An Evaluation of Potential Associations between Arsenic Concentrations in Ground Water and 2000 - 2004 Cancer Incidence Rates in Idaho by Zip Code

January 2007

CDRI Cancer Data Registry of Idaho

Department of Environmental Quality

Prepared by Christopher J. Johnson, MPH, CDRI
and Tonia S. Mitchell, P.G., DEQ

Contents

Abstract 1
Background 1
Methods 2
  Ground Water Data Sources 2
  Ground Water Arsenic Results 2
  Cancer Incidence Rates 4
  Linkage of Cancer Incidence and Ground Water Arsenic Databases 4
  Behavioral Risk Factor Data 4
Statistical Methods 5
Results 5
Discussion 7
References and Data Sources 8

Figure and Tables

Figure A. Mean ZIP Code Arsenic Ground Water Concentrations through 2004 3

Table A. Proportion of CDRI ZIP Code Cases Included in Arsenic Study. 4
Table B. Numbers of Cases and Crude Cancer Incidence Rates by Site-category and Mean ZIP Code Arsenic Concentration, Idaho, 2000 - 2004. 6
Table C. Modeled Effect Sizes for 10.0 µg/l Increase in Mean ZIP Code Arsenic Concentration. 6

An Evaluation of Potential Associations Between Arsenic Concentrations in Ground Water and 2000-2004 Cancer Incidence Rates in Idaho by Zip Code

Abstract

Arsenic is a known human carcinogen, and exposure to arsenic through drinking water has been associated with several site-categories of cancer. The Cancer Data Registry of Idaho and the Idaho Department of Environmental Quality partnered to investigate the relationship between arsenic concentrations in ground water and cancer incidence rates in Idaho at the geographic level of ZIP Codes.

Counts of cancer cases in eight site-categories within ZIP Code areas were modeled as Poisson random variables in multilevel models with arsenic exposure at the ZIP Code level and cancer risk factor and/or screening behaviors at the county level. Eight cancer site-categories were selected for analysis on the basis of having known or suspected relationships to arsenic exposure: bladder, colorectal, kidney and renal pelvis, leukemia, liver and intrahepatic bile duct, lung and bronchus, myeloma, and urinary system (includes bladder). This investigation did not yield evidence linking Idaho's cancer incidence rates with mean ground water arsenic concentration at the ZIP Code level. For none of the eight primary cancer site-categories investigated was there a statistically significant relationship between ZIP Code level cancer incidence and mean ground water arsenic concentration, adjusting for county-level cancer risk factor and screening prevalence. Several study design caveats limit the generalizability of the results.

Background

In 2002, the Ground Water Program of the Idaho Department of Environmental Quality (DEQ) collaborated with the Cancer Data Registry of Idaho (CDRI) to evaluate whether a correlation exists between arsenic concentrations in ground water and cancer incidence rates by county. Using information updated through 2004, CDRI and DEQ recently revisited the project at the finer geographic scale of ZIP Code.

Arsenic is a human carcinogen: the U.S. Department of Health and Human Services has determined that inorganic arsenic is a known carcinogen, and the International Agency for Research on Cancer and the Environmental Protection Agency (EPA) have determined that inorganic arsenic is carcinogenic to humans. People can be exposed to both natural and man-made arsenic and arsenic-containing compounds through drinking water, food, soil, and air. There is convincing evidence from a large number of epidemiological studies and case reports that ingestion of inorganic arsenic increases the risk of developing skin cancer. In addition to the risk of skin cancer, there is mounting evidence that ingestion of arsenic may increase the risks of several internal cancers.

In Idaho, 95% of the population relies on ground water as a source of drinking water. Regulated Public Water Systems (PWSs) supply drinking water to customers from approximately 2,700 wells and 145 springs as ground water sources (J. Henry, DEQ, 2006, Personal Communication). Public Water Systems are required under the Safe Drinking Water Act to sample for water quality analysis and comply with regulatory standards. Private wells are not subject to water quality regulations under the Safe Drinking Water Act. The Idaho Department of Water Resources (IDWR) estimates conservatively that there are approximately 170,000 private wells used for domestic purposes in the state (J. Sharkey, 2006, Personal Communication).

The current maximum contaminant level (MCL) for arsenic established by the EPA is 10.0 micrograms per liter (μg/l) [or parts per billion (ppb)]. The current rule for arsenic (effective January 2006) requires both community water systems and non-community non-transient water systems to be in compliance with the arsenic MCL of 10.0 μg/l. A community water system is one that has 15 or more connections or serves 25 or more residents year round. A non-community non-transient water system regularly serves 25 or more of the same individuals for at least six months of the year (e.g., schools and offices). There is no requirement to test for arsenic in non-community transient systems (e.g., restaurants, rest areas, and campgrounds) or in private wells.

Methods

Ground Water Data Sources

Ground water monitoring results for arsenic from various agencies were compiled by DEQ. For this evaluation, data collected through 2004 were used. For wells with multiple sampling dates, the most recent sample was selected. The data were drawn from Statewide Ground Water Quality Monitoring Program (Statewide Program) studies and other United States Geological Survey (USGS) studies. The Statewide Program is administered by the IDWR, with sampling performed by the USGS, and consists of approximately 1,600 statistically distributed sampling locations throughout the State. The objectives of the Statewide Program are to characterize the ground water quality of the state's major aquifers, to identify trends and changes in ground water quality, and to identify potential ground water quality problem areas. The majority of the Statewide Program locations are sampled on five-year cycles, with about 100 locations sampled annually. The arsenic results from the Statewide Program were combined with those from DEQ regional/local monitoring projects and from PWS ground water sources sampled at the spring or well and reported by the PWS. DEQ regional/local monitoring projects for arsenic (generally from individual private wells) include a 1995 Arsenic Study in Washington County, a 1998 Follow-Up Study, and the 1996-1997 Arena Valley Study. In 2003, DEQ collaborated with the Idaho State Department of Agriculture (ISDA) to collect samples for arsenic analysis while on a routine agricultural ground water quality monitoring project near Weiser in Washington County. ISDA does not usually sample for arsenic in their projects.

Ground Water Arsenic Results

The ZIP Code database used by DEQ contained 250 ZIP Code areas. For each ZIP Code, the mean ground water arsenic concentration (μg/l) value was calculated using the results of the most recent water analysis from wells with multiple sampling events. The number of individual well values used to calculate the mean ZIP Code value ranged from 0 to 106. For ZIP Codes with no individual well values, data were imputed using ESRI Spatial Analyst (weighted by inverse distance; 12 closest; power (2); pixel size 1 km). Because data entry for detection limits and analytic methods varied across agencies and time periods, the value 0.0 μg/l was imputed for "below detection limits" for individual well values. Nineteen ZIP Code areas had mean ground water arsenic concentration values greater than or equal to the MCL (10.0 μg/l), and 25 ZIP Code areas had mean ground water arsenic concentration values between 5.0 and 9.9 μg/l (Figure A). No background or natural arsenic level has been established.

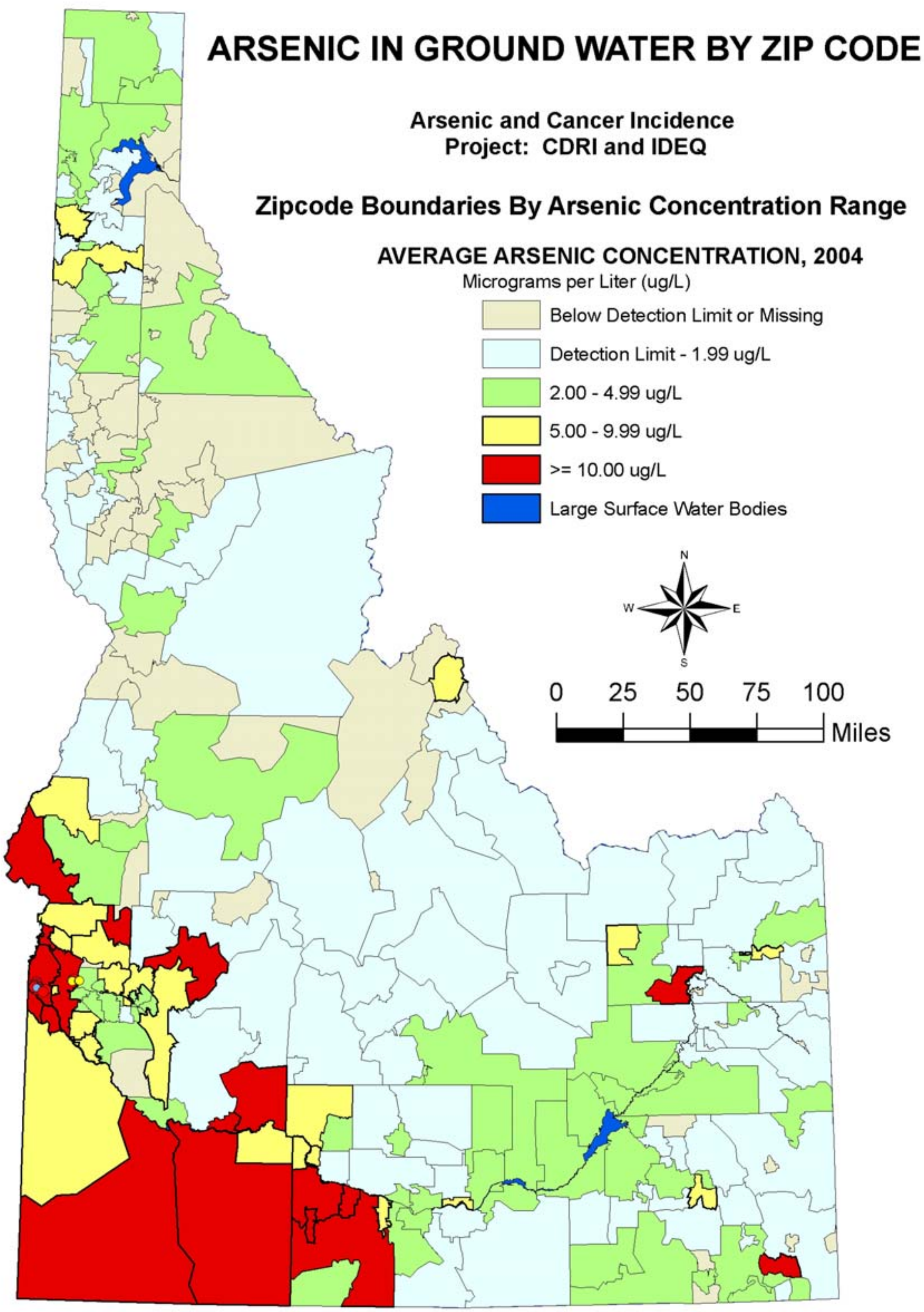

Figure A. Mean ZIP Code Arsenic Ground Water Concentrations through 2004.

Cancer Incidence Rates

Cancer incidence rates were provided by CDRI. CDRI, a program of the Idaho Hospital Association, is a statewide cancer registry that collects incidence and survival data on all cancer patients who reside in the state of Idaho or who are diagnosed and/or treated for cancer in the state of Idaho. CDRI is a North American Association of Central Cancer Registries gold standard registry on the basis of high levels of quality, completeness, and timeliness. Age- and sex-specific cancer incidence rates were calculated for Idaho residents by ZIP Code of residence at the time of diagnosis for the time period 2000 - 2004. Population counts by age and sex were obtained from Census 2000 data for Idaho Zip Code Tabulation Areas (ZCTAs). Environmental Systems Research Institute 2004 ZIP Code population estimates were used in combination with the Census 2000 data to linearly interpolate annual population counts for 2000 - 2004, which were used as denominators in the rate calculations. There were 271 ZIP Codes in the CDRI cancer incidence database.

Eight cancer site-categories were selected for analysis on the basis of having known or suspected relationships to arsenic exposure: bladder, colorectal, kidney and renal pelvis, leukemia, liver and intrahepatic bile duct, lung and bronchus, myeloma, and urinary system (includes bladder). Invasive and bladder in situ cases were used in the calculation of incidence rates, in accordance with the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program. Incidence statistics were calculated with SEER*Stat software (Information Management Service, Inc., Silver Springs, Maryland).

Linkage of Cancer Incidence and Ground Water Arsenic Databases

Seven ZIP Codes with ground water arsenic concentrations but no cancer incidence data were excluded from the study. Mean arsenic concentrations ranged from 0.0 to 5.5 μg/l in these ZIP Codes, but only two of the ZIP Codes had concentrations actually detected (the remaining five used imputed values). Twenty-eight (28) ZIP Codes with cancer incidence data but no ground water arsenic measured were excluded from the study. In total, these ZIP Codes accounted for 0.0% - 1.8% of total incident cases by cancer site-category (Table A). A total of 243 ZIP Codes included cancer incidence and arsenic concentration data and were used in the statistical analysis.

Table A. Proportion of CDRI ZIP Code Cases Included in Arsenic Study.

Primary Site-category	Number of Cases in CDRI ZIP Code Database	Number with ZIP Code covered by As Study	Proportion of ZIP Code Cases in As Study
Bladder	1,311	1,287	98.2%
Colorectal	2,776	2,750	99.1%
Kidney and Renal Pelvis	799	792	99.1%
Leukemia	853	846	99.2%
Liver and Bile Duct	206	206	100.0%
Lung and Bronchus	3,480	3,438	98.8%
Myeloma	324	321	99.1%
Urinary System	1,446	1,424	98.5%

Behavioral Risk Factor Data

County-level prevalence rates of risk factors and screening behaviors for the eight cancer site-categories included in the study were calculated by CDRI using aggregated Idaho Behavioral

Risk Factor Surveillance System (BRFSS) data for the years 1997 - 2005 provided by the Bureau of Health Policy and Vital Statistics, Idaho Department of Health and Welfare. County-level covariates included current smoking prevalence, proportion of adults aged less than 65 with health insurance, obesity prevalence, and blood stool testing (one type of screening for colorectal cancer) among adults aged 50 and older. For counties with fewer than 30 respondents for any BRFSS measure, county-level data were imputed with health district results for that measure.

Statistical Methods

Counts of cancer cases in each of the eight site-categories within each included ZIP Code area were modeled as Poisson random variables in multilevel models with arsenic concentration at the ZIP Code level and BRFSS risk factor and/or screening behavior factors at the county level. Known risk factors and screening behaviors were selected as appropriate for each cancer site-category (Lenhard et al. 2001, Schottenfeld et al. 1996). The statistical software SAS (version 9.1) Proc GLIMMIX (November 2005 release) was used to run the generalized linear mixed models. County was included as a random effect and mean arsenic concentration by ZIP Code and BRFSS variables by county as fixed effects. Spatial autocorrelation was modeled using a radial smoother covariance structure based on latitude and longitude coordinates of ZIP Code centroids. The models were optimized using the Newton-Raphson technique with ridging, starting from generalized linear model estimates.

The number of expected cases was calculated by site-category for each ZIP Code based on age- and sex-specific rates for the total of all ZIP Codes and the population size by age and sex for each ZIP Code. The natural log of the number of expected cases was used as an offset variable in the statistical models. This model parameterization results in e (beta for arsenic concentration) being interpreted as the standardized incidence ratio (SIR) associated with a unit change in arsenic concentration at the ZIP Code level, adjusting for covariates at the county level. An SIR with the value of 1.0 means that there is no relationship between arsenic concentration and cancer incidence. SIR values less than 1.0 mean that as arsenic concentration goes up, cancer incidence goes down. SIR values greater than 1.0 mean that as arsenic concentration goes up, so does cancer incidence. P-values associated with the SIR values show the probability that an association of this magnitude, or further away from 1.0, would be found by chance alone. P-values less than 0.05 are statistically significant and unlikely to be due to chance alone.

Results

Table B shows the numbers of cases and crude cancer incidence rates for the primary site-categories studied, categorized by mean ZIP Code arsenic concentration. Crude cancer incidence rates were typically higher in the ZIP Codes with mean arsenic concentrations greater than or equal to 10.0 μg/l.

Table B. Numbers of Cases and Crude Cancer Incidence Rates by Site-category and Mean ZIP Code Arsenic Concentration, Idaho, 2000 - 2004.

	Mean Ground Water Concentration in ZIP Code area			
	> = 10.0 μg/l (19 ZIP Codes)		< 10.0 μg/l (224 ZIP Codes)	
Primary Site-category	Cases	Crude Rate per 100,000	Cases	Crude Rate per 100,000
Bladder	126	21.8	1,161	19.0
Colorectal	275	47.5	2,475	40.5
Kidney and Renal Pelvis	70	12.1	722	11.8
Leukemia	77	13.3	769	12.6
Liver and Bile Duct	29	5.0	177	2.9
Lung and Bronchus	374	64.6	3,064	50.1
Myeloma	21	3.6	300	4.9
Urinary System	128	22.1	1,296	21.2

For none of the eight primary cancer site-categories investigated was there a statistically significant relationship between ZIP Code level cancer incidence and mean ground water arsenic concentration, adjusted for county-level cancer risk factors and screening prevalence (Table C).

Table C. Modeled Effect Sizes for 10.0 μg/l Increase in Mean ZIP Code Arsenic Concentration.

Primary Site-category	Estimated Adjusted Standardized Incidence Ratio (SIR)	p-value
Bladder	0.98	0.665
Colorectal	0.99	0.743
Kidney and Renal Pelvis	1.02	0.783
Leukemia	1.05	0.460
Liver and Bile Duct	1.17	0.196
Lung and Bronchus	1.01	0.825
Myeloma	0.88	0.341
Urinary System	0.99	0.862

- Adjusted for county-level current smoking prevalence, there was not a statistically significant relationship between urinary bladder cancer incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.665$).
- Adjusted for county-level physical activity, colorectal cancer screening, and obesity prevalence, there was not a statistically significant relationship between colorectal cancer incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.743$).
- Adjusted for county-level current smoking prevalence, there was not a statistically significant relationship between kidney and renal pelvis cancer incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.783$).
- Adjusted for county-level health insurance coverage, there was not a statistically significant relationship between leukemia incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.460$).

- Adjusted for county-level acute and chronic alcohol drinking prevalence, there was not a statistically significant relationship between liver and intrahepatic bile duct cancer incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.196$).
- Adjusted for county-level current smoking prevalence, there was not a statistically significant relationship between lung and bronchus cancer incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.825$).
- Adjusted for county-level health insurance coverage, there was not a statistically significant relationship between myeloma incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.341$).
- Adjusted for county-level current smoking prevalence, there was not a statistically significant relationship between urinary system cancer incidence and mean ground water arsenic concentration (μg/l) at the ZIP Code level ($p = 0.862$).

Discussion

This evaluation did not yield evidence linking Idaho's cancer incidence rates and mean ground water arsenic concentration at the ZIP Code level. The results were insensitive to parameterization of ZIP Code ground water arsenic concentration; models using arsenic concentration measured as a dichotomous variable (whether or not mean ZIP Code arsenic concentration was ≥ 10.0 μg/l) yielded the same pattern of non-significant results.

There are several potential caveats to the results. First, while several specific cancer site-categories have been related to arsenic exposure, the proportion of cases attributable to arsenic exposure has not been well documented. This complicates the choice of cancer site-categories for study and the interpretation of results, particularly negative results. Second, a demonstrated relationship (or lack thereof) between arsenic concentration and cancer incidence at the ZIP Code level may not be used to infer a relationship (or lack thereof) at the individual level (the "ecological fallacy"). Third, the use of address of residence at time of diagnosis as a de facto proxy for exposure in that ZIP Code may be inaccurate. ZIP Code cancer incidence rates do not account for in-migration of cases exposed elsewhere, or out-migration of cases diagnosed with cancer after leaving a given ZIP Code; this information is not available. Finally, results from the 2005 BRFSS show that approximately 14% of Idahoans drink bottled or vended water, meaning that ground water arsenic concentration may not be an appropriate measure of their exposure.

In 2002 and 2003, DEQ compiled ground water quality information for arsenic by county. Out of 44 counties in Idaho, 38 counties have had at least one incident of an arsenic concentration greater than or equal to the MCL of 10.0 μg/l. DEQ has also correlated the ground water quality data with topography, land use, geology, and hydrogeology to identify areas of arsenic detections for further investigation that may identify the extent and occurrence of arsenic (http://www.deq.idaho.gov/water/data_reports/ground_water/reports.cfm).

Based on preliminary review, some public water supply wells that have been sampled in consistent intervals over several years (from approximately 1990 to 2002) and generally at the same time of year have exhibited high and low cycles or fluctuations in arsenic concentrations. Further investigations should include measuring the ground water level or elevation of water in a well. Fluctuations of ground water levels such as those brought on by seasonal and/or drought conditions could possibly alter the oxidation state of aquifer materials and may be associated with changes or fluctuations of arsenic concentrations in a well. Additional investigations should also provide information to rule out anomalous results and to determine the extent of the arsenic in an area, the subsurface lithologic zones associated with arsenic concentrations (including oxidation states), and potential sources of arsenic in the groundwater. Correlations of arsenic concentrations with aquifer materials may provide information necessary to design and construct wells to avoid lithologic zones associated with high arsenic concentrations. Evaluations may be used in determining if an area may be considered a potential Area of Drilling Concern as designated by the Idaho Department of Water Resources.

The drinking water standard for arsenic (10.0 μg/l) is measured as total arsenic. Analytical methods have been developed for arsenic speciation that differentiate the various forms of arsenic and may provide for more detailed health risk evaluations.

References and Data Sources

Environmental Protection Agency, 1974, Safe Drinking Water Act including Amendments of 1986 and 1994. http://www.epa.gov/safewater/sdwa/index.html.

Henry, Jerri, 2006, Idaho Department of Environmental Quality, Personal Communication.

Idaho Dept. of Environmental Quality, Ground Water Reports http://www.deq.idaho.gov/water/data_reports/ground_water/reports.cfm.

Idaho Dept. of Water Resources, Statewide Monitoring Program http://www.idwr.idaho.gov/hydrologic/info/statewide/.

Idaho State Dept. of Agriculture http://www.agri.idaho.gov/Categories/Environment/indexEnvironment.php.

Lenhard RE, Osteen RT, Gansler T (eds). *Clinical Oncology*. The American Cancer Society, Inc.: Atlanta; 2001.

Schottenfeld D, Fraumeni JF Jr. (eds). *Cancer Epidemiology and Prevention*. New York: Oxford University Press; 1996.

Sharkey, John, 2006, Idaho Department of Water Resources, Personal Communication.