

GIS Analysis of Brain Cancer Incidence near National Priorities List Sites in New Jersey

Oleg I Muravov, MD, PhD,* Wendy E Kaye, PhD, C Virginia Lee, MD, MPH, Paul A Calame, MS, Kevin S Liske, MA

Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services, Atlanta, GA

Abstract

This analysis was done to test the hypothesis that living close to hazardous waste sites included on the US Environmental Protection Agency's National Priorities List (NPL) might be associated with an increased risk for brain cancer. A total of 2,556 cases of primary brain cancer (code 191 of the *International Classification of Diseases* [ICD-9]) received from the New Jersey State Cancer Registry for the years 1986 through 1990, were geocoded using the Matchmaker 2000 address-matching program from Geographic Data Technology. Of the 2,556 cases, 178 (6.96%) could not be geocoded, 1 (0.04%) was found to be in another state, and 226 cases (8.84%) reported from death certificates only were excluded, leaving 2,151 cases (84.15%). The NPL sites in the state were mapped using a geographic information system (GIS), and 1-mile buffers were created around each of them. These areas were analyzed for excess brain cancer. Also, the average distance between cases and the nearest NPL site was determined. There were 177 cases (8.23%) within 1 mile of an NPL site. Using total population data from 112 NPL sites in New Jersey, there were 1,031,504 (13%) persons living within 1 mile of an NPL site. No elevated cancer incidence rates were found in the analyzed areas. Also, the sites were classified according to known off-site contamination. No statistically significant differences were found among either cases' age or distance from the nearest site in relation to the primary site contaminant. This analysis can be useful as a tool for developing more in-depth environmental health hypotheses.

Keywords: brain cancer, cancer epidemiology, National Priorities List (NPL) sites, spatial analysis, surveillance system

Introduction

The federal Agency for Toxic Substances and Disease Registry (ATSDR) is developing a surveillance system using cancer registry data from states to identify potential patterns between the occurrence of brain cancers in those states with US Environmental Protection Agency National Priorities List (NPL) sites and possible exposures to hazardous substances (1). Selected cancers is one of ATSDR's priority health conditions (2). Exposures to several chemicals that have been associated with an increased incidence of primary brain cancer (3) might be occurring at hazardous waste sites included on the NPL and a concern has been expressed about the rates of primary brain cancer around some of these sites. Initially, six states—California, Florida, Massachusetts, New York, Pennsylvania, and Virginia—were included in the project. Virginia was dropped,

* Oleg I Muravov, MD, PhD, Agency for Toxic Substances and Disease Registry, US Department of Health and Human Services, 1600 Clifton Rd., MS E-31, Atlanta, GA 30333 USA; (p) 404-639-5131, (f) 404-639-6219; E-mail: oim0@cdc.gov

however, because address data were not available, and New Jersey was added in response to a reported cluster of childhood brain cancer and leukemia in the Tom's River area. New Jersey brain cancer incidence data from 1986 through 1990 are used for this analysis. Census data for 1990 were used to obtain denominator data for census tracts and counties. Residence location at the time of diagnosis was used for the geographic analysis of cases. The analysis used year, age, residence at diagnosis, type of tumor, race, and sex of the cases.

Goal

The goal of the analysis was to test the hypothesis that living close to hazardous waste sites included on the NPL might be associated with an increased risk for brain cancer.

Objectives

The objectives of the analysis were:

- To compare the incidence rate of brain cancer among residents living within 1 mile of an NPL site with that of all New Jersey residents.
- To compare brain cancer incidences according to the off-site contamination of the nearest NPL site to find out whether there is an association between environmental contamination and brain cancer occurrence.

Data

Case Data Source

The Cancer Registry Program of the New Jersey Department of Health and Senior Services provided for analysis the street addresses of 2,556 cases of primary brain cancer (code 191 of the *International Classification of Diseases* [4]) diagnosed in New Jersey from 1986 through 1990.

NPL Sites in New Jersey

The information on NPL sites in New Jersey was obtained from the HazDat database on ATSDR's Internet home page (<http://atsdr1.atsdr.cdc.gov:8080/hazdat.html>). In particular, the HazDat Sensitive State Map (<http://atsdr1.atsdr.cdc.gov:8080/haz-usa1.html>) was used to obtain the details of site location, chemical content, and extent of on- and off-site contamination. Where NPL sites were found to be within 1 mile of any of the primary brain cancers included in the analysis, the sites were classified by whether there was known off-site contamination.

Methods

Outline

For the purpose of this analysis, the incidence rate of brain cancer among New Jersey residents living within 1 mile of an NPL site was compared with that of all New Jersey residents. The observed number of cases was compared with the expected

number by the standardized incidence ratio (SIR) and its 95% confidence interval. Also, an additional analysis was implemented using cancer cases grouped by the nearest NPL site's off-site contamination to find out whether there was an association between environmental contamination and brain cancer occurrence. Student's t-test was used to assess the statistical significance of the difference between the above-mentioned groups of cases.

Geocoding

The file obtained from the New Jersey Cancer Registry Program was cleaned to ensure consistency of town and street names. The plus-4 codes were added to the zip codes of streets from the US Postal Service's *Zip+4 State Directory* for New Jersey (5). Street addresses were geocoded using the Matchmaker 2000 address-matching program from Geographic Data Technology (Lebanon, NH). Cases whose addresses were missing or were just post office box numbers were removed from the file before geocoding. The geocoded file was exported into a dBASE format to use in ARC/INFO (Environmental Systems Research Institute, Redlands, CA). The dBASE file was converted to an Info file and projected onto a coverage of New Jersey. Half-mile and 1-mile buffers were created around the NPL sites in the state (6). Cases falling within those buffers were extracted and a new file containing the sites within buffers was created. In addition, the NEAR command in ARC/INFO was run to determine the average distance between cases and the nearest NPL site. A data file was created with the cases, the nearest NPL site for each case, and the distance to that site.

Expected Number of Cancer Cases

The expected numbers of cancer cases were calculated for the 1990 population using the stratum-specific incidence rates observed in New Jersey (the standard population) during the period 1986 through 1990. The expected number of cases was calculated for each stratum first, as a product of the New Jersey incidence rate and the size of the stratum within 1 mile of an NPL site, and then summed over the strata. The number of cases expected annually from 1986 through 1989 was assumed to be equal to that estimated for 1990.

The Standardized Incidence Ratios Estimate

Standardized incidence ratios (SIRs) were used for quantitative analysis of brain cancer incidence in the 1-mile areas around the NPL sites. An SIR is calculated by dividing the observed number of cases by an expected number for the investigated population over the time period reviewed. The observed number of cancer cases for this analysis was provided by the New Jersey State Cancer Registry. The expected number of cancer cases was calculated using average annual State of New Jersey age- and sex-specific incidence rates from 1986 through 1990. The comparison rates were provided to ATSDR by the New Jersey State Cancer Registry. The lower and upper limits of the 95% confidence interval were calculated for each SIR using the Poisson distribution (7).

Results

There were 2,556 cases in the original data set with one duplicate case found. Of that total, 145 (5.7%) did not include any address information and 33 (1.3%) were post office

box numbers. Those 178 cases (7.0%) were removed from the file before address matching was done. An additional 83 cases had no street address and 3 cases had unidentifiable street addresses. After the address matching, there were 2,114 matches to street address (82.7%), 37 (1.5%) matched to the zip+4 centroid, 30 (1.17%) matched to the zip+2 centroid, and 195 (7.6%) matched to the zip code centroid. One of the zip code matches (0.04%) was removed when it was found to be in New York. Thus, 2,377 cases were in the final match for a match rate of 93%. In addition, 226 cases (8.84%) reported from death certificates only were excluded, leaving 2,151 cases (84.15%). Exclusion of the cases reported from death certificates only was based on the assumption that they were not primary brain cancer cases but likely were metastases. These 226 cases had an unspecified histological code 8000/3 ("malignant neoplasm," *International Classification of Diseases for Oncology, Morphology* [8]). In addition, 74% of these cases were 45 years of age or older, while 62% of them were 55 years of age or older. Exclusion of these cases from this analysis did not influence the findings because just 3 (1.32%) of them were within a half-mile of an NPL site. Another 22 cases (9.74%) were within 1 mile of an NPL site but were further than a half-mile, while 110 cases (47.8%) were further than 3 miles from an NPL site.

There were 177 cases (8.23%) within 1 mile of an NPL site and 54 cases (2.51%) within a half-mile of an NPL site. Using the total population data from 112 NPL sites in New Jersey, there were 1,031,504 persons living within 1 mile of an NPL site in the state. The total population of New Jersey was 7,730,188, so 13% of the total population lived within 1 mile of one of those 112 sites. The average distance between the nearest NPL site and any of the cases was $6,265.55 \pm 4,324.48$ meters (3.89 ± 2.69 miles). Of the 112 NPL sites in New Jersey, 58 (51.79%) were found to be within 1 mile of at least one brain cancer case included in the analysis.

The most frequent histologic types of cancer among the cases within 1 mile of an NPL site were glioblastoma multiforme and astrocytoma (39.53% and 37.85%, respectively). The rarest types were ependymoma (0.56%), medulloblastoma (1.13%), and meningioma (1.69%). No nerve sheath tumors were diagnosed in this population during the period studied (Table 1).

Table 1 Histologic Types of Brain Cancers among New Jersey Residents Living within 1 Mile of an NPL Site, by Year of Diagnosis

Histologic Type	ICD-O ^a Codes	1986	1987	1988	1989	1990	Total
Astrocytoma	(9400–9421)	9	10	16	13	19	67
Glioblastoma multiforme	(9440–9442)	12	11	17	12	18	70
Oligodendroglioma	(9450–9460)	1	—	1	2	—	4
Medulloblastoma	(9470–9472)	—	2	—	—	—	2
Ependymoma	(9391–9394)	1	—	—	—	—	1
Other gliomas	(9380–9383)	1	10	2	5	5	23
	(9422–9430)	—	—	1	1	1	3
Meningioma	(9530–9539)	2	—	1	2	2	7
Other brain cancers	—	—	—	—	—	—	—
Total		26	33	38	35	45	177

^a International Classification of Diseases for Oncology, Morphology (8)

The highest number of brain cancers within 1 mile of an NPL site was diagnosed among Caucasian males (61.02%), while the lowest was found among non-Caucasian females and males (1.13% and 3.95%, respectively).

Table 2 presents by age group the observed and expected numbers of cases, SIRs, and lower and upper limits of the 95% Poisson confidence interval within 1 mile of an NPL site. The expected numbers in this table were based on the incidence rates observed in New Jersey from 1986 through 1990. The values for each age group, as well as for all ages combined, were smaller than expected.

Table 2 Standardized Incidence Rates (SIRs) and 95% Confidence Intervals (CI) for Brain Cancer, 1 Mile from NPL Sites, New Jersey, 1986–1990

Age Groups	Number of Cases		SIR	95% CI Lower–Upper
	Observed	Expected		
0–14	19	25	0.76	0.46–1.20
15–44	41	60	0.68	0.49–0.93
45–64	50	83	0.60	0.45–0.75
65+	67	109	0.62	0.48–0.78
Total	177	277	0.64	0.55–0.74

Table 3 shows distribution of the cases within 1 mile of an NPL site according to the primary contamination at the closest NPL site. A total of 51 (88%) NPL sites had known off-site contamination. Of those, 30 sites (52%) were contaminated by volatile organic compounds (VOCs) and 21 sites (36%) were contaminated by metals, polychlorinated biphenyls (PCBs), or radiation. A total of 143 cases (81%) were found within 1 mile of an NPL site with known off-site contamination. Of those, 74 cases (42%) were found in proximity to VOC-contaminated sites and 69 (39%) were in proximity to sites characterized by other contaminants.

Table 3 Brain Cancer Cases within 1 Mile of NPL Sites, by Primary Off-Site Contamination, New Jersey, 1986–1990

	Known Off-Site Contamination				No Known Off-Site Contamination
	VOCs	Metals	PCBs	Radiation	
Cancer cases	74	26	15	28	34
NPL sites	30	13	2	6	7

Discussion

This analysis did not indicate that residence near an NPL site in New Jersey at time of diagnosis increased the incidence of brain cancer. The observed numbers of brain cancer within a 1-mile radius of an NPL site were lower than expected in total and in each age category. Also, histologic types of the brain cancers diagnosed in these residents

and their age, sex, or racial distributions did not differ from those of other New Jersey residents. However, there are many limitations in this type of analysis that should be considered when interpreting these results. One of the major limitations in projects involving GIS methods is the quality of the geocoding of cases. Typical address matching rates range from 20% up to 95% for rural states (9). This particular investigation had an extraordinary geocoding rate of 93%. At the same time, however, a group of 178 cases (6.96%) whose addresses were missing or just post office box addresses were removed from the dataset prior to analysis. In addition, addresses of 261 cases (10%) were incomplete, so they were geocoded to either zip+ or zip code centroids. Given the relatively small number of cases, this could have had an impact on the findings. In particular, there is a possibility that some of these cases were within a 1-mile buffer zone around an NPL site but were excluded because of geocoding errors. Another limitation in this analysis was the use of addresses available only at the time of diagnosis. Such information might not have reflected where a person got his or her exposure due to a latency period in the development of cancers and the high mobility in the US population. Should these issues be resolved, an association between living close to an NPL site and brain cancer occurrence could be either stronger or weaker than was found.

Also, the sites were classified and analyzed by known off-site contamination. A limited number of the brain cancer cases lived within 1 mile of an NPL site in New Jersey (177 cases, 8.23%), making it impossible to look for associations with specific chemicals or agents, such as ionizing radiation, and forcing investigators to group them into VOC and non-VOC cases with near equal numbers of cases in each of the groups. No statistically significant differences were found among cases' age, histological type of tumor, or distance from the nearest site in relation to the primary site contamination. No differences were found either when comparing the cases within the area of sites with known off-site contamination versus those with unknown off-site contamination.

It should be noted that the overall impact of residential proximity to NPL sites is unknown. No clear association has been found between health effects in humans and hazardous waste sites either (10,11,12). Overall, small sample size, lack of individual exposure data, poor hazardous site selection for analysis, and inappropriate health effects for the toxic substances being studied could have led to negative findings in some cases, as well as possible erroneous positive findings (11). The 1-mile radius buffer zone was chosen for this analysis as the smallest geographic area (with the shortest proximity to possible sources of exposure) in which the number of cases was large enough to provide measurable statistical power for analysis of such a rare health event as brain cancer. At the same time, estimation of relative risk (brain cancer incidence rates within a half-mile versus 1 mile) could be useful and prove a valuable addition, as could comparison of local rates to the state and national cancer rates. The small number of brain cancers diagnosed within a half-mile of NPL sites in New Jersey (54 cases, 2.5%) made it impossible to implement this approach for this particular analysis. However, it should be considered for future investigation when brain cancer incidence data from several states are available.

References

1. Shy C, Greenberg R, Winn D. 1994. Sentinel health events of environmental contamination: A consensus statement. *Environmental Health Perspectives* 102(3):316-7.

2. Lybarger JA, Spengler RF, DeRosa CT. 1993. *Priority health conditions: An integrated strategy to evaluate the relationship between illness and exposure to hazardous substances*. US Department of Health and Human Services. July. NTIS publication #PB93-203529.
3. Inskip PD, Linet MS, Heineman EF. 1995. Etiology of brain tumors in adults. *Epidemiologic Reviews* 17(2):382–414.
4. World Health Organization (WHO). 1977–78. *Manual of the international statistical classification of diseases, injuries, and causes of death: Based on the recommendations of the Ninth Revision Conference, 1975, and adopted by the Twenty-Ninth World Health Assembly*. Geneva: WHO.
5. US Postal Service. 1994. *Zip+4 state directory*. Washington, DC: US Postal Service.
6. Briggs DJ, Elliott P. 1995. The use of geographical information systems in studies on environment and health. *World Health Statistics Quarterly* 48:85–94.
7. Breslow NE, Day NE. 1987. *Statistical methods in cancer research: Vol. II—the design and analysis of cohort studies*. Ed. E Heseltine. Lyon, France: International Agency for Research on Cancer. IARC Scientific Publication #82.
8. Berg J, Maguin P, Muir C, et al. 1987. *International classification of diseases for oncology, morphology*. Ed. C Persy, VV Holten. Field Trial Edition. Lyon, France: International Agency for Research on Cancer.
9. Vine MF, Degnan D, Hanchette C. Geographic information systems: Their use in environmental epidemiologic research. *Environmental Health Perspectives*. In press.
10. Dayal H, Gupta S, Trieff N, Maierson D, Reigh D. 1995. Symptom clusters in a community with chronic exposure to chemicals in two Superfund sites. *Archives in Environmental Health* 50(2):108–11.
11. Johnson BL. 1995. Nature, extent, and impact of Superfund hazardous waste sites. *Chemosphere* 31(1):2415–28.
12. Najem GR, Strunck T, Feuerman M. 1994. Health effects of a Superfund hazardous chemical waste disposal site. *American Journal of Preventive Medicine* 10(3):151–5.