Childhood Cancer Incidence in Pennsylvania Counties in Relation to Living in Counties With Hydraulic Fracturing Sites

Jon Fryzek, PhD, Susan Pastula, MPH, Xiaohui Jiang, MPH, and David H. Garabrant, MD, MPH

Objective: Evaluate whether childhood cancer incidence is associated with counties with hydraulic fracturing (HF). Methods: We compared cancer incidence in children in Pennsylvania counties before and after HF drilling began, using standardized incidence ratios (SIRs) and 95% confidence intervals (CIs). Results: The total number of cancers observed was close to expected both before drilling began (SIR = 0.94; 95% CI, 0.90 to 0.99) and after drilling (SIR = 1.02; 95% CI, 0.98 to 1.07) for counties with oil and natural gas wells. Analyses for childhood leukemia were also unremarkable (SIR for leukemia before drilling = 0.97 [95% CI, 0.88 to 1.06]; SIR for leukemia after drilling = 1.01 [95% CI, 0.92 to 1.11]). A slightly elevated SIR was found for central nervous system tumors after drilling (SIR = 1.13; 95% CI, 1.02 to 1.25). This was because of a slight excess in those counties with the fewest number of wells. Conclusions: This study offers comfort concerning health effects of HF on childhood cancers.

Despite recent attention, hydraulic fracturing (HF) is not a new technology, having been in use for more than 60 years. Although it has long been known that natural gas was trapped in shale formations, it was not economically feasible to try to recover the natural gas until recently when horizontal drilling practices became technically feasible. Basically, the process involves the injection of a highly pressurized mixture of water and sand to create new channels in shale rock formations to increase the extraction rates and ultimate recovery of natural gas. The HF process was first developed in Texas in the 1950s and first used in large-scale production there in 1986. The Marcellus shale is the largest shale play in the United States, covering six states in the northeast, including much of Pennsylvania. In the 1980s, the HF process was brought to this shale, but was not economically feasible until the early 2000s, and has been steadily increasing since. Potential hazards from exposures associated with HF include chemicals, radionuclides, odors, traffic, noise, seismic activity, and fires and explosions from gas handling. Although specific exposures to HF workers and area residents are not well characterized, a number of potential exposure mechanisms as a result of processes associated with HF have been proposed. These include exposure potentials through transportation of materials to HF sites; the handling of fracturing fluids, which may lead to spills and contamination of ground- and surface water; injection of fracturing fluids into the wells, possibly leading to potential contamination of groundwater; the disposal of used HF or drilling waste; and gas collection, compression, storage, and transport, which may lead to air emissions, noise, and odors. In 2011, the US Committee on Energy and Commerce investigated the use of chemicals in the HF process by asking the leading oil and gas companies to disclose the types and volumes of products used in their HF fluids. The top chemicals used included methanol, isopropanol, crystalline silica, ethylene glycol monobutyl ether, and ethylene glycol.

In 2008, 14,346 adolescents and children developed cancer in the United States. Leukemia represents 26% (3664) of all childhood cancers. The majority of these childhood leukemias (approximately 75%) are acute lymphocytic leukemia. Central nervous system (CNS) tumors are the second most common tumors in children, accounting for about 20% of all childhood cancers. Although the association between many environmental or exogenous exposures and childhood cancer has been studied, few risk factors are confirmed. Established risk factors for childhood cancer include diagnostic x-rays in pregnancy, diethylstilbestrol, Epstein Bar Virus, hepatitis B, and human immunodeficiency virus. Many environmental factors have been proposed as causative for leukemia, but only ionizing irradiation and cytotoxic agents (DNA alkylators and topoisomerase II inhibitors) have been confirmed. Benzene is an established risk factor for acute myelogenous leukemia in adults. It is therefore hypothesized that benzene may play a causal role in childhood leukemia as well. Except for therapeutic radiation to the head, risk factors for CNS tumors remain ill defined.

Because of the potential of residential exposure to chemicals and radionuclides, we investigated whether living in a county with HF increased the rate of childhood cancer; in particular, the rate of childhood leukemia and childhood CNS tumors.

MATERIALS AND METHODS

All data in this report are publicly available. Analyses are presented at the county level because this is the smallest area for which both cancer data and population estimates are offered from the Pennsylvania Department of Health and the United States Census Bureau, respectively.

Pennsylvania Well Data

Information on oil and gas well drilling was ascertained from Spud data reports available through the Pennsylvania Department of Environmental Protection. The Spud data reports indicate the date when drilling of a well commences at a well site. This date must occur after the drill and operate well permit has been issued to the well operator and the well operator has a paper copy of the permit at the well site. Data elements abstracted include county where drilling occurred, dates when wells were drilled, the number of wells drilled, and the type of wells drilled (gas, oil, or other; vertical vs horizontal). The earliest Spud date in each county was set as the date the first well was drilled in the county. Although well data are available through 2011, our analysis was limited to those wells drilled between 1990 and 2009, the dates for which the cancer data are also available.

Pennsylvania Cancer Data

Analyses were conducted for all cancers in children younger than 20 years, with a special emphasis on leukemia and CNS tumors. The number of new cases of childhood cancer, childhood leukemia, and CNS tumors between 1990 and 2009 was obtained from the Pennsylvania Cancer Registry. Cancer information is available by tumor type, year, county, age group, gender, and race. It must be noted that these data were provided by the Bureau of Health Statistics and Research, Pennsylvania Department of Health. The department
specifically disclaims responsibility for any analyses, interpretations, or conclusions.

**Statistical Analyses**

Standardized incidence ratios (SIRs) were calculated to compare the observed number of cancers with that expected on the basis of rates of cancer in the general population. The observed numbers of all childhood cancers, childhood leukemia, and CNS tumors were determined for county, sex, race, and 5-year age groups for each year from 1990 to 2009. The population of each county for each study year and rates of cancer in the general population were obtained from the Surveillance, Epidemiology, and End Results (SEER) program. Expected numbers of cases were calculated by multiplying the estimated population for each county for each year of study (1990 to 2009) by annual SEER cancer rates, stratified by 5-year age groups, race, and sex. Observed and expected counts were then determined for each county for two periods or time: (1) from 1990 to the year before the first well was drilled in a county and (2) from the year the first well was drilled in a county to 2009. Observed and expected counts were then summed across all counties for each period of time, and SIRs were calculated by dividing the observed number by the expected number. Ninety-five percent confidence intervals were determined using standard techniques. Counties were further characterized by the number of wells, type of wells (gas vs oil), horizontal drilling, and Marcellus shale wells so that stratified analyses could be conducted within each of these subgroups.

**RESULTS**

More than 29,000 wells were drilled in Pennsylvania between 1998 and 2009. The number of wells drilled throughout Pennsylvania steadily increased from October 2003 through October 2008 (Fig. 1) for all types of wells (gas, oil, or other). The majority of these were gas wells (64%) and nonhorizontal wells (98%) (Table 1). Figure 2 shows the number of wells, grouped into five categories of the increasing number of wells, drilled between 1998 and 2009 in each county. As can be seen in the figure, most wells were drilled in western Pennsylvania.

We observed 1874 cancers before any type of drilling commenced in the counties of interest compared with 1996 cancers after (Table 2). Overall, SIRs for total childhood cancers combined were similar, with a SIR of 0.94 (95% confidence interval [CI], 0.90 to 0.99) before drilling and a SIR of 1.02 (95% CI, 0.98 to 1.07) after drilling. The SIRs for the time after drilling by the number of wells revealed a borderline statistically significantly elevated SIR for counties with 500 wells or fewer (SIR = 1.09; 95% CI, 1.03 to 1.15) but not for counties with more than 500 wells (SIR for counties with 501 to 1000 wells = 0.93 [95% CI, 0.79 to 1.09]; SIR for counties with 1001 to 2000 wells = 0.87 [95% CI, 0.76 to 1.00]; SIR for counties with 2001 or more wells = 0.96 [95% CI, 0.86 to 1.07]). There was no evidence of increasing SIRs as the number of wells increased. The SIRs for childhood leukemia were close to expected both before and after drilling (SIR = 0.97; 95% CI, 0.88 to 1.06) and after drilling (SIR = 1.01; 95% CI, 0.92 to 1.11), and there was no evidence of increasing SIRs as the number of wells increased. Although a slightly elevated SIR was found for CNS tumors after drilling (SIR = 1.13; 95% CI, 1.02 to 1.25), this was because of a slight excess in those counties with the fewest number of wells (SIR for counties with 1 to 500 wells = 1.22 [1.07 to 1.37]). The SIRs did not increase for CNS tumors with increasing number of wells drilled.

Because the majority of the wells were gas wells, it is not surprising that analyses restricted to counties with gas wells (Table 3) were similar to the analyses for all types of wells (Table 2). Observed numbers of totals cancers (before drilling SIR = 0.95 [95% CI, 0.91 to 0.99]; after drilling SIR = 1.02 [95% CI, 0.98 to 1.07]) and of leukemia (before drilling SIR = 0.96 [95% CI, 0.88 to 1.06]; after

![Figure 1. Gas, oil, and other types of well drilling in Pennsylvania from January 1, 1998, to December 31, 2009.](image)

**TABLE 1. Numbers and Percentages of Types of Wells Drilled in Pennsylvania and Drilling Process Used, 1998 to 2009**

<table>
<thead>
<tr>
<th>Well Type</th>
<th>Nonhorizontal Wells, n (%)</th>
<th>Horizontal Wells, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coalbed methane</td>
<td>442 (99.1)</td>
<td>4 (0.9)</td>
</tr>
<tr>
<td>Combination of oil and gas</td>
<td>3,650 (99.8)</td>
<td>8 (0.2)</td>
</tr>
<tr>
<td>Gas</td>
<td>18,180 (96.2)</td>
<td>726 (3.8)</td>
</tr>
<tr>
<td>Injection</td>
<td>48 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Oil</td>
<td>6,673 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Unknown</td>
<td>5 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Total</td>
<td>28,998 (97.5)</td>
<td>738 (2.5)</td>
</tr>
</tbody>
</table>
TABLE 2. Observed Number of Cancers, SIRs,* and 95% CIs for Childhood Cancer, Childhood Leukemia, and Childhood CNS Tumors in Pennsylvania County of Residence, Between 1990 and 2009, Overall and by Frequency of Well Drilling

<table>
<thead>
<tr>
<th>Cancer Type</th>
<th>Frequency of Drilling†</th>
<th>Before Drilling</th>
<th>After Drilling</th>
<th>Total, 1990–2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>SIR (95% CI)</td>
<td>Observed</td>
<td>SIR (95% CI)</td>
</tr>
<tr>
<td>All Cancers</td>
<td>No wells</td>
<td>6,838</td>
<td>0.99 (0.97–1.02)</td>
<td>6,838</td>
</tr>
<tr>
<td></td>
<td>1–500 wells</td>
<td>1,403</td>
<td>0.97 (0.92–1.02)</td>
<td>1,291</td>
</tr>
<tr>
<td></td>
<td>501–1,000 wells</td>
<td>114</td>
<td>0.90 (0.75–1.09)</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>1,001–2,000 wells</td>
<td>129</td>
<td>0.83 (0.69–0.98)</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>&gt;2,001 wells</td>
<td>228</td>
<td>0.91 (0.80–1.04)</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td>Total for all counties with wells</td>
<td>1,874</td>
<td>0.94 (0.90–0.99)</td>
<td>1,996</td>
</tr>
<tr>
<td>Leukemia</td>
<td>No wells</td>
<td>1,640</td>
<td>0.98 (0.93–1.03)</td>
<td>1,640</td>
</tr>
<tr>
<td></td>
<td>1–500 wells</td>
<td>339</td>
<td>0.98 (0.87–1.08)</td>
<td>303</td>
</tr>
<tr>
<td></td>
<td>501–1,000 wells</td>
<td>31</td>
<td>1.04 (0.71–1.47)</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>1,001–2,000 wells</td>
<td>35</td>
<td>0.96 (0.67–1.33)</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>&gt;2,001 wells</td>
<td>52</td>
<td>0.89 (0.66–1.16)</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Total for all counties with wells</td>
<td>457</td>
<td>0.97 (0.88–1.06)</td>
<td>471</td>
</tr>
<tr>
<td>CNS tumors</td>
<td>No wells</td>
<td>1,225</td>
<td>0.97 (0.92–1.03)</td>
<td>1,225</td>
</tr>
<tr>
<td></td>
<td>1–500 wells</td>
<td>242</td>
<td>0.90 (0.79–1.02)</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td>501–1,000 wells</td>
<td>22</td>
<td>0.93 (0.58–1.41)</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>&gt;2,001 wells</td>
<td>24</td>
<td>0.82 (0.53–1.22)</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Total for all counties with wells</td>
<td>327</td>
<td>0.89 (0.79–0.99)</td>
<td>392</td>
</tr>
</tbody>
</table>

*Adjusted for age, sex, and race.
†Frequency of drilling was determined by the number of wells drilled in each county in Pennsylvania. County of residence was then used to place each observed cancer case into categories: all childhood cancers, ICD-0-3 (ICD-0-2): C000-C809 (C000-C809); leukemia, ICD-0-3 (ICD-0-2): 9733, 9742, 9800–9920, 9963–9964 (9800–9941); CNS tumors, ICD-0-3: (ICD-0-2): C700-C729/excluding: 9590–9989 (C700-C729/excluding: 9590–9989).

CIs, confidence intervals; CNS, central nervous system; ICD, International Classification of Diseases; SIRs, standardized incidence ratios.
for horizontally drilled gas wells (those most likely to involve HF) showed that SIRs before and after drilling were not meaningfully different for all childhood cancers (before drilling SIR = 0.99 [95% CI, 0.95 to 1.03]; after drilling SIR = 0.94 [95% CI, 0.84 to 1.05]), for childhood leukemia (before drilling SIR = 1.01 [95% CI, 0.93 to 1.09]; after drilling SIR = 0.93 [95% CI, 0.73 to 1.18]), or for CNS tumors (before drilling SIR = 1.01 [95% CI, 0.92 to 1.11]; after drilling SIR = 1.06 [95% CI, 0.82 to 1.36]).

### DISCUSSION

To our knowledge, this was the first study to examine childhood cancer in all Pennsylvania counties where HF was used in gas recovery. Although it is difficult to draw etiological inferences from county-level data, the results of this study are comforting because they suggest that the incidence of childhood cancer, childhood leukemia, or childhood CNS tumors did not increase after HF drilling.

A number of chemicals may be associated with HF drilling and fracturing. After a well has been drilled and cemented, HF typically involves injection under high pressure of 2 million to 6 million gallons of fluid that is more than 99% water and sand, with chemical additives that modify the rheology and lubricity of the fluids, gels, biocides, scale inhibitors, and surfactants.1-26 Flow-back fluids, a mixture that may include drilling fluids, rock cuttings, hydrocarbons, radioactive materials, heavy metals, water, and sand return to the surface.25,26 As the gas is recovered from the well, a number of organic compounds may be released that have to be removed from the gas, including benzene, toluene, ethylbenzene, and xylene. Nevertheless, the actual exposure levels of residents to these chemicals is unknown. Of particular interest to this research is the possibility of residential exposures to hydrocarbons (specifically, benzene) and radionuclides. Although benzene is an established carcinogen for acute myelogenous leukemia in adults, its relation to childhood leukemia is less well established. On the contrary, ionizing radiation is related to childhood leukemia in many studies.17,18 Levels of residential exposure to benzene or to ionizing radiation associated with HF are unknown. Regardless, we found no evidence that childhood leukemia was elevated in any county after HF commenced. In fact, those counties that were most impacted by HF activities had a similar number of total cancer cases and a reduced, not statistically significant number of leukemia cases after drilling compared with the period of time before drilling.  

We conducted a review of the published literature to identify any case reports and studies that reported a link between living near a HF site and health effects, but few studies were found. A health survey of a sample of 16 Pavillion, Wyoming, residents was conducted after results of 2009 EPA study of drinking water contamination in the area, in which 11 of 39 drinking water wells were found to be contaminated with methane, volatile organic compounds, and other chemicals, possibly from HF activities in the area.24 The health impacts that were asked about in relation to odor events were nonspecific symptoms such as headaches, sore throat, and burning eyes and nose, lasting from 5 minutes to ongoing. There was no assessment of cancer risks. The Texas state department responded to community concerns over potential health effects from natural gas drilling in Dish, Texas, in 2010.6 Biological samples were collected from 28 residents. Blood volatile organic compound levels did not show a consistent pattern of elevation in the samples, indicating airborne community exposure. In addition to HF activities, other sources of contamination were suggested to explain the elevated levels of compounds in the blood samples, including smoking, public water chlorination by-products, and other consumer products. Urine samples did not show contamination, and 1 of 27 tap water samples was contaminated with high levels of disinfectant from the public water system.

One human health assessment has been conducted for residents of Garfield County, Colorado, to estimate the cancer and noncancer risk associated with living near a HF site. On the basis of several assumptions, including that a resident lives, works, and otherwise remains within the town with the HF site 24 hours a day per 7 days a week for 30 years and that lifetime of a resident is otherwise remains within the town with the HF site 24 hours a day per 7 days a week for 30 years and that lifetime of a resident is
water. Specifically, these were counties that were nominated by EPA-invited stakeholders through informational public meetings and the submission of electronic or written comments. Of the 48 counties nominated, EPA listed 4 counties in Pennsylvania as finalists for their study (Bradford, Greene, Susquehanna, and Washington counties), and these counties were the focus of the analysis by Moller Hikel and colleagues. We also calculated the SIRs for these four counties for the years 1998 to 2007 and found a not significant SIR of 1.13 (95% CI, 0.80 to 1.53), indicating that the rate of childhood leukemia in the counties chosen by Hikel et al is not significantly increased over what is expected on the basis of the SEER registry. It must be noted that Hikel et al conducted their analysis for the years 1998 to 2007. In contrast, we began follow-up in each county the year that drilling began and followed them through 2009, the date when we calculated the SIRs for cancer before drilling. We therefore began drilling in 1998. All of the other counties began drilling after that time (Bradford in 2002; Susquehanna in 2006; and Washington in 1999). Thus, Hikel et al included in their analyses periods of time before drilling began in the chosen counties.

This study has several limitations common to this type of investigation. First, we had no individual-level information on environmental, lifestyle, medical, or other characteristics on the population in Pennsylvania. Our data included only area-wide cancer statistics. The estimated area population sizes during the study period from 1990 to 2009 were based on the 1990 and 2000 census, with estimates of intercensus population calculated by the US Census Bureau. If the population of the HF counties had changed dramatically since 1990 and these changes were not accounted for by the US census, then the SIRs could be affected. Finally, we had no information on any exposures from HF for any individual in this study. Therefore, our SIRs could include HF-exposed and HF-unexposed individuals.

Nevertheless, we found no evidence that persons living in HF counties experienced higher childhood cancer rates overall or for childhood leukemia specifically after HF drilling commenced. We calculated indirectly standardized SIRs that do not allow a direct comparison between two groups. Indirectly standardized SIRs are useful for comparing the study population (ie, SIR for cancer before drilling) with a standard (US SEER rates). To compare the SIR before drilling to the one after drilling, we would need to conduct direct standardization techniques. Nevertheless, direct standardization of childhood leukemia runs into sparse data problems. When we calculate crude incidence rates in year-, age-, sex-, race-, and county-specific cells, the rates are zero in all the cells in which there were no observed cases and are high in all the cells in which observed cases occurred. The rates are very unstable and as a result the rates have high variiances, which increases the probability that we will miss an important finding if there was one (because of unnecessarily wide confidence intervals). As Schoenberg23 has commented, “when sample populations are so small that their strata contain mostly unstable rates and zeroes, the direct standardization procedure may not be appropriate and an alternate procedure becomes desirable.” Therefore, we believe that indirect standardization is preferable and gives a more accurate representation of the cancer risks related to HF activities than directly standardized rates.

CONCLUSIONS

The observed number of childhood cancers both before and after drilling were as expected, on the basis of SEER cancer incidence rates. This research does not support a conclusion that populations living in the vicinity of HF activities are at increased risk of childhood cancer, childhood leukemia, or childhood CNS tumors.

REFERENCES


