

Risk of Internal Cancers from Arsenic in Drinking Water

Knashawn H. Morales,¹ Louise Ryan,^{1,2} Tsung-Li Kuo,³ Meei-Maan Wu,⁴ and Chien-Jen Chen⁵

¹Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, USA; ²Dana-Farber Cancer Institute, Boston, Massachusetts, USA; ³Department of Forensic Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan; ⁴Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan; ⁵Graduate Institute of Epidemiology, National Taiwan University, Taipei, Taiwan

The U.S. Environmental Protection Agency is under a congressional mandate to revise its current standard for arsenic in drinking water. We present a risk assessment for cancers of the bladder, liver, and lung from exposure to arsenic in water, based on data from 42 villages in an arseniasis-endemic region of Taiwan. We calculate excess lifetime risk estimates for several variations of the generalized linear model and for the multistage-Weibull model. Risk estimates are sensitive to the model choice, to whether or not a comparison population is used to define the unexposed disease mortality rates, and to whether the comparison population is all of Taiwan or just the southwestern region. 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. Despite all of these sources of uncertainty, however, our analysis suggests that the current standard of 50 µg/L is associated with a substantial increased risk of cancer and is not sufficiently protective of public health. **Key words:** bladder cancer, generalized linear model, lifetime death risk, lung cancer, margin of exposure, multistage-Weibull. *Environ Health Perspect* 108:655–661(2000). [Online 5 June 2000]

<http://ehpnet1.niehs.nih.gov/docs/2000/108p655-661morales/abstract.html>

A metal found in rocks and mineral formations in the earth's crust, arsenic has long been associated with the development of cancer in humans. Exposure can occur via inhalation, primarily in industrial settings, or through ingestion. Because drinking water is one of the primary routes of exposure, standards set in 1942 established a maximum contaminant level (MCL) of 50 µg/L in drinking water. In 1975, 50 µg/L was adopted as the interim standard in response to the 1974 Safe Drinking Water Act (1). In a 1984 health assessment, the U.S. Environmental Protection Agency (EPA) classified arsenic as a class A human carcinogen, based primarily on epidemiologic evidence, and produced quantitative risk estimates for both ingestion and inhalation routes of exposure (2). Although the EPA assessment for the inhalation route is well accepted, the risk assessment for ingestion remains controversial. The 1984 risk assessment for arsenic in drinking water was based on an epidemiologic study in Taiwan that examined an association between arsenic exposure via drinking water and skin cancer (non-melanoma) (3). EPA investigators estimated that the lifetime risk of skin cancer for individuals who consumed 2 L water per day at 50 µg/L could be as high as 2 in 1,000. This high value prompted questions about the 1984 risk assessment, including applicability of the risk assessment to the U.S. population, the role of arsenic as an essential nutrient, the relevance of skin lesions as the basis for the risk assessment, and the role of arsenic intake via food. In 1988, the EPA Risk Assessment Forum published a revised

skin cancer risk assessment and focused attention on these questions (4). The EPA is currently under a congressional mandate to finalize a new rule for arsenic in drinking water by 1 January 2001 (5).

There has been substantial focus on the association between arsenic and skin cancer, and there is also substantial evidence that exposure to arsenic in drinking water increases the mortality risk for several internal cancers. Increases in bladder and lung cancer mortality were found in a region of northern Chile (6). An association was also found between bladder cancer mortality and arsenic in drinking water in Argentina (7). Significant increased mortality was observed for males and females in Taiwan due to lung, liver, skin, kidney, and bladder cancer (8). The National Research Council presents a more detailed summary of the evidence linking arsenic exposure to internal cancer (1).

The purpose of this article is to present a risk assessment for mortality due to several internal cancers based on a reanalysis of the data reported by Chen et al. (8). Brown (9) discussed the limitations of the data available for analysis when the current EPA risk assessment (4) was prepared. For several reasons, it can be argued that the risk assessment of internal cancers presented in this paper yields more convincing results than the previous EPA assessment based on skin cancer. First, the current study focuses on mortality from bladder, lung, and liver cancers identified through national death records. In addition, unlike the Tseng et al. (3) study that was used in the EPA analysis, which grouped data into three broad

exposure intervals [low (< 300 µg/L), medium (300–600 µg/L), and high (> 600 µg/L)], data now available provide exposure at the individual village level.

This paper is a follow-up to a preliminary study that focused only on bladder cancer and examined model sensitivity (10). The current analysis is expanded to include lung and liver cancers and examines issues of dose–response modeling by Poisson regression, in addition to application of the multistage-Weibull (MSW) model, in more detail.

Materials and Methods

Internal cancer data. Data used in this analysis were derived from a study in an arseniasis-endemic area of Taiwan (11–13). Cancer mortality data were collected from death certificates of residents of 42 villages during 1973–1986. These data were originally collected in 1987, so only records up to 1986 were available. Causes of death were classified according to the *Eighth Revision of International Classification of Diseases, 1965 Revision* (ICD) (14). The data consisted of person-years at risk and the number of deaths due to bladder (ICD code 188), lung (ICD code 162), and liver (ICD code 155) cancer in 5-year age increments for both males and females. Table 1 summarizes the internal cancer data and provides person-years at risk and observed number of cancer deaths by age, sex, and arsenic level. Although individual village arsenic levels are available and will be used in subsequent analyses, exposure levels are grouped in Table 1 for convenience of presentation. The numbers of bladder, liver, and lung cancers are given, along with the number of person-years at risk. For example, males between the ages of 50 and 69 contributed 21,040 person-years at risk and 6, 17, and 12 deaths were observed from bladder, liver, and lung cancer, respectively.

Address correspondence to L. Ryan, Department of Biostatistics, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115 USA. Telephone: (617) 632-3602. Fax: (617) 632-2444. E-mail: ryan@jimmy.harvard.edu

Support was received from the National Institutes of Health (grants 5F31GM18906, ES0002, and CA48061), the David and Lucile Packard Foundation, and the Department of Health, Executive Yuan, ROC (DOH88-HR-503).

Received 29 December 1999; accepted 14 March 2000.

Exposure data. Drinking water samples were collected from wells of 42 villages in 1964–1966 (12). The artesian wells were gradually closed; the last one closed in mid-1970. Although mortality data were collected for a later time period (1973–1986), it is likely that arsenic levels in well water remained relatively unchanged over this time period. It could also be argued that because of the long latency of the cancers of interest, it is appropriate for exposure to be based on a time period 10 to 20 years before death. A strength of the currently available exposure data is that individual well concentration levels are available for each village. Physical and chemical characteristics of drinking water such as pH value and levels of arsenic, sodium, calcium, magnesium, manganese, iron, mercury, chromium, lead, nitrite and nitrate nitrogen, fluoride, and bicarbonate have been intensively studied in both Blackfoot disease-endemic and -nonendemic areas (15,16). Arsenic level was the only level that was significantly higher than the maximal allowable limit and strikingly different in water from shallow wells and artesian wells. The data also have some limitations. The drinking water was not tested for levels of dissolved radon and other α -emitters. Fluorescent compounds, especially humic acids, have been found in the well water. These fluorescent substances result from the decomposition of organic matter, particularly dead plants. However, it is unlikely that their presence causes confounding in this analysis because widespread contamination is not confined to the arsenic-endemic area.

Standardized mortality ratio. We used standardized mortality ratios (SMRs) to summarize the observed patterns of mortality in data. SMRs provide a popular approach to comparing mortality in a specific population with mortality from a suitable comparison population (17). SMRs correspond to ratios of observed and expected number of events and are calculated by $\Sigma O_i / \Sigma E_i$, where O_i is the observed number of deaths in the i^{th} age group and E_i is the corresponding expected number of deaths, calculated by multiplying the study population size (P_i) by the age-specific cancer death rate (M_i) in a comparison population (i.e., $E_i = P_i \times M_i$). Usually SMRs are expressed as a percentage so that the value $100 \times \Sigma O_i / \Sigma E_i$ is the number reported. There are concerns with using all of Taiwan as a comparison population because of the potential for bias associated with differences in the populations (e.g., rural vs. urban). For this reason, we considered two comparison populations in this analysis: all of Taiwan and the southwestern region of Taiwan (18). The latter is expected to provide a more suitable comparison basis for the study population, which is largely rural and fairly poor.

Table 2 contains the data from the two comparison populations. The number of deaths due to bladder, lung, and liver cancers and person years at risk (PYR) were extracted by age group and sex for 1973–1986.

Generalized linear model. Poisson modeling is often used in epidemiologic analysis, particularly for rare events such as cancer deaths. In fact, SMRs can correspond to maximum likelihood estimates of risk ratios from a Poisson model (17). In our analysis, we assumed that the number of deaths due to cancer follows a Poisson distribution with parameter equal to the person-years at risk multiplied by the hazard of dying of cancer. The hazard is often modeled as a function of age (t) and exposure (x). As described by Breslow and Day (17), a broad class of models can be characterized using the following general form,

$$h(x,t) = h_0(t) \times g(x), \quad [1]$$

where $h_0(t)$ denotes the baseline hazard function that only depends on age, t , and describes the instantaneous hazard of dying of cancer for the unexposed population. The risk ratio attributed to exposure level x is denoted by $g(x)$. Of course, it is likely that a variety of factors, including cigarette smoking, use of bottled water, and dietary intake of inorganic arsenic, could influence or even confound the model. The model described in Equation 1 will allow consideration of other covariates. Unfortunately, measurements for these and other potentially important factors were not available for our study. Rather, this is an ecologic study wherein only relatively simple exposure and population characteristics could

be measured. It will be important to consider this and other sources of uncertainty when interpreting the results. Although not discussed extensively here, it is possible for the risk ratio $g(x)$ to also depend on age, t . For example, older people may be more susceptible to exposure. We did in fact explore such age-dependent risk models and found that in general, it was adequate to model the relative risk as a function of exposure only. A wide range of models was obtained by varying a) the use of comparison populations; b) the way age is modeled in $h_0(t)$, e.g., linear, quadratic, or the use of regression splines; c) transformations of exposure concentrations; and d) the way exposure is modeled. Table 3 summarizes the various modeling options considered in this analysis. Each model corresponds to choosing one option from each column. For example, the model excluding the comparison population, with a linear age effect, an exponential linear dose effect, and no transformation on dose, is characterized by $h_0(t) = \exp(\alpha_0 + \alpha_1 t)$ and $g(x) = \exp(\beta_1 x)$. Note that the linear and quadratic dose-effect models (generally referred to as additive models) do not fit into the usual class of generalized linear models (GLMs) and require special programming. Exponential linear and exponential quadratic models fall under the general class of multiplicative models. The spline age effect was modeled using natural splines because of the ease of obtaining predicted values (19). There are three options for the baseline hazard: model the hazard without including a comparison population, treat the comparison population as an unexposed group, or replace the baseline hazard function with empirical

Table 1. Person-years at risk by age, sex, and arsenic level with observed number of deaths from cancer (bladder, liver, and lung).

Sex, arsenic level ($\mu\text{g/L}$)	Age (years) ^a				Total
	20–30	30–49	50–69	≥ 70	
Male					
< 100	35,818 (0, 0, 0)	34,196 (1, 10, 2)	21,040 (6, 17, 12)	4,401 (10, 4, 14)	95,455 (17, 31, 28)
100–299	18,578 (0, 0, 0)	16,301 (0, 4, 3)	10,223 (7, 15, 14)	2,166 (2, 4, 13)	47,268 (9, 23, 30)
300–599	27,556 (0, 3, 0)	25,544 (5, 7, 9)	15,747 (15, 23, 30)	3,221 (12, 6, 14)	72,068 (32, 39, 53)
≥ 600	16,609 (0, 0, 1)	15,773 (4, 12, 3)	8,573 (15, 15, 23)	1,224 (8, 2, 6)	42,179 (27, 29, 33)
Total	98,561 (0, 3, 1)	91,814 (10, 33, 17)	55,583 (43, 70, 79)	11,012 (32, 16, 47)	256,970 (85, 122, 144)
Female					
< 100	27,901 (0, 0, 0)	32,471 (3, 1, 5)	21,556 (9, 6, 18)	5,047 (9, 5, 5)	86,975 (21, 12, 29)
100–299	13,381 (0, 0, 0)	15,514 (0, 3, 4)	11,357 (9, 6, 10)	2,960 (2, 5, 5)	43,212 (11, 14, 19)
300–599	19,831 (0, 0, 0)	24,343 (0, 5, 6)	16,881 (19, 6, 20)	3,848 (11, 2, 10)	64,903 (30, 13, 36)
≥ 600	12,988 (0, 0, 0)	15,540 (0, 4, 6)	9,084 (21, 7, 28)	1,257 (7, 1, 4)	38,869 (28, 12, 38)
Total	74,101 (0, 0, 1)	87,868 (3, 13, 21)	58,878 (58, 25, 76)	13,112 (29, 13, 24)	233,959 (90, 51, 122)

^aValues in parentheses are number of deaths from bladder, liver, and lung cancer, respectively.

estimates based on the comparison population (not included in Table 1). The third option can be accomplished by fitting a Poisson model containing indicators corresponding to the age categories observed in the comparison population. This approach essentially corresponds to the traditional SMR approach. Because there were no villages with zero concentration levels, the method used to model the baseline hazard had a fairly strong influence on the results. In particular, the choice of whether to include a comparison population had a strong influence. The use of an unexposed comparison population has the potential to provide more information about the shape of the model at low concentrations.

Although not a member of the usual GLM class, the MSW model was also

considered because it was used in the previous risk assessment (4). The MSW corresponds to letting $g(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ and $h_0(t) = C(t - T_0)_+^{(k)}$ (10), where t denotes age and x denotes exposure concentration. The plus sign (+) indicates a truncation on the $(t - T_0)$ term (i.e., if $T_0 > t$ then the term is set to zero). Results based on the MSW model are only presented for comparison. The GLM approach has several advantages over the MSW model. First, the MSW model appears to be more sensitive to outliers than the GLM model (10). Also, the hazard function for the MSW model involves a truncation in t that complicates estimation. Finally, the inclusion of the power parameter k (for our purposes, $k = 2$) tends to give the fitted model a relatively sublinear shape that leads,

in general, to higher benchmark doses than the GLM models.

Quantitative risk assessment. Because the risk of dying from cancer is age dependent, it is common to base risk assessment on the excess risk of dying from cancer over the course of a typical lifetime. The adjusted lifetime death risk can be calculated by integrating the death hazard over the typical lifetime in the population of interest,

$$ldr(x) = \int_0^\infty h(x,t)S(t)dt,$$

where $S(t)$ is the probability of surviving until age t and $h(x,t)$ is the hazard for dying of the cancer of interest at age t for someone exposed at level x . Applying integration by parts, $ldr(x)$ can also be written as

$$ldr(x) = 1 - \int_0^\infty \lambda(t)S(t)dt,$$

where $\lambda(t)$ denotes the hazard of dying at age t from causes other than the cancer of interest. This function can be approximated by

$$ldr(x) \approx 1 - \sum_t q_t \exp\left[-5 \sum_{s=t} h(x,s)\right],$$

where \sum_t denotes the sum over all 5-year age groups in the study and q_t is the probability of dying during the 5-year time interval indicated by t . Values for q_t were taken from the 1996 U.S. population life table for males and females (Table 4) (20).

Traditionally, standards for carcinogenic compounds have been set by finding the exposure level that yields a rate of 10^{-6} over background. As suggested by Brown (9) and discussed by Smith and Sharp (21), this estimate is probably unreliable for epidemiologic data, where exposure is not typically measured accurately enough to extrapolate to such low risk levels. The new EPA guidelines for cancer risk assessments (22) suggest the use of a point-of-departure analysis for settings where the mode of action is supportive of linearity or there is insufficient support for a nonlinear mode of action. The idea is to

Table 2. Comparison population data, 1973–1986.

Sex, age (years)	All-Taiwan				Southwestern region			
	PYR	Deaths (n)			PYR	Deaths (n)		
Bladder		Lung	Liver	Bladder		Lung	Liver	
Male								
20–25	13,271,386	3	45	206	2,956,638	2	14	43
25–30	11,054,191	4	86	426	2,175,046	3	26	88
30–35	8,628,516	8	144	782	1,580,019	2	33	140
35–40	6,793,545	20	217	1,351	1,320,637	6	38	245
40–45	6,375,466	50	447	2,030	1,327,866	18	89	403
45–50	6,384,052	91	951	3,145	1,334,769	34	181	565
50–55	6,062,515	164	1,852	4,140	1,214,443	52	323	716
55–60	5,018,542	213	2,882	4,562	977,820	61	478	832
60–65	3,666,535	345	3,557	4,030	739,460	103	595	722
65–70	2,443,367	413	3,569	3,259	520,965	126	607	704
70–75	1,480,126	418	2,658	2,107	320,158	130	465	463
75–80	720,375	305	1,318	1,170	158,750	88	230	246
80–85	287,294	146	512	436	63,236	32	80	103
85+	105,411	66	152	188	22,651	15	22	33
Female								
20–25	12,612,276	0	39	81	2,595,529	0	7	15
25–30	10,548,089	2	70	134	1,846,189	2	19	34
30–35	8,210,507	2	102	168	1,402,764	0	17	39
35–40	6,458,620	5	205	247	1,215,899	2	41	53
40–45	5,802,856	20	365	396	1,191,615	8	75	75
45–50	5,157,821	41	525	590	1,111,810	14	112	138
50–55	4,335,755	76	730	763	957,985	36	160	169
55–60	3,517,193	124	1,018	1,018	774,836	52	200	255
60–65	2,776,622	153	1,224	1,039	634,758	77	258	243
65–70	2,106,715	173	1,280	1,039	492,203	68	230	235
70–75	1,490,659	185	1,062	875	342,767	70	190	199
75–80	888,468	157	707	602	199,630	43	108	127
80–85	433,245	81	330	300	96,293	21	45	59
85+	217,590	41	136	153	46,089	9	10	31

PYR, person-years. Data from the Department of Health (18).

Table 3. Poisson modeling options.

Comparison population	Age effect $h_0(t)$	Dose transformation	Dose effect $g(x)$
None	Linear $\exp(\alpha_0 + \alpha_1 t)$	Linear $x = \text{ppb}^a$	Linear $\beta_1 x$
Southwestern Taiwan	Quadratic $\exp(\alpha_0 + \alpha_1 t + \alpha_2 t^2)$	Logarithmic $x = \log(1 + \text{ppb})$	Quadratic $\beta_1 x + \beta_2 x^2$
All of Taiwan	Regression spline $\exp[\alpha_0 + \alpha_1 ns(t)]^b$	Square root $x = \sqrt{\text{ppb}}$	Exponential linear $\exp(\beta_1 x)$ Exponential quadratic $\exp(\beta_1 x + \beta_2 x^2)$

^aRepresents exposure concentration in parts per billion, which is equivalent to micrograms per liter. ^b $ns(t)$ represents a natural spline applied to t .

Table 4. U.S. death probabilities by age and sex, 1996.

Age (years)	Probability of death (q_t)	
	Male	Female
20–25	0.00742	0.00239
25–30	0.00755	0.00307
30–35	0.00962	0.00423
35–40	0.01227	0.00595
40–45	0.01621	0.00834
45–50	0.02182	0.01224
50–55	0.03144	0.01938
55–60	0.04622	0.02938
60–65	0.06966	0.04577
65–70	0.09278	0.06417
70–75	0.12183	0.09207
75–80	0.14149	0.12267
80–85	0.15457	0.16036
85+	0.24949	0.41813

Data from *Vital Statistics of the United States, 1996* (20).

estimate a point on the exposure response curve within the observed range of the data and then extrapolate linearly to lower doses. The dose associated with 10% excess risk (ED_{10}) is the standard point of departure, but often in epidemiologic studies, an excess risk of 10% is fairly large and occurs only at relatively high doses. We will use both 1% and 5% excess risks for the point of departure (ED_{01} and ED_{05} , respectively). We computed confidence intervals for excess lifetime risk using the Delta method (23). Bootstrap methods were also used for models with non-parametric age effects, yielding similar results (24). The new guidelines also suggest a margin-of-exposure analysis (MOE), defined as the point of departure divided by the environmental exposure of interest. This approach is the proposed default mode of action when linearity is not the most reasonable assumption (22). For subsequent discussion we will use MOE_{01} (50) to represent the margin of exposure using the ED_{01} point of departure and 50 $\mu\text{g/L}$ as the environmental exposure of interest.

Results

Table 5 contains a descriptive summary of the internal cancer data, showing person-years at

risk, observed number of cancers, and the SMRs for age, sex, and exposure grouped into the same intervals used by the EPA in the skin cancer risk assessment. As in Wu et al. (12), the analysis is limited to persons ≥ 20 years of age because there were essentially no cancer deaths observed in those younger than 20 years of age. Note that the entire Taiwanese population was used to calculate the expected deaths used in the computation of SMRs in Table 5. Although the computed SMRs display a large amount of noise, there appear to be higher SMRs at high exposure levels compared to exposures in the lower range, especially for bladder and lung cancer. There is no observed tendency in SMRs with respect to age, which suggests no age dependency on the risk ratio, $g(x)$, defined in Equation 1. Overall, females have higher SMRs than males. Liver cancer mortality is generally higher than expected, although there is no particularly strong exposure–response relationship.

The GLM analysis began by fitting all possible models and comparing the Akaike information criterion (AIC) to narrow the model choice (25). The AIC is a commonly used criterion for comparing nonnested models. It penalizes the model deviance by

adding twice the number of parameters to it. Thus a model with a low AIC will be one that describes the observed data well (a low deviance) yet with relatively few parameters (small penalty). The models with the lowest AIC provide the best fit. Because it was not appropriate to compare AIC for the models based on different data sets, we provide separate analyses with and without comparison populations.

Table 6 identifies models 1–9 and the MSW model that appear in Tables 7–10 and Figures 1–3. For simplicity we will refer to the model numbers. Table 7 compares the four top performing models based on AIC for the models with and without comparison populations for male bladder cancer. Several other models fit reasonably well, but we chose to present only four (see “Discussion”). It is important to note, however, that models including exposure concentration were highly significant compared to models excluding concentration. We also present the MSW model. Although detailed results are shown here only for male bladder cancer, the same general patterns apply to females and to all cancer outcomes except for the combined analysis (see “Discussion”). In general, models with no transformation on dose and an exponential linear dose effect fit well when we used no comparison population. When we used population data from the southwestern region of Taiwan or the entire Taiwanese population, models with the square root and log transformation fit well. This is most likely

Table 5. Summary statistics.

	Villages (n)	PYR ^b	Observed no. of cancer deaths (SMR ^a × 100)			
			Bladder	Lung	Liver	Combined
Overall	42	490,929	175 (1,327)	266 (266)	173 (134)	614 (254)
Exposure ($\mu\text{g/L}$)						
0–50	8	92,920	26 (1,002)	30 (156)	29 (118)	85 (183)
50–100	6	102,797	12 (415)	31 (143)	18 (65)	61 (116)
100–200	4	40,679	12 (1,047)	21 (243)	19 (174)	52 (251)
200–300	3	36,514	8 (766)	24 (308)	14 (144)	46 (247)
300–400	4	28,870	6 (744)	12 (197)	6 (77)	24 (163)
400–500	3	28,655	22 (2,968)	21 (365)	12 (160)	55 (393)
500–600	7	79,446	34 (1,490)	56 (332)	34 (159)	124 (306)
600+	7	81,048	55 (3,271)	71 (514)	41 (217)	167 (486)
Age (years)						
20–40	42	258,789	2 (1,446)	5 (178)	18 (169)	25 (184)
40–60	42	164,549	21 (730)	116 (365)	76 (130)	213 (228)
60+	42	67,591	121 (1,189)	145 (222)	79 (133)	345 (256)
Sex						
Male	42	256,970	85 (1,005)	144 (220)	122 (127)	351 (206)
Female	42	233,959	90 (1,904)	122 (354)	51 (158)	263 (368)

^aDefinition from Breslow and Day (17). ^b1973–1986.

Table 6. Model description.

Model	Dose transformation	Dose effect ^a	Age effect
1	Identity	Linear	Quadratic
2	Identity	Linear	Spline
3	Identity	Quadratic	Spline
4	Log	Linear	Quadratic
5	Log	Quadratic	Quadratic
6	Log	Quadratic	Spline
7	Sqrt	Linear	Quadratic
8	Sqrt	Quadratic	Quadratic
9	Sqrt	Quadratic	Spline
MSW	Identity	Quadratic	Truncated

^aExponential linear or exponential quadratic.

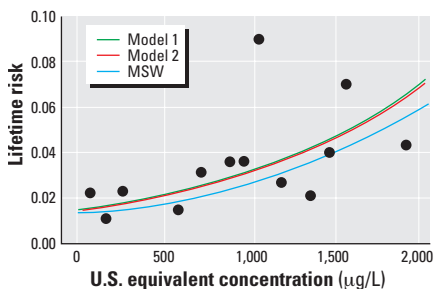


Figure 1. Estimated lifetime death risk for male bladder cancer without comparison population. For a description of models, see Table 6.

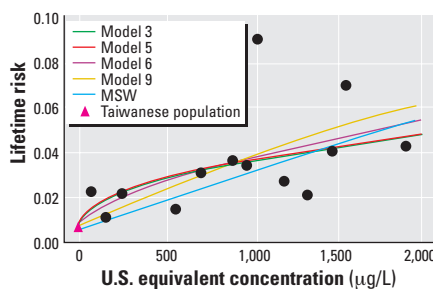


Figure 2. Estimated lifetime death risk for male bladder cancer with Taiwanese-wide comparison population. For a description of models, see Table 6.

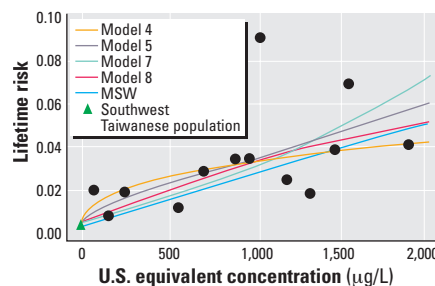


Figure 3. Estimated lifetime death risk for male bladder cancer with southwestern Taiwanese region comparison population. For a description of models, see Table 6.

due to the relatively low cancer death rates in the comparison population. The log-transformation allowed the fitted curve to rise more quickly from zero to accommodate this difference. Using the log-transformation without the comparison population gave a good model fit, according to AIC, but risk estimates were not easy to interpret because of instability of the fitted model at low dose. For this reason, we chose not to pursue the log-transformation without the comparison population any further. A few additive models gave a good fit, but in most cases, the multiplicative models did a better job, so we chose not to continue with the additive models. Also note that the MSW model fit reasonably well (Figures 1–3 show graphical representations). Each dot in Figures 1–3 corresponds to the estimated lifetime risk of dying of bladder cancer for villages, grouped by 50- $\mu\text{g}/\text{L}$ exposure levels (0–50, 50–100, etc.). The grouping is for presentation purposes only because village-specific estimates were highly variable. The idea of grouping for the purpose of graphical presentation of a fitted model has been widely used in the logistic regression context as well (26). Fitted curves for the models without the comparison population are very similar in shape, whereas there is a considerable amount of variability in the models with a comparison population.

Tables 8–10 contain risk statistics for the best-fitting GLM models and the MSW model with and without comparison population data. 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. For models 1 and 2, which have no transformation on dose, ED_{01} estimates equal 395 and 351 $\mu\text{g}/\text{L}$, respectively, for male bladder cancer. For models 3, 4, and 5, which have a log-transformed dose effect, ED_{01} estimates for male bladder cancer range from 21 to 54 $\mu\text{g}/\text{L}$. Models 7 and 8, which have a square root dose effect, give higher estimates (156 and 108 $\mu\text{g}/\text{L}$, respectively). Results for model 9 are similar to models 7 and 8. When

Table 7. AIC for best-fitting models.

Model	None	All of Taiwan	Southwestern area
1	302.1655	—	—
2	302.5547	—	—
3	—	334.8289	326.9948
4	302.9700	—	326.6287
5	—	330.0863	—
6	303.3353	330.9968	—
7	—	—	326.1207
8	—	—	327.1098
9	—	333.8307	—
MSW	302.0293	348.4275	334.3308

a comparison population is used (Tables 9 and 10), there is more variability in the predicted lifetime risk from model to model. It appears that the inclusion of a large unexposed comparison population had a relatively strong influence on estimation of risk.

Estimates of ED_{01} and ED_{05} based on using the southwestern region of Taiwan tended to be much lower than those based on using the Taiwanese-wide population. The MSW model implies a lower risk when no comparison population was used ($\text{ED}_{01} = 633 \mu\text{g}/\text{L}$

Table 8. Concentrations ($\mu\text{g}/\text{L}$) for different measures of risk (without comparison population).

Model no. ^a	Bladder		Lung		Liver		Combined	
	M	F	M	F	M	F	M	F
<i>1^b</i>								
ED_{01}	395	252	364	258	573	673	169	121
LED_{01}	326	211	294	213	437	410	148	105
$\text{MOE}_{01}(50)$	7.9	5.0	7.3	5.2	11.5	13.5	3.4	2.4
ED_{05}	1,277	813	1,345	885	— ^c	—	720	493
LED_{05}	1,076	690	1,086	733	—	—	629	430
$\text{MOE}_{05}(50)$	25.54	16.3	26.9	17.7	—	—	14.4	9.9
<i>2^d</i>								
ED_{01}	351	244	343	256	585	657	164	120
LED_{01}	296	209	279	215	451	405	144	106
$\text{MOE}_{01}(50)$	7.0	4.9	6.9	5.1	11.7	13.1	3.3	2.4
ED_{05}	1,181	796	1,288	879	—	—	703	492
LED_{05}	1,005	683	1,045	735	—	—	617	433
$\text{MOE}_{05}(50)$	23.6	15.9	25.8	17.6	—	—	14.1	9.8
<i>MSW^e</i>								
ED_{01}	633	365	227	396	864	824	163	267
$\text{MOE}_{01}(50)$	12.7	7.3	4.5	7.9	17.3	16.5	3.3	5.3
ED_{05}	1,439	828	1,171	898	—	—	706	605
$\text{MOE}_{05}(50)$	28.8	16.6	23.4	18.0	—	—	14.1	12.1

^aDose transformation, dose effect, and age effect, respectively. ^bIdentity, linear, and quadratic. ^c ED_{05} outside the observable range of data. ^dIdentity, linear, and spline. ^eIdentity, quadratic, and truncated.

Table 9. Concentrations ($\mu\text{g}/\text{L}$) for different measures of risk (with Taiwanese comparison population).

Model no. ^a	Bladder		Lung		Liver		Combined	
	M	F	M	F	M	F	M	F
<i>3^b</i>								
ED_{01}	22	21	11	8	254	331	3	2
LED_{01}	18	17	8	6	54	63	3	2
$\text{MOE}_{01}(50)$	0.4	0.4	0.2	0.2	5.1	6.6	0.1	0.0
ED_{05}	504	330	1,145	448	— ^c	—	111	54
LED_{05}	355	248	514	260	—	—	76	42
$\text{MOE}_{05}(50)$	10.1	6.6	22.9	9.0	—	—	2.2	1.1
<i>5^d</i>								
ED_{01}	23	19	11	8	239	339	3	2
LED_{01}	19	16	8	6	51	65	3	2
$\text{MOE}_{01}(50)$	0.5	0.4	0.2	0.2	4.8	6.8	0.1	0.0
ED_{05}	539	304	1,276	476	—	—	113	56
LED_{05}	380	231	564	274	—	—	77	43
$\text{MOE}_{05}(50)$	10.8	6.1	25.5	9.5	—	—	2.3	1.1
<i>6^e</i>								
ED_{01}	41	17	128	33	608	404	86	9
LED_{01}	18	9	42	10	337	87	35	3
$\text{MOE}_{01}(50)$	0.8	0.3	2.6	0.7	12.2	8.1	1.7	0.2
ED_{05}	611	293	925	491	—	—	389	125
LED_{05}	416	185	684	346	—	—	278	75
$\text{MOE}_{05}(50)$	12.2	5.9	18.5	9.8	—	—	7.8	2.5
<i>9^f</i>								
ED_{01}	100	72	76	68	895	511	45	17
LED_{01}	65	52	32	34	542	148	19	10
$\text{MOE}_{01}(50)$	2.0	1.4	1.5	1.4	17.9	10.2	0.9	0.3
ED_{05}	708	407	978	579	—	—	499	228
LED_{05}	516	309	659	433	—	—	337	160
$\text{MOE}_{05}(50)$	14.2	8.1	19.6	11.6	—	—	10.0	4.6
<i>MSW^g</i>								
ED_{01}	164	88	196	116	480	551	106	53
$\text{MOE}_{01}(50)$	3.3	1.8	3.9	2.3	9.6	11.0	2.1	1.1
ED_{05}	852	455	1,014	579	1,089	—	544	273
$\text{MOE}_{05}(50)$	17.0	9.1	20.3	11.6	21.8	—	10.9	5.5

^aDose transformation, dose effect, and age effect, respectively. ^bLog, linear, and quadratic. ^c ED_{05} outside the observable range of data. ^dLog, linear, and spline. ^eLog, quadratic, and spline. ^fSqrt, quadratic, and spline. ^gIdentity, quadratic, and truncated.

Table 10. Concentrations (µg/L) for different measures of risk (southwestern Taiwanese comparison population).

Model no. ^a	Bladder		Lung		Liver		Combined	
	M	F	M	F	M	F	M	F
<i>4^b</i>								
ED ₀₁	21	19	10	10	119	467	3	2
LED ₀₁	17	16	8	8	37	76	2	2
MOE ₀₁ (50)	0.4	0.4	0.2	0.2	2.4	9.3	0.1	0.0
ED ₀₅	649	452	768	522	— ^c	—	93	63
LED ₀₅	422	313	403	312	—	—	66	48
MOE ₀₅ (50)	13.0	9.0	15.4	10.4	—	—	1.9	1.3
<i>5^d</i>								
ED ₀₁	54	25	76	27	503	455	62	9
LED ₀₁	21	12	22	9	247	110	22	3
MOE ₀₁ (50)	1.1	0.5	1.5	0.5	10.1	9.1	1.2	0.2
ED ₀₅	723	464	780	520	—	—	330	132
LED ₀₅	508	315	558	362	—	—	226	79
MOE ₀₅ (50)	14.5	9.3	15.6	10.4	—	—	6.6	2.6
<i>7^e</i>								
ED ₀₁	156	136	79	76	309	485	21	20
LED ₀₁	131	117	62	63	174	242	17	17
MOE ₀₁ (50)	3.1	2.7	1.6	1.5	6.2	9.7	0.4	0.4
ED ₀₅	917	624	880	608	—	—	347	250
LED ₀₅	786	548	705	510	—	—	292	219
MOE ₀₅ (50)	18.3	12.5	17.6	12.2	—	—	6.9	5.0
<i>8^f</i>								
ED ₀₁	108	85	50	63	779	559	31	18
LED ₀₁	65	56	25	35	400	168	14	10
MOE ₀₁ (50)	2.2	1.7	1	1.3	15.6	11.2	0.6	0.4
ED ₀₅	817	536	778	582	—	—	416	238
LED ₀₅	594	406	489	431	—	—	275	167
MOE ₀₅ (50)	16.3	10.7	15.6	11.6	—	—	8.3	4.8
<i>MSW</i>								
ED ₀₁	185	101	181	113	709	597	98	55
MOE ₀₁ (50)	3.7	2.0	3.6	2.3	14.2	11.9	2.0	1.1
ED ₀₅	959	520	936	583	1,608	—	506	284
MOE ₀₅ (50)	19.2	10.4	18.7	11.7	32.2	—	10.1	5.7

^aDose transformation, dose effect, and age effect, respectively. ^bLog, linear, and quadratic. ^cED₀₁ outside the observable range of data. ^dLog, quadratic, and quadratic. ^eSqrt, linear, and quadratic. ^fSqrt, quadratic, and quadratic.

for male bladder cancer) compared to estimates when a comparison population is used (164 and 185 µg/L).

Discussion

In contrast to the 1988 EPA risk assessment that focused on skin cancer incidence (4), this study examines cancer mortality in a setting where exposure is measured at village level. Although there is an advantage to having individual village measurements, there also appears to be variability in the exposure assessment, causing high variability in the risk estimates. Depending on the model and whether or not a comparison population is used in the analysis, ED₀₁ estimates range in value from 21 to 633 µg/L for male bladder cancer. For males, the lung cancer risk tends to be slightly higher than the risk for bladder cancer, with ED₀₁ values ranging from 10 to 364 µg/L. Although this result seems in contrast to the high SMRs for bladder cancer in Table 5, the risk estimates are calculated on an additive scale and are influenced by background cancer rates. Hence, even though bladder cancer has high SMRs, the number of excess bladder cancer deaths associated with exposure is only moderate because of

the low bladder cancer death rate in the general population. In contrast, because lung cancer is more prevalent in the general population, even a moderate SMR can lead to high numbers of excess deaths. There does not appear to be high risk associated with liver cancer in males with the exception of estimates based on three models that used a log-transformation of exposure (models 3, 4, and 5). ED₀₁ estimates range from 309 to 895 µg/L for models apart from the latter, which yields values that range from 199 to 254 µg/L. The risk associated with female cancers tends to be higher than that of males for each cancer type. For bladder cancer, ED₀₁ estimates for females range from 17 to 365 µg/L. For lung and liver cancer, female ED₀₁ estimates range from 8 to 396 µg/L and 331 to 824 µg/L, respectively. The best models according to AIC for bladder, lung, and liver cancer combined did not exactly correspond to the models presented in Tables 8–10. For males, the best model with no comparison population is model 1, which has a linear untransformed dose effect and a quadratic age effect (Table 8). For females the best model for combined cancer has a square root transformation on dose with a

quadratic dose and age effect. The ED₀₁ estimate based on this model equals 844 µg/L. When a comparison population is used (either all of Taiwan or the southwestern region of Taiwan), the best model for both males and females has a square root transformation on dose with a linear dose effect and spline age effect. ED₀₁ estimates based on this model with the entire Taiwanese population equal 22 and 18 µg/L for males and females, respectively. When the southwestern region was used, ED₀₁ estimates equal 21 and 20 µg/L for males and females, respectively.

Our results show that exposure–response assessments depend highly on the choice of model, as well as whether or not a comparison population is used in the analysis. As discussed by Morales et al. (10), one possible explanation is the uncertainty associated with an ecologic study design. We assumed the same arsenic concentration for all persons in the same village and individual exposures can vary widely in a village. Mortality records are available for individuals, but their individual exposures are not. The National Academy of Sciences (1) provides a good discussion on this subject.

Although one might argue that the appropriate strategy would be to select the best model based on accepted statistical criteria, several models gave essentially the same quality fit (as measured by AIC), yet yielded substantial differences in risk estimates. For example, for the models without a comparison population, the MSW model gave a fit comparable to some of the GLM models, but produced ED₀₁ estimates almost twice as high. Despite the comparably good fit, we preferred the GLM models to the MSW model. For example, sensitivity analysis revealed that the MSW model was influenced strongly by the removal of various subsets of villages, whereas the GLM was not (10). The poor nutritional status of the Taiwanese in the Blackfoot disease region could be another contributing factor of uncertainty. We could not account for dietary intake of inorganic arsenic in food for either population, or for other confounders in this analysis.

Differences in ED₀₁ estimates were particularly affected by whether or not a comparison population was used. There is reason to believe that the urban Taiwanese population is not a comparable population for the poor rural population used in this study. Thus, risk estimates using the Taiwanese population may be biased. As an alternative, we used the southwestern region of Taiwan; we found very different risk estimates based on the two different comparison populations (Tables 9 and 10). We could have done other analyses. For example, we could have calculated lifetime death rates for the unexposed

group [$ldr(0)$] using U.S. population data. It would be of interest to see how the unexposed death rates in the United States compare to the death rates in Taiwan.

Despite the considerable variation in estimated ED_{01} , the results are sobering and indicate that current standards are not adequately protective against cancer. For the combined analysis with no comparison population and identity transformation on dose, the $MOE_{01}(50)$ values range from 0.4 to 16.9 for both males and females. When we include a comparison population, the $MOE_{01}(50)$ values range from 0.2 to 3.4. The current arsenic standard of 50 $\mu\text{g/L}$ (4) is actually below the estimated ED_{01} , which suggests that the risk at the current standard is higher than 1 in 100. Note, however, this estimate is likely to be overly conservative because the data suggest that the log-transformations lead to somewhat unstable results. Even considering the identity transformation, which tended to give less extreme results, the risk associated with a concentration of 50 $\mu\text{g/L}$ is approximately 1 in 300, based on linear extrapolation from the point of departure. Risks of a similar magnitude were reported by Smith et al. (27). This is an extremely high value. We could argue that if indeed the risk were this high, we would expect to find epidemiologic evidence even within the U.S. population. The *SEER Cancer Statistics Review* (28) estimates that the age-adjusted U.S. mortality rates for bladder, lung, and liver cancer are 3.2, 49.5, and 2.8 per 100,000, respectively. It is also estimated that approximately 5% of large and small regulated water supply systems in the United States have arsenic concentrations > 20 $\mu\text{g/L}$ (29). Thus, if the excess cancer risk associated with 50 $\mu\text{g/L}$ arsenic is on the order of magnitude 1 in 1,000, we would expect an increase of approximately 0.05 per 1,000 or 5 per 100,000 in the population. It is not surprising that epidemiologic studies in the United States have not so far been able to identify clear associations. Thus, we conclude that arsenic in drinking water may indeed be contributing to excess cancer mortality in the United States.

REFERENCES AND NOTES

- National Research Council. Arsenic in Drinking Water. Washington, DC:National Academy Press, 1999.
- U.S. EPA. Health Assessment Document for Inorganic Arsenic. EPA 600/8-83/021F. Cincinnati, OH:U.S. Environmental Protection Agency, 1984.
- Tseng WP, Chu HM, How SW, Fong JM, Lin CS, Yeh S. Prevalence of skin cancer in an endemic area of chronic arsenicism in Taiwan. *J Natl Cancer Inst* 40:453–463 (1968).
- U.S. EPA. Special Report of Inorganic Arsenic: Skin Cancer; Nutritional Essentiality. EPA 625/3-87/013. Washington, DC:U.S. Environmental Protection Agency, 1988.
- Safe Drinking Water Act Amendments. 42 U.S.C. 300g-1 (b) (12) (A) (iv), 1996.
- Smith AH, Goycolea M, Haque R, Biggs ML. Marked increase in bladder and lung cancer mortality in a region of northern Chile due to arsenic in drinking water. *Am J Epidemiol* 147:660–669 (1998).
- Hopenhayn-Rich C, Biggs ML, Fuchs A, Bergoglio R, Tello EE, Nicolli H, Smith AH. Bladder cancer mortality associated with arsenic in drinking water in Argentina. *Epidemiology* 7:117–124 (1996).
- Chen CJ, Chuang YC, Lin TM, Wu HY. Malignant neoplasms among residents of a Blackfoot disease-endemic area in Taiwan: high-arsenic artesian well water and cancers. *Cancer Res* 45:5895–5899 (1985).
- Brown KG. Assessing risk of inorganic arsenic in drinking water in the United States. *Hum Ecol Risk Assess* 4:1061–1070 (1998).
- Morales KH, Ryan L, Brown KG, Kuo TL, Chen CJ, Wu MM. Model sensitivity in an analysis of arsenic exposure and bladder cancer in southwestern Taiwan. In: *Proceedings of the Third International Conference on Arsenic Exposure and Health Effects*, 12–15 July 1998, San Diego, California. New York:Elsevier, 1999:207–215.
- Chen CJ, Kuo TL, Wu MM. Arsenic and cancers [Letter]. *Lancet* ii:414–415 (1988).
- Wu MM, Kuo TL, Hwang YH, Chen CJ. Dose-response relation between arsenic concentration in well water and mortality from cancers and vascular diseases. *Am J Epidemiol* 130: 1123–1132 (1989).
- Chen CJ, Chen CW, Wu MM, Kuo TL. Cancer potential in liver, lung, bladder and kidney due to ingested inorganic arsenic in drinking water. *Br J Cancer* 66:888–892 (1992).
- WHO. *International Classification of Diseases: Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, 1965 Revision, Vol 1 and 2*. Geneva:World Health Organization, 1967.
- Chen KP, Wu HY, Wu TC. Epidemiologic studies on Blackfoot disease in Taiwan. 3: physicochemical characteristics of drinking water in endemic Blackfoot disease areas. *Mem Coll Med Natl Taiwan Univ* 8:115–129 (1962).
- Yeh SJ, Yang MH. Preliminary report on trace element content in water samples from the Blackfoot disease endemic area and in biopsy samples from Blackfoot disease patients. In: *Report on Blackfoot Disease Research, Vol. 8*. Taichung, Taiwan:Taiwan Provincial Department of Health, 1980:22–28.
- Breslow NE, Day NE. *Statistical Methods in Cancer Research. Vol II: The Design and Analysis of Cohort Studies*. New York:Oxford University Press, 1987.
- Department of Health. *Health Statistics. Vol II: Vital Statistics, 1973–1986*. Taipei, Taiwan:Department of Health, 1974–1987.
- Venables WN, Ripley BD. *Modern Applied Statistics with S-PLUS*. New York:Springer-Verlag, 1994.
- Vital Statistics of the United States, 1996. Hyattsville, MD:National Center for Health Statistics, 1996.
- Smith AH, Sharp DS. A standardized benchmark approach to the use of cancer epidemiology data for risk assessment. *Toxicol Ind Health* 1:205–212 (1985).
- U.S. Environmental Protection Agency. Proposed guidelines for carcinogen risk assessment; notice. *Fed Reg* 61:17959–18011 (1996).
- Stuart A, Ord K. *Kendall's Advanced Theory of Statistics. Vol 1: Distribution Theory*. London:Edward Arnold, 1994.
- Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York:Chapman & Hall, 1993.
- Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. New York:Chapman & Hall, 1990.
- Hosmer DW, Lemeshow S. *Applied Logistic Regression*. New York:John Wiley & Sons, 1989.
- Smith AH, Hopenhayn-Rich C, Bates MN, Goeden HM, Hertz-Picciotto I, Duggan HM, Wood R, Kosnett MJ, Smith MT. Cancer risks from arsenic in drinking water. *Environ Health Perspect* 97:259–267 (1992).
- Ries LAG, Kosary CL, Hankey BF, Miller BA, Clegg L, Edwards BK, eds. *SEER Cancer Statistics Review, 1973–1996*. Bethesda, MD:National Cancer Institute, 1999.
- Welch AH, Helsel DR, Focazio MJ, Watkins SA. Arsenic in ground water supplies of the United States. In: *Proceedings of the Third International Conference on Arsenic Exposure and Health Effects*, 12–15 July 1998, San Diego, CA. New York:Elsevier, 1999:9–17.