Lung Cancer Mortality Is Elevated in Coal Mining Areas of Appalachia

RESEARCH PAPER 2008-1

Michael Hendryx, Ph.D.
Kathryn O’Donnell, B.S.
Kimberly Horn, Ed.D.

Department of Community Medicine
West Virginia University

Abstract: Previous research has documented increased lung cancer incidence and mortality in Appalachia. The current study tests whether residence in coal mining areas of Appalachia is a contributing factor. We conducted a national county-level analysis to identify contributions of smoking rates, socioeconomic variables, coal mining intensity and other variables to age-adjusted lung cancer mortality. Results demonstrate that lung cancer mortality for the years 2000-2004 is higher in areas of heavy Appalachian coal mining after adjustments for smoking, poverty, education, age, sex, race and other covariates. Higher mortality may be the result of exposure to environmental contaminates associated with the coal mining industry, although smoking and poverty are also contributing factors. The knowledge of the geographic areas within Appalachia where lung cancer mortality is higher can be used to target programmatic and policy interventions. The set of socioeconomic and health inequalities characteristic of coal mining areas of Appalachia highlights the need to develop more diverse, alternative local economies.

Key Words: Lung cancer; coal mining; mortality; Appalachia; social inequalities

Address correspondence to: Michael Hendryx, Ph.D., Associate Professor, Department of Community Medicine, West Virginia University, One Medical Center Drive, PO Box 9190, Morgantown, WV 26506; mhendryx@hsc.wvu.edu; (304) 293-9206; (304) 293-6685 (fax).

The authors are also affiliated with the Institute for Health Policy Research (Hendryx), the Regional Research Institute (Hendryx and O'Donnell), and the Mary Babb Randolph Cancer Center (Horn) at West Virginia University.
Lung Cancer Mortality Is Elevated in Coal Mining Areas of Appalachia

Introduction

Smoking is the primary cause of lung cancer, but about 10% of lung cancer cases occur in persons who are lifetime never smokers [1], and other cases may result from the interactive effects of smoking and exposure to environmental risks. Environmental causes of lung cancer include exposure to second-hand smoke [2], airborne particulates from urban traffic or fossil fuel combustion [1,3-5], and exposure to ambient metals including zinc, chromium, copper, cadmium and nickel [6-8]. Arsenic exposure is a clear risk factor [9], including exposure through contaminated water supplies [10-12]. Other environmental risks include exposure to asbestos, polycyclic aromatic hydrocarbons [1] and radon [13-15].

Previous research identified higher lung cancer incidence and mortality in Appalachia compared to the rest of the country [16-18]. Furthermore, lung cancer incidence in rural portions of Appalachia is higher than in other rural areas of the United States [16]. Since Appalachia is primarily rural, higher lung cancer incidence and mortality is not attributable to factors unique to urban areas such as automobile exhaust or urban industry. Higher lung cancer incidence and mortality in Appalachia is thought to result from higher smoking rates and correlates of poor socioeconomic conditions characteristic of the region such as limited access to health care.

However, another factor to consider is the impact of Appalachian coal mining on the health of the resident population. Coal provides 40% of the world’s electricity [19] and its mining constitutes a major industrial activity for eight Appalachian states.
(Alabama, Kentucky, Maryland, Ohio, Pennsylvania, Tennessee, Virginia and West Virginia), where 389.9 million tons were mined in 2004 [20]. Residents of Appalachian coal mining communities report exposure to contaminated air and water from coal mining activities and express concerns for resulting illnesses [21], but empirical evidence on community health risks from coal mining activities is limited [22-24]. Coal contains carcinogenic impurities including zinc, cadmium, nickel, arsenic and many others [25], and the mining and cleaning of coal at local processing sites creates large quantities of ambient particulate matter and contaminated water [26-28]. Shiber [29] reports elevated arsenic levels in drinking water sources in coal mining areas of central Appalachia.

Elevated lung cancer mortality rates previously identified within Appalachia may result from behaviors such as smoking and other correlates of the poor socioeconomic conditions prevalent in the area, but may also result from exposure to environmental contaminants. The objective of the current study was to determine whether elevated lung cancer mortality in Appalachia is attributable to smoking, poverty, education, and other demographics, or whether there is an additional effect linked to residence in intense coal mining areas.

**Methods**

This study investigated lung cancer mortality rates for Appalachia and the nation for the years 2000-2004. Data were obtained from the Centers for Disease Control and Prevention (CDC) on lung cancer mortality rates. Mortality rates are measured at the county level per 100,000 population, age-adjusted using the 2000 U.S. standard population for mortality from cancer of the trachea, bronchus and lung (ICD-10 group
Coal production data were obtained from the Energy Information Administration, measured as tons of coal mined in the county from surface and underground mining combined. The primary analyses compared Appalachian coal mining areas to other areas of Appalachia and to non-coal mining counties outside Appalachia; 97 non-Appalachian coal mining counties were excluded from analysis unless otherwise specified.

Levels of coal mining were not normally distributed across counties. Two primary analyses examined mortality effects based on alternative methods of measuring coal mining exposure. The first grouped counties into three dummy variables: Appalachian coal mining up to three million tons combined over the six years 2000-2004, Appalachian coal mining greater than three million tons, and other counties (the latter used as the referent in regression models.) The choice of three million tons divides Appalachian coal mining counties approximately in half. The second estimated per capita exposure in Appalachia by dividing county tons mined by the county population from the 2000 Census; counties were grouped into three levels: per capita exposure up to 100 tons per person, per capita exposure greater than 100 tons, and other counties (used as the referent.)

A series of supplementary analyses were conducted to test for the robustness of findings across conditions. One set of analyses examined coal mining effects based on alternative dummy variables at integer levels from one to six million tons. A second set correspondingly examined per capita exposure effects at increasing levels. A third examined whether differences in mortality rates were related to surface mining versus underground mining. A fourth examined whether mortality rates were elevated only in
Appalachian coal mining areas or in coal mining areas outside of Appalachia, and whether differences in population density may be related to national variation.

Covariates were taken from the 2005 Area Resource File [36], CDC BRFSS smoking rate data [37], and the Appalachian Regional Commission (ARC) [38].

Covariates included percent male population, college and high school education rates, poverty rates, race/ethnicity rates, health uninsurance rates, physician supply, rural-urban continuum code, smoking rates, Southern state (yes or no), and Appalachian county.

Selection of covariates was based on previously identified risk factors or correlates of lung cancer incidence or mortality [39-44]. Specific race/ethnicity groups included percent of the population who were African American, Native American, Non-white Hispanic, and Asian American (using White as the referent category in regression models). Rural-urban continuum is scored on a nine point scale from least to most rural; because the effects of this measure may be non-linear [45] this measure was recoded into three dummy variables representing metropolitan, micropolitan, and rural or non-core areas (the latter used as the referent.) Physician supply is the number of active MDs and DOs per 1,000 population. Because residence in the South is associated with poorer health status and higher mortality risk [46-47] a dichotomous Southern variable was created to capture regional effects that partially overlap with Appalachia; Southern states included Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. CDC BRFSS smoking rates were available for states and some county-based metropolitan areas, supplemented with county rates available from some state public health websites; the state average was used...
when the county rate was not available. Appalachian counties included the 417 counties and independent cities in 13 states as defined by the ARC [38].

Analyses were conducted using bivariate correlations, general linear models and ordinary least squares regression models to test for the association between residence in coal mining areas and lung cancer mortality, without and with control for covariates. Post-hoc tests employed the Ryan-Einot-Gabriel-Welsch test to adjust for Type 1 error. The study is an analysis of anonymous, secondary data sources and met university Internal Review Board standards for an exception from human subjects review.

**Results**

First, we confirmed that age-adjusted lung cancer mortality was in fact significantly higher in Appalachia compared to the nation: 67.06 vs. 56.55 per 100,000 (two-tailed $t=12.67$, df=3026, $p<.0001$). There was also a significant gradient effect comparing three groups: lung cancer mortality was highest in heavy coal mining areas (74.21), followed by all other areas of Appalachia (65.70) and the nation (56.55); $F=88.91$, df=2,3025, $p<.0001$; post-hoc tests correcting for Type I error found all means significantly different at $p<.05$ (see Figure 1).

Table I summarizes study variables for heavy coal areas, other Appalachian areas, and the rest of the nation, based on the definition of coal mining as Appalachian counties with greater than 3 million tons mined from 2000-2004, and deleting coal mining areas outside Appalachia. Poverty, college education, and smoking rates show the same relationship: these measures are least favorable in heavy coal mining areas, and intermediate in the rest of Appalachia, compared to the rest of the country. Coal mining
areas are characterized by proportionately small race/ethnicity minority populations.

Rates of health insurance were different in the omnibus F test but post-hoc tests showed that means were not significantly different. Rural-urban status and per capita supply of doctors did not differ across these groups of counties.

Bivariate correlations were examined to test for multicollinearity among independent variables. The county poverty rate was highly correlated to percent of the population without health insurance ($r=0.82$); because preliminary regression models revealed that poverty was related to lung cancer mortality but insurance was not, the insurance variable was dropped from further analysis. Heavy coal mining areas correlated significantly with other risk factors but not at levels indicating multicollinearity, including smoking ($r=0.23$), high school education ($r=-0.12$), college education ($r=-0.10$), and poverty ($r=0.12$).

The results of the two primary regression models are shown in Table II. The table includes both unstandardized coefficients and standardized Betas (B). The results of the two coal exposure specifications were very similar. Living in Appalachian areas where lower levels of coal mining took place did not increase lung cancer mortality risk, but areas of heavy mining were associated with significantly higher adjusted lung cancer mortality. Higher age-adjusted lung cancer mortality was associated with smoking rates, urban residence, poverty, living in the South, percent male population and lower education. Residence in Appalachia was related to lower lung cancer mortality after adjusting for poverty, coal mining and other variables. A greater supply of physicians was related to higher mortality rates. The African-American and Hispanic variables were related to lower adjusted mortality. Examination of standardized Betas indicates that the
The strongest effects were for smoking, poverty, lack of high school education, residence in metropolitan counties, and the Hispanic population variable.

The race effect was examined further. In a forward inclusion regression model of national data beginning with the four race/ethnicity variables, the variable measuring percent Native Americans was not related to lung cancer mortality, a higher percent of African Americans was related to higher rates, and Asian American and Hispanic variables were related to lower rates of lung cancer mortality. The effect for the African American variable became significant and negative after adding high school education, poverty rate, smoking rate, and metropolitan county status. That is, the apparent lower lung cancer mortality rate among African American minorities is due to the confound of socioeconomic variables with race variables.

The sensitivity of the Table II results was examined by running regression models based on different levels of coal mining and per capita exposure. A summary of these models is provided in Table III. The table shows the unstandardized coal mining beta coefficient and p level based on alternative specifications of high levels of coal mining exposure. (The full regression model results for these various specifications are not shown, but they are almost identical to the Table II results.) The effect of the coal mining exposure variable was significant for all levels and both specifications, except for the lowest level of exposure measured in tons. Furthermore, the size of the beta coefficient increases with greater exposure, indicating an increasing number of adjusted deaths per 100,000.

To estimate number of deaths, the population of Appalachian coal mining areas was found from the 2000 Census (N=3,875,656 based on counties with more than 3
Translating the age-adjusted death rate from Table I into population figures, the difference between Appalachian coal-mining areas and the national rate equates to 684 excess lung cancer deaths in coal mining areas. Most of the Appalachian coal mining disparity is the result of factors such as poverty and smoking, but after adjusting for all covariates, translating the Table II beta coefficient (3.72) into number of deaths per 100,000 indicates that Appalachian coal mining counties are still associated with an excess of 144 deaths from lung cancer over the years 2000-2004.

Exposure to Appalachian coal mining activity was also significantly related to lung cancer mortality when coal mining was measured separately for surface and underground mines. Elevated mortality was found to be specific to Appalachia; mortality was not significantly higher in non-Appalachian areas where heavy coal mining took place. Table IV shows the coal mining beta coefficients and p levels for these tests, controlling for other covariates, and based on the definition of more than 3 million tons of coal. The largest coefficient was found for Appalachian surface mining. We examined whether the distinction between Appalachian and non-Appalachian mining might be related to population density. We found that population density was significantly higher in Appalachian coal mining areas (95.5 people per square mile) than in other coal mining areas (43.0 people per square mile; Satterthwaite correction for unequal variances t=4.44, df=117, p<.0001).

Some research suggests that coal miners may be at elevated risk for lung cancer, although the evidence is equivocal [48]. To address the possibility that our results are due to current or former miners who live in coal mining areas, we conducted an additional regression model limited to the heavy Appalachian coal mining counties.
(N=66). This model is based on the fact that almost all coal miners are men. Within these counties, percent male population was not related to lung cancer mortality (t=-0.71, p<.48). The fact that populations with higher percentages of males are not at higher risk suggests that the effect in coal mining locations is likely not the result of current or former miners who live in the area and who were directly exposed through occupational hazards. In addition, based on employment figures provided by the Energy Information Administration [49], coal miners constitute only about 1% of the Appalachian population in heavy coal mining areas.

**Discussion**

Lung cancer mortality is higher in Appalachia because of smoking and the correlates of poverty and low education, but an additional risk factor is living in heavy coal mining areas. Living in these areas may expose residents to pollution from the coal mining industry, or may be associated with additional behavioral or demographic characteristics not captured through other covariates. Access to health care as measured by insurance rates and doctor supply is not an explanation for higher lung cancer mortality, consistent with other research showing that coal mining areas with an adequate supply of primary care providers still experience increased health problems [50]. To eliminate lung cancer mortality disparity in Appalachia, it is necessary to continue efforts to reduce smoking and improve socioeconomic conditions; however, because coal mining location is an independent risk factor, and because coal mining overlaps with other known risks including smoking, education, and poverty, targeting anti-smoking and socioeconomic improvement interventions to these areas may be a cost-effective strategy.
Policies that would improve environmental quality in coal mining areas are also suggested by these results.

The possibility that environmental contamination from the coal mining industry causes lung cancer is consistent with known risks linked to coal. Toxins found in coal are well established carcinogens [51]. The release of particulate matter and toxins from burning coal is a lung cancer risk factor [1,52-55]. There is also an abundance of information on the deleterious health consequences of working as a coal miner, including increased risk for pneumoconiosis, heart disease, chronic obstructive pulmonary disease and perhaps lung cancer [49,56-57]. Exposure to particulate matter or toxic impurities from the coal mining industry may extend to the general population. The coal mining industry includes not only the mining of coal, but also its processing, storage and transport, and the resulting local water and air pollution can be severe [26-29, 58] and may result in increased lung cancer among community residents. The suggestion that the results may be stronger for exposure to surface mining operations relative to underground mining suggests the likelihood of greater exposure to airborne particulates from surface mining operations.

Limitations of the study include the reliance on secondary county-level data and the limited measures of coal mining exposure. Causes of individual lung cancer cases cannot be identified, and the precise pathway between residence in coal mining areas and lung cancer is unknown. Smoking rates were imprecisely measured and smoking effects, including exposure to second-hand smoke linked to poorer socioeconomic conditions, may be underestimated. Demographic or cultural variables not captured through available covariates may be contributing factors; these variables might include
Appalachian cultural beliefs such as fatalism [59] that increase risk for poor health behaviors or lack of early health care intervention, or weak tobacco control policies that increase second-hand smoke exposure. Future research should improve measures of coal mining exposure by distinguishing aspects of the mining industry, including post-mining processing facilities, and mountaintop removal mining from other forms of surface mining, and relating these aspects to health indicators. Additional research is also needed to identify exposure routes (i.e., air, water and soil), exposure levels and biological mechanisms of action that can account for higher lung cancer mortality in Appalachian coal mining areas.

The results of this study may be linked to a growing body of evidence demonstrating increased health risks across a spectrum of indicators associated with residence in Appalachian coal mining areas. This evidence includes higher mortality rates for all causes and for cardiopulmonary conditions [60], increased hospitalization risk for hypertension and chronic obstructive pulmonary disease [23], and increased rates of self-reported chronic illness and lower health status [22]. These findings are not simply the result of poverty or other demographic variables, although poverty is a contributing factor.

Regardless of whether causes are environmental, behavioral or economic, it is clear that populations in coal mining areas are at risk for a host of health problems. Those areas of Appalachia where poverty has been most persistent over time are characterized by single source economies including tobacco and coal [38]. Based on social inequalities models [61], addressing the health disparities of coal mining communities requires developing economies that offer more diverse job opportunities at
lower environmental cost, enacting and enforcing environmental protection policies, improving support for educational development, and creating built environments that are conducive to health and wellness.
Acknowledgement

The study was funded in part by a grant to the first author from the Regional Research Institute, West Virginia University. The sponsor had no role in study design; data collection, analysis, or interpretation; writing; or journal submission decision.

Conflict of Interest Statement

None declared.
References

Table I. Summary of study independent variables by geographic location.

<table>
<thead>
<tr>
<th></th>
<th>Heavy Appalachian coal (N=66)</th>
<th>Other Appalachian (N=347)</th>
<th>Rest of Nation (N=2615)</th>
<th>F or $\chi^2$ p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking rate $^a$</td>
<td>27.7</td>
<td>25.2</td>
<td>21.7</td>
<td>.0001</td>
</tr>
<tr>
<td>Percent male $^b$</td>
<td>49.1</td>
<td>49.5</td>
<td>49.9</td>
<td>.0001</td>
</tr>
<tr>
<td>Percent African American $^c$</td>
<td>3.1</td>
<td>7.3</td>
<td>9.5</td>
<td>.0002</td>
</tr>
<tr>
<td>Percent Native American $^d$</td>
<td>0.2</td>
<td>0.4</td>
<td>2.1</td>
<td>.0001</td>
</tr>
<tr>
<td>Percent Hispanic $^b$</td>
<td>0.7</td>
<td>1.8</td>
<td>7.3</td>
<td>.0001</td>
</tr>
<tr>
<td>Percent Asian American $^b$</td>
<td>0.4</td>
<td>0.5</td>
<td>1.0</td>
<td>.0001</td>
</tr>
<tr>
<td>High school education $^b$</td>
<td>70.0</td>
<td>71.4</td>
<td>78.3</td>
<td>.0001</td>
</tr>
<tr>
<td>College education $^a$</td>
<td>11.4</td>
<td>13.5</td>
<td>17.1</td>
<td>.0001</td>
</tr>
<tr>
<td>Poverty rate $^a$</td>
<td>18.2</td>
<td>14.9</td>
<td>13.3</td>
<td>.0001</td>
</tr>
<tr>
<td>Percent metropolitan counties</td>
<td>27.3</td>
<td>34.0</td>
<td>35.4</td>
<td>.36</td>
</tr>
<tr>
<td>Percent micropolitan counties</td>
<td>21.2</td>
<td>24.2</td>
<td>21.8</td>
<td>.58</td>
</tr>
<tr>
<td>Percent counties in the South $^a$</td>
<td>33.3</td>
<td>68.6</td>
<td>20.0</td>
<td>.0001</td>
</tr>
<tr>
<td>Percent without health insurance</td>
<td>14.3</td>
<td>13.7</td>
<td>14.8</td>
<td>.0005</td>
</tr>
<tr>
<td>Mean physicians per 1,000</td>
<td>1.49</td>
<td>1.31</td>
<td>1.32</td>
<td>.62</td>
</tr>
</tbody>
</table>

$^a$ Post-hoc tests significantly different between all three means.

$^b$ Coal mining and Appalachian areas significantly different from the nation.

$^c$ Coal mining areas significantly different from Appalachia and the nation.

$^d$ Post-hoc differences between means not significant.
Table II. Ordinary least squares regression model, age-adjusted lung cancer mortality rate.

<table>
<thead>
<tr>
<th>Coal exposure measured in tons</th>
<th>Coal exposure measured per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coefficient</td>
</tr>
<tr>
<td>Intercept</td>
<td>58.60</td>
</tr>
<tr>
<td>Coal mining up to 3 million tons</td>
<td>-0.15</td>
</tr>
<tr>
<td>Coal mining ≥ 3 million tons</td>
<td>3.72</td>
</tr>
<tr>
<td>Coal mining up to 100 tons per person</td>
<td>--</td>
</tr>
<tr>
<td>Coal mining ≥ 100 tons per person</td>
<td>--</td>
</tr>
<tr>
<td>Appalachia</td>
<td>-2.96</td>
</tr>
<tr>
<td>Smoking rate</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>Mean 1</td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Percent male</td>
<td>0.25</td>
</tr>
<tr>
<td>Percent African American</td>
<td>-0.07</td>
</tr>
<tr>
<td>Percent Native American</td>
<td>-0.04</td>
</tr>
<tr>
<td>Percent Hispanic</td>
<td>-0.45</td>
</tr>
<tr>
<td>Percent Asian American</td>
<td>0.15</td>
</tr>
<tr>
<td>High school education</td>
<td>-0.49</td>
</tr>
<tr>
<td>College education</td>
<td>-0.30</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.52</td>
</tr>
<tr>
<td>Metropolitan</td>
<td>9.11</td>
</tr>
<tr>
<td>Micropolitan</td>
<td>4.03</td>
</tr>
<tr>
<td>South</td>
<td>2.16</td>
</tr>
<tr>
<td>Primary care physicians per 1,000</td>
<td>0.85</td>
</tr>
</tbody>
</table>

a F=122.6 (16, 3010), p<.0001; adjusted R^2=.39.

b F=122.8 (16, 3010), p<.0001, adjusted R^2=.39.
Table III. Effect of high level of Appalachian coal mining exposure on adjusted lung cancer mortality, based on alternate specifications of exposure.

<table>
<thead>
<tr>
<th>Tons of coal in millions</th>
<th>beta</th>
<th>P&lt;</th>
<th>Per capita exposure in tons</th>
<th>beta</th>
<th>P&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.41</td>
<td>.13</td>
<td>50</td>
<td>4.12</td>
<td>.014</td>
</tr>
<tr>
<td>2</td>
<td>3.42</td>
<td>.041</td>
<td>100</td>
<td>4.49</td>
<td>.015</td>
</tr>
<tr>
<td>3</td>
<td>3.72</td>
<td>.036</td>
<td>150</td>
<td>3.90</td>
<td>.044</td>
</tr>
<tr>
<td>4</td>
<td>3.63</td>
<td>.044</td>
<td>200</td>
<td>5.34</td>
<td>.010</td>
</tr>
<tr>
<td>5</td>
<td>4.05</td>
<td>.036</td>
<td>250</td>
<td>5.54</td>
<td>.009</td>
</tr>
<tr>
<td>6</td>
<td>4.71</td>
<td>.017</td>
<td>300</td>
<td>5.59</td>
<td>.009</td>
</tr>
</tbody>
</table>
Table IV. Adjusted regression coefficients and p values based on type of mining, and Appalachian or non-Appalachian coal mining areas.

<table>
<thead>
<tr>
<th></th>
<th>Surface Mining</th>
<th>Underground Mining</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appalachia coal mining</td>
<td>5.60 (p&lt;.008)</td>
<td>4.55 (p&lt;.024)</td>
<td>3.72 (p&lt;.036)</td>
</tr>
<tr>
<td>Non-Appalachian coal mining</td>
<td>1.11 (p&lt;.57)</td>
<td>1.79 (p&lt;.47)</td>
<td>2.04 (p&lt;.21)</td>
</tr>
</tbody>
</table>
Figure 1. Age-adjusted lung cancer mortality per 100,000, years 2000-2004.