Published in final edited form as:

Environ Sci Technol. 2013 April 2; 47(7): 3410–3416. doi:10.1021/es305295d.

Association of Carcinogenic Polycyclic Aromatic Hydrocarbon Emissions and Smoking with Lung Cancer Mortality Rates on a Global Scale

Oleksii Motorykin¹, Melissa M. Matzke², Katrina M. Waters², and Staci L. Massey Simonich^{1,3,*}

¹Department of Chemistry, Oregon State University, Corvallis, Oregon USA 97331

²Computational Biology and Bioinformatics, Pacific Northwest National Laboratory, Richland, Washington USA 99352

³Environmental and Molecular Toxicology, Oregon State University, Corvallis, Oregon USA, 97331

Abstract

The objective of this research was to investigate the relationship between lung cancer mortality rates, carcinogenic polycyclic aromatic hydrocarbons (PAHs) emissions, and smoking on a global scale, as well as for different socioeconomic country groups. The estimated lung cancer deaths per 100,000 people (ED100000) and age standardized lung cancer death rate per 100,000 people (ASDR100000) in 2004 were regressed on PAH emissions in benzo[a]pyrene equivalence (BaPeq), smoking prevalence, cigarette price, gross domestic product per capita, percentage of people with diabetes, and average body mass index using simple and multiple linear regression for 136 countries. Using stepwise multiple linear regression, a statistically significant positive linear relationship was found between $\log_e(ED100000)$ and $\log_e(BaPeq)$ emissions for high (p-value<0.01) and for the combination of upper middle and high (p-value<0.05) socioeconomic country groups. A similar relationship was found between $\log_e(ASDR100000)$ and $\log_e(BaPeq)$ emissions for the combination of upper middle and high (p-value<0.01) socioeconomic country groups. Conversely, for $\log_e(ED100000)$ and $\log_e(ASDR100000)$, smoking prevalence was the only significant independent variable in the low socioeconomic country group (p-value<0.001). These results suggest that reducing BaPeq emissions in the U.S., Canada, Australia, France, Germany, Brazil, South Africa, Poland, Mexico, and Malaysia could reduce ED100000, while reducing smoking prevalence in Democratic People's Republic of Korea, Nepal, Mongolia, Cambodia, and Bangladesh could significantly reduce the ED100000 and ASDR100000.

Introduction

Cancer is the second leading cause of death worldwide [1]. While one in four people die from cancer in industrialized countries, one in eight die worldwide [1, 2]. Worldwide, twelve percent of all cancer cases are lung cancer and lung cancer is the first leading cause of cancer deaths in men and the second leading cause of cancer deaths in women [1, 3]. In addition, lung cancer incidence and mortality rates vary among different geographic regions

*Address correspondence to S.L.M. Simonich, 1141 Agricultural and Life Sciences Corvallis, OR 97331-7301 USA. Telephone: (541) 737-9194. Fax: (541) 737-0497. staci.simonich@orst.edu.

Supporting Information

Supporting information is available via <http://pubs.acs.com> and contains the table with all independent and dependent variables, interpolation method, and figures that were not included in main article.

of the world and ethnicities [2, 4]. For most populations, approximately 80% of all lung cancer cases are associated with tobacco smoking and the remaining 20% have been attributed to exposure to secondhand smoke, radon (and its decay products), asbestos, some metals (including arsenic, beryllium, cadmium), organic compounds (including carcinogenic polycyclic aromatic hydrocarbons (PAHs)), radiation, and genetic susceptibility [1, 2]. In addition, ambient air pollution has been estimated to account for 5% of male cancer deaths and 3% of female cancer deaths in the U.S. [5].

The primary anthropogenic sources of PAHs to the atmosphere include incomplete combustion of coal, wood, and other liquid and solid fuels [6]. Natural processes, including forest fires and volcanic eruptions, also contribute to global PAH emissions [7, 8]. Some PAHs are known to be mutagens [9, 10] and animal carcinogens [6, 11] and are ranked as probable human carcinogens by the U.S. Environmental Protection Agency Integrated Risk Information System [12] and by the International Agency for Research on Cancer (IARC) (1983). In 2010, IARC concluded that indoor emissions from household combustion of coal are carcinogenic to humans and classified biomass burning emissions as probable carcinogens [13].

Cigarette smoking contributes the most to lung cancer occurrence and mortality [1]. Farrelly et al [14] found that increasing the cigarette price led to a decrease in the number of cigarettes smoked daily. Other studies have shown that obesity [15, 16] and diabetes [17–22] may affect lung cancer development and mortality. Gross Domestic Product (GDP) and GDP per capita were found to positively correlate with per capita health care expenditure and life expectancy [23, 24]. In addition, Ultraviolet B (UVB) irradiance has been shown to play a protective role in the development of some cancers, including lung cancer [25, 26].

Zhang et al [27] found that lung cancer mortality was highly correlated with ambient carcinogenic PAH exposure in China, while Grant [5] found a similar relationship for the U.S. The emission of particulate matter less than 2.5 μm in diameter, on which most carcinogenic PAH are sorbed, can also lead to an increased risk of lung cancer [5, 28]. Carcinogenic PAH concentrations are often reported in benzo[a]pyrene equivalents (BaP_{eq}) because benzo[a]pyrene (BaP) is a well-studied PAH that is known to be carcinogenic. BaP_{eq} is a measure of the concentration of carcinogenic PAHs in a mixture relative to pure BaP [29].

The objective of this research was to investigate the relationship between lung cancer mortality rates, carcinogenic PAH emissions, and smoking prevalence on a global scale, as well as for different socioeconomic country groups. Multiple linear regression (MLR) models were built to determine the influence of BaP_{eq} and smoking prevalence on lung cancer mortality rates after taking into account other factors that may contribute to lung cancer mortality rates. The estimated lung cancer deaths per 100,000 people in 2004 (ED100000) and age standardized lung cancer mortality rate per 100,000 people in 2004 (ASDR100000) were regressed on smoking prevalence, cigarette price, gross domestic product (GDP) per capita, the percentage of people with diabetes, average body mass index (BMI) and carcinogenic PAH emissions (in BaP_{eq}) using MLR and data for 136 different countries. The contribution of carcinogenic PAH emissions and smoking prevalence on lung cancer mortality rates was assessed for countries in different socioeconomic country groups. This is the first study to link carcinogenic PAH emissions, as a measure of air pollution, with lung cancer mortality rates on a global scale.

Experimental Section

Variables and Data Used

The most recent lung cancer mortality rate (LCMR) data available, including ED100000 and ASDR100000, were taken from the World Health Organization for 2004 [30]. The ASDR100000 is a weighted average of the age-specific lung cancer mortality rates per 100,000 persons, where the weights are the proportion of persons in the corresponding age groups of the WHO standard population [31]. In 2004, 1.28 million people died of lung cancer and, for the 136 countries we had entire datasets for, the highest ASDR100000 were in Maldives (67 deaths per 100,000 people), Hungary (52 deaths per 100,000 people) and Poland (42 deaths per 100,000 people) (Table S1). The U.S. ranked fifth with 39 deaths per 100,000 people (Table S1). The lowest rates were in Fiji (1 death per 100,000 people), Federated States of Micronesia (2.4 deaths per 100,000 people), and Mozambique (3.1 deaths per 100,000 people) (Table S1).

Population and GDP data for 2004 were obtained from Zhang and Tao [32] (Table S1). The total population for the 136 countries modeled was 5.93 billion people, with the highest populations in China (1.33 billion), India (1.08 billion), and the United States (0.29 billion). To account for the socioeconomic level of the country, GDP per capita was used to divide the countries into four groups recommended by the World Bank in 2004 [33]: 1) low income: \$825 USD; 2) low middle income: \$826 to \$3,255 USD; 3) upper middle income: \$3256 to \$10,065 USD; and 4) high income: more than \$10,065 USD (Table S1). We also tested the combinations of low and low middle income countries, as well as upper middle and high income countries.

The most recent smoking prevalence and the cigarette price per pack data available, and closest to 2004, were compiled for all 136 countries [34]. The available smoking prevalence data was from 1997 to 2008, while the closest available cigarette price data to 2004 was from 2008 [34]. Cigarette price per pack does not accurately represent the price per cigarette because the number of cigarettes per pack (from 10 to 50) varies for different regions of the world [35, 36]. The highest smoking prevalence was in Nauru (47%), Russia (44%), and Austria (41%), while the lowest smoking prevalence was in Suriname (1%), Ethiopia (3%), and Ghana (4%) (Table S1). The highest cigarette prices were in Iceland (\$11.27 USD per pack), Norway (\$10.14 USD per pack), and Singapore (\$8.06 USD per pack), while the lowest cigarette prices were in Democratic People's Republic of Korea (\$0.14 USD per pack), Paraguay (\$0.20 USD per pack), and Pakistan (\$0.23 USD per pack) (Table S1).

Data on the percentage of people with diabetes (for 2000 and 2006) and average BMI (for 2000 and estimated for 2030), were taken from WHO [37, 38]. These data were interpolated to 2004 for each country (see Supporting Information). BMI is a relative measure of human body fat based on an individual's weight and height, and is equal to body mass (kg) divided by body height squared (m^2). The highest percentages of people with diabetes in 2004 were in Nauru (17.4%), Seychelles (11.1%), and Malta (10.3%), while the lowest percentages were in Uganda (0.5%), Nigeria (0.1%), and the Democratic Republic of the Congo (0.03%) (Table S1). The highest BMIs in 2004 were in Nauru (33.6), Cook Island (32.3), and Tonga (31.9), while the lowest BMIs were in the Democratic Republic of the Congo (20.5), Ethiopia (20.18), and Bangladesh (20.13) (Table S1).

The carcinogenic PAH emissions (in BaPeq) data for the 136 countries in 2004 were taken from Zhang and Tao [32] and were calculated using the BaPeq emission factors for different fuel types, in different countries. The highest BaPeq emission rates were in China (110,000 tons/year), India (90,000 tons/year), and United States (32,000 tons/year), while the lowest

emission rates were in Dominica (0.74 tons/year), St. Vincent and the Grenadines (1.1 tons/year), and Nauru (1.3 tons/year) (Table S1).

Although there is a 20 year lag time for developing measurable cancer [39], we did not account for this lag time in the independent variable datasets because the dependent variables (ED100000 and ASDR100000) are rates. By doing this, we assumed that the ED100000 and ASDR100000 rates would not change significantly over this 20 year period. In addition, the independent variables used in the models were not available from 20 years ago for the 136 countries.

Statistical Analysis and Models Tested

The ED100000 and ASDR100000 were independently modeled using a stepwise multiplicative multiple linear regression model to determine the statistically significant association between lung cancer mortality rates, carcinogenic PAH emissions (BaPeq), and smoking prevalence after accounting for other potentially influential factors:

$$LCMR = BaPeq^{\beta_1} * SP^{\beta_2} * Price^{\beta_3} * GDP.CAP^{\beta_4} * BMI^{\beta_5} * Diabetes^{\beta_6} \quad [1]$$

where LCMR was either ED100000 or ASDR100000, *BaPeq* was the carcinogenic PAH emissions in BaPeq (Mt/year), *SP* was the smoking prevalence (%), *Price* was the cigarette price (\$USD per pack), *GDP.CAP* was GDP per capita (\$US'000), *BMI* was the average body mass index (kg/m²), *Diabetes* was the percentage of people with diabetes, and $\beta_0 \dots \beta_6$ were coefficients in the model. MLR models were independently generated for each of the four socioeconomic groups, as well as the combination of low and low middle country groups and upper middle and high country groups (Table S4).

To investigate the percent change in LCMR as a function of a percent change in a given independent variable, equation [2] was used:

$$\% \text{ change in LCMR} = \left(\left(1 + \frac{z}{100} \right)^{\beta_x} - 1 \right) * 100\% \quad [2]$$

where β_x is the variable's coefficient estimate and *Z* is the percent increase in the variable.

Standard regression model checking and refinement techniques, including residual plots, Q-Q plot, Cook's distance and stepwise variable selection using Akaike Information Criteria (AIC) [40], were used. Effects were deemed statistically significant for a p-value < 0.05. The statistical package R v. 2.12.0 (Free Software Foundation, Inc., Boston, MA) was used for all modeling.

Results

An initial graphical review of the data, including LCMRs and independent variables, revealed the need for a logarithmic transformation (see Supporting Information, Figures S1 and Figure S2). Specifically, these data tend to be highly skewed. Data from the LCMRs and independent variables were transformed to the log_e scale to resolve this issue and allowed the data to be modeled linearly, rather than multiplicatively. The resulting MLR model is written as:

$$\log_e(LCMR) = \beta_0 + \beta_1 * \log_e(BaPeq) + \beta_2 * \log_e(SP) + \beta_3 * \log_e(Price) + \beta_4 * \log_e(GDP.CAP) + \beta_5 * \log_e(BMI) + \beta_6 * \log_e(Diabetes) \quad [3]$$

where *LCMR* is either ED100000 or ASDR100000 and the independent variables are consistent with previous descriptions.

Co-linearity among the $\log_e$ independent variables, for the entire dataset, was explored (Figure S2 and Table S2). There were statistically significant linear relationships (p -value<0.05) between $\log_e$ (SP), $\log_e$ (Diabetes), $\log_e$ (Price), $\log_e$ (BMI), and $\log_e$ (GDP.CAP). The PAH emission variable, $\log_e$ (BaPeq), had a statistically significant negative linear relationship with $\log_e$ (Diabetes) ($r^2 = 0.09$), $\log_e$ (BMI) ($r^2 = 0.22$), and $\log_e$ (Price) ($r^2 = 0.05$) (Figure S2 and Table S2).

The linear relationships among the $\log_e$ LCMR and $\log_e$ independent variables for the entire dataset were explored using simple linear regression (SLR) (Figure S2 and Table S3). Linear relationships were further explored by individual socioeconomic country group (low, low-middle, upper-middle, and high), as well as for the combination of low and low middle country groups and upper middle and high country groups using SLR (Figure 1, Table S3, Figures S3–S13). Table S3 shows the regression coefficients, standard error, and percent of the total regression sum-of-squares due to β_n for the SLRs. Table 1 shows the percent change in the median LCMR, given a 10% increase in the mean of the independent variable from the SLRs, for the entire dataset, as well as the different socioeconomic country groups and groupings.

The SLR analysis for the socioeconomic country group subsets showed that smoking prevalence ($\log_e$ (SP)) was significantly positively related to the LCMRs in the low, low-middle and upper-middle country groups, as well as for the combination of low and low middle country groups, upper middle and high country groups, and the entire dataset (Table 1, Figures S3 and S4). The price of a pack of cigarettes ($\log_e$ (Price)) was significantly negatively related to the LCMRs for the low (both ED100000 and ASDR100000) and low-middle (ED100000 only) socioeconomic country groups (Figures S5 and S6). However, the price of a pack of cigarettes ($\log_e$ (Price)) was significantly positively related to both ED100000 and ASDR100000 for the combination of upper middle and high socioeconomic country groups and the entire dataset (Table S3). The GDP per capita ($\log_e$ (GDP.CAP)) and percentage of people with diabetes ($\log_e$ (Diabetes)) were significantly positively related to the LCMRs for only the low socioeconomic country group (Figures S7–S10). The average BMI ($\log_e$ (BMI)) was significantly negatively and positively related to $\log_e$ (ED100000) and $\log_e$ (ASDR100000) for different socioeconomic country groups and groupings (Figures S11 and S12). The BaP equivalents emission ($\log_e$ (BaPeq)) was significantly positively related to $\log_e$ (ED100000) for the upper-middle and high socioeconomic country groups, as well as the combination of upper middle and high socioeconomic country groups (Table 1 and Figure 1). However, $\log_e$ (BaPeq) was significantly positively related to $\log_e$ (ASDR100000) only for the upper-middle socioeconomic country group and the combination of upper middle and high socioeconomic country groups (Table 1 and Figure S13).

The relationships among the LCMRs and the independent variables were modeled using equation [3] and stepwise multiple linear regression (MLR) for the entire dataset, as well as the different socioeconomic country groups and groupings. Table 2 shows the percent change in the median LCMR, given a 10% increase in the mean of the independent variable in the MLR models, for the entire dataset, as well as the different socioeconomic country groups and groupings. Table S4 shows the regression coefficients, standard error, and percent of the total regression sum-of-squares due to β_n for the MLR models.

The stepwise procedure primarily selected the smoking prevalence ($\log_e$ (SP), positive relationship) and body mass index ($\log_e$ (BMI), negative relationship) variables as the most predictive of the LCMRs for the entire dataset and the various socioeconomic country

groups and groupings. However, $\log_e(BMI)$ was not selected by the stepwise procedure for the low and high socioeconomic country group MLR models. The BaP equivalents emission ($\log_e(BaP_{eq})$) was selected by the stepwise procedure for ED100000 for the high socioeconomic country group and the combination of the upper middle and high socioeconomic country groups in the MLR (positive relationship). However, the BaP equivalents emission ($\log_e(BaP_{eq})$) was selected by the stepwise procedure for ASDR100000 only for the combination of the upper middle and high socioeconomic country groups in the MLR (positive relationship). The entire dataset MLR models also included the average price of a pack of cigarettes (for ASDR100000), and GDP per capita (for ED100000 and ASDR100000). Price of cigarettes ($\log_e(Price)$) was significantly positively correlated with smoking prevalence ($\log_e(SP)$) only for the low socioeconomic country group (Figure S14). The variance inflation factors test was used to evaluate cross-correlation between $\log_e(Price)$ and $\log_e(SP)$ in the models where both were statistically significant. The results showed that there was no case of cross-correlation between these independent variables in any of the models. Table S4 shows the percent of total sum-of-squares due to each of the variables in the models.

Discussion

Smoking prevalence and body mass index have been shown to be contributing factors to lung cancer mortality rates [1, 15, 16]. In this study, smoking prevalence was shown to be a significant predictor of the lung cancer mortality rates, with the exception of high socioeconomic countries status for the age standardized lung cancer death rate per 100,000 people. The lack of correlation for these countries may be explained by smoking habits in those individuals with higher incomes. Townsend et al. [41] found that smokers with high incomes tend to smoke less than smokers with lower incomes. Additionally, over the past few decades, smoking prevalence among people with high incomes has decreased significantly [42–45] and the quit rates are higher within the highest socioeconomic group [46]. The smoking prevalence data used in this study does not account for the frequency of smoking. It would be interesting to evaluate the association of carcinogenic PAH emissions and smoking with male and female lung cancer mortality rates separately. However, this could not be done because the lung cancer mortality rates, and independent variables, are not reported separately for males and females.

The negative relationship between body mass index and lung cancer mortality rates, after accounting for the other independent factors, is consistent with previous studies that found reduced lung cancer mortality with increasing BMI (especially when $BMI > 28$) [15, 16, 47, 48]. Figure S15 shows $\log_e(BMI)$ by socioeconomic country group. The highest median BMI was for upper-middle socioeconomic country group, followed by the high and low-middle socioeconomic country groups. Only the low-middle and upper-middle socioeconomic country groups had individual countries with average BMI's greater than 28. This may explain the significant negative correlation of $\log_e(ED100000)$ and $\log_e(ASDR100000)$ with $\log_e(BMI)$ for these groups. Another possible explanation for this negative relationship is that smokers tend to have lower BMI than non-smokers [49]. Additionally, decreased levels of vitamin A and carotene (which may be protective against lung cancer) were observed in lean men compared to obese men [50]. Other confounding factors, not accounted for in this study, may also play a role in this negative relationship. Further medical research is necessary to understand the relationship between BMI and lung cancer.

In addition to these prominent factors, PAH emissions, as measured by BaP equivalents, were significantly positively correlated with the estimated lung cancer deaths per 100,000 people in high socioeconomic countries, as well as with both ED100000 and ASDR100000

for the combination of upper-middle and high socioeconomic country groups. When inhaled, PAHs diffuse into lung cells and bind to the aryl hydrocarbon receptor [51]. After metabolic reactions with CYP-enzymes (P450, CYP1A1, CYP1B1, and CYP3A4) and microsomal epoxide hydrolase (EPHX1), PAHs form DNA-adducts [51, 52]. DNA-adducts have been shown to correlate with exposure to PAHs via smoking [51, 53] or exposure to environmental pollution [52, 54, 55]. DNA-PAH adducts may initiate carcinogenesis [56, 57] and their levels may be predictive of cancer risk [55, 56].

This suggests that reducing PAH emissions in high socioeconomic countries and with high PAH emissions, including the U.S., Canada, Australia, France and Germany, could reduce the estimated lung cancer deaths per 100,000 people. In addition, there is strong evidence of the same trend in upper-middle socioeconomic countries with high PAH emissions, including Brazil, South Africa, Poland, Mexico, and Malaysia (Figure 1). Conversely, smoking prevalence was the only significant independent variable in the low socioeconomic country group for the estimated lung cancer deaths per 100,000 people and age standardized lung cancer mortality rate per 100,000 people. This suggests that reducing smoking prevalence in countries in low socioeconomic countries and with high smoking prevalence, including Democratic People's Republic of Korea, Nepal, Mongolia, Cambodia, and Bangladesh, could significantly reduce the estimated lung cancer deaths per 100,000 people, as well as the age standardized lung cancer mortality rate per 100,000 people.

The lack of a correlation between estimated lung cancer deaths per 100,000 people and age standardized lung cancer mortality rate per 100,000 people with PAH emissions for the low and low-middle socioeconomic country groups may be due to the different life expectancy for these two socioeconomic country groups, compared to the upper-middle and high socioeconomic country groups. In 2004, the mean life expectancy for the low, low-middle, upper-middle and high socioeconomic country groups was 53.9, 69.3, 72.1, and 79.0 years, respectively (Figure S16) [58]. Lung cancer incidence and mortality increase with age [59]. For a cohort of 22,874 people with lung cancer, 42.8% were 65 years old, or older, at diagnosis [59]. In addition, the development of clinically detectable cancer takes more than 20 years [39]. Therefore, people with a relatively shorter life expectancy may not develop detectable lung cancer in their lifetimes. If the population in all of the socioeconomic country groups had high life expectancy, there might be a correlation between estimated lung cancer deaths per 100,000 people and age standardized lung cancer mortality rate per 100,000 people with PAH emissions for all of the different socioeconomic country groups. In the MLR, after accounting for smoking prevalence, $\log_e(\text{BaPeq})$ remained significantly correlated with $\log_e(\text{ED100000})$ for the high socioeconomic country group and with both $\log_e(\text{ED100000})$ and $\log_e(\text{ASDR100000})$ for the combination of upper-middle and high socioeconomic country groups (Table 2 and Table S4). These results suggest BaPeq emissions influence lung cancer mortality rates in upper-middle and high socioeconomic country groups more than in the other socioeconomic country groups. Although this study found an association between PAH emissions and lung cancer mortality rates, no causal relationship can be proven and other factors (including other air pollutants) cannot be ruled out. Finally, future studies should test for the potential association between average UVB irradiance for a given country and lung cancer mortality rates as this data becomes available [25, 26].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This publication was made possible in part by grant number P30ES00210 from the National Institute of Environmental Health Sciences (NIEHS), NIH and NIEHS Grant P42 ES016465. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the NIEHS, NIH. We thank John Molitor of Oregon State University for helpful suggestions.

References

1. Garcia, M.; Jemal, A.; Ward, E.; Center, M.; Hao, Y.; Siegel, R.; Thun, M. Global Cancer Facts & Figures 2007. American Cancer Society; 2007.
2. WHO. The world health report 2008: primary health care now more than ever. World Health Organization; 2008.
3. Jemal A, Murray T, Ward E, Samuels A, Tiwari R, Ghafoor A, Feuer E, Thun M. Cancer statistics 2005. *CA Cancer J Clin.* 2005; 55(10–30)
4. Thun MJ, Hannan LM, Adams-Campbell LL, Boffetta P, Buring JE, Feskanich D, Flanders WD, Jee SH, Katanoda K, Kolonel LN, Lee IM, Marugame T, Palmer JR, Riboli E, Sobue T, Avila-Tang E, Wilkens LR, Samet JM. Lung Cancer Occurrence in Never-Smokers: An Analysis of 13 Cohorts and 22 Cancer Registry Studies. *PLoS Med.* 2008; 5(9):1357–1371.
5. Grant WB. Air Pollution in Relation to US Cancer Mortality Rates: An Ecological Study; Likely Role of Carbonaceous Aerosols and Polycyclic Aromatic Hydrocarbons. *Anticancer Res.* 2009; 29(9):3537–3545. [PubMed: 19667146]
6. Tonne CC, Whyatt RM, Camann DE, Perera FP, Kinney PL. Predictors of personal polycyclic aromatic hydrocarbon exposures among pregnant minority women in New York City. *Environ Health Perspect.* 2004; 112(6):754–759. [PubMed: 15121521]
7. Baek SO, Field RA, Goldstone ME, Kirk PW, Lester JN, Perry R. A review of atmospheric polycyclic aromatic-hydrocarbons - sources, fate and behavior. *Water Air Soil Pollut.* 1991; 60(3–4):279–300.
8. Xu SS, Liu WX, Tao S. Emission of polycyclic aromatic hydrocarbons in China. *Environ Sci Technol.* 2006; 40(3):702–708.
9. Bostrom CE, Gerde P, Hanberg A, Jernstrom B, Johansson C, Kyrklund T, Rannug A, Tornqvist M, Victorin K, Westerholm R. Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environ Health Perspect.* 2002; 110:451–488. [PubMed: 12060843]
10. Hainaut P, Hollstein M. p53 and human cancer: The first ten thousand mutations. *Adv Cancer Res.* 2000; 77:81–137. [PubMed: 10549356]
11. Neff, JM. Polycyclic aromatic hydrocarbons in the aquatic environment and cancer risk to aquatic organisms and man. Symposium: Carcinogenic Polynuclear Aromatic Hydrocarbons in the Marine Environment; 1978. p. 385–409.
12. EPA. U S Integrated Risk Information System (IRIS). Jul 25. http://cfpub.epa.gov/ncea/iris/index.cfm?fuseaction=iris.showSubstanceList&list_type=alpha&view=B
13. WHO; IARC . IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. World Health Organization; 2010.
14. Farrelly MC, Nimsch CT, Hyland A, Cummings M. The effects of higher cigarette prices on tar and nicotine consumption in a cohort of adult smokers. *Health Econ.* 2004; 13(1):49–58. [PubMed: 14724893]
15. Kollarova H, Machova L, Horakova D, Cizek L, Janoutova G, Janout V. Is obesity a preventive factor for lung cancer? *Neoplasma.* 2008; 55(1):71–73. [PubMed: 18190245]
16. Leung CC, Lam TH, Yew WW, Chan WM, Law WS, Tam CM. Lower lung cancer mortality in obesity. *Int J Epidemiol.* 2011; 40(1):174–182. [PubMed: 20709687]
17. Hack KM, Meza JL, Kessinger MA, Ganti AK. Effect of diabetes mellitus on the epidemiology of lung cancer. *J Thorac Oncol.* 2009; 4(9):S429–S429.
18. Hatlen P, Gronberg BH, Langhammer A, Carlsen SM, Amundsen T. Prolonged Survival in Patients with Lung Cancer with Diabetes Mellitus. *J Thorac Oncol.* 2011; 6(11):1810–1817. [PubMed: 21964531]

19. Turan O, Akkoclu A. May diabetes mellitus affect the development of lung cancer? *J Thorac Oncol.* 2009; 4(9):S652–S653.
20. Lee MY, Lin KD, Hsiao PJ, Shin SJ. The association of diabetes mellitus with liver, colon, lung, and prostate cancer is independent of hypertension, hyperlipidemia, and gout in Taiwanese patients. *Metab-Clin Exp.* 2012; 61(2):242–249. [PubMed: 21820134]
21. Luo JH, Chlebowski R, Wactawski-Wende J, Schlecht NF, Tinker L, Margolis KL. Diabetes and Lung Cancer Among Postmenopausal Women. *Diabetes Care.* 2012; 35(7):1485–1491. [PubMed: 22619084]
22. Shieh SH, Probst JC, Sung FC, Tsai WC, Li YS, Chen CY. Decreased survival among lung cancer patients with co-morbid tuberculosis and diabetes. *BMC Cancer.* 2012; 12:174. [PubMed: 22578056]
23. Chen MY, Lin FL, Chang CK. Relations between health care expenditure and income: an application of local quantile regressions. *Appl Econ Lett.* 2009; 16(2):177–181.
24. Swift R. The relationship between health and gdp in oecd countries in the very long run. *Health Econ.* 2011; 20(3):306–322. [PubMed: 20217835]
25. Chen WQ, Clements M, Rahman B, Zhang SW, Qiao YL, Armstrong BK. Relationship between cancer mortality/incidence and ambient ultraviolet B irradiance in China. *Cancer Causes & Control.* 2010; 21(10):1701–1709. [PubMed: 20552265]
26. Mohr SB, Garland CF, Gorham ED, Grant WB, Garland FC. Could ultraviolet B irradiance and vitamin D be associated with lower incidence rates of lung cancer? *J Epidemiol Community Health.* 2008; 62(1):69–74. [PubMed: 18079336]
27. Zhang YX, Tao S, Shen HZ, Ma JM. Inhalation exposure to ambient polycyclic aromatic hydrocarbons and lung cancer risk of Chinese population. *Proc Natl Acad Sci U S A.* 2009; 106(50):21063–21067. [PubMed: 19995969]
28. Lewtas J. Air pollution combustion emissions: Characterization of causative agents and mechanisms associated with cancer, reproductive, and cardiovascular effects. *Mutat Res-Rev Mutat Res.* 2007; 636(1–3):95–133.
29. Jung KH, Yan BZ, Chillrud SN, Perera FP, Whyatt R, Camann D, Kinney PL, Miller RL. Assessment of Benzo(a)pyrene-equivalent Carcinogenicity and Mutagenicity of Residential Indoor versus Outdoor Polycyclic Aromatic Hydrocarbons Exposing Young Children in New York City. *Int J Environ Res Public Health.* 2010; 7(5):1889–1900. [PubMed: 20622999]
30. WHO. The global burden of disease: 2004 update. World Health Organization; Geneva: 2008.
31. WHO. Age-standardized death rates per 100,000 by cause. Aug 21. <http://www.who.int/whosis/whostat2006AgeStandardizedDeathRates.pdf>
32. Zhang YX, Tao S. Global atmospheric emission inventory of polycyclic aromatic hydrocarbons (PAHs) for 2004. *Atmos Environ.* 2009; 43(4):812–819.
33. World Bank. World Bank Analytical Classifications. Feb 03. <http://data.worldbank.org/about/country-classifications/a-short-history>
34. WHO. Report on the Global Tobacco Epidemic, 2009: Implementing smoke-free environments. World Health Organization; 2009.
35. Franks P, Jerant AF, Leigh JP, Lee D, Chiem A, Lewis I, Lee S. Cigarette prices, smoking, and the poor: Implications of recent trends. *Am J Public Health.* 2007; 97(10):1873–1877. [PubMed: 17761576]
36. Kozlowski LT, Heatherton TF, Ferrence RG. Pack size, reported cigarette-smoking rates, and the heaviness of smoking. *Can J Public Health-Rev Can Sante Publ.* 1989; 80(4):266–270.
37. WHO. Global Database on Body Mass Index. An interactive surveillance tool for monitoring nutrition transition. Dec 5. <http://apps.who.int/bmi/index.jsp>
38. WHO. Country and regional data on diabetes. Dec 5. http://www.who.int/diabetes/facts/world_figures/en/
39. Loeb LA. A mutator phenotype in cancer. *Cancer Res.* 2001; 61(8):3230–3239. [PubMed: 11309271]
40. Akaike H. Citation classic - a new look at the statistical-model identification. *Current Contents/Engineering Technology & Applied Sciences.* 1981; (51):22–22.

41. Townsend J, Roderick P, Cooper J. Cigarette-smoking by socioeconomic group, sex, and age - effects of price, income, and health publicity. *Br Med J*. 1994; 309(6959):923–927. [PubMed: 7950662]
42. Etter JF. Smoking prevalence, cigarette consumption and advice received from physicians: Change between 1996 and 2006 in Geneva, Switzerland. *Addict Behav*. 2010; 35(4):355–358. [PubMed: 19919891]
43. Filion KB, Steffen LM, Duval S, Jacobs DR, Blackburn H, Luepker RV. Trends in Smoking Among Adults From 1980 to 2009: The Minnesota Heart Survey. *Am J Public Health*. 2012; 102(4):705–713. [PubMed: 21852651]
44. Fukuda Y, Nakamura K, Takano T. Socioeconomic pattern of smoking in Japan: Income inequality and gender and age differences. *Ann Epidemiol*. 2005; 15(5):365–372. [PubMed: 15840550]
45. Nagelhout GE, de Boer DD, Kunst AE, van der Meer RM, de Vries H, van Gelder BM, Willemssen MC. Trends in socioeconomic inequalities in smoking prevalence, consumption, initiation, and cessation between 2001 and 2008 in the Netherlands. Findings from a national population survey. *BMC Public Health*. 2012; 12
46. Whitlock G, VanderHoorn S, Davis P, Jackson R, Norton R. Socioeconomic distribution of smoking in a population of 10 529 New Zealanders. *N Z Med J*. 1997; 110(1051):327–330. [PubMed: 9323370]
47. Kabat GC, Wynder EL. Body-mass index and lung-cancer risk. *Am J Epidemiol*. 1992; 135(7):769–774. [PubMed: 1595676]
48. Knekt P, Heliovaara M, Rissanen A, Aromaa A, Seppanen R, Teppo L, Pukkala E. Leanness and lung-cancer risk. *Int J Cancer*. 1991; 49(2):208–213. [PubMed: 1879966]
49. Albanes D, Jones DY, Micozzi MS, Mattson ME. Associations between smoking and body-weight in the united-states population - analysis of NHANES-II. *Am J Public Health*. 1987; 77(4):439–444. [PubMed: 3493709]
50. Garn SM, Rosenberg KR, Schaefer AE. Relationship between fatness level and size attainment in central america. *Ecol Food Nutr*. 1983; 13(3):157–165.
51. Mollerup S, Berge G, Baera R, Skaug V, Hewer A, Phillips DH, Stangeland L, Haugen A. Sex differences in risk of lung cancer: Expression of genes in the PAH bioactivation pathway in relation to smoking and bulky DNA adducts. *Int J Cancer*. 2006; 119(4):741–744. [PubMed: 16557573]
52. Billet S, Abbas I, Le Goff J, Verdin A, Andre V, Lafargue PE, Hachimi A, Cazier F, Sichel F, Shirali P, Garcon G. Genotoxic potential of Polycyclic Aromatic Hydrocarbons-coated onto airborne Particulate Matter (PM(2.5)) in human lung epithelial a549 cells. *Cancer Lett*. 2008; 270(1):144–155. [PubMed: 18554780]
53. Phillips DH, Hewer A, Martin CN, Garner RC, King MM. Correlation of dna adduct levels in human-lung with cigarette-smoking. *Nature*. 1988; 336(6201):790–792. [PubMed: 3205307]
54. Kriek E, Vanschooten FJ, Hillebrand MJX, Vanleeuwen FE, Denengelse L, Delooff AJA, Dijkmans APG. DNA adducts as a measure of lung-cancer risk in humans exposed to polycyclic aromatic-hydrocarbons. *Environ Health Perspect*. 1993; 99:71–75. [PubMed: 8319662]
55. Shields PG, Bowman ED, Harrington AM, Doan VT, Weston A. Polycyclic aromatic hydrocarbon-dna adducts in human lung and cancer susceptibility genes. *Cancer Res*. 1993; 53(15):3486–3492. [PubMed: 8339251]
56. Veglia F, Matullo G, Vineis P. Bulky DNA adducts and risk of cancer: A meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2003; 12(2):157–160. [PubMed: 12582026]
57. Schwartz AG, Prysak GM, Bock CH, Cote ML. The molecular epidemiology of lung cancer. *Carcinogenesis*. 2007; 28(3):507–518. [PubMed: 17183062]
58. CIA . The World Factbook. Mar 29. <https://www.cia.gov/library/publications/the-world-factbook/>
59. Orourke MA, Feussner JR, Feigl P, Laszlo J. Age trends of lung-cancer stage at diagnosis - implications for lung-cancer screening in the elderly. *JAMA-J Am Med Assoc*. 1987; 258(7):921–926.

Figure 1.

Scatter plot between lung cancer mortality rate (ED100000) and BaPeq . Significance of estimates: * p-value<0.05; **p-value<0.01, *** p-value <0.001.

Table 1

The associated change in median LCMR (%), given a 10% increase in mean independent variable, in the simple linear regression. The 95% confidence interval is given in parenthesis.

LCMR	Socioeconomic Group	BaPeq	SP	Price	GDP.CAP	BMI	Diabetes
ED100000	Low	-1.1 (-2.3, 0.17)	9.8 *** (5.9, 13.8)	-6.4 *** (-9.8, -2.8)	8.0 ** (2.7, 13.7)	27.4 (-8.9, 78.3)	4.1 ** (1.5, 6.7)
	Low Middle	1.3 (-0.11, 2.7)	6.1 ** (1.9, 10.4)	-4.4 * (-8.4, -0.28)	0.16 (-9.1, 10.3)	-30.2 (-53.5, 4.7)	7.1 (-0.71, 15.4)
	Upper Middle	1.0 * (0.01, 2.0)	9.6 *** (5.4, 14.0)	-0.22 (-7.1, 7.2)	-1.9 (-12.1, 9.6)	-37.7 * (-60.5, -1.8)	-1.7 (-8.1, 5.1)
	High	0.77 * (0.14, 1.4)	4.2 (-0.13, 8.8)	-0.99 (-4.6, 2.8)	0.58 (-2.2, 3.4)	-4.0 (-21.9, 18.0)	-0.69 (-3.4, 2.1)
	Low + Low middle	-0.31 (-1.4, 0.75)	9.0 *** (6.1, 12.0)	-2.0 (-5.1, 1.2)	6.3 *** (3.8, 8.8)	29.7 * (3.2, 63.0)	6.5 *** (4.0, 9.0)
	Upper Middle + High	1.55 *** (0.73, 2.4)	11.8 *** (7.9, 15.9)	8.0 *** (4.2, 11.9)	4.7 ** (2.0, 7.6)	-50.0 *** (-63.7, -31.0)	-3.6 (-8.6, 1.6)
	World	0.27 (-0.50, 1.0)	11.3 *** (8.7, 13.9)	4.8 *** (2.8, 6.9)	4.7 *** (3.7, 5.7)	30.6 *** (7.8, 58.2)	6.9 *** (4.7, 9.2)
	Low	-0.98 (-2.0, 0.01)	7.3 *** (4.1, 10.6)	-4.8 ** (-7.7, -1.9)	6.0 ** (1.6, 10.5)	20.7 (-8.0, 58.0)	3.0 ** (0.94, 5.1)
ASDR100000	Low Middle	0.76 (-0.33, 1.9)	4.0 * (0.74, 7.3)	-2.7 (-5.9, 0.60)	0.85 (-6.3, 8.5)	-27.0 * (-46.2, -1.0)	2.4 (-3.6, 8.6)
	Upper Middle	0.76 * (0.18, 1.4)	5.3 *** (2.6, 8.0)	-1.4 (-5.7, 2.9)	-3.2 (-9.4, 3.5)	-20.2 (-40.2, 6.5)	-1.4 (-5.3, 2.8)
	High	0.25 (-0.32, 0.82)	1.7 (-2.0, 5.6)	0.63 (-2.4, 3.8)	0.44 (-1.8, 2.8)	5.5 (-10.9, 24.9)	-0.54 (-2.8, 1.7)
	Low + Low middle	-0.29 (-1.1, 0.48)	5.8 *** (3.6, 8.0)	-1.7 (-4.0, 0.62)	3.6 *** (1.7, 5.5)	12.3 (-5.3, 33.0)	3.7 *** (1.8, 5.6)
	Upper Middle + High	0.91 *** (0.48, 1.3)	5.4 *** (3.2, 7.6)	2.9 *** (0.82, 5.1)	2.1 ** (0.62, 3.6)	-24.3 ** (-37.0, -9.0)	-1.7 (-4.5, 1.1)
	World	0.14 (-0.35, 0.62)	6.3 *** (4.7, 8.0)	2.0 *** (0.66, 3.3)	2.4 *** (1.7, 3.1)	15.1 * (1.9, 30.0)	3.8 *** (2.4, 5.2)

* p-value<0.05,
 ** p-value<0.01,
 *** p-value<0.001

Table 2

The associated change in the median LCMR (%), given a 10% increase in mean independent variable, in the stepwise multiple linear regression. The 95% confidence interval is given in parenthesis.

LCMR	Socioeconomic Group	BaPeq	SP	Price	GDP.CAP	BMI	R ²
ED100000	Low	D	9.8*** (5.9, 13.8)	D	D	D	0.45***
	Low Middle	D	6.6** (2.6, 10.8)	D	D	-34.6* (-54.5, -6.1)	0.30**
	Upper Middle	D	10.6*** (7.2, 14.2)	D	D	-45.6*** (-59.5, -26.9)	0.66***
	High	0.81** (0.23, 8.5)	4.5* (0.68, 8.5)	D	D	D	0.37**
	Low + Low middle	D	6.0*** (3.2, 8.9)	-3.2* (-5.7, -0.70)	5.4*** (2.9, 7.8)	D	0.49***
	Upper Middle + High	0.70* (0.13, 1.3)	8.8*** (5.7, 12.1)	3.8** (1.2, 6.5)	D	-38.4*** (-51.0, -22.5)	0.68***
	World	D	8.8*** (6.7, 10.9)	D	4.5*** (3.5, 5.5)	-24.0*** (-34.4, -11.9)	0.63***
	Low	D	7.3*** (4.2, 10.6)	D	D	D	0.40***
	Low Middle	D	4.4** (1.4, 7.5)	D	D	-30.2* (-47.2, -7.5)	0.27**
	Upper Middle	D	5.8*** (3.3, 8.3)	D	D	-26.0*** (-40.5, -8.0)	0.51***
ASDR100000	High	D	D	D	D	D	D
	Low + Low middle	D	4.3*** (2.1, 6.6)	-2.1* (-4.0, -0.04)	4.6*** (2.2, 7.0)	-18.5* (-32.5, -1.5)	0.38***
	Upper Middle + High	0.58** (0.20, 0.96)	4.5*** (2.6, 6.5)	D	D	-16.1* (-27.8, -2.4)	0.47***
	World	D	4.9*** (3.4, 6.4)	-1.8* (-3.2, -0.33)	3.1*** (2.1, 4.1)	-13.8*** (-22.6, -4.0)	0.50***

D= stepwise procedure dropped variable from model

* p-value<0.05,

** p-value<0.01,

*** p-value<0.001