AJRCCM Articles in Press. Published on July 1, 2010 as doi:10.1164/rccm.201002-0222OC
Media embargo until 2 weeks after above posting date; see thoracic.org/go/embargo

Wood Smoke Exposure and Gene Promoter Methylation are Associated with Increased Risk for COPD in Smokers

Akshay Sood, M.D., M.P.H.[1], Hans Petersen, M.S.[2], Christopher M. Blanchette, Ph.D. [2], Paula Meek, RN, Ph.D. [3], Maria A. Picchi, M.P.H. [2], Steven A. Belinsky, Ph.D. [2], Yohannes Tesfaigzi, Ph.D. [2].

Institutional Affiliations: [1]University of New Mexico School of Medicine (AS) and [2]Lovelace Respiratory Research Institute (HP, CMB, MAP, SAB, YT), Albuquerque, NM, U.S.A.; and University of Colorado at Denver (PM), Denver, CO, U.S.A.

Corresponding author: Yohannes Tesfaigzi, Ph.D., Lovelace Respiratory Research Institute, 2425 Ridgecrest Dr. S.E., Albuquerque, NM 87108; Telephone 001-505-348-9495; Fax 001-505-348-8567; E-mail address: YTesfaig@lrri.org

Sources of Support: This work was supported from funding by the State of New Mexico (appropriation from the Tobacco Settlement Fund), and from the National Institutes of Health (K23 HL 094531-01 to AS, RO1 ES015482 to YT, and U01 CA 097356 to SB).

Running Title: Wood Smoke and Methylation Increase COPD Risk

Descriptor: 9.6: COPD: Epidemiology

Word Count: 3073

"This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org"

ABSTRACT

Rationale: Wood smoke-associated chronic obstructive pulmonary disease (COPD) is common in women in developing countries but has not been adequately described in developed countries.

Objectives: Our objective was to determine whether wood smoke exposure was a risk factor for COPD in a population of smokers in the United States and whether aberrant gene promoter methylation in sputum may modify this association.

Methods: For this cross sectional study 1,827 subjects were drawn from the Lovelace Smokers' Cohort, a predominantly female cohort of smokers. Wood smoke exposure was self-reported. Post-bronchodilator spirometry was obtained and COPD outcomes studied included percent predicted FEV_1, airflow obstruction, and chronic bronchitis. Effect modification of wood smoke exposure with current cigarette smoke, ethnicity, sex, and promoter methylation of lung cancer-related genes in sputum on COPD outcomes were separately explored. Multivariable logistic and poisson regression models were used for binary and rate-based outcomes, respectively.

Measurements and Main Results: Self-reported wood smoke exposure was independently associated with lower percent predicted FEV_1 (point estimate -0.03 ± 0.01 {S.E.}) and higher prevalence of airflow obstruction and chronic bronchitis (ORs 1.96 {95%C.I. 1.52, 2.52} and 1.64 {95%C.I. 1.31, 2.06} respectively). These associations were stronger among current cigarette smokers, non-Hispanic whites, and men. Furthermore, wood smoke exposure interacted in a multiplicative manner with aberrant promoter methylation of the p16 or GATA4 genes on lower percent predicted FEV_1.

Conclusions: These studies identify a novel link between wood smoke exposure and gene promoter methylation that synergistically increases the risk for reduced lung function in cigarette smokers.

Word Count: 250

KEY WORDS

Wood smoke, hispanic ethnicity, cigarette smokers, airflow obstruction, chronic bronchitis, gene promoter methylation in sputum DNA

INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is the fourth leading cause of death, affecting approximately 25-30% of all cigarette smokers. Because of the large number of cigarette smokers, the incidence of this disease is rising and approaching worldwide epidemic proportions (1). Although COPD is clinically defined as airflow obstruction that is not fully reversible, it is a heterogeneous disease with different phenotypes - airflow obstruction due to parenchymal lung destruction resulting in the loss of elastic recoil (emphysema) and small airway obstruction resulting from peribronchial inflammation and chronic mucous hypersecretion. It is clear that COPD is primarily attributable to cigarette smoking; however, other factors including bacterial and viral infections and environmental air pollutants may affect the development of COPD.

From the global perspective, the magnitude of the health consequences of indoor air pollution rivals that of tobacco use (2). Epidemiologic studies have linked exposure to wood smoke with increased prevalence of respiratory symptoms and illness in children and adults. Increased respiratory symptoms, decreased pulmonary function, and increased prevalence of asthma, COPD, respiratory failure and cor pulmonale, as well as increased mortality have been associated with wood smoke exposure in various countries across Asia, Africa, and the Americas (3). These studies from developing countries and one from Spain (4) have established that smoke from biomass fuel during cooking activities can cause COPD in women who are non smokers. Although studies of children in the United States suggest that exposure to wood smoke leads to cough, wheeze, and respiratory tract infections (5, 6) association of wood smoke exposure with COPD has not been previously studied in the United States. This association has not been studied among cigarette smokers, a population subgroup already vulnerable to the development of this disease. Further, the interaction of host factors (such as

concurrent cigarette smoke exposure, ethnicity, sex, and genetic factors) with wood smoke exposure on this association has not been well-described.

The objective of the present study was to determine whether wood smoke exposure was a significant risk factor for COPD in a cohort of ever-smokers living in an urban area of Southwestern United States and if this association was modified by other host characteristics such as cigarette smoking, ethnicity, sex, or promoter methylation of genes in sputum DNA.

Some of the results of these studies have been reported in the form of an abstract (7).

MATERIALS AND METHODS

Study Population:

Study subjects (n=1,861) were drawn from eligible participants enrolled between 2001 and 2007 in a cohort study in New Mexico, *i.e.*, Lovelace Smokers Cohort that has been described previously (8). This large cohort disproportionately enrolled women ever-smokers to study the susceptibility to the development of COPD since women are underrepresented in most studies of airflow obstruction in the United States (1). The catchment area was Albuquerque, New Mexico and its surrounding communities, comprising a population of approximately 700,000 persons living at altitudes of approximately 1,500 meters above sea level. Most participants were recruited through newspaper or television advertisements and were paid a small stipend for their participation. This study was approved by the Western Institutional Review Board (Olympia, WA).

Inclusion Criteria:

Subjects were included in the study if they were 40 to 75 years of age, former or current smokers with at least 10 pack-years of smoking history, and able to understand English. In addition, subjects were required to undergo spirometric testing in the same timeframe as obtaining the medical history. Of the 2,004 subjects enrolled in the original cohort, those with missing data (n = 35); and those with < 10 pack-years of smoking (n = 108) were excluded from the analysis.

Study Measurements:

All tests were conducted at Lovelace Scientific Resources (Albuquerque, NM). Information related to demographics, respiratory diseases, and smoking was obtained by self-report from all

study participants via a questionnaire. Wood smoke exposure was also self-reported as part of the general health survey although no additional details about type, intensity, and duration of wood smoke exposure were obtained. Smoking-related variables included heavy smoking (> 40 pack-years, based on the mean cut point of 39.4) and current smoking status at the time of testing. Body mass index (BMI) was measured using standardized methods (9).

Pre- and post-bronchodilator spirometry were obtained on all subjects by registered respiratory therapists, strictly adhering to the 1994 American Thoracic Society (ATS) guidelines (10). After completion of pre-bronchodilator spirometry, all subjects were given two puffs of albuterol (90 mcg/spray metered dose inhaler) with a LiteAire® dual valve spacer (Thayer Medical Corporation, Tucson, AZ), and spirometry was repeated after 15 minutes. Participants were requested not to take any inhalers for four hours prior to their appointment. Vmax Encore 22 (Viasys Respiratory Care, Yorba Linda, CA) and KoKo (Ferraris Respiratory, Louisville, CO) spirometers were used. Both machines met the 1994 ATS recommendations and were calibrated daily and checked at three different injection speeds, as per the ATS guidelines. Additionally, respiratory therapists were monitored and periodically re-credentialed, as part of a standardized laboratory proficiency testing plan. Only spirometric tests that meet the ATS criteria for within-maneuver and between-maneuver acceptability were included in the analyses.

Induced sputum was collected and stored in Saccomanno's fixative. Three slides were made for each sputum sample to check for adequacy, as defined by the presence of lung macrophages or Curschmann's spirals (11). From each study subject adequate sputum samples was taken for DNA isolation by protease digestion followed by phenol chloroform extraction and ethanol precipitation. A methylation specific polymerase chain reaction assay was performed. A panel of eight genes (p16, MGMT, DAPK, RASSF1A, PAX5 α, PAX5 β, GATA4 and GATA5) were selected for analysis of aberrant gene promoter methylation in sputum based on our previous studies establishing their association with risk for lung cancer

(12). Nested methylation-specific polymerase chain reaction was used to detect methylated alleles in DNA recovered from the sputum samples, as previously described (13). Methylation of cytosines within PCR products was verified by various methods; first, the PCR products amplified from tumor samples have been sequenced by our group and others and show that they capture methylated cytosines. Second, the methylation-specific PCR assays are optimized to be highly specific for the region amplified. Lastly, a subset of sputum samples that gave positive methylation products were also analyzed by a second method using restriction enzyme digestion that can discriminate methylation status of CpGs within the resulting PCR product (12), (14).

Outcomes:

Outcome measures and assessment of the exposure variables were obtained at the same visit for all participants. COPD was defined by measurements of percent lung function, presence of airflow obstruction, and chronic bronchitis.

Percent lung function: Percent lung function: Our primary outcome was percent predicted value of post-bronchodilator FEV_1. The normative lung function tables used for Hispanics and non-Hispanic whites (NHW) in our study were from the Mexican Americans and Caucasian Americans, respectively, in the third National Health and Nutrition Examination Survey or NHANES III sampled population (15).

Post-bronchodilator airflow obstruction: Airflow obstruction was defined by a post-bronchodilator FEV_1/FVC ratio of < 70%, as defined by the GOLD criteria (16).

Chronic bronchitis: Participants with self-reported cough productive of phlegm for at least 3 months per year for at least 2 consecutive years were considered to have chronic bronchitis.

Statistical Analysis:

Summary statistics, including means, standard deviations (S.D.), medians, and interquartile ranges for continuous variables and proportions for categorical variables, were obtained. Chi-square test was used for analysis of categorical variables and the Wilcoxon rank sum test was used for continuous variables. Percent predicted FEV_1 was not normal in distribution and was therefore analyzed as a proportion by nonparametric tests, including multivariable generalized linear models with a Poisson distribution and a log-link [with parametric sensitivity analyses conducted additionally using multivariable ordinary least squares (OLS) regression models]. Similarly, multivariable binary logistic regression models were used for binary outcomes. Covariates considered in the adjusted models included sex, age, heavy smoking, current smoking, obesity, educational status (at least high school or not) and Hispanic ethnicity. All covariates were treated as categorical variables. In Tables 3, 4, 5 and 7, effect sizes for associations (*i.e.* Odds ratios and point estimates) are presented for combined risk of two exposures in the third row and individual exposures in the first two rows, as compared to no exposures. An additive effect between the two exposures was judged to be present when the effect size in row 3 was approximately the sum of the individual effect sizes in rows 1 and 2, which is equivalent to no additive interaction. Two-way regression analyses for multiplicative interactions between exposures on outcomes were also performed. All analyses were conducted in SAS 9.1 (Cary, NC). A two-sided p-value of < 0.05 was considered statistically significant.

Role of the funding source:

The study sponsor played no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

RESULTS

Demographic characteristics

This study involved 1,861 eligible subjects in New Mexico that included 1,503 women (80.8%); 317 Hispanics (17.0%); and 1,079 current smokers (58.0%). Wood smoke exposure was reported by 515 subjects (27.7%); subjects exposed to wood smoke were more likely to be younger than 50 years of age, Hispanic, and current cigarette smokers (Table 1).

Wood smoke exposure was independently associated with greater odds of respiratory disease, particularly among current smokers, non-Hispanic whites and men, as compared to those who were former smokers, Hispanics, and women respectively:

Wood smoke exposure was a significant predictor for all COPD outcomes studied (for all analyses $p < 0.001$, Table 2), after adjustment for cigarette smoking and other covariates in a multivariable analysis. The effect sizes related to wood smoke exposure were generally similar to that of current cigarette smoke exposure, except for chronic bronchitis (Table 2). Non-Hispanic whites had higher effect sizes for each of the COPD outcomes studied compared to Hispanics (for all analyses $p \leq 0.03$, Table 2). In addition, men had higher effect sizes than women for each of the COPD outcomes studied (for all analyses $p \leq 0.002$, Table 2) except for percent predicted FEV_1 ($p = 0.07$).

There was an additive effect with current exposure to cigarette smoke and wood smoke for all COPD outcome measures (Table 3). While a significant multiplicative interaction was noted between non-Hispanics and wood smoke exposure on percent predicted FEV_1, an additive effect was observed for the remaining outcome measures (Table 4). Similarly, there was a multiplicative interaction of sex with wood smoke exposure on FEV_1 but an additive effect for the remaining outcome measures (Table 5). Therefore, three-way interactions between sex or

ethnicity, wood smoke and cigarette smoke exposures on percent predicted FEV_1 were analyzed. These analyses demonstrated that men had significantly lower percent predicted FEV_1 than women currently exposed to both wood smoke and cigarette smoke (p for three-way interaction = 0.02). Further, irrespective of cigarette smoke exposure, non-Hispanic whites exposed to wood smoke had lower percent predicted FEV_1 compared to those without exposure to wood smoke. Wood smoke exposure in Hispanics was not associated with a lower percent predicted FEV_1 (p for three-way interaction = 0.001).

Wood smoke exposure was more strongly associated with selected COPD outcomes in the presence of methylated p16 or GATA4 genes in sputum:

Sputum samples of 1,267 subjects (68.1% of eligible cohort) were analyzed for gene promoter methylation of a panel of eight lung cancer-related genes. Analysis of individual genes revealed that methylation of p16 gene was significantly associated with lower percent predicted FEV_1 and possibly higher odds of chronic bronchitis while GATA4 methylation was associated with lower percent predicted FEV_1 after correction for multiple comparisons (Table 6, Table S2).

Further, we noted that wood smoke exposure was associated with significantly lower percent predicted FEV_1 in the presence of methylated p16 gene in the sputum (multiplicative interaction p = 0.002; Table 7 footnote). For chronic bronchitis, the odds ratio for the combined risk of p16 methylation and wood smoke was larger than that of each separately. Similarly, wood smoke exposure was associated with lower percent predicted FEV_1 and higher odds of airflow obstruction in the presence of aberrantly methylated GATA4 in the sputum (for all analysis multiplicative interaction $p \leq 0.04$; Table 7).

Finally, the association between methylation index and each outcome was assessed. Analysis of methylation index, the number of methylated genes in each sputum sample, dichotomized

into low (fewer than two genes methylated) and high (two or more genes methylated) showed that high methylation index was significantly associated with lower percent predicted FEV_1 and airflow obstruction (Table S1).

Alternate analytic strategies:

Our low prevalence of bronchodilator reversibility (approximately 11%), use of post-bronchodilator lung function measures, and exclusion of subjects with < 10 pack-year smoking history decreased the misclassification bias resulting from asthma as a cause of airway obstruction. Nevertheless, additional analyses were performed after excluding participants with self-reported asthma (n = 224 or 11%) and similar results were seen. Furthermore, similar results were also observed when airflow obstruction was defined by a post-bronchodilator FEV_1/FVC ratio below the 5th percentile of the NHANES III predicted value instead of the GOLD criteria or when absolute post-bronchodilator FEV_1/FVC ratio was analyzed instead of percent predicted FEV_1.

DISCUSSION

The present study shows that exposure to wood smoke is associated with all COPD phenotypes studied (*i.e.* low lung function, airflow obstruction, and chronic bronchitis) in a cohort of smokers living in an urban area of Southwestern United States; independent of cigarette smoking. These associations are stronger among current cigarette smokers, non-Hispanic whites and men, as compared to former cigarette smokers, Hispanics and women respectively. In addition, smokers with aberrant promoter methylation of the p16 and GATA4 genes in sputum demonstrate stronger associations between wood smoke exposure and lower lung function than those without these epigenetic changes.

In developed countries, people are exposed to wood smoke in a variety of ways, including smoke from residential heating, cooking stoves, campfires, forest fires, and prescribed fires (17). Wood burning is an important contributor to particle and gaseous material in ambient air, and in some locations accounts for up to 80% of the airborne particle concentrations during the winter (18). Measurements in homes heated with wood show that the total particulate matter (TPM) concentrations range from 0.05-0.1 mg/m^3 (19). Wood burning not only increases indoor but also outdoor 'neighborhood' pollution; thereby exposing many non-users to wood smoke components (20). Based on seasonal variations in $PM_{2.5}$ the local air agency confirmed that wood smoke may be an important contributor to the pollution in the Albuquerque area. Wood smoke is a complex mixture of numerous volatile and particulate substances constituted by different organic and inorganic compounds known to be toxic or irritating to the respiratory system. Its composition varies with the wood type and the conditions of combustion. More than 200 chemical and compound groups have been identified, most of which are in the inhalable size range, generally smaller than 1 μm (21) and often include ultrafine particles (less than 100 μm). Exposure to wood smoke in developed countries tends to be at sustained low-levels unlike

exposure to cigarette smoke that is short-term but intense with a single cigarette introducing 15-40 mg TPM into the respiratory tract.

Our study contrasts with most studies conducted outside the United States that have focused on non-smokers. Our population of relatively older smokers may be particularly susceptible to the adverse respiratory effects of wood smoke exposure, compared to the general population. This conclusion is supported by the observed additive effect between current cigarette smoke and wood smoke exposures on COPD phenotypes. Furthermore, these epidemiological findings are substantiated by our laboratory findings in which pulmonary inflammation and pathological changes were enhanced in mice concurrently exposed to wood smoke and cigarette smoke compared to cigarette smoke alone (Tesfaigzi, unpublished).

Our findings that New Mexican non-Hispanic whites are at greater risk for wood smoke-associated COPD than Hispanics is generally consistent with previous studies by our group and others showing that non-Hispanic whites in New Mexico may be at greater risk for COPD (22-24). Although the bases for these findings is not known, possible explanations include ethnic differences in the metabolism of wood smoke products, genetic susceptibility to the effects of wood smoke, type of wood burnt in homes, and prevalence of obesity.

Interestingly, our study suggests that men may be at higher risk than women with respect to wood smoke-associated COPD. This may reflect the fact that men may have greater involvement with loading, lighting, and maintaining wood stoves than women in developed countries, resulting in greater wood smoke-exposure.

Aberrant promoter methylation of genes in sputum of smokers was associated with various COPD phenotypes, particularly with reduced lung function. The majority of participants with COPD phenotypes in our cohort have mild to moderate (Stage I and II) disease based on the GOLD criteria. The observed association between high methylation index in sputum and

reduced pulmonary function suggests that gene promoter methylation in sputum may be an early biomarker for COPD. However, additional longitudinal studies, including those using COPD phenotypes defined by high resolution computed tomography, are needed to confirm this hypothesis.

In this study, we report a synergistic association for lower lung function between wood smoke exposure and aberrant promoter methylation of the p16 and GATA4 genes in sputum of smokers. We did not find that promoter methylation caused or explained away the wood smoke association. However, both wood smoke and promoter methylation were independent predictors of low lung function. GATA4 is a transcriptional regulator of numerous cell cycle genes (25) and p16 mediates cell cycle arrest and senescence (26) suggesting that these pathways may be disrupted during the development of COPD. Interestingly, each of these variables, i.e., low lung function (27), wood smoke exposure (28), and methylation of the above-mentioned genes (14) are independently associated with increased risk for lung cancer. Therefore, one would postulate that exposure to wood smoke may enhance the risk for aberrant gene promoter methylation and the development of lung cancer in cigarette smokers. Due to the large number of people exposed to wood smoke world-wide this hypothesis has great public health importance and needs further investigation.

The strengths of our study include its analysis of interactions between wood smoke exposure and cigarette smoke exposures, ethnicity, sex, and epigenetic changes in sputum on COPD outcomes. Additional strengths include use of post-bronchodilator spirometry to define obstruction, strict adherence to the 1994 ATS guidelines in the performance of spirometry, use of NHANES III reference standards, and similar results whether a fixed ratio (*i.e.* $FEV_1/FVC < 70\%$) or a statistically-defined lower limit of FEV_1/FVC ratio (29) was used to define obstruction (16).

We also recognize several limitations to our study. We cannot exclude differences in alpha-1 antitrypsin deficient status as an alternative explanation to our findings. However, severe alpha-1 antitrypsin deficiency only accounts for 1-2% of cases of COPD. Our study cohort may not be representative of all smokers in New Mexico and in other parts of the United States. However, the smoking behavior in this study is consistent with that observed in representative surveys of the state of New Mexico (30). Finally, obtaining a binary exposure variable, based on whether or not subjects were exposed to wood smoke from a self-report, overlooks the potential for large variability of exposures and could introduce information bias. We recognize the need to better measure exposure to wood smoke constituents by either validated questionnaire instruments or home exposure monitoring devices to obtain the type, unit amount, and duration of wood smoke the people are exposed to. Therefore, additional research on wood smoke-associated COPD should be performed in cigarette smokers with particular emphasis on understanding the characteristics and dose-response relationship of wood smoke exposure. Although DNA methylation is generally considered to be a stable epigenetic mark, longitudinal studies need to establish the stability of the epigenetic changes in sputum DNA. Furthermore, we analyzed genes that are believed to be primarily associated with lung cancer. Future studies designed to identify genes methylated specifically in COPD will be necessary to develop better biomarkers for this disease. Having successfully established a mouse model that shows enhanced inflammation when exposed to cigarette and wood smoke compared to cigarette or wood smoke alone, studies on identifying epigenetic changes in DNA isolated from lung cells will help identify the genes modified by cigarette and wood smoke exposure.

ACKNOWLEDGEMENTS

None of the authors are involved with organization(s) with financial interest in the subject matter or have any actual or potential conflicts of interest or have received any compensation for this work. The State of New Mexico and the National Institutes of Health had no influence in each of the following: design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

REFERENCES

1. Silverman EK, Weiss ST, Drazen JM, Chapman HA, Carey V, Campbell EJ, Denish P, Silverman RA, Celedon JC, Reilly JJ, et al. Gender-related differences in severe, early-onset chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000;162:2152-2158.

2. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: Systematic analysis of population health data. *Lancet* 2006;367:1747-1757.

3. Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *Lancet* 2009;374:733-743.

4. Orozco-Levi M, Garcia-Aymerich J, Villar J, Ramirez-Sarmiento A, Anto JM, Gea J. Wood smoke exposure and risk of chronic obstructive pulmonary disease. *Eur Respir J* 2006;27:542-546.

5. Honicky RE, Osborni JS, Akpom CA. Symtoms of respirtatory illness in young children and the use of wood-burning stoves for indoor heating. *Pediatrics* 1985;75:587-593.

6. Koenig JQ, Larson TV, Hanley QS, Rebolledo V, Dumler K, Checkoway H, Wang SZ, Lin D, Pierson WE. Pulmonary function changes in children associated with fine particulate matter. *Environ Res* 1993;63:26-38.

7. Sood A, Petersen H, Blanchette C, Meek P, Belinsky SA, Picchi M, Tesfaigzi Y. Wood smoke–associated chronic obstructive pulmonary disease (copd) – underappreciated in the united states? *Am J Resp and Crit Care Med* 2009;179:A4742.

8. Hunninghake GM, Cho MH, Tesfaigzi Y, Soto-Quiros ME, Avila L, Lasky-Su J, Stidley C, Melen E, Soderhall C, Hallberg J, et al. Mmp12, lung function, and copd in high-risk populations. *The New England journal of medicine* 2009;361:2599-2608.

9. Gordon CC, Chumlea WC, Roche AF. Anthropometric stnadardization of reference manual. 1988.

10. Standaridization, of, Spirometry. 1994 update. *Am J Resp and Crit Care Med* 1995;152:1107-1136.

11. Saccomanno G, Archer VE, Auerbach O, Saunders RP, Brennan LM. Development of carcinoma of the lung as reflected in exfoliated cells. *Cancer* 1974;33:256-270.

12. Palmisano WA, Divine KK, Saccomanno G, Gilliland FD, Baylin SB, Herman JG, Belinsky SA. Predicting lung cancer by detecting aberrant promoter methylation in sputum. *Cancer research* 2000;60:5954-5958.

13. Belinsky SA, Palmisano WA, Gilliland FD, Crooks LA, Divine KK, Winters SA, Grimes MJ, Harms HJ, Tellez CS, Smith TM, et al. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer Res* 2002;62:2370-2377.

14. Belinsky SA, Liechty KC, Gentry FD, Wolf HJ, Rogers J, Vu K, Haney J, Kennedy TC, Hirsch FR, Miller Y, et al. Promoter hypermethylation of multiple genes in sputum precedes lung cancer incidence in a high-risk cohort. *Cancer research* 2006;66:3338-3344.

15. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general u.S. Population. *American journal of respiratory and critical care medicine* 1999;159:179-187.

16. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Nhlbi/who global initiative for chronic obstructive lung disease (gold) workshop summary. *Am J Respir Crit Care Med* 2001;163:1256-1276.

17. Migliaccio CT, Bergauff MA, Palmer CP, Jessop F, Noonan CW, Ward TJ. Urinary levoglucosan as a biomarker of wood smoke exposure: Observations in a mouse model and in children. *Environmental health perspectives* 2009;117:74-79.

18. Larson TV, Koenig JQ. Wood smoke: Emissions and noncancer respiratory effects. *Annu Rev Public Health* 1994;15:133-156.

19. Ward TJ, Palmer CP, Houck JE, Navidi WC, Geinitz S, Noonan CW. Community woodstove changeout and impact on ambient concentrations of polycyclic aromatic hydrocarbons and phenolics. *Environmental science & technology* 2009;43:5345-5350.

20. Naeher LP, Brauer M, Lipsett M, Zelikoff JT, Simpson CD, Koenig JQ, Smith KR. Woodsmoke health effects: A review. *Inhal Toxicol* 2007;19:67-106.

21. Kleeman MJ, Shauer JJ, Cass GR. Size and composition distribution of fine particulate matter emitted from wood burning, meat charboiling and cigarettes. *Environ Sci Technol* 1999;33:3516-3523.

22. Samet JM, Schrag SD, Howard CA, Key CR, Pathak DR. Respiratory disease in a new mexico population sample of hispanic and non-hispanic whites. *The American review of respiratory disease* 1982;125:152-157.

23. Samet JM, Wiggins CL, Key CR, Becker TM. Mortality from lung cancer and chronic obstructive pulmonary disease in new mexico, 1958-82. *American journal of public health* 1988;78:1182-1186.

24. Sood A, Stidley CA, Picchi MA, Celedon JC, Gilliland F, Crowell RE, Belinsky SA, Tesfaigzi Y. Difference in airflow obstruction between hispanic and non-hispanic white female smokers. *Copd* 2008;5:274-281.

25. Rojas A, Kong SW, Agarwal P, Gilliss B, Pu WT, Black BL. Gata4 is a direct transcriptional activator of cyclin d2 and cdk4 and is required for cardiomyocyte proliferation in anterior heart field-derived myocardium. *Molecular and cellular biology* 2008;28:5420-5431.

26. Takahashi A, Ohtani N, Yamakoshi K, Iida S, Tahara H, Nakayama K, Nakayama KI, Ide T, Saya H, Hara E. Mitogenic signalling and the p16ink4a-rb pathway cooperate to enforce irreversible cellular senescence. *Nature cell biology* 2006;8:1291-1297.

27. McCaughan F, Spiro SG. Management of patients with copd. *Lancet* 2004;364:2015.

28. Ramanakumar AV, Parent ME, Siemiatycki J. Risk of lung cancer from residential heating and cooking fuels in montreal, canada. *American journal of epidemiology* 2007;165:634-642.

29. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948-968.

30. Centers for Disease Control and Prevention (CDC). Behavioral risk factor surveillance system user's guide. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.; 1998.

TABLES

Table 1: Descriptive characteristics associated with wood smoke exposure in the eligible cohort

Characteristic	Wood Smoke exposure present n = 515 (27.7%) N (%) or mean(SD)	Wood Smoke exposure absent n = 1,346 (72.3%) N (%) or mean(SD)	p value
Sex - Female	425 (82.5%)	1,078 (80.1%)	0.23
Age (years)			
40-49	165 (32.0%)	360 (26.7%)	0.02
50-59	196 (38.1%)	456 (33.9%)	0.09
60-69	119 (23.1%)	373 (27.7%)	0.04
≥ 70	35 (6.8%)	157 (11.7%)	0.002
Hispanic ethnicity	115 (22.3%)	202 (15.0%)	<0.001
Obese (BMI ≥ 30 kg/m^2)	156 (30.3%)	410 (30.5%)	0.94
≥ High school education	349 (67.8%)	969 (72.0%)	0.07
Current cig. smoker	318 (62.7%)	761 (57.5%)	0.04
Cig. smoking in pack years	39.0 (18.9)	39.5 (21.1)	0.88
Cig. smoking > 40 pack years	181 (35.1%)	485 (36.0%)	0.72
Chronic bronchitis	199 (38.6%)	373 (27.7%)	<0.001
Airflow obstruction (NHANES III)	150 (29.1%)	277 (20.6%)	<0.001
Airflow obstruction (GOLD)	164 (31.8%)	323 (24.0%)	0.001
FEV_1/FVC ratio %	70.6 (10.8)	71.9 (10.5)	0.01
FVC % predicted	100.6 (17.5)	100.9 (16.3)	0.92
FEV_1 % predicted	90.1 (20.8)	91.9 (18.5)	0.09

Note 1: Categorical and continuous variables were analyzed using chi-square and Wilcoxon rank sum tests respectively. N and percent values for categorical variables and mean ± SD for continuous variables are provided.

Table 2: Multivariable analysis of wood smoke as predictor for COPD outcomes (n=1,861)

Exposure variable	FEV_1 % predicted		Airflow Obstruction (GOLD)		Chronic Bronchitis	
	PE (SE)	p value	OR (95% CI)	p value	OR (95% CI)	p value
Wood smoke n = 515 (27.7%)	-0.03 (0.01)	<0.001	1.96 (1.52-2.52)	<0.001	1.64 (1.31-2.06)	<0.001
Current Cigarette Smoke n = 1,079 (58.0%)	-0.03 (0.01)	<0.001	1.34 (1.05-1.72)	0.02	3.48 (2.72-4.45)	<0.001
Non-Hispanic White n = 1,544 (83.0%)	-0.04 (0.01)	<0.001	2.20 (1.52–3.17)	<0.001	1.39 (1.04-1.86)	0.03
Male sex n = 358 (19.2%)	0.01 (0.01)	0.07	2.33 (1.78–3.06)	<0.001	1.50 (1.16–1.94)	0.002

Note 1: All models were adjusted for sex, age categories, Hispanic ethnicity, obesity (BMI ≥ 30 kg/m2), educational status (at least high school or not), heavy smoking history (pack-years > 40 or not), current cigarette smoke exposure and wood smoke exposure, where relevant.
Note 2: Exclusion of those with self-reported asthma (n = 224 or 11%) did not significantly alter the results.
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: Nonparametric tests using GLM regression models with a Poisson distribution and log link were used to evaluate the continuous FEV_1 percent predicted outcome and logistic regression models were used for airflow obstruction and chronic bronchitis.

Table 3: Additive effect of exposures to cigarette smoke and wood smoke on COPD outcomes (n=1,861)

Exposure variable	FEV_1 % predicted		Airflow Obstruction (GOLD)		Chronic Bronchitis	
	PE (SE)	p value	OR (95% CI)	p value	OR (95% CI)	p value
Current Cigarette Smoke only n = 761 (40.9%)	-0.03 (0.01)	<0.001	1.25 (0.94-1.67)	0.13	3.92 (2.92-5.26)	<0.001
Wood Smoke only n = 197 (10.6%)	-0.03 (0.01)	0.001	1.70 (1.15-2.49)	0.007	2.12 (1.41-3.18)	<0.001
Both Smoke n = 318 (17.1%)	-0.06 (0.01)	<0.001	2.71 (1.89-3.89)	<0.001	5.74 (4.05-8.13)	<0.001

Note 1: The referent group included those currently exposed to neither smoke (n = 585 or 31.4%).
Note 2: All models were adjusted for sex, age categories, Hispanic ethnicity, obesity (BMI ≥ 30 kg/m2), educational Status (at least high school or not), and heavy smoking history (pack-years > 40 or not).
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: Nonparametric tests using GLM regression models with a Poisson distribution and log link were used to evaluate the continuous FEV_1 percent predicted outcome and logistic regression models were used for airflow obstruction and chronic bronchitis.
Note 5: Two-way regression analyses showed that multiplicative interactions between wood smoke and cigarette smoke exposures were not significant for any of the above outcome measures (all $p \geq 0.14$). However, the effect sizes in row 3 were approximately the sum of the individual effect sizes for rows 1 and 2, indicating an additive effect.

Table 4: Additive effect of non-Hispanic ethnicity and exposure to wood smoke on COPD outcomes (n=1,861)

Exposure variable	FEV_1 % predicted		Airflow Obstruction (GOLD)		Chronic Bronchitis	
	PE (SE)	p value	OR (95% CI)	p value	OR (95% CI)	p value
Hispanics with Wood Smoke; n = 115 (6.2%)	0.02 (0.01)	0.21	1.97 (1.00-3.86)	0.05	1.22 (0.73–2.04)	0.44
Non-Hispanics without Wood Smoke; n = 1,144 (61.5%)	-0.02 (0.01)	0.004	2.20 (1.37–3.52)	0.001	1.21 (0.85-1.73)	0.28
Non-Hispanics with Wood Smoke; n = 400 (21.5%)	-0.07 (0.01)	<0.001	4.30 (2.59-7.12)	<0.001	2.14 (1.45-3.16)	<0.001

Note 1: The referent group included Hispanics without current exposure to wood smoke (n = 202 or 10.9%).
Note 2: All models were adjusted for sex, age categories, current smoking, obesity (BMI ≥ 30 kg/m^2), educational status (at least high school or not), and heavy smoking history (pack-years > 40 or not).
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: Nonparametric tests using GLM regression models with a Poisson distribution and log link were used to evaluate the continuous FEV_1 percent predicted outcome and logistic regression models were used for airflow obstruction and chronic bronchitis.
Note 5: Among the various COPD outcome measures, the only statistically significant multiplicative interaction using two-way regression analyses between non-Hispanic ethnicity and wood smoke exposure was on FEV_1 % predicted ($p < 0.001$). For remaining outcome measures, the effect sizes in row 3 were approximately the sum of the individual effect sizes for rows 1 and 2, indicating an additive effect.

Table 5: Additive effect of male sex and exposure to wood smoke on COPD outcomes (n=1,861)

Exposure variable	FEV_1 % predicted		Airflow Obstruction (GOLD)		Chronic bronchitis	
	PE (SE)	p value	OR (95% CI)	p value	OR (95% CI)	p value
Women with wood smoke; n = 425 (22.8%)	-0.03 (0.01)	<0.001	1.93 (1.45-2.56)	<0.001	1.72 (1.34-2.22)	<0.001
Men without wood smoke; n = 268 (14.4%)	0.02 (0.01)	0.01	2.29 (1.67-3.15)	<0.001	1.60 (1.19-2.16)	0.002
Men with wood smoke; n = 90 (4.8%)	-0.04 (0.01)	0.001	4.72 (2.88-7.71)	<0.001	2.14 (1.34-3.43)	0.002

Note 1: The referent group included women not exposed to wood smoke (n = 1,078 or 57.6%).
Note 2: All models were adjusted for age categories, Hispanic ethnicity, obesity (BMI ≥ 30 kg/m2), educational status (at least high school or not), current cigarette smoking, and heavy smoking history (pack-years > 40 or not).
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: Nonparametric tests using GLM regression models with a Poisson distribution and log link were used to evaluate the continuous FEV_1 percent predicted outcome and logistic regression models were used for airflow obstruction and chronic bronchitis.
Note 5: Among the various COPD outcome measures, the only statistically significant multiplicative interaction using two-way regression analyses between sex and wood smoke exposure was on FEV_1 % predicted ($p = 0.04$). For remaining outcome measures, the effect sizes in row 3 were approximately the sum of the individual effect sizes for rows 1 and 2, indicating an additive effect.

Table 6: Associations between promoter methylation of genes in sputum and COPD outcomes (n=1,267)

		FEV_1 % predicted		Airflow Obstruction (GOLD)		Chronic Bronchitis	
Exposure	**N**	**PE (SE)**	**p value**	**OR (95% CI)**	**p value**	**OR (95% CI)**	**p value**
P16	216/1267	-0.02 (0.01)	0.006	1.12 (0.78-1.60)	0.553	1.56 (1.13-2.16)	0.007
GATA4	485/1267	-0.02 (0.01)	0.003	1.39 (1.05-1.85)	0.021	1.15 (0.89-1.49)	0.290

Note 1: The referent group included those participants who did not have the specific methylated gene in sputum (n=1,051 for p16 and 782 for GATA4).
Note 2: All models were adjusted for sex, age categories, Hispanic ethnicity, obesity (BMI ≥ 30 kg/m2), educational status (at least high school or not), heavy smoking history (pack-years > 40 or not), and current cigarette smoke exposure and wood smoke exposure.
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: Nonparametric tests using GLM regression models with a Poisson distribution and log link were used to evaluate the continuous FEV_1 percent predicted outcome and logistic regression models were used for airflow obstruction and chronic bronchitis.
Note 5: Two of the panel of eight lung cancer-related genes (p16, MGMT, DAPK, RASSF1A, PAX5 α, PAX5 β, GATA4 and GATA5) were selected for analysis of promoter methylation in sputum. RASSF1A was also significantly and inversely associated with FEV_1/FVC ratio and FEV_1 % predicted but was not included in the model because of small sample size (10/1267).
Note 5: To adjust for multiple comparisons, the p values in the table are to be compared to an α of 0.05 divided by 8 (the number of genes studied); the results for FEV_1 % predicted remained significant under Bonferroni correction - a conservative approach for explaining these number of comparisons.

Table 7: Effect of wood smoke exposure and promoter methylation of genes in sputum on COPD outcomes (n=1,267)

	FEV_1 % predicted [Note 4]		**Airflow Obstruction (GOLD)**		**Chronic Bronchitis**	
Exposure	**PE (SE)**	**p value**	**OR (95% CI)**	**p value**	**OR (95% CI)**	**p value**
methylated p16 gene in sputum						
Unmethylated p16 with wood smoke (n=296)	-0.04 (0.01)	<0.001	1.90 (1.41-2.57)	<0.001	1.81 (1.37-2.38)	<0.001
p16 methylation without wood smoke (n=151)	-0.02 (0.01)	0.01	1.22 (0.81-1.85)	0.35	1.75 (1.20-2.53)	0.003
p16 methylation with wood smoke (n=65)	-0.07 (0.01)	<0.001	2.27 (1.26-4.07)	0.006	3.07 (1.78-5.27)	<0.001
methylated GATA4 gene in sputum						
Unmethylated GATA4 with wood smoke (n=215)	-0.05 (0.01)	<0.001	1.95 (1.37-2.78)	<0.001	2.28 (1.66-3.13)	<0.001
GATA4 methylation without wood smoke (n=339)	-0.02 (0.01)	0.001	1.65 (1.23-2.22)	0.001	1.62 (1.23-2.13)	0.001
GATA4 methylation with wood smoke (n=146)	-0.06 (0.01)	<0.001	2.54 (1.70-3.78)	<0.001	1.85 (1.27-2.70)	0.002

Note 1: The referent group included those participants who did not have the specific methylated gene in sputum and were not exposed to wood smoke (n=755 for p16 and 567 for GATA4).
Note 2: All models were adjusted for sex, age categories, Hispanic ethnicity, obesity (BMI ≥ 30 kg/m2), educational status (at least high school or not), heavy smoking history (pack-years > 40 or not), current cigarette smoke exposure and wood smoke exposure.
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: Nonparametric tests using GLM regression models with a Poisson distribution and log link were used to evaluate the continuous FEV_1 percent predicted outcome and logistic regression models were used for airflow obstruction and chronic bronchitis.
Note 5: Statistically significant multiplicative interaction using two-way regression analyses between methylated p16 gene and wood smoke exposure on FEV_1 % predicted was present (p = 0.002). Similarly, significant multiplicative interactions between methylated GATA4 gene and wood smoke exposure on airflow obstruction, using GOLD criteria (p = 0.04) and on FEV_1 % predicted (p = 0.002) were seen. Of note, multiplicative interactions cannot be obtained in this table by multiplying individual effect sizes from rows 1 and 2.

Supplementary Material

Wood Smoke Exposure and Gene Promoter Methylation are Associated with Increased Risk for COPD in Smokers

Authors: Akshay Sood, M.D., M.P.H., Hans Petersen, M.S., Christopher M. Blanchette, Ph.D. , Paula Meek, RN, Ph.D., Maria A. Picchi, M.P.H., Steven A. Belinsky, Ph.D., Yohannes Tesfaigzi, Ph.D.

Table S1: Associations between aberrant promoter methylation status of genes in sputum and COPD outcomes (n=1,267)

		FEV_1 % predicted		Airflow Obstruction (GOLD)		Chronic Bronchitis	
Exposure	**N**	**PE (SE)**	**p value**	**OR (95% CI)**	**p value**	**OR (95% CI)**	**p value**
≥2 genes methylated	507/1267	-0.02 (0.01)	0.01	1.39 (1.05-1.85)	0.02	1.26 (0.97-1.63)	0.08

Note 1: The referent group included those participants who had <2 methylated gene in sputum (n=760).
Note 2: All models were adjusted for sex, age categories, Hispanic ethnicity, obesity (BMI ≥ 30 kg/m2), educational status (at least high school or not), heavy smoking history (pack-years > 40 or not), and current cigarette smoke exposure and wood smoke exposure.
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: A panel of eight lung cancer-related genes (p16, MGMT, DAPK, RASSF1A, PAX5 α, PAX5 β, GATA4 and GATA5) were selected for analysis of promoter methylation in sputum.

Table S2: Associations between promoter methylation of all eight genes in sputum and COPD outcomes (n=1,267)

Exposure	N	FEV_1 % predicted		Airflow Obstruction (GOLD)		Chronic Bronchitis	
		PE (SE)	p value	OR (95% CI)	p value	OR (95% CI)	p value
p16	216/1267	-0.02 (0.01)	0.006	1.12 (0.78-1.60)	0.553	1.56 (1.13-2.16)	0.007
GATA4	485/1267	-0.02 (0.01)	0.003	1.39 (1.05-1.85)	0.021	1.15 (0.89-1.49)	0.290
GATA5	225/1267	0.01 (0.01)	0.46	0.91 (0.63-1.30)	0.59	1.27 (0.92-1.75)	0.15
MGMT	338/1267	0.01 (0.01)	0.20	1.10 (0.81-1.50)	0.55	1.11 (0.84-1.47)	0.47
DAPK	227/1267	0.0008 (0.01)	0.92	1.22 (0.86-1.73)	0.26	1.16 (0.84-1.60)	0.37
RASSF1A	10/1267	-0.18 (0.04)	<0.001	2.63 (0.63-10.93)	0.18	2.59 (0.68-9.82)	0.16
PAX5α	191/1267	0.01 (0.01)	0.43	1.17 (0.81, 1.69)	0.41	0.83 (0.58-1.19)	0.30
PAX5β	119/1267	-0.01 (0.01)	0.29	1.02 (0.64-1.61)	0.94	1.03 (0.67-1.59)	0.88

Note 1: The referent group for each analysis included those participants who did not have the specific methylated gene in sputum.
Note 2: All models were adjusted for sex, age categories, Hispanic ethnicity, obesity (BMI ≥ 30 kg/m2), educational status (at least high school or not), heavy smoking history (pack-years > 40 or not), and current cigarette smoke exposure and wood smoke exposure.
Note 3: Results for absolute postbronchodilator FEV_1/FVC ratio and for airflow obstruction defined by NHANES III criterion were similar to those for FEV_1 percent predicted and airflow obstruction defined by GOLD criterion respectively.
Note 4: Data on two of the panel of eight genes (p16 and GATA4) are also shown in a condensed form in Table 6 in the text. These genes were significantly associated with selected COPD outcomes and had sufficiently large sample size. RASSF1A showed significant association with lower percent predicted FEV_1 but was excluded from further analysis because a total of only 10 people showed gene methylation.
Note 5: To adjust for multiple comparisons, the p values in the table are to be compared to an α of 0.05 divided by 8 (the number of genes studied) or 0.0063 under Bonferroni correction - a conservative approach for explaining these number of comparisons.