S. Environmental Protection Agency.

U.S. EPA. 2005b. Toxicological review of ingested inorganic arsenic. Available at: http://www.epa.gov/waterscience/sab/AsDraft_SAB.pdf. U.S. Environmental Protection Agency.

U.S. EPA. 2005c. Charge to the EPA Science Advisory Board Arsenic Review Panel. July 26, 2005. Available at: www.epa.gov/sab/pdf/arsenic_review_panel_final_charge_7-25-05.pdf. U.S. Environmental Protection Agency.

Wu, M.M., T.L. Kuo, Y.H. Hwang, and C.J. Chen. 1989. Dose-response relation between arsenic concentration in well water and mortality from cancers and vascular diseases. Am. J. Epidemiol. 130:1123–1132.