

LUNG CANCER IN OREGON

John Hart □ Sherman College of Chiropractic

□ Factors thought to be related to lung cancer include smoking, radon, and educational attainment. These factors were analyzed in the present ecological study for Oregon with correlation and linear regression statistics. A moderate, inverse, and statistically significant correlation was found with educational attainment while surprisingly, negligible and statistically insignificant correlations were found with smoking and radon. More rigorous research such as case-control study designs, are indicated to verify or refute these findings.

Keywords: Oregon, lung neoplasms, radon, educational status, smoking

INTRODUCTION

Among factors related to lung cancer are smoking, which is said to be its “number one cause,” radon (NCI, 2010a), and level educational attainment (Albano et al 2007). It is said that radon “causes” lung cancer in nonsmokers and smokers (Samet, 1989). Rates of lung cancer may vary by race, with black persons having a higher rate than white persons (CDC, 2009). The Iowa Radon Lung Cancer Study (IRLCS), a case controlled study which compared lung cancer patients (cases) to controls (no lung cancer at beginning of study), concluded that radon is a significant factor in lung cancer (Field et al, 2000). Although IRLCS found that cases had a slightly greater median time spent in their home, slightly less education, more previous lung disease, and almost three times greater percentage of smokers, adjustment for age, smoking and education revealed odds in excess of 0.50 and 0.83 in all cases and in live cases (Field et al, 2000).

While Raabe (2010) notes that inhaled radionuclides tend to be more carcinogenic when their distribution is uniform within the lungs, Cohen (2000) found that lung cancer and radon levels are inversely correlated, regardless of the inclusion of 54 socioeconomic potentially confounding variables, whether analyzed separately or in combination. However, Puskin (2003), analyzing Cohen’s data, claims that Cohen’s findings failed to show a protective effect from radon, and that the negative correlation between lung cancer and radon is due to confounding factors. Lubin (1998) and Smith et al (1998) have suggested Cohen’s findings should be rejected because they are based on an ecological confounding.

Address correspondence to John Hart, DC, MHSc, Assistant Director of Research, Sherman College of Chiropractic, P.O. Box 1452, Spartanburg, SC 29304; Phone: 864-578-8770, ext. 232; E-mail: jhart@sherman.edu

Cohen (1998a) responded to Lubin (1998) by pointing out that such correlations do not account for the discrepancy between Cohen’s findings and findings predicted by the linear no-threshold theory. In responding to Smith et al (1998), Cohen (1998b) states that many of these authors’ criticisms are not supported by evidence, and that their criticisms are more applicable to case control studies. In addition, Seiler and Alvarez (2000) defend the use of ecological studies such as Cohen’s, stating that raw data from ecological studies are appropriate for assessing risk.

The present ecological study assesses the relationship between lung cancer and exposures to radon and smoking in Oregon. The selection of this state was based primarily on availability of smoking data several years prior to available cancer data by county provided by the National Cancer Institute (NCI), currently for years 2002–2006.

METHODS

The response variable in this study was age-adjusted lung and bronchus cancer (hereafter referred to as *lung cancer*) mortality rates, both genders, all races, under age 65 (NCI, 2010b; Figure 1). The age <65 years old was selected to study death rates below the age of life expectancy, which was 77.8 in 2004 (80.4 for females and 75.2 for males) (United States Life Tables, 2007). In addition, the median age at death for all sites cancer was 73 during 2002–2006 (NCI, 2009).

The predictor variables were radon, smoking, and education. The variable *radon* consisted of a percent of samples by county in measured

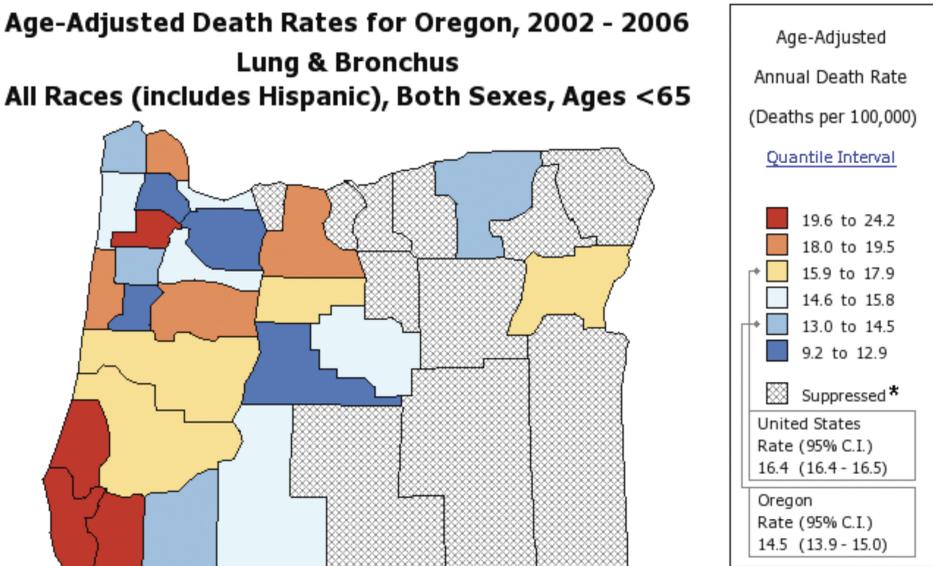


FIGURE 1. Lung cancer map of Oregon adapted from NCI (2010b), cited 2-12-2010 at <http://state-cancerprofiles.cancer.gov>. Gray shaded counties have suppressed data due to low counts.

Pico-curies per liter by homeowners during the latter part of the 1990s (Oregon Public Health Division, 1998). No details were provided regarding consistency of measurement locations, e.g., basement versus main floor. The variable *smoking* consisted of percent of current adult smokers (Oregon Health Division, 1999). No details were provided as to duration and amount of smoking. The variable *education* consisted of percent of persons 25 years and older having a bachelor’s degree or higher in 2000 (USA Counties, 2010). Counties not having reportable data due to too few counts according the NCI obviously were not included. For smoking data, Wasco county also included Sherman county but lung cancer data was reported only for Wasco county. Consequently, Wasco and Sherman counties were also not included. The raw data by county is provided in Table 1. The years for these variables were selected based on their preceding cancer data, as cancer may take years to develop, while not having too many years to minimize the effect of population mobility. A radon map from the U.S. Environmental Protection Agency is provided in Figure 2 although the actual radon data in the present study was obtained, as previously noted, from the Oregon Public Health Division (1998).

TABLE 1. Summary data of means, percents, and standard deviations by Oregon county.

| | Death | Radon | Smoke | Education |
|--------------------|-------|-------|-------|-----------|
| County | | | | |
| Baker | 17.9 | 3.6 | 20.0 | 16.4 |
| Benton | 9.8 | 1.2 | 23.0 | 47.4 |
| Clackamas | 12.9 | 2.3 | 30.0 | 28.4 |
| Clatsop | 14.1 | 1.7 | 28.0 | 19.1 |
| Columbia | 19.4 | 2.4 | 23.0 | 14.0 |
| Coos | 22.7 | 1.9 | 26.0 | 15.0 |
| Crook | 15.2 | 2.2 | 40.0 | 12.6 |
| Curry | 24.2 | 0.8 | 24.0 | 16.4 |
| Deschutes | 10.2 | 0.8 | 25.0 | 25.0 |
| Douglas | 17.9 | 1.3 | 22.0 | 13.3 |
| Jackson | 14.2 | 1.2 | 24.0 | 22.3 |
| Josephine | 20.5 | 1.0 | 33.0 | 14.1 |
| Klamath | 15.3 | 0.9 | 29.0 | 15.9 |
| Lane | 16.0 | 1.7 | 26.0 | 25.5 |
| Lincoln | 18.9 | 1.4 | 30.0 | 20.8 |
| Linn | 18.5 | 1.3 | 25.0 | 13.4 |
| Marion | 14.7 | 3.3 | 27.0 | 19.8 |
| Multnomah | 15.1 | 3.1 | 24.0 | 30.7 |
| Polk | 14.3 | 1.8 | 27.0 | 25.3 |
| Tillamook | 15.8 | 3.0 | 23.0 | 17.6 |
| Umatilla | 14.5 | 2.7 | 24.0 | 16.0 |
| Wasco | 19.5 | 1.6 | 27.0 | 15.7 |
| Washington | 9.2 | 2.7 | 22.0 | 34.5 |
| Yamhill | 19.8 | 2.2 | 29.0 | 20.6 |
| Mean | 16.28 | 1.92 | 26.29 | 20.83 |
| Standard deviation | 3.81 | 0.83 | 4.24 | 8.19 |

Data analysis for assessing the relationship between cancer and predictors consisted of correlation and linear regression, performed in SAS 9.2 (Cary, NC). Normal distribution was considered present if skew values, obtained from spreadsheet analysis, were between -1.96 and $+1.96$ (Pett, 1997). Strengths of the correlation coefficients were interpreted as follows: <0.2 = negligible; 0.20 to 0.40 = low; 0.40 to 0.70 = moderate; 0.7 to 0.90 = high; >0.90 = very high (DHS, 2006). Correlation coefficients were considered to be statistically significant if their p-value was ≤ 0.05 . For linear regression, linearity was assessed in PASW 17 (SPSS, Chicago, IL) while homogeneity between variables was considered present if White's test (via the /spec procedure in SAS) was >0.05 (UCLA, 2009). Relative strength of prediction among predictors was assessed by noting individual predictor standardized estimates (also known as beta coefficients) and their p-values. For linear regression, two models were run; one with all three predictors ("Model 1") and another with an interaction between predictors smoking and radon ("Model 2").

RESULTS

Skew values indicated a distribution not significantly different statistically from normal. Linearity and homogeneity were found between cancer and predictors. Correlations with cancer mortality rates with the three predictors were as follows. Radon: -0.141 , $p = 0.5$; Smoking: 0.097 , $p = 0.6$; education: -0.683 , $p = 0.0002$. A scatterplot for the statistically significant correlation is provided in Figure 2. In multiple linear regression, Model 1 revealed statistical significance ($p = 0.003$) and an adjusted R-squared of 0.4175 (Table 2) while Model 2 revealed statistical non-significance ($p = 0.5$) and an adjusted R-squared of -0.0289 (Table 3). Standardized estimates and individual predictor p-values in Model 1 (lung cancer and the three predictors) were as follows: radon: -0.143 , $p = 0.3$; smoking: -0.110 , $p = 0.5$; education: -0.707 , $p = 0.0003$ (Table 2).

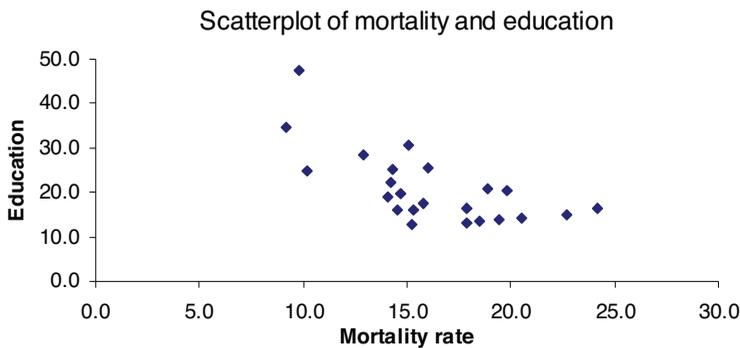


FIGURE 2. Scatterplot of correlation between mortality and education exhibiting a fairly strong linear relationship. Pearson correlation coefficient = -0.677 , $p = 0.0002$

TABLE 2. Multiple linear regression Model 1

| Predictor | Standardized parameter estimate | Predictor p-value | VIF |
|-----------|---------------------------------|-------------------|------|
| Radon | -0.143 | 0.387 | 1.04 |
| Smoking | -0.110 | 0.517 | 1.11 |
| Education | -0.707 | <0.001 | 1.06 |

VIF = variance inflation factor. Model p-value = 0.003. Model R-squared = 0.4935

TABLE 3. Linear regression Model 2

| Predictor | Predictor p-value |
|-------------|-------------------|
| Interaction | 0.5 |

Model p-value = 0.558. Model adjusted R-squared = -0.0289

DISCUSSION

Limitations of this study include: a) the uncertainty of radon and smoking exposure, as noted in the Methods section and b) this is an ecological study, where specific exposures to specific individuals is not known. Still, this study signals a need for further inquiry to verify or refute these findings. The most surprising findings of this inquiry consisted of the negligible relationships between lung cancer, smoking and radon. The negligible relationship with smoking certainly goes against conventional wisdom, and may be an only an anomaly considering the plethora of literature supporting the smoking – lung cancer link. It is possible that there is significant variation of smoking duration and amount, and this could account for the anomaly. In the case of radon however, the present study’s finding joins with existing literature, such as Cohen (2000) that challenges the notion that radon causes lung cancer.

The question of population mobility may be a factor. However, Cohen (2000) found that removing radon data for states with higher percentages of retired folks, who received their radon exposures elsewhere, did not “appreciably” change results.

Because this is an ecological study, cause-and-effect inferences are not possible (Grimes and Schulz, 2002). Determining cause-and-effect relationships would seem to be a daunting task even with the more rigorous case-control studies as establishing cause-and-effect relationships are difficult (Hill, 1965). Ecological studies can however be the spark for more rigorous research that seeks to make causal inferences (Grimes and Schulz, 2002).

CONCLUSION

In this ecological study, radon and smoking did not reveal statistically significant relationships with lung cancer. Level of educational attain-

ment showed the strongest relationship with lung cancer, having a moderate, inverse, and statistically significant correlation and statistically significant predictive power. Further study is indicated to verify or refute these findings.

REFERENCES

- Albano JD, Ward E, Jemal A, Anderson R, Cokkinides VE, Murray T, Henley J, Liff J, Thun MJ. 2007. Cancer mortality in the United States by education level and race. *Journal of the National Cancer Institute*. 99:1384-1394.
- Centers for Disease Control and Prevention (CDC). 2009. Lung cancer rates by race and ethnicity. [Cited 2010 Feb 10]. Available from: <http://www.cdc.gov>
- Cohen BL. 1998a. Response to Lubin's proposed explanations of our discrepancy. *Health Physics*. 75(1):18-22.
- Cohen BL. 1998b. Response to criticisms of Smith et al. *Health Physics*. 75(1):23-28
- Cohen B. 2000. Updates and extensions to tests of the linear-no threshold theory. *Technology*. 7:657-672.
- Field RW, Steck DJ, Smith BJ, Burs CP, Fisher EL, Neuberger JS, Platz CE, Robinson RA, Woolson RF, Lynch CF. 2000. Residential radon gas exposure and lung cancer. *American Journal of Epidemiology*. 151: 1091-1102.
- Grimes DA, Schulz KF. 2002. Descriptive studies: what they can and cannot do. *The Lancet*. 359: 145-149.
- Hill AB. 1965. The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*. 58: 295-300.
- Lubin JH. 1998. On the discrepancy between epidemiologic studies in individuals of lung cancer and residential radon and Cohen's ecologic regression. *Health Physics*. 75 (1): 4-10.
- National Cancer Institute (NCI). 2010a. Cancer risk: understanding the puzzle. [Cited 2010 Feb 10]. Available from: <http://understandingrisk.cancer.gov>
- National Cancer Institute (NCI). 2010b. Age-adjusted death rates for Oregon, 2002-2006. [Cited 2010 Feb 10]. Available from: <http://statecancerprofiles.cancer.gov>
- National Cancer Institute (NCI). 2009. SEER Stat Fact Sheets. Surveillance Epidemiology and End Results. National Cancer Institute. [Cited 2009 Oct 5]. Available from: www.seer.cancer.gov
- Oregon Health Division. 1999. Executive Summary. Oregon's Tobacco Prevention and Education Program. [Cited 2010 Feb 9]. Available from: <http://www.oregon.gov>
- Oregon Public Health Division. 1998. Radon levels listed by county. [Cited 2010 Feb 9]. Available from: <http://www.oregon.gov>
- Pett MA. 1997. *Nonparametric statistics for health care research*. Thousand Oaks, CA; Sage Publications.
- Puskin JS. 2003. Smoking as a confounding factor in ecological correlations of cancer mortality rates with average county radon levels. *Health Physics*. 84(4):526-532.
- Raabe OG. 2010. Concerning the Health Effects of Internally Deposited Radionuclides. *Health Physics*. 98(3):515-536.
- Samet JM. 1989. Radon and lung cancer. *Journal of the National Cancer Institute*. 81(10):745-758.
- Seiler FA, Alvarez JL. 2000. Is the "Ecological Fallacy" a Fallacy? *Human and Ecological Risk Assessment*. 6(6):921 - 941.
- Smith BJ, Field RW, Lynch CF. 1998. Residential 222Rn Exposure and Lung Cancer: Testing the Linear No-Threshold Theory with Ecologic Data. *Health Physics*. 75(1):11-17.
- UCLA: Academic Technology Services, Statistical Consulting Group. 2009. Regression with SAS. Available from: <http://www.ats.ucla.edu/stat/sas/webbooks/reg/chapter2/sasreg2.htm>
- USA Counties. 2010. U.S. Census Bureau. [Cited 2010 Feb 10]. Available from <http://censtats.census.gov>
- U.S. Department of Health and Human Services (DHS). 2006. Guilford's suggested interpretation for correlation coefficient values. (2006). [Internet]. [cited 2007 May 11]. Administration for Children and Families. Available from: www.acf.hhs.gov/programs/cb/pubs/cwo02/appendix/appendixG.htm
- United States Life Tables, 2004. 2007. *National Vital Statistics Report*. 56(9). Centers for Disease Control and Prevention. [Cited 2009 Oct 5]. Available from: www.cdc.gov