Ohio Drinking Water Source and Cancer Rates

RONALD J. KUZMA, MS, CECILIA M. KUZMA, MS, AND C. RALPH BUNCHER, ScD

Abstract: The 88 counties of Ohio were classified as either ground water or surface water counties based on the source of the drinking water used by a majority of the county residents included in the 1963 U.S. Public Health Service Inventory of Municipal Water Facilities. Average cancer mortality rates for surface and ground water counties were compared using analysis of covariance. Mortality rates for stomach, bladder, and all malignant neoplasms were higher for white males in counties served by surface water supplies than in counties served by ground water supplies. Mortality rates for stomach neoplasms were higher for white females in surface water counties. These differences in mortality rates were not attributable to other factors known to be associated with cancer death rates including urbanization, median income, population size, manufacturing activity, and agriculture-forestry-fishery activity. (Am. J. Public Health 67:725–729, 1977)

Concern has recently been expressed over the existence of trace amounts of organic chemicals, some of which may be hazardous to human health, in the nation's drinking water supplies. The Environmental Protection Agency (EPA) has conducted two surveys to determine the source and concentration of some of these chemicals.[1] The first was a survey of 80 cities across the United States to determine the levels of four trihalomethanes, carbon tetrachloride, and 1,2-dichloroethane in the drinking water. Simultaneously, in a substudy, the organic content of the drinking water of ten cities was characterized as completely as possible under existing analytical techniques. The second survey was of the drinking water of 83 cities in the states of Ohio, Indiana, Illinois, Wisconsin, Minnesota, and Michigan (EPA Region V) with an emphasis on the trihalomethanes. The results of these surveys indicate that the problem of organic chemicals in public water supplies exists throughout the country.

More than 300 different specific organic chemicals have been identified in the drinking water in the United States. Most of them have not yet been tested for carcinogenic potential. However, some of the organic chemicals identified have been shown to cause cancer in animals or man and are considered known or suspected carcinogens. With the presence of known or suspected carcinogens in the drinking water, the need arises to delineate more clearly the extent to which these contaminants and any other constitute a human health hazard.

The results of both EPA studies have shown that the concentration of the surveyed organic chemicals was higher on the average in water supplies with surface water sources than in supplies with ground water sources. Other research has suggested that the levels of carcinogenic trace metals and of viruses may also be higher in surface water sources than in ground water sources.[2, 3] The extensive use of the Ohio River, its tributaries, and Lake Erie as sources of water for residents of the State of Ohio suggests the hypothesis that residents served by these sources may be exposed to a higher level of chemicals than those served by water from wells. If the chemicals in the drinking water are affecting the cancer death rates of the residents of Ohio, then we might expect to see higher cancer death rates among residents who drink from surface water sources than in residents whose drinking water comes from wells. As a first step in assessing the hazards to residents of the State of Ohio, we compared the cancer mortality patterns in counties with surface water supplies to counties with ground water supplies.

Methods

Ohio counties were classified as having surface water supplies or ground water supplies based on the 1963 United States Public Health Service (PHS) Inventory of Municipal Water Facilities.[4] The Inventory contains the source of water and the estimated population served in 1960, but not the county served, for approximately 700 Ohio facilities. The county location of each facility was determined and the estimated

From the Department of Environmental Health, University of Cincinnati. Address reprint requests to Dr. C. Ralph Buncher, Dept. of Environmental Health, Division of Epidemiology and Biostatistics, University of Cincinnati, Cincinnati, OH 45267. This paper, submitted to the Journal December 3, 1976, was revised and accepted for publication March 7, 1977.

population served was allocated to that county. In cases where a facility involved more than one county, the estimated population served was allocated to the counties on the basis of the land area served. Sources listed solely as wells were considered ground water sources. All other sources or combinations of sources (e.g. rivers, lakes) were considered surface water sources. The 88 counties were then classified according to whether the majority of the estimated users listed were served by ground water or surface water. Forty-two of the 88 were classified as ground water counties and 46 as surface water counties.

For each county, average annual age-adjusted cancer mortality rates (per 100,000 population) during the 1950–69 period were obtained for the more common cancer sites with emphasis on the gastrointestinal tract.[5] These sites included stomach (ICD 151), large intestine (ICD 153), rectum (ICD 154), biliary passages and liver (ICD 155), pancreas (ICD 157), bladder and other urinary organs (ICD 181), and all cancers combined for the white male and for the white female population, as well as lung, trachea, and bronchus (ICD 162, 163) for the white male population and breast (ICD 170) for the white female population.[6] The cancer mortality rates were calculated for counties listed on the death certificates as the county of usual residence. The 1960 population age distribution was used as the standard for the 1950–69 period. The following five variables were also recorded for each county: per cent urbanization,[7] median income,[8] population,[7] an index of the manufacturing activity,* and an index of the agriculture-forestry-fishery activity.* These variables represent social and demographic characteristics which have been reported to be associated with cancer mortality rates.[9] The manufacturing activity and the agriculture-forestry-fishery activity indices measure the per cent of the male population of each county engaged in the manufacturing industry and in the agricultural-forestry-fishery industries, respectively.

Analysis of covariance was used to test the hypothesis that the cancer mortality rates are no different in counties with ground water sources than in counties with surface water sources. The purpose of the analysis was to determine if the data would support the alternative hypothesis that the cancer mortality rates are different (higher) in counties with surface water sources than in counties with ground water sources. The analysis of covariance model included the five social and demographic variables as covariables and the water classification variable as the two-level factor of interest.[10] In general, the surface water counties are more populated, more urbanized, and more industrialized than the ground water counties and can be expected to have higher cancer death rates on the basis of these factors alone. In order to adequately compare surface water with ground water counties, these factors should first be accounted for in the analysis. The use of the analysis of covariance model permitted a comparison of the ground and surface water counties after the cancer mortality rates were adjusted for all five social and demographic characteristics. A separate analysis was done for each cancer site and for males and females. The statistical details of this analysis are given in the Appendix.

Results

Table 1 shows the average age-adjusted cancer rates from this analysis in surface water counties and ground water counties for each cancer site for white males and for white females. The average age-adjusted rates are given both prior to and after adjustment for the five covariables and are denoted as mortality rates and adjusted mortality rates, respectively. The magnitude and direction of the difference between the adjusted rates for ground and surface water sources is a measure of the effect of the drinking water source on cancer mortality. The results in Table 1 show that after adjustment for the covariables the cancer mortality rates for all cancers, and for stomach and bladder cancers were significantly ($p < .01$) higher for white males in surface water counties than in ground water counties. The rate for all cancers is about 5 per cent higher for white males in surface water counties, while the rates for stomach and bladder cancers are 19 per cent and 18 per cent higher, respectively. Table 1 also shows the mortality rates after adjustment for the covariables for stomach cancer are significantly ($p = .038$) higher for white females in surface water counties than in ground water counties. No other differences reached conventional significance levels.

The classification of source for the 88 counties as either surface or ground water counties was straightforward. In 62 (70 per cent) of the counties, the classified source represents at least 80 per cent of the residents served by municipal water supplies listed in the PHS Inventory. In this sense, the classification can be considered valid. However, five other points should be mentioned: 1) Some people may have lived in more than one county with different water sources during their lifetime; 2) It is recognized that some ground water supplies are contaminated by seepage from nearby surface water; 3) Some of the municipal supplies may be incorrectly classified as surface or ground water in the 1963 PHS Inventory because of a mistake or misinterpretation; 4) Many private and small public wells may not be in the Inventory; 5) The Inventory does not account for all persons listed in the 1960 census.

Consequently, additional analyses were calculated using subsets of the 88 counties for which the classification of the water source is more precise. The first subset included those 30 ground and 39 surface water counties for which the PHS Inventory accounted for at least 30 per cent of the 1960 census population for the county. The second subset included counties for which at least 50 per cent of the census population was accounted for by the Inventory. There were 16 ground water and 23 surface water counties in this subset.

The results of the analysis for all, stomach, and bladder cancer for these two subsets are shown in Table 2. As before, cancer mortality rates for all cancers and for stomach and bladder cancers for white males in surface water counties are higher than in ground water counties. The mortality rates are 6 per cent higher in the surface water counties in both the first

*From unpublished work on occupational indices as predictors of cancer mortality by Joseph Thomasino, C. Ralph Buncher, George Anstadt, and Shulom Wacholder

TABLE 1—Average Adjusted Cancer Mortality Rates per 100,000 Persons for Surface and Ground Water Counties

Cancer Site	County Water Source	White Males Mortality Rates	White Males Adj Mortality Rates	White Males p-Value	White Females Mortality Rates	White Females Adj Mortality Rates	White Females p-Value
All	Ground	148.2	151.2	(.0008)	128.1	128.8	(.15)
	Surface	161.5	159.1		131.7	131.0	
Stomach	Ground	11.7	11.7	(.0005)	6.1	6.2	(.038)
	Surface	13.9	13.9		6.9	6.9	
Large Intestine	Ground	15.1	15.6	(.43)	17.2	17.2	(.85)
	Surface	16.4	16.1		16.9	17.2	
Rectum	Ground	6.4	6.7	(.44)	4.8	4.8	(.79)
	Surface	6.9	6.9		4.8	4.8	
Liver	Ground	5.2	5.2	(.93)	6.2	6.2	(.20)
	Surface	5.2	5.2		6.7	6.7	
Pancreas	Ground	7.9	7.9	(.24)	4.9	5.0	(.59)
	Surface	8.5	8.5		5.2	5.2	
Bladder	Ground	4.9	5.0	(.0046)	2.2	2.2	(.53)
	Surface	6.1	5.9		2.4	2.4	
Lung	Ground	28.2	29.6	(.47)	•	•	•
	Surface	31.7	30.3				
Breast	Ground	•	•	•	23.7	23.7	(.16)
	Surface				24.6	24.6	

46 Surface water counties
42 Ground water counties

subset and the second subset. Stomach cancer mortality rates are 19 per cent higher in the first subset and 32 per cent higher in the second subset while bladder cancer mortality rates are 19 per cent higher in the first subset and 10 per cent higher in the second. We note that the bladder cancer difference in the second subset does not reach statistical significance. Also, as in Table 1, the stomach cancer mortality rate for white females is higher in surface water counties than in ground water counties. In addition, the significance level for the comparison of the mortality rates for all cancers is $p = .07$ in the second subset and the rates are 2-3 per cent higher for white females in surface water counties.

Discussion

The results in Table 1 indicate that lung cancer in white males is not associated with a difference in ground and surface

TABLE 2—Average Adjusted Cancer Mortality Rates per 100,000 Persons for Surface and Ground Water Counties

Cancer Site	County Water Source	White Males Mortality Rates	White Males Adj. Mortality Rates	White Males p-Value	White Females Mortality Rates	White Females Adj. Mortality Rates	White Females p-Value
FIRST SUBSET*							
All	Ground	151.2	154.4	(.0028)	128.8	129.5	(.12)
	Surface	164.7	163.1		133.2	132.4	
Stomach	Ground	11.9	11.9	(.0014)	6.2	6.4	(.12)
	Surface	14.2	14.2		7.1	7.1	
Bladder	Ground	5.0	5.2	(.0036)	2.3	2.4	(.62)
	Surface	6.2	6.2		2.5	2.5	
SECOND SUBSET**							
All	Ground	151.2	159.1	(.012)	128.8	131.0	(.07)
	Surface	173.8	168.8		136.1	134.6	
Stomach	Ground	11.9	11.4	(.0011)	6.2	6.1	(.014)
	Surface	14.6	15.1		7.4	7.6	
Bladder	Ground	5.3	5.8	(.25)	2.2	2.4	(.60)
	Surface	6.7	6.4		2.5	2.3	

*FIRST SUBSET = 39 Surface water counties, 30 Ground water counties
**SECOND SUBSET = 23 Surface water counties, 18 Ground water counties

water sources after adjustment. Thus, in spite of its large contribution to the total number of cancer deaths, lung cancer is not a major factor in the difference between ground and surface water sources for all cancers combined in white males. An analysis of the mortality rates for all cancers except lung shows essentially the same difference as does the analysis of all cancers combined. Similarly, in white females breast cancer has no pattern of relationship to the drinking water. An analysis of all cancers except breast shows essentially the same differences as the analysis of all cancers. These findings provide reassurance that the omission of variables such as cigarette smoking, known to be a causative factor for lung cancer, or fat consumption in the diet, known to be associated with breast cancer, have probably not caused any major problems in this analysis.

Tables 1 and 2 show that, in general, the effect of the covariables in the analysis is to decrease the difference between the cancer mortality rates in surface and ground water counties. Nevertheless, after the adjustment there is a statistically significant difference between ground and surface water counties in the mortality rates for stomach, bladder, and all cancers for white males and in stomach for white females. The inclusion of additional covariables in the analysis may in fact serve to further decrease these differences to the point that they are no longer significant. However, the risk of not adjusting for other possibly relevant covariables must be tempered by the risk of diminishing a real difference by including irrelevant covariables. The anlysis has already considered major characteristics that have been reported to be related to cancer mortality rates. Obviously, there are a number of other variables which can reasonably be expected to be associated with cancer mortality. For example, stomach cancer varies with ethnicity and social class and bladder cancer varies with occupational exposure to chemicals. However, comparable data were not available to us on these characteristics. Fortunately, such characteristics are likely to be correlated with the covariables already in the model and in that sense are at least partially accounted for in the analysis. On the other hand, the failure of the major characteristics to totally explain the difference between surface and ground water suggests that the difference is not due to social and demographic factors alone.

One likely explanation for the differences is the presence in the surface drinking water supplies of substances which can cause cancer. The Environmental Protection Agency surveys have shown that carcinogens are present in drinking water and that the concentration of organic chemicals is higher in surface water supplies. This study confirms the hypothesis that cancer mortality rates are higher among residents of Ohio counties whose drinking water comes from surface water supplies than among the residents whose drinking water comes from ground water supplies. This is the first study to demonstrate such a difference between ground and surface drinking water supplies. While the study is only a first step in assessing the health affects of these chemicals and by no means establishes a causative factor, it suggests that organic chemicals or other substances in the water supply may relate to increased cancer mortality in communities served by surface water supplies. It further indicates the need for more specific epidemiologic studies using more direct water quality assessments.

STATISTICAL APPENDIX

The model used in the analysis of covariance was:

$$Y_{ij} = \mu + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \beta_3 X_{ij3} + \beta_4 X_{ij4} + \beta_5 X_{ij5} + \alpha_i + e_{ij}$$

where Y_{ij} represents the cancer mortality rate for the i-th water source ($i = 1,2$) and the j-th county ($j = 1, \ldots, 88$) and μ represents the intercept.

$\beta_1, \ldots, \beta_5$ are the regression coefficients for $X_{ij1}, \ldots, X_{ij5}$ respectively.

X_{ij1} is the logarithm of the county population.

X_{ij2} represents the per cent urbanization.

X_{ij3} corresponds to the median income.

X_{ij4} is the index of manufacturing activity.

X_{ij5} represents the index of agriculture-forestry-fishery activity.

α_i corresponds to the effect of the ith water source.

and e_{ij} represents the residual error and is assumed to be distributed $N(O,\sigma^2_e)$.

This model was used in the separate analysis of the cancer mortality rate for each site and sex. The five covariables were identical in each analysis. Because the cancer mortality rates are proportions, the arcsine square root transformation was applied to the rates prior to analysis as a variance stabilizing measure.[11] The adequacy of this model was investigated through plots of the observed, predicted, and residual values versus the covariables and comparison of the within-group regression equations for each water source. These analyses indicated no appreciable deviations from the model or assumptions. In this sense, the model can be considered adequate for the Ohio data. The analysis was computed using the general linear model format in Procedure REGR in the 1972 version of the Statistical Analysis System.[12] The adjusted water source means from the analysis were backtransformed to the mortality rate per 100,000 persons. It is these backtransformed rates that appear in Tables 1 and 2.

REFERENCES

1. Preliminary Assessment of Suspected Carcinogens in Drinking Water, Report to Congress, U.S. Environmental Protection Agency, Washington, DC 20460, December 1975.
2. Berg, J. W. and Burbank, F., Correlations between carcinogenic trace metals in water supplies and cancer mortality, Ann. NY Acad Sci 199:249-264, 1972.
3. Gilbert, R. G., Rice, R. C., Bouwer, H., Gerba, C. P., Wallis, C., and Melnick, J. L., Wastewater renovation and reuse: virus removal by soil filtration, Science 192:1004-1005, 1976.
4. 1963 Inventory of Municipal Water Facilities, U.S. Department of Health, Education, and Welfare, Public Health Service Publication No. 775, 1964.
5. Mason, T. J. and McKay, F. W., U.S. Cancer Mortality by County: 1950-1969, U.S. Department of Health, Education, and Welfare, Publication No. 74-615, National Cancer Institute, Bethesda, MD, 1974.
6. World Health Organization Manual of International Statistical Classification of Disease, Injuries and Causes of Death, 7th Revision, Geneva, 1957.
7. County and City Data Book 1972: A Statistical Abstract Supplement, U.S. Department of Commerce, Bureau of Census, U.S Government Printing Office, Washington, DC, 1973.

8. United States 1970 Census, Fourth Count Tapes, Census Work File, 1970.
9. Lilienfeld, A. M., Levin, M. L., and Kessler, I. I., Cancer in the United States, MA: Harvard University Press, 1972.
10. Winer, B. J., Statistical Principles in Experimental Design, Second Edition, McGraw-Hill Book Company, pp. 809–812, 1971.
11. Snedecor, G. W. and Cochran, W. G., Statistical Meth Sixth Edition, IA: The Iowa State University Press, pp. 327– 1969.
12. Barr, J. and Goodnight, J., Statistical Analysis System, N Carolina State University, 1972.