Association between household air pollution exposure and chronic obstructive pulmonary disease outcomes in 13 low- and middle-income country settings

Authors: Trishul Siddharthan* MD (1, 2), Matthew R Grigsby* MSPH (1, 2), Dina Goodman MSPH (1, 2), Muhammad Chowdhury MBBS (3), Adolfo Rubinstein MD PhD (4), Vilma Irazola MD (4), Laura Gutierrez MD (4), J Jaime Miranda MD PhD (5, 6), Antonio Bernabe-Ortiz MD MPH (5, 6), Dewan Alam MBBS PhD (7), Bruce Kirenga MBChB (8), Rupert Jones MBBS MD (9), Frederick van Gemert MD (10), Robert A Wise MD (1), William Checkley MD PhD (1, 2).

1. Division of Pulmonary and Critical Care, School of Medicine, Johns Hopkins University, Baltimore, USA.
2. Center for Global Non-Communicable Diseases, School of Medicine, Johns Hopkins University, Baltimore, USA.
3. Centre for Control of Chronic Diseases, icddr,b, Dhaka, Bangladesh.
4. Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina.
5. CRONICAS Centre of Excellence in Chronic Diseases, Universidad Peruana Cayetano Heredia, Lima, Peru.
6. Departamento de Medicina, Facultad de Medicina, Universidad Peruana Cayetano Heredia, Lima, Peru.
7. School of Kinesiology and Health Science, Faculty of Health, York University, Toronto, Ontario, Canada M3J 1P3
8. Makerere Lung Institute, Makerere University, Kampala, Uganda
9. Plymouth University, Plymouth, United Kingdom
10. University of Groningen, University Medical Centre Groningen, Harlingen, Netherlands

Correspondence:
William Checkley, MD, PhD
Division of Pulmonary and Critical Care
School of Medicine
Johns Hopkins University
1830 Monument St, Room 555
Baltimore, MD 21205
E-mail: wcheckl1@jhmi.edu

Word count: 3,988

* Joint first authorship.

Conception and design: TS, MG, WC; Analysis and Interpretation: TS, MG, DG, WC; Drafting

the manuscript for important intellectual content: TS, MRG, DG, LG, JJM, AB, DA, BK, RJ, FVG,

RAW, WC

1

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 2 of 45

ABSTRACT

Rationale: Forty percent of households worldwide burn biomass fuels for energy, which may be most the important contributor to household air pollution.

Objective: To examine the association between household air pollution exposure and COPD outcomes in 13 resource-poor settings.

Methods: We analyzed data from 12,396 adult participants living in 13 resource-poor, population-based settings. Household air pollution exposure was defined as using biomass materials as the primary fuel source in the home. We used multivariable regressions to assess the relationship between household air pollution exposure and COPD outcomes, evaluated for interactions, and conducted sensitivity analyses to test the robustness of our findings.

Results: Average age was 54.9 years (44.2-59.6 years across settings), 48.5% were women (38.3%-54.5%), prevalence of household air pollution exposure was 38% (0.5%-99.6%), and 8.8% (1.7%-15.5%) had COPD. Participants with household air pollution exposure were 41% more likely to have COPD (adjusted OR=1.41, 95% CI 1.18 to 1.68) than those without the exposure, and 13.5% (6.4% to 20.6%) of COPD prevalence may due to household air pollution exposure, compared with 12.4% due to cigarette smoking. The association between household air pollution exposure and COPD was stronger in women (1.70, 1.24 to 2.32) than in men (1.21, 0.92 to 1.58).

Conclusions: Household air pollution exposure was associated with a higher prevalence of COPD, particularly among women, and it is likely a leading population attributable risk factor for COPD in resource-poor settings.

Keywords: COPD; Air Pollution, Indoor/adverse effects; Biomass

2

INTRODUCTION

Approximately three billion people rely on the burning of solid fuels such as wood, dung, agricultural crop waste, and coal for energy, and biomass fuels are the main source of domestic energy for ~40% of households worldwide (1). Households in many low- and middle-income countries (LMICs) often burn biomass inefficiently and with poor ventilation, resulting in exposure to a range of pollutants (2). The resulting household air pollution (HAP) accounts for an estimated 2.9 million deaths and 85.6 million disability-adjusted life years lost based on the Global Burden of Disease Study 2015, making it the eighth leading risk factor for the global burden of disease (3).

Current evidence supports an association between HAP exposure and a range of respiratory diseases including pneumonia, COPD and lung cancer (4-7). COPD, in particular, is a salient consequence of HAP exposure since it poses a considerable socioeconomic burden and disproportionally affects impoverished populations in LMICs (8). Previous studies have demonstrated that the relationship between HAP exposure and respiratory health outcomes is strongest among women and children who have the most intense exposure (9). Two recent population-based studies in Latin America found that women with HAP exposure were more likely to have COPD than those who did not have the exposure (10, 11).

Few population-based studies have evaluated the attributable risk for COPD due to HAP exposure. Here we describe the relationship between HAP exposure and COPD in 13 LMIC settings. These settings represent a diversity of geographies, ethnicities, variations in altitude, and degrees of urbanization in resource-poor settings of Latin America, Sub-Saharan Africa and Southeast Asia.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 4 of 45

METHODS

Study setting

We pooled data from five population-based studies spanning six countries and 13 settings in Latin America, Sub-Saharan Africa and Southeast Asia. Included studies were sponsored by United States National Institutes of Health/National Heart, Lung, and Blood Institute and UnitedHealth Chronic Disease Initiative (http://www.nhlbi.nih.gov/about/org/globalhealth/centers), the Fogarty International Center of the United States National Institutes of Health, and the FRESH AIR Study Group (http://www.theipcrg.org/freshair). To be eligible, studies had to contribute data with the following specifications: 1) adults aged ≥ 18 years, 2) site located in a World Bank defined LMIC participating within the above described networks, 3) conducted a population-based study, 4) performed post-bronchodilator spirometry in those with obstruction, and willing to share data for pooled analysis. Specifically, data were compiled from the Pulmonary Risk in South America (PRISA) study, conducted by the Institute for Clinical Effectiveness and Health Policy (IECS) in two sites in Argentina (Marcos Paz and Bariloche), one in Chile (Temuco), and one in Uruguay (Canelones), the CRONICAS Cohort Study in Peru, conducted by the CRONICAS Centre of Excellence for Chronic Diseases at Universidad Peruana Cayetano Heredia and Johns Hopkins University, and a longitudinal study in Bangladesh, conducted by the Centre for Control of Chronic Diseases at the International Centre for Diarrheal Disease Research, Bangladesh (icddr,b), Lung Function Study in Nakaseke and Uganda (LiNK) and the FRESH AIR Uganda, conducted by the Makerere Lung Institute. Both PRISA and CRONICAS studies are prospective longitudinal studies with multiple years of follow up that started in 2010 (12, 13). icddr,b conducted a longitudinal study from 2011 to 2012 (14). LiNK is an ongoing longitudinal study with baseline data collected in 2015. FRESH AIR Uganda is a cross-sectional study conducted in 2012 in rural Masindi (15).

4

Study Design

PRISA and CRONICAS used age- and sex-stratified random sampling whereas the Bangladesh study used simple random sampling of available census data at each site. LiNK used a sampling technique outlined by the WHO, while FRESH AIR used a multi-level sampling approach (10, 13-18). We limited our analysis to participants aged 35-95 years to match reference equation upper age limits (19). All studies obtained informed consent from local and international ethics boards, and all investigators required confidentiality training for field workers. Details can be found elsewhere (10, 13-16).

Spirometry

All sites followed joint ATS/ERS recommendations when performing and grading spirometry. PRISA, CRONICAS, LiNK, and the Bangladesh study used similar spirometry devices (ndd, Zurich, Switzerland), while FRESH AIR used Pneumotrac spirometers (10, 13-15). Pre- and post-bronchodilator forced expiratory volumes (FEVs) were measured for all individuals in PRISA and CRONICAS, whereas other studies only took post-bronchodilator measurements on those who screened positive for obstruction on pre-bronchodilator spirometry ($FEV_1/FVC \leq 0.7$ in FRESH AIR and the Bangladesh studies, and $FEV_1/FVC \leq$ lower limit of normal in LiNK).

Definitions

We defined COPD as having a post-bronchodilator FEV_1/FVC Z-score ≤ -1.64 SDs of the Global Lung Function Initiative (GLI2012) mixed ethnic reference population (19). COPD severity was categorized according to the GOLD strategy (20, 21). Pack-years of smoking was defined as the number of packs smoked per day multiplied by the number of years smoking. Participants were considered to have symptomatic COPD if they had wheeze, cough, or phlegm currently or in the last 12 months. We defined restrictive spirometric pattern as a pre-bronchodilator FVC Z-score < -1.64 and no spirometric evidence of COPD (22); daily smoking as having ≥ 1 cigarette/day;

5

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 6 of 45

and, HAP exposure if biomass was the primary source of fuel in the home. We defined lung function reversibility as the difference between post- and pre-bronchodilator FEVs was >200 mL and/or the percent increase was >12%.

Biostatistical analysis

Our primary analytical plan was to characterize the association between HAP exposure and COPD. We conducted secondary analyses to assess the association between HAP exposure and other COPD outcomes, namely severity and the presence of concomitant respiratory symptoms, and pre-bronchodilator FEVs; and, between HAP exposure and restrictive spirometric pattern.

For our primary analysis, we used multivariable alternating logistic regressions (ALR) to model the association between HAP exposure and COPD, adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis and secondary education (i.e., confounders). ALR is a variant of generalized estimating equations where the association between pairs of participants for a particular site is modeled with log odds ratios instead of correlations (23). In sensitivity analyses, we used the Mantel-Haenszel method to estimate unadjusted odds ratios weighted by site, and multivariable random effects logistic regression to determine if our findings were robust to the approach chosen to model heterogeneity across settings (Online Supplement) (24, 25). We also examined for effect modification by sex, age (≥55 or <55 years), self-reported daily cigarette smoking, and having secondary education. We used adjusted odds ratio estimates to calculate the population attributable fraction of COPD due to HAP exposure using Levin's formula, i.e., $PAR = \frac{p \times (OR - 1)}{1 + p \times (OR - 1)}$ (26).

6

For secondary analyses, we used multivariable random effects ordinal logistic regressions to examine the association between HAP exposure and COPD severity (none, mild, moderate or severe/very severe COPD) or symptomatic COPD (none, asymptomatic COPD and symptomatic COPD) adjusted for confounders (*vide supra*). To graphically assess for proportionality of odds, we compared the ORs and corresponding 95% CIs obtained for each variable from a logistic regression model evaluated at each of the threshold points for each of the above ordinal scales (Online Supplement). We used ALR to examine the association between HAP exposure and either having restrictive spirometric pattern or reversibility adjusted for the above mentioned confounders. We used multivariable linear mixed-effects models with a random intercept by site to study the association between HAP exposure and pre-bronchodilator FEVs accounting for an interaction with age and adjusted for the above mentioned confounders (24, 27, 28).

In sensitivity analyses, we used the GLI2012 Caucasian reference values to determine if our estimates were consistent regardless of the reference chosen; used pack-years smoked instead of daily smoking to rule out residual confounding by not adjusting for frequency of smoking exposure; conducted leave-one-site-out and ten-fold cross validation analyses to determine if one site or subset of data heavily influenced exposure-outcome relationships, respectively; limited analyses to sites with a prevalence of HAP exposure <95% and >5%, or with at least 5 participants in each category of the contingency table between HAP exposure and COPD to determine if these sites heavily influenced exposure-outcome relationships (Online Supplement).

Analyses were performed in R using the `lme4, gmodels, ggplot2, alr, doParallel` and `ordinal` packages (29, 30).

7

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 8 of 45

RESULTS

Population Characteristics

The 13 sites contributed data on 13,023 participants but 12,396 met eligibility criteria and had complete data for analysis (Figure E1). Fifty-eight percent of the study sample lived in Latin America, 28% in Southeast Asia, and 14% in Sub-Saharan Africa (Figure E2). Average age among participants in the study sample was 54.9 years (range of mean age across settings 44.2-59.6 years; p<0.001 for differences between sites), 48.5% were women (range of proportions across settings 38.3%-54.5%; p<0.001) and the overall prevalence of HAP exposure was 38% (0.5%-99.6%; p<0.001). There was no difference in COPD prevalence (8.9% vs. 8.8%; p=0.97) between excluded and included participants; however, those who were excluded had a higher prevalence of HAP exposure (92% vs 38%; p<0.001), were younger (46.0 years vs. 54.9 years; p<0.001), and were more likely to be women (54.9% vs. 48.5%; p=0.002). Self-reported biomass use ranged from 0.5% in Marcos Paz, Argentina to 99.6% in Nakaseke, Uganda. Daily cigarette smoking ranged from 0.2% in rural Puno, Peru to 36.2% in Masindi, Uganda. All sites were located in resource-poor settings in LMICs with a variety of kitchen layouts (Figure 1).

Epidemiology of COPD

The overall prevalence of COPD was 8.8% with a range of 1.7% in Kampala, Uganda to 15.5% in Masindi, Uganda. Men had a higher prevalence of COPD than women (10.3% vs. 7.2%; p<0.001), however, there was significant heterogeneity in the prevalence of COPD by sex across sites (Figure 2). Among those with COPD, 394 (36.1%) were mild, 524 (48.0%) were moderate, and 173 (15.9%) were severe/very severe; however, there was also substantial heterogeneity in severity profiles across settings (Figure 2). For example, the prevalence of severe COPD ranged from 0% in both rural and urban Puno, Peru to 26.5% in Dhaka, Bangladesh. Men also had a higher prevalence of moderate (53.4% vs. 39.8%; p <0.001) or

8

severe/very severe COPD (19.3% vs. 10.6%; p<0.001) than women. In site-weighted analyses, daily cigarette smokers were more likely to have COPD than participants who did not smoke daily (Mantel-Haenszel OR=2.55; 95% CI 2.17 to 3.00). When stratified by sex, both men (Mantel-Haenszel OR[OR_{MH}]=2.30, 1.89 to 2.80) and women (OR_{MH}=1.83, 1.35 to 2.48) who were daily smokers were more likely to have COPD than those who were not smokers.

HAP exposure and COPD outcomes

Participants with HAP exposure had a higher prevalence of COPD than those without the exposure (10.8% vs. 7.6%; p<0.001). While the prevalence of COPD between participants with HAP exposure was higher at any age when compared to those without the exposure, the difference by HAP exposure status was greater at older ages (Figure 3). In site-weighted analyses, participants with HAP exposure were more likely to have COPD than those without the exposure (OR_{MH}=1.68, 95% 1.29 to 2.19). In multivariable-adjusted analyses accounting for clustering by site, participants with HAP exposure were more likely (adjusted OR=1.41, 95% CI 1.18 to 1.68) to have COPD than those without the exposure (Figure 4). Participants with HAP exposure also had a greater odds of having more severe disease (adjusted OR=1.51, 95% CI 1.16 to 1.96) or having symptoms (adjusted OR=1.50, 95% CI 1.15 to 1.97) when compared to those without the exposure. For models where ordinal scales were used, we visually confirmed that the proportionality of odds assumption was reasonable (Figure E3 and Figure E4).

We plotted interaction effects between HAP exposure and potential effect modifiers on the odds ratio of having COPD (Figure 4), and found that the association between HAP exposure and COPD was stronger in women (adjusted OR=1.70, 95% CI 1.24 to 2.32) than in men (1.21, 0.92 to 1.58). We also found that HAP exposure was associated with a higher odds of having COPD among participants aged ≥55 years (1.43, 1.09 to 1.87) and a marginally higher odds of having COPD for those aged <55 years (1.32, 0.97 to 1.78).

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 10 of 45

We estimated that 13.5% (95% CI 6.4% to 20.6%) of the COPD prevalence in our study sample was due to HAP exposure, in contrast to 12.4% due to daily cigarette smoking, 9.4% due to lower education and 6.6% due to post-treatment pulmonary tuberculosis. When stratified by sex, the PAF was higher in women (21.0%, 95% CI 8.4% to 33.5%) than in men (7.3%, -3.1% to 18.0%). When stratified by region, the PAF was highest in Sub-Saharan Africa (28.2%, 14.6% to 39.6%), followed by Southeast Asia (17.8%, 8.7% to 26.6%) and Latin America (6.4%; 2.9% to 10.3%).

In sensitivity analyses, we found that using the GLI2012 Caucasian reference population for FEV_1/FVC did not affect the direction or magnitude of reported exposure-outcome associations when compared to a GLI mixed ethnic reference population (Table E1 and Table E2). Similarly, analyses using pack-years smoked instead of daily smoking showed almost identical results (Table E3). Both leave-one-site-out and ten-fold cross-validation analyses (Table E4 and Table E5) revealed that no single site or small groups of participants appeared to have heavily influenced the association between HAP exposure and COPD. Moreover, the association between HAP exposure and COPD was consistent in magnitude and direction when we limited our data to sites with <95% and >5% prevalence of HAP exposure, or sites with at least 5 participants in each category of the contingency table between HAP exposure and COPD (Online Supplement).

HAP exposure and lung function

We plotted unadjusted Z-scores of pre-bronchodilator FEV_1 by deciles of age and stratified by HAP exposure (Figure 5). On average, participants with HAP exposure had lower pre-bronchodilator FEV_1, FVC and FEV_1/FVC Z-scores at any age when compared to those without the exposure. There were notable differences in the trends with age, however. Specifically,

10

across all age deciles, participants with HAP exposure had a consistently lower pre-bronchodilator FEV_1 Z-score when compared to participants who did not have the exposure. In contrast, differences in FVC Z-scores between participants with and without HAP exposure were greater in younger ages whereas differences in FEV_1/FVC Z-scores were greater at older ages. These trends remained consistent in multivariable regression analyses that accounted for heterogeneity across sites. Specifically, participants with HAP exposure had: a marginally lower adjusted pre-bronchodilator FEV_1 Z-score (-0.11 SD, 95% CI -0.24 to 0.03 SD) than those who did not have the exposure across all ages, with no interaction effect between HAP exposure and age (p=0.07); a lower pre-bronchodilator FVC Z-score at age 35 years (-0.14 SD, -0.28 to -0.004 SD) but not at age 60 years (-0.02 SD, -0.13 to 0.10 SD); and, no difference in pre-bronchodilator FEV_1/FVC Z-score at age 35 years (0.02 SD, -0.09 to 0.14 SD) but a lower pre-bronchodilator FEV_1/FVC Z-score at 60 years (-0.25 SD, -0.34 to -0.15 SD). In subset analysis of studies with both pre- and post-bronchodilator spirometry, we found that participants with HAP exposure were more likely to have lung function reversibility than those without the exposure at younger ages (adjusted OR at 35 years=1.62, 95% CI 1.18 to 2.24) but not at older ages (adjusted OR at 60 years=1.27, 0.93 to 1.74). Sensitivity analyses showed comparable results when using Caucasian reference values (Online Supplement).

HAP exposure and restrictive spirometric patterns

Participants with HAP exposure had a lower prevalence of restrictive spirometric patterns than those who did not have the exposure (11.4% vs. 14.8%; p<0.001) and this difference was consistent across all ages (Figure 3). In site-weighted analysis, participants with HAP exposure had similar odds of having restrictive spirometric pattern (OR_{MH}=0.83, 95% CI 0.64 to 1.06) when compared to those without the exposure. There was a similar effect of HAP exposure when using multivariable analyses accounting for clustering by site (adjusted OR=0.86, 95%

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 12 of 45

12

0.72 to 1.04). Site-specific analysis showed a wide range of adjusted odds ratios, ranging from

0.46 in Nakaseke, Uruguay to 3.92 in Masindi, Uganda (Table E6).

12

DISCUSSION

We used data from pooled population-based cohorts to examine the association between HAP exposures and COPD outcomes across 13 diverse LMICs. We found a positive association between HAP exposure and COPD outcomes in 12,396 participants, namely a higher overall prevalence and worse disease severity both in terms of symptoms and lung function. This was especially true among women and in participants from Sub-Saharan Africa, for whom 21% and 28% of COPD prevalence was attributed to HAP exposure, respectively. Our data suggests that HAP exposure is likely the leading population attributable risk factor for COPD in our resource-poor settings, even above that of cigarette smoking.

The association between HAP exposure and COPD outcomes have been well studied but with variable results. A meta-analysis of 11 studies found that women and men over age 30 years with HAP exposure had 3.2 and 1.8 times the risk of having COPD than those without the exposure, respectively (6). A more recent meta-analysis of 25 studies found that women with HAP exposure had 2.4 times the odds of having COPD when compared to those without the exposure (7). The reported relationships between HAP exposure and COPD in our pooled analyses for resource-poor settings in LMICs were positive but were more modest in magnitude when compared to the findings of previous two meta-analyses. There are several potential reasons for these different results. First, the above meta-analyses included case-control studies or convenience samples whereas our studies were all population-based. Second, these meta-analyses only included participants who lived in households with high particulate matter concentrations. Third, previous studies used fixed cutoffs to identify COPD, which may under-estimate its prevalence in younger individuals and over-estimate it in older individuals (10, 31, 32). In our analysis, we used the lower fifth percentile of post-bronchodilator FEV_1/FVC to identify COPD, which may capture a more accurate prevalence in the general population.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 14 of 45

In a recent analysis, BOLD investigators did not find an association between biomass fuel smoke exposure and COPD among non-smokers using data from 12 countries when evaluated for the overall study population or when stratified intro high-income countries (HICs) and LMICs (33). In contrast, our study sample is limited to resource-poor settings in LMICs only and did not include studies conducted in HICs. Specifically, HAP exposure in HICs traditionally does not result in same levels of smoke exposure as that observed in LMICs (34). This may be because homes in HICs use biomass fuels mostly for heating with well-ventilated chimneys or stoves in contrast to poorly ventilated open-fire stoves used in LMICs. This may result in important mis-classification of exposure, which could be non-differential in nature. For example, the analysis of BOLD data revealed that 71% of households in Lexington, Kentucky (USA) used biomass fuels when compared to only 46% of households in Ile-Ife (Nigeria). There are other similar examples of a disconnect between HAP exposure in HICs and LMICs.

Pollutants caused by incomplete burning of biomass fuel have been linked to abnormal inflammatory response of the lungs and, thus, COPD (35). HAP exposure triggers a lung-specific and systemic inflammatory state that heightens mechanisms of cell damage, such as oxidative stress (36). Particulate matter, for instance, has been thought to stimulate an inflammatory response involving airway macrophages, neutrophils, and the respiratory epithelium (2). Beyond the direct effect of toxic pollutants on the lungs, HAP exposure affects lung function across the lifespan of an individual. Proposed mechanisms during intrauterine development include deposition of particulate matter in maternal lung tissue resulting in impaired fetal growth, and carbon monoxide exposure may result in reduced oxygen delivery to the fetal placenta (9). HAP exposure may also be associated with a higher prevalence of childhood pneumonia (9). These early life events result in lower baseline lung function in early

14

adulthood and accelerated lung function decline which predisposes individuals to COPD (37, 38).

We found that participants with HAP exposure had a lower pre-bronchodilator FEV_1 at all ages, a lower pre-bronchodilator FEV_1/FVC that became more pronounced at older ages, and a higher odds of having lung function reversibility at younger ages but not at older ages when compared to those without the exposure. These findings support the notion that HAP exposure has deleterious effects on lung function and worsen airflow obstruction that may become non-reversible with older age. The effect of HAP exposure on FVC was not as pronounced, explaining the overall decrease in pre-bronchodilator FEV_1/FVC among those exposed to biomass overtime. A number of cross-sectional studies have found exposure-response relationships between HAP and lung function (39, 40). Longitudinal studies, however, have not demonstrated improved FEV_1 with reduction in HAP exposure, although analysis of lung function has so far been limited due to short follow-up periods (41). Individuals exposed to biomass fuel smoke had higher chances of airway reversibility at younger ages but not at older ages, suggesting that chronic inflammation from HAP exposure is associated with the development of chronic airway disease.

Our analysis has some important strengths. First, we used large and diverse population-based sample with harmonized variables, allowing for the adjustment of *a priori* known risk factors for COPD. Second, we only included studies conducted in LMICs where biomass are commonly burned in poorly ventilated areas (34). Third, we used the lower limit of normal to diagnose COPD instead of a fixed cutoff which could lead to over-diagnosis especially in older participants (42). Fourth, our sensitivity analyses did not identify a single site or subgroup of participants that heavily influenced exposure-outcome relationships. Fifth, the prevalence of

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 16 of 45

HAP exposure in our study sample is consistent with previously published reports of worldwide prevalence (1).

Our analysis also has some potential shortcomings. Our inferences are based on observational data that may be affected by unmeasured confounding. Longitudinal studies with repeated assessments of lung function and exposure to HAP or randomized control trials including experimental elimination of HAP are needed to establish temporal relationships and ultimately causality. As with previous studies, we were unable to quantify direct exposure to biomass beyond self-reported questionnaires. Some of the included study sites, however, have previously published HAP concentrations among those with and without HAP exposure (6). Second, the GLI2012 mixed ethnic reference population may not accurately represent all individuals in our study, which may help explain inconsistent findings. To mitigate this concern, we conducted sensitivity analyses with other reference populations. Third, we were unable to ascertain subject-specific time period of biomass fuel smoke exposure using the available pooled data. However, previous studies in LMICs have reported that number of years of biomass exposure are closely linked to age, particularly among women who use biomass fuels daily for cooking (10). Time or dose-dependent relationships may be reflected in the higher odds of having COPD among women vs. than of men with HAP exposure when compared to those without the exposure. Fourth, we did not have data on occupational exposure history or pack-years of tobacco smoking, which may result in residual confounding. Fifth, HAP exposure is closely linked to a lower socioeconomic status (SES), which is also a known risk factor for COPD (16). For this analysis, we used secondary education which is a proxy for SES but could not harmonize across other factors, which again may result in residual confounding.

In this analysis, we also calculated the population attributable fractions, i.e., the proportional reduction in disease if exposure to a risk factor were mitigated. Accordingly, we estimated that

there would be a 13.5% reduction in COPD prevalence if HAP exposure were eliminated compared to 12.4% if cigarette smoking were eliminated. This finding emphasizes the importance of HAP reduction strategies as public health intervention to reduce the burden of COPD among LMICs. However, to date, a number of trials using cleaner biomass-burning cookstoves aimed at reducing HAP exposure have failed to produce meaningful reductions in HAP exposure. Future intervention trials with clean fuels will ultimately be needed to determine the effect of HAP exposure on multiple health outcomes, including those on lung function and COPD.

Conclusions

We found that HAP exposure was associated with COPD in resource-poor settings of LMICs, and it was associated with both severity of disease and overall lung function. Women were the most affected, and regions like Sub-Saharan Africa may share a disproportionate share of the global burden from this risk factor.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 18 of 45

Acknowledgments

18

References

1. Bonjour S, Adair-Rohani H, Wolf J, Bruce NG, Mehta S, Pruss-Ustun A, Lahiff M, Rehfuess EA, Mishra V, Smith KR. Solid fuel use for household cooking: country and regional estimates for 1980-2010. *Environ Health Perspect* 2013; 121: 784-790.

2. Torres-Duque C, Maldonado D, Pérez-Padilla R, Ezzati M, Viegi G. Biomass fuels and respiratory diseases: a review of the evidence. *Proc Am Thorac Soc* 2008; 5: 577-590.

3. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016; 388: 1659-1724.

4. Hosgood HD, 3rd, Wei H, Sapkota A, Choudhury I, Bruce N, Smith KR, Rothman N, Lan Q. Household coal use and lung cancer: systematic review and meta-analysis of case-control studies, with an emphasis on geographic variation. *Int J Epidemiol* 2011; 40: 719-728.

5. Jary H, Simpson H, Havens D, Manda G, Pope D, Bruce N, Mortimer K. Household Air Pollution and Acute Lower Respiratory Infections in Adults: A Systematic Review. *PLoS One* 2016; 11: e0167656.

6. Kurmi OP, Semple S, Simkhada P, Smith WC, Ayres JG. COPD and chronic bronchitis risk of indoor air pollution from solid fuel: a systematic review and meta-analysis. *Thorax* 2010; 65: 221-228.

7. Po JY, FitzGerald JM, Carlsten C. Respiratory disease associated with solid biomass fuel exposure in rural women and children: systematic review and meta-analysis. *Thorax* 2011; 66: 232-239.

8. Mannino DM, Braman S. The epidemiology and economics of chronic obstructive pulmonary disease. *Proc of the Am Thoracic Soc* 2007; 4: 502-506.

9. Desai MA, Mehta S, Smith KR. Indoor smoke from solid fuels: assessing the environmental burden of disease at national and local levels. Geneva: World Health Organization: 2004 (WHO Environmental Burden of Disease Series, No. 4).

10. Jaganath D, Miranda JJ, Gilman RH, Wise RA, Diette GB, Miele CH, Bernabe-Ortiz A, Checkley W. Prevalence of chronic obstructive pulmonary disease and variation in risk factors across four geographically diverse resource-limited settings in Peru. *Respir Res* 2015; 16: 40.

11. Caballero A, Torres-Duque CA, Jaramillo C, Bolívar F, Sanabria F, Osorio P, Orduz C, Guevara DP, Maldonado D. Prevalence of COPD in five Colombian cities situated at low, medium, and high altitude (PREPOCOL study). *Chest* 2008; 133: 343-349.

12. Miranda JJ, Bernabe-Ortiz A, Smeeth L, Gilman RH, Checkley W, Group CCS. Addressing geographical variation in the progression of non-communicable diseases in Peru: The CRONICAS Cohort Study Protocol. *BMJ Open* 2012; 2: e000610.

13. Rubinstein AL, Irazola VE, Bazzano LA, Sobrino E, Calandrelli M, Lanas F, Lee AG, Manfredi JA, Olivera H, Ponzo J, Seron P, He J. Detection and follow-up of chronic obstructive pulmonary disease (COPD) and risk factors in the Southern Cone of Latin America: the pulmonary risk in South America (PRISA) study. *BMC Pulm Med* 2011; 11: 34.

14. Alam DS, Chowdhury MA, Siddiquee AT, Ahmed S, Clemens JD. Prevalence and Determinants of Chronic Obstructive Pulmonary Disease (COPD) in Bangladesh. *COPD* 2015: 1-10.

15. van Gemert F, Kirenga B, Chavannes N, Kamya M, Luzige S, Musinguzi P, Turyagaruka J, Jones R, Tsiligianni I, Williams S, de Jong C, van der Molen T. Prevalence of chronic obstructive pulmonary disease and associated risk factors in Uganda (FRESH AIR Uganda): a prospective cross-sectional observational study. *Lancet Glob Health* 2015; 3: e44-e51.

16. Grigsby M, Siddharthan T, Chowdhury MA, Siddiquee A, Rubinstein A, Sobrino E, Miranda JJ, Bernabe-Ortiz A, Alam D, Checkley W. Socioeconomic status and COPD among low- and middle-income countries. *Int J Chron Obstruct Pulmon Dis* 2016; 11: 2497-2507.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 20 of 45

17. Bostoen K, Chalabi Z. Optimization of household survey sampling without sample frames. *Int J of Epidemiol* 2006; 35: 751-755.

18. Chao LW, Szrek H, Peltzer K, Ramlagan S, Fleming P, Leite R, Magerman J, Ngwenya GB, Pereira NS, Behrman J. A Comparison of EPI Sampling, Probability Sampling, and Compact Segment Sampling Methods for Micro and Small Enterprises. *J Dev Econ* 2012; 98: 94-107.

19. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J, Initiative ERSGLF. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324-1343.

20. Vestbo J, Hurd SS, Agusti AG, Jones PW, Vogelmeier C, Anzueto A, Barnes PJ, Fabbri LM, Martinez FJ, Nishimura M, Stockley RA, Sin DD, Rodriguez-Roisin R. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013; 187: 347-365.

21. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, Van Weel C. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007; 176: 532-555.

22. Godfrey MS, Jankowich MD. The Vital Capacity Is Vital: Epidemiology and Clinical Significance of the Restrictive Spirometry Pattern. *Chest* 2016; 149: 238-251.

23. Carey V, Zeger SL, Diggle P. Modelling multivariate binary data with alternating logistic regressions. *Biometrika* 1993; 80: 517-526.

24. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982; 38: 963-974.

25. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; 22: 719-748.

26. Levin ML. The occurrence of lung cancer in man. *Acta Unio Int Contra Cancrum* 1953; 9: 531-541.

27. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw* 2014.

28. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *J Am Stat Assoc* 1993; 88: 9-25.

29. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.; 2016.

30. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag 2016.

31. Buist AS, McBurnie MA, Vollmer WM, Gillespie S, Burney P, Mannino DM, Menezes AM, Sullivan SD, Lee TA, Weiss KB. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet* 2007; 370: 741-750.

32. Menezes AMB, Perez-Padilla R, Jardim JB, Muiño A, Lopez MV, Valdivia G, de Oca MM, Talamo C, Hallal PC, Victora CG. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. *Lancet* 2005; 366: 1875-1881.

33. Lamprecht B, McBurnie MA, Vollmer WM, Gudmundsson G, Welte T, Nizankowska-Mogilnicka E, Studnicka M, Bateman E, Anto JM, Burney P. COPD in never smokers: results from the population-based burden of obstructive lung disease study. *Chest* 2011; 139: 752-763.

34. Bryden KM, Bruce N, Peck M. WHO indoor air quality guidelines: household fuel combustion. 2015.

35. Mannino D, Watt G, Hole D, Gillis C, Hart C, McConnachie A, Smith GD, Upton M, Hawthorne V, Sin D. The natural history of chronic obstructive pulmonary disease. *Eur Respir J* 2006; 27: 627-643.

36. Silva R, Oyarzún M, Olloquequi J. Pathogenic mechanisms in chronic obstructive pulmonary disease due to biomass smoke exposure. *Arch Bronconeumol* 2015; 51: 285-292.

37. Manuck TA, Levy PT, Gyamfi-Bannerman C, Jobe AH, Blaisdell CJ. Prenatal and perinatal determinants of lung health and disease in early life: a National Heart, Lung, and Blood Institute Workshop Report. *JAMA Pediatr* 2016; 170: e154577-e154577.

20

38. Lange P, Celli B, Agustí A, Boje Jensen G, Divo M, Faner R, Guerra S, Marott JL, Martinez FD, Martinez-Camblor P. Lung-function trajectories leading to chronic obstructive pulmonary disease. *N Eng J Med* 2015; 373: 111-122.

39. Regalado J, Pérez-Padilla R, Sansores R, Paramo Ramirez JI, Brauer M, Paré P, Vedal S. The effect of biomass burning on respiratory symptoms and lung function in rural Mexican women. *Am J Respir Crit Care Med* 2006; 174: 901-905.

40. Dave M, Ahankari AS, Myles PR, Arokiasamy P, Khobragade P, Mortimer K, Fogarty AW. Household air pollution and lung function in Indian adults: a cross-sectional study. *Int J Tuberc Lung Dis* 2017; 21: 702-704.

41. Clark ML, Peel JL, Burch JB, Nelson TL, Robinson MM, Conway S, Bachand AM, Reynolds SJ. Impact of improved cookstoves on indoor air pollution and adverse health effects among Honduran women. *Int J Environ Health Res* 2009; 19: 357-368.

42. Swanney MP, Ruppel G, Enright PL, Pedersen OF, Crapo RO, Miller MR, Jensen RL, Falaschetti E, Schouten JP, Hankinson JL, Stocks J, Quanjer PH. Using the lower limit of normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction. *Thorax* 2008; 63: 1046-1051.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 22 of 45

Table 1: Sociodemographic characteristics stratified by site. Mean and SD pack-years smoked was calculated among individuals who smoked.

	Bariloche	Canelones	Dhaka	Kampala	Lima	Marcos Paz	Masindi	Matlab	Nakaseke	Rural Puno	Temuco	Tumbes	Urban Puno	Total
Number of people	1099	851	1672	596	997	1236	414	1824	721	500	1038	945	503	12396
COPD positive, % (n)	7.1 (78)	9.3 (79)	9.7 (162)	1.7 (10)	5.6 (56)	13.8 (171)	15.5 (64)	15.3 (279)	7.4 (53)	8.8 (44)	5.2 (54)	1.7 (16)	5.0 (25)	8.8 (1091)
Age in years, mean (SD)	57.6 (7.8)	59.6 (8.5)	51.8 (9.3)	44.2 (8.9)	54.9 (11.8)	58.8 (8.2)	49.7 (12.5)	55.1 (10.6)	49.1 (11.2)	55.5 (12.5)	59.2 (8.5)	55.5 (13.1)	55.2 (12.1)	54.9 (10.9)
Height in centimeters, mean (SD)	161.2 (9.8)	162.2 (9.3)	155.5 (9.0)	162.1 (8.7)	154.8 (8.5)	161.7 (9.4)	161.3 (9.3)	153.9 (8.3)	159.9 (8.0)	155.4 (8.0)	160.3 (9.1)	158.4 (8.7)	156.9 (9.0)	158.2 (9.4)
Number of males, % (n)	61.2 (673)	61.7 (525)	45.5 (761)	48.7 (290)	49.2 (491)	59.9 (740)	50.0 (207)	46.0 (839)	45.9 (331)	47.0 (235)	54.7 (568)	50.2 (474)	49.5 (249)	51.5 (6383)
Education ≥ secondary, % (n)	49.9 (548)	49.8 (424)	44.1 (738)	20.8 (124)	45.7 (456)	34.3 (424)	16.9 (70)	18.8 (342)	7.1 (51)	33.4 (167)	70.7 (734)	36.0 (340)	65.2 (328)	38.3 (4746)
Biomass as primary source of fuel, % (n)	23.1 (254)	5.1 (43)	3.6 (61)	93.5 (557)	2.4 (24)	0.5 (6)	92.8 (384)	98.1 (1789)	99.6 (718)	95.4 (477)	22.4 (233)	16.7 (158)	2.0 (10)	38.0 (4714)
Daily smokers, % (n)	23.9 (263)	21.7 (185)	6.5 (108)	8.7 (52)	3.2 (32)	22.7 (281)	36.2 (150)	9.0 (164)	6.9 (50)	0.2 (1)	14.9 (155)	5.6 (53)	2.2 (11)	12.1 (1505)
Pack-years smoked, mean (SD)	23.0 (18.3)	35.2 (30.0)	20.6 (18.0)	8.7 (9.4)	2.7 (9.0)	31.4 (22.1)	6.7 (13.1)	19.8 (15.9)	6.5 (7.7)	1.3 (4.3)	13.0 (14.2)	5.3 (12.9)	2.7 (7.2)	4.6 (12.7)
BMI (kg/m^2), mean (SD)	29.0 (5.8)	29.6 (6.2)	24.6 (4.9)	26.0 (5.3)	28.4 (4.4)	30.1 (5.8)	22.9 (4.3)	20.6 (3.6)	23.8 (4.5)	25.2 (3.7)	29.1 (4.9)	28.3 (4.7)	27.9 (4.4)	
BMI (kg/m^2), % (n)														26.4 (5.8)
0-18.5	0.9 (10)	1.1 (9)	10.3 (173)	3.4 (20)	0.1 (1)	0.5 (6)	9.4 (39)	30.5 (556)	8.2 (59)	1.6 (8)	0.5 (5)	0.4 (4)	0.6 (3)	7.2 (893)
18.5-25	25.0 (275)	20.0 (170)	44.7 (748)	46.5 (277)	22.0 (219)	17.9 (221)	67.1 (278)	58.3 (1063)	61.0 (440)	50.4 (252)	16.7 (173)	23.7 (224)	23.3 (117)	36.0 (4457)
25-30	36.2 (398)	36.0 (306)	32.7 (546)	28.9 (172)	45.4 (453)	35.1 (434)	17.4 (72)	9.6 (176)	21.1 (152)	37.2 (186)	46.2 (480)	44.3 (419)	49.7 (250)	32.6 (4044)
30+	37.9 (416)	43.0 (366)	12.3 (205)	21.3 (127)	32.5 (324)	46.5 (575)	6.0 (25)	1.6 (29)	9.7 (70)	10.8 (54)	36.6 (380)	31.5 (298)	26.4 (133)	24.2 (3002)

Figure 1. Typical kitchen and stoves in selected sites. Top row, from left to right: Puno Peru; Lima, Peru; Tumbes, Peru; and Nakaseke, Uganda. Bottom row, from left to right: Kampala, Uganda; Dhaka, Bangladesh; Matlab, Bangladesh; and Temuco, Chile.

23

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 24 of 45

Figure 2. Sex-, site- and severity-stratified prevelances of COPD. The prevalence of COPD was stratified by sex (women on the left, men on the right) and severity as defined by lung function (in shades of grey) across the 13 LMIC sites. Sites were ordered according to the overall prevalence of COPD from lowest (top) to highest (bottom). Overall sex-stratified site-specific prevalences are given next to each bar.

Figure 3. Prevalence and corresponding 95% confidence intervals of COPD and restricted spirometric pattern by deciles of age stratified by household air pollution (HAP) exposure. We calculated point prevalences of COPD (left panel) and restricted spirometric pattern (right panel) at each age decile and by HAP exposure status. Values on the x-axis represent the starting age for each decile. HAP exposure status was stratified according to participants who reported using biomass as the predominant fuels (B) and those who did not use biomass as the predominant fuel (C). We used smoothing splines to describe the relationship between HAP exposure status and COPD or restricted spirometry pattern prevalence across age deciles. The broken lines summarize trends for participants with HAP exposure (B) and continuous lines summarize trends for those without the exposure (C).

Figure 4. Associations between household air pollution (HAP) exposure and COPD outcomes obtained from multivariable regression models, and interaction effects with sex, smoking status, age, and educational attainment. Panel A shows estimates using data from all sites, while panel B shows site-specific estimates. In panel A, odds ratios and the corresponding 95% confidence intervals are represented by diamonds and lines, respectively. We also tabulated numerical values for the odds ratios and the corresponding 95% confidence intervals. In panel B, site-specific odds ratios are presented by triangles. In the overall model, we evaluated the association between HAP exposure and COPD prevalence adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education. We then evaluated for interaction effects between HAP exposure and either sex, smoking status, age, or educational attainment on COPD outcomes. Models stratified by sex were adjusted for age, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education. Models stratified by smoking status were adjusted for age, sex, body mass index, post-treatment pulmonary tuberculosis, and secondary education. Models stratified by age were adjusted for sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education. Models stratified by educational attainment were adjusted for age, sex, daily cigarette smoking, body mass index, and post-treatment pulmonary tuberculosis. Models with severity and symptom status of COPD as outcomes were adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education.

Figure 5. Mean pre-bronchodilator FEV$_1$, FVC and FEV$_1$/FVC Z-scores with corresponding 95% confidence intervals, stratified by deciles of age and by household air pollution (HAP) exposure status. We calculated average Z-scores for FEV$_1$ (left panel), FVC (middle panel) or FEV$_1$/FVC (right panel) at each age decile and by household air pollution exposure status. Values on the x-axis represent the starting age for each decile. HAP exposure status was stratified according to participants who reported using biomass as the predominant fuel (B) and those who did not use biomass as the predominant fuel (C). The broken lines summarize trends for participants with HAP exposure (B) while the continuous lines summarize trends for those without the exposure (C).

27

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 28 of 45

Figure 1. Typical kitchen and stoves in selected sites. Top row, from left to right: Puno Peru; Lima, Peru; Tumbes, Peru; and Nakaseke, Uganda. Bottom row, from left to right: Kampala, Uganda; Dhaka, Bangladesh; Matlab, Bangladesh; and Temuco, Chile.

109x66mm (600 x 600 DPI)

Figure 2. Sex-, site- and severity-stratified prevelances of COPD. The prevalence of COPD was stratified by sex (women on the left, men on the right) and severity as defined by lung function (in shades of grey) across the 13 LMIC sites. Sites were ordered according to the overall prevalence of COPD from lowest (top) to highest (bottom). Overall sex-stratified site-specific prevalences are given next to each bar.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 30 of 45

Figure 3. Prevalence and corresponding 95% confidence intervals of COPD and restricted spirometric pattern by deciles of age stratified by household air pollution (HAP) exposure. We calculated point prevalences of COPD (left panel) and restricted spirometric pattern (right panel) at each age decile and by HAP exposure status. Values on the x-axis represent the starting age for each decile. HAP exposure status was stratified according to participants who reported using biomass as the predominant fuels (B) and those who did not use biomass as the predominant fuel (C). We used smoothing splines to describe the relationship between HAP exposure status and COPD or restricted spirometry pattern prevalence across age deciles. The broken lines summarize trends for participants with HAP exposure (B) and continuous lines summarize trends for those without the exposure (C).
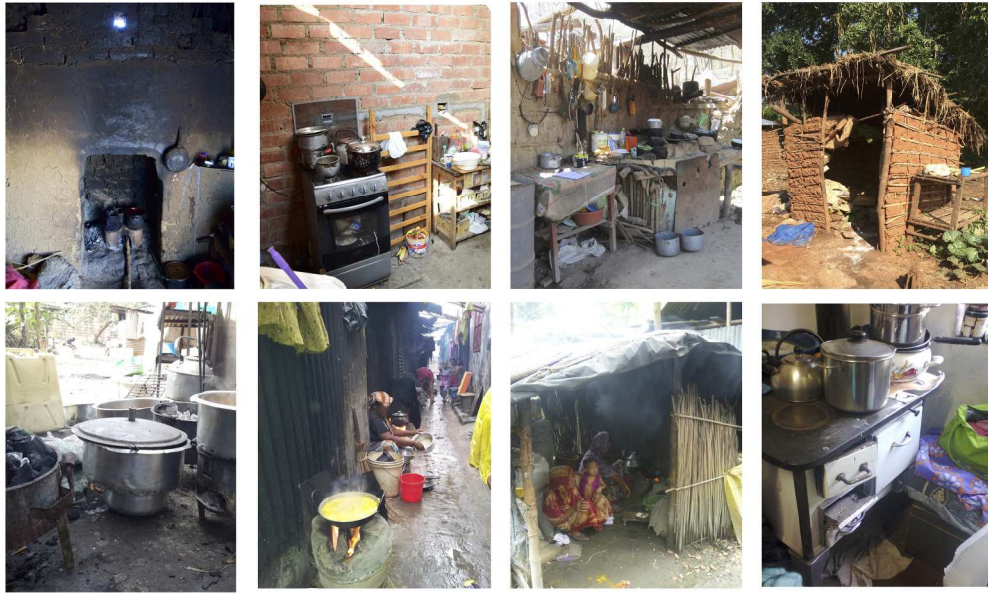
A All-sites · **B** Site-specific

Overall	1.41 (1.18, 1.68)	
Women	1.70 (1.24, 2.32)	
Men	1.21 (0.92, 1.58)	
Smoke daily	1.36 (1.01, 1.84)	
Do not smoke daily	1.38 (1.09, 1.73)	
Age ≥ 55 years	1.43 (1.09, 1.87)	
Age ≤ 55 years	1.32 (0.97, 1.78)	
Secondary education or higher	1.62 (1.23, 2.12)	
Primary education or lower	1.25 (0.99, 1.57)	
Severity of COPD	1.49 (1.15, 1.95)	
Symptomatic COPD	1.51 (1.16, 1.97)	

Odds Ratio of having COPD

Figure 4. Associations between household air pollution (HAP) exposure and COPD outcomes obtained from multivariable regression models, and interaction effects wit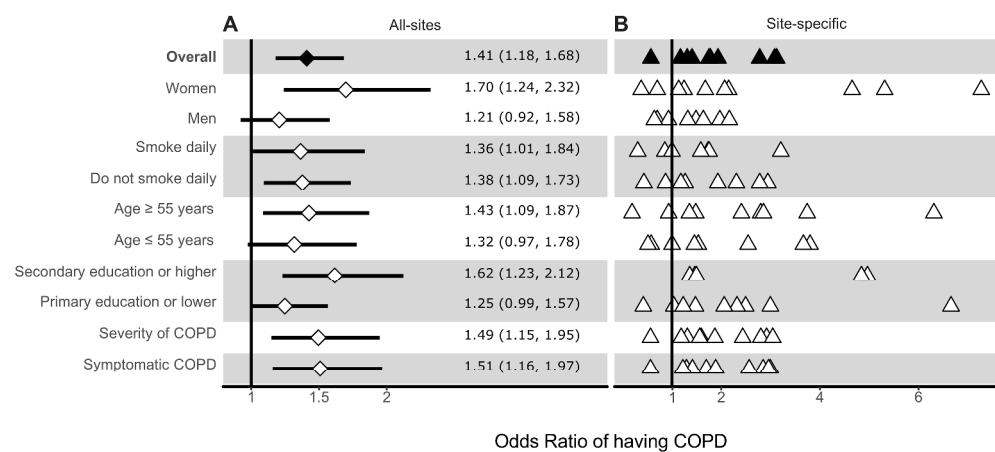h sex, smoking status, age, and educational attainment. Panel A shows estimates using data from all sites, while panel B shows site-specific estimates. In panel A, odds ratios and the corresponding 95% confidence intervals are represented by diamonds and lines, respectively. We also tabulated numerical values for the odds ratios and the corresponding 95% confidence intervals. In panel B, site-specific odds ratios are presented by triangles. In the overall model, we evaluated the association between HAP exposure and COPD prevalence adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education. We then evaluated for interaction effects between HAP exposure and either sex, smoking status, age, or educational attainment on COPD outcomes. Models stratified by sex were adjusted for age, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education. Models stratified by smoking status were adjusted for age, sex, body mass index, post-treatment pulmonary tuberculosis, and secondary education. Models stratified by age were adjusted for sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education. Models stratified by educational attainment were adjusted for age, sex, daily cigarette smoking, body mass index, and post-treatment pulmonary tuberculosis. Models with severity and symptom status of COPD as outcomes were adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis, and secondary education.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 32 of 45

Figure 5. n pre-bronchodilator FEV1, FVC and FEV1/FVC Z-scores with corresponding 95% confidence intervals, stratified by deciles of age and by household air pollution (HAP) exposure status. We calculated average Z-scores for FEV1 (left panel), FVC (middle panel) or FEV1/FVC (right panel) at each age decile and by household air pollution exposure status. Values on the x-axis represent the starting age for each decile. HAP exposure status was stratified according to participants who reported using biomass as the predominant fuel (B) and those who did not use biomass as the predominant fuel (C). The broken lines summarize trends for participants with HAP exposure (B) while the continuous lines summarize trends for those without the exposure (C).

Page 33 of 45

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Online supplement for: "Association between household air pollution exposure and chronic obstructive pulmonary disease outcomes in 13 low- and middle-income country settings"

Authors: Trishul Siddharthan MD, Matthew R Grigsby MSPH, Dina Goodman MSPH, Muhammad Chowdhury MBBS, Adolfo Rubinstein MD PhD, Vilma Irazola MD, Laura Gutierrez MD, J Jaime Miranda MD PhD, Antonio Bernabe-Ortiz MD MPH, Dewan Alam MBBS PhD, Bruce Kirenga MBChB, Rupert Jones MBBS MD, Frederick van Gemert MD, Robert A Wise MD, William Checkley MD PhD.

Study participants and settings

A total of 13,023 participants were enrolled at 13 sites (Figure E1) but 878 were excluded due to having an age < 35 or > 95 years, missing lung function or missing information on self-reported biomass fuel use. A total of 12,396 participants were included in final analysis. An additional 251 individuals had missing information on age, sex, site, BMI, secondary education, daily cigarette use or history of post-treatment tuberculosis (n.b., some individuals were missing more than one variable). Participants were enrolled in studies in Latin America, Africa, and Asia (Figure E2).

HAP exposure and COPD outcomes

Pooled analyses: Overall, the majority of odds ratios in Table E1-E4 show a positive association between HAP exposure and the odds of having COPD, suggesting that pooling odds ratios is a reasonable approach for our study. We used alternating logistic regression (ALR) to adjust for correlated outcomes within site. ALR approach uses (log) odds ratio instead of correlations as does GEE to adjust for lack of independent units.

Using different approaches to account for heterogeneity: Because of the inherent clustered nature of these data, we performed sensitivity analysis by comparing three techniques designed for such situations. In addition to results from ALR models presented in the manuscript, we also confirmed our results using multivariable random-effects (RE) logistic regression and Mantel-Haenszel (MH) techniques. We found consistent positive associations between HAP and COPD across all modeling methods. Specifically, individuals who were exposed to HAP had higher odds of COPD by 1.41 (95% CI 1.18 to 1.68) using ALR, by 1.50 (95% CI 1.15 to 1.95) using RE, and 1.68 (95% 1.29 to 2.19) using MH, when compared to people without HAP exposure.

Using different reference populations: When using Global Lung Function Initiative 2012 Caucasian reference populations, we found that those with HAP exposure were 1.39 times more likely (95% CI 1.21 - 1.60) to have COPD when compared to those without the exposure, even after adjusting for age, sex, daily cigarette smoking, body mass index, po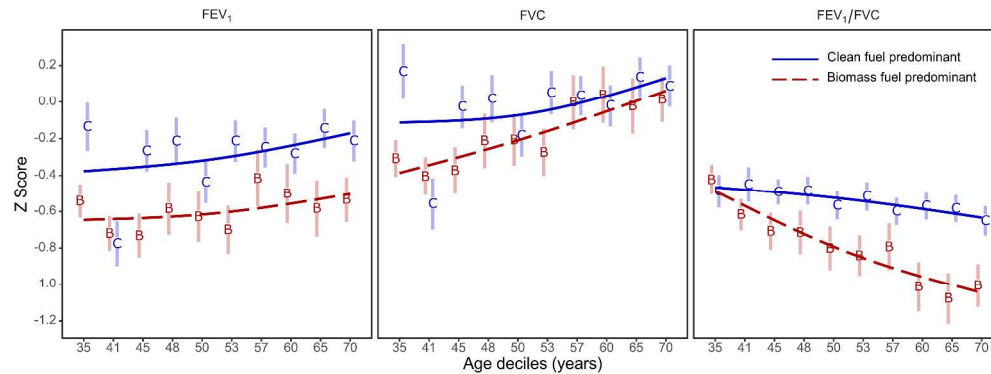st-treatment pulmonary tuberculosis and secondary education. We also calculated site-specific ORs using Fisher's exact test and the MH test for both GLI mixed ethnic and GLI Caucasian in Table E1 and Table E2. These estimates show consistency and speak to the robustness of the association between HAP exposure and COPD.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 34 of 45

Adjusting for pack-years smoked: Using daily smoking as an indicator for smoking exposure leaves the possibility of residual confounding because it is not a frequency-based measure. To make sure residual confounding was not present in our analysis, we replaced daily smoking with pack-years smoked and analyzed the relationship between HAP exposure and COPD while also adjusting for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis and secondary education (Table E3). Participants exposed to HAP were more likely to have COPD (OR_{ALR}=1.42, 95% CI 1.18 to 1.72) compared to those who were not exposed. In further sex-stratified analysis, HAP exposure was associated with higher odds of having COPD in women (OR_{ALR}=1.73, 95% CI 1.25 to 2.40) but not in men (OR_{ALR}=1.20, 95% CI 0.91 to 1.57). These outcomes are highly consistent to our results when using daily smoking, adding validity our hypothesis of HAP exposure causing increased risk of COPD and suggesting that our results are not affected by residual confounding due to frequency of cigarette smoking exposure. In addition, we found that for every five pack-years smoked, the overall odds of someone having COPD increased by 12.7% (95% CI 1.0% to 15.2%).

Leave-one-out sites and subgroup analysis: To assess the potential that some sites a have heavier influence on COPD prevalence than others, we subset our data to sites with prevalence of HAP exposure between 5% and 95%. Sites that met these criteria were Bariloche, Canelones, Kampala, Masindi, Temuco, and Tumbes. In this limited subgroup analysis, we found that participants with HAP exposure were 1.40 times more likely (95% CI 1.06 - 1.84) to have COPD when compared to those without the exposure. We also excluded sites with less than five observations in any cell when calculating site-specific contingency tables for COPD and biomass use (Table E2). Sites that met this criterion were Bariloche, Canelones, Dhaka, Masindi, Temuco, and Tumbes. Adjusted analysis with these sites showed that participants with HAP exposure were 1.43 (95% CI 1.17 - 1.74) times more likely to have COPD than those without the exposure. These findings are consistent with our overall adjusted estimate (adjusted OR = 1.41; 95% CI 1.18 - 1.68) and add strength to the argument of an association between HAP exposure and COPD.

Testing the proportionality of odds: To test whether the proportionality of odds assumption held true in our ordinal logistic regression models, we graphically assessed odds ratios with 95% confidence intervals for each independent variable regressed at each dependent variable cutoff point of each ordinal scale for COPD severity (i.e., none, mild, moderate and severe/severe COPD according to the GOLD strategy) and symptomatic COPD (none, asymptomatic COPD, symptomatic COPD). As seen in Figures E3 and E4, there is clear overlap between 95% confidence intervals between cutoffs for all variables, suggesting the proportionality of odds assumption is reasonable for our analysis.

HAP exposure and lung function
When using Global Lung Function Initiative 2012 Caucasian reference population equations instead of using those for the mixed ethnic reference population, participants with HAP exposure had: a lower adjusted pre-bronchodilator FEV_1 Z-score (-0.10 SD,

Page 35 of 45

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

95% CI -0.20 to 0.03 SD) than those without exposure across all ages, with no interaction effect between HAP exposure and age (p=0.08); a lower pre-bronchodilator FVC Z-score at age 35 years (-0.13 SD, -0.25 to -0.002 SD) and slightly larger at age 60 years (0.10 SD, 0.006 to 0.21 SD); and, no difference in pre-bronchodilator FEV_1/FVC Z-score at age 35 years (0.02 SD, -0.09 to 0.13 SD) but a lower pre-bronchodilator FEV_1/FVC Z-score at 60 years (-0.26 SD, -0.35 to -0.17 SD).

HAP exposure and restrictive spirometric patterns

While the overall adjusted (OR=0.86, 95% 0.72 to 1.04) and site weighted (OR_{MH}=0.83, 95% CI 0.64 to 1.06) odds of restrictive spirometric patterns were similar, sensitivity analysis revealed a wide range of results across sites. Specifically, in Table E6 we show that adjusted odds ranged from 0.46 in Nakaseke, Uganda, to 3.92 in Masindi, Uganda.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 36 of 45

Table E1: Two-by-two contingency tables of the association between household air pollution (HAP) exposure and COPD when using the Global Lung Function Initiative 2012 mixed ethnic reference population. Fisher exact odds ratios (OR) and 95% confidence interval calculations were calculated per site, while the overall estimate and 95% confidence interval was calculated using the Cochran-Mantel-Haenszel (MH) chi-squared test. COPD was defined as having $FEV_1/FVC \leq 5^{th}$ percentile (lower limit of normal) based on the Global Lung Function Initiative 2012 mixed ethnic reference populations. We were unable to calculate the odds of COPD for some sites with too few observations in cells of the contingency table.

Site	HAP+ve, COPD-ve	HAP+ve, COPD+ve	HAP-ve, COPD-ve	HAP-ve, COPD+ve	Fisher exact OR	Exact 95% CI
Bariloche	233	21	788	57	1.25	0.70 - 2.14
Canelones	34	9	738	70	2.79	1.13 - 6.24
Dhaka	49	12	1,461	150	2.38	1.13 - 4.67
Kampala	547	10	39	0	NA	NA
Lima	22	2	919	54	1.55	0.17 - 6.57
Marcos Paz	4	2	1,061	169	3.13	0.28 - 22.1
Masindi	327	57	23	7	0.57	0.22 - 1.66
Matlab	1,513	276	32	3	1.95	0.60 - 10.0
Nakaseke	665	53	3	0	NA	NA
Rural Puno	433	44	23	0	NA	NA
Temuco	217	16	767	38	1.49	0.76 - 2.79
Tumbes	152	6	777	10	3.06	0.90 - 9.47
Urban Puno	9	1	469	24	2.17	0.05 - 16.8
					OR_{MH}	95% CI
Overall	4205	509	7100	582	1.68	1.29 - 2.19

Page 37 of 45

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Table E2: Two-by-two tables for the association between household air pollution (HAP) exposure and COPD when using the Global Lung Function Initiative 2012 Caucasian reference population. Fisher exact odds ratio (OR) and 95% confidence interval calculations were done per site, while the overall estimate and 95% confidence interval was calculated using Cochran-Mantel-Haenszel (MH) chi-squared test for count data. COPD was defined as having $FEV_1/FVC \leq 5^{th}$ percentile (lower limit of normal) based on the Global Lung Function Initiative 2012 Caucasian reference populations. We were unable to calculate the odds of COPD for some sites with too few observations in the cells of the contingency table.

Site	HAP+ve, COPD-ve	HAP+ve, COPD+ve	HAP-ve, COPD-ve	HAP-ve, COPD+ve	Fisher exact OR	Exact 95% CI
Bariloche	237	17	805	40	1.44	0.75 - 2.66
Canelones	37	6	747	61	1.98	0.66 - 5.00
Dhaka	50	11	1,475	136	2.38	1.09 - 4.78
Kampala	548	9	39	0	NA	NA
Lima	23	1	937	36	1.13	0.03 - 7.38
Marcos Paz	4	2	1,091	139	3.92	0.35 - 27.6
Masindi	342	42	25	5	0.61	0.22 - 2.17
Matlab	1,541	248	32	3	1.72	0.53 - 8.83
Nakaseke	675	43	3	0	NA	NA
Rural Puno	449	28	23	0	NA	NA
Temuco	220	13	777	28	1.64	0.77 - 3.34
Tumbes	155	3	782	5	3.02	0.46 - 15.7
Urban Puno	9	1	474	19	2.76	0.06 - 21.8
					OR$_{MH}$	MH 95% CI
Overall	4290	424	7210	472	1.72	1.27 - 2.32

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 38 of 45

Table E3. Odds ratio of COPD among those exposed to HAP versus those where who were not, stratified by variable used to adjust for smoking status (daily smoking and pack years smoked) and characteristics of interest (sex, age, education, severity of COPD, and symptomatic COPD).

Association between biomass in	Odds ratios adjusted for daily smoking	Odds ratios adjusted for pack-years smoked
Overall	1.41 (1.18 – 1.68)	1.42 (1.18 – 1.72)
Women	1.70 (1.24 – 2.32)	1.73 (1.25 – 2.40)
Men	1.21 (0.92 – 1.58)	1.20 (0.91 – 1.57)
Age ≥55 years	1.43 (1.09 – 1.87)	1.43 (1.09 – 1.87)
Age <55 years	1.32 (0.97 – 1.78)	1.41 (1.02 – 1.95)
Secondary education or higher	1.62 (1.23 – 2.12)	1.65 (1.21 – 2.25)
Primary education	1.25 (0.99 – 1.57)	1.28 (1.02 – 1.61)
Severity of COPD	1.51 (1.16 – 1.96)	1.50 (1.16 – 1.96)
Symptomatic COPD	1.50 (1.15 – 1.97)	1.51 (1.16 – 1.97)

Table E4: Odds ratios (OR) for having COPD among participants with HAP exposure when compared to those without the exposure, stratified by which site was left out of the analysis. The alternating logistic regressions (ALR) approach was used to calculate adjusted odds ratios accounting for clustering and Cochran-Mantel-Haenszel (MH) chi-squared tests for count data were used to calculate the common odds ratio to investigate the possibility of a single site having heavy influence on the overall primary outcome. ALR models were adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis and secondary education.

Site removed	OR_{ALR}	95% CI	OR_{MH}	95% CI
Bariloche	1.49	1.22 - 1.82	1.87	1.37 - 2.56
Canelones	1.36	1.14 - 1.63	1.59	1.20 - 2.11
Dhaka	1.45	1.18 - 1.78	1.59	1.19 - 2.12
Kampala	1.45	1.22 - 1.71	1.66	1.27 - 2.17
Lima	1.39	1.16 - 1.67	1.68	1.28 - 2.21
Marcos Paz	1.44	1.18 - 1.76	1.66	1.27 - 2.17
Masindi	1.46	1.22 - 1.75	1.86	1.41 - 2.45
Matlab	1.42	1.18 - 1.71	1.66	1.26 - 2.18
Nakaseke	1.43	1.20 - 1.70	1.87	1.28 - 2.19
Rural Puno	1.37	1.15 - 1.63	1.59	1.24 - 2.13
Temuco	1.38	1.12 - 1.69	1.59	1.28 - 2.33
Tumbes	1.31	1.09 - 1.58	1.66	1.23 - 2.13
Urban Puno	1.40	1.17 - 1.68	1.68	1.28 - 2.19

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 40 of 45

Table E5: Odds ratios (OR) for having COPD among participants with HAP exposure when compared to those without the exposure, stratified by which fold of data was removed from the analysis. The alternating logistic regressions (ALR) approach was used to calculate adjusted odds ratios and Cochran-Mantel-Haenszel (MH) chi-squared tests for count data were used to calculate the common odds ratio to investigate the possibility of a single subset of individuals having heavy influence on the overall primary outcome. ALR models were adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis and secondary education.

Fold removed	OR_{ALR}	95% CI	OR_{MH}	95% CI
1	1.42	1.20 - 1.69	1.67	1.26 - 2.20
2	1.39	1.14 - 1.69	1.63	1.23 - 2.15
3	1.42	1.15 - 1.74	1.69	1.26 - 2.26
4	1.43	1.16 - 1.75	1.66	1.25 - 2.2
5	1.47	1.20 - 1.82	1.75	1.33 - 2.32
6	1.41	1.13 - 1.77	1.60	1.21 - 2.11
7	1.50	1.23 - 1.83	1.80	1.36 - 2.39
8	1.39	1.16 - 1.67	1.65	1.25 - 2.19
9	1.43	1.19 - 1.73	1.69	1.28 - 2.24
10	1.44	1.19 - 1.74	1.66	1.25 - 2.20

Table E6: Site-specific odds ratios for having restrictive spirometric pattern among participants with HAP exposure when compared to those without the exposure. Multivariable logistic regression was used to calculate adjusted odds ratios. All models were adjusted for age, sex, daily cigarette smoking, body mass index, post-treatment pulmonary tuberculosis and secondary education. We were unable to calculate the odds of restrictive spirometric pattern for some sites with too few observations.

Site	Odds Ratio	95% CI
Masindi	3.92	0.50 - 30.5
Tumbes	1.41	0.61 - 3.27
Bariloche	1.01	0.60 - 1.68
Temuco	0.87	0.41 - 1.81
Kampala	0.69	0.27 - 1.76
Matlab	0.55	0.24 - 1.26
Dhaka	0.55	0.29 - 1.03
Canelones	0.49	0.19 - 1.26
Nakaseke	0.46	0.04 - 5.37
Urban Puno	NA	NA
Marcos Paz	NA	NA
Lima	NA	NA

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 42 of 45

Figure E1. Flow chart of study participants included in pooled analysis.

Participants enrolled in all
13 sites
(n=13,023)

Exclusion Criteria:
- Age <35 years (n=192)
- Age >95 years (n=3)

Participants eligible for
analysis
(n=12,828)

Missing lung function values (n=13)

Participants with complete
outcome data
(n=12,815)

Missing biomass fuel use data (n=213)

Participants with complete
exposure data
(n=12,602)

Missing other information (n=206)
- BMI (n=2)
- Secondary education (n=5)
- Daily cigarette use (n=3)
- Post treatment pulmonary tuberculosis
(n=199)

Participants included in
analyses
(n=12,396)

Figure E2: World map with study sites in Bariloche and Marcos Paz, Argentina; Canelones, Uruguay; Dhaka and Matlab, Bangladesh; Kampala, Masindi and Nakaseke, Uganda; Lima, Puno and Tumbes, Peru; and, Temuco, Chile.

AJRCCM Articles in Press. Published on 11-January-2018 as 10.1164/rccm.201709-1861OC

Page 44 of 45

Figure E3: Evaluation of the proportionality of odds assumption across categories of COPD severity by lung function. Odds ratios and the corresponding 95% confidence intervals are represented by diamonds and lines, respectively. These odds ratios were obtained from single variable logistic regression models using alternating logistic regression and calculated at each ordinal cutoff (mild, moderate and severe/very severe COPD vs. none as reference). Sex, educational attainment, household air pollution (HAP) exposure and daily cigarette smoking were included as dichotomous variables, whereas body mass index and age were included as continuous variable (and interquartile odds ratios were calculated).
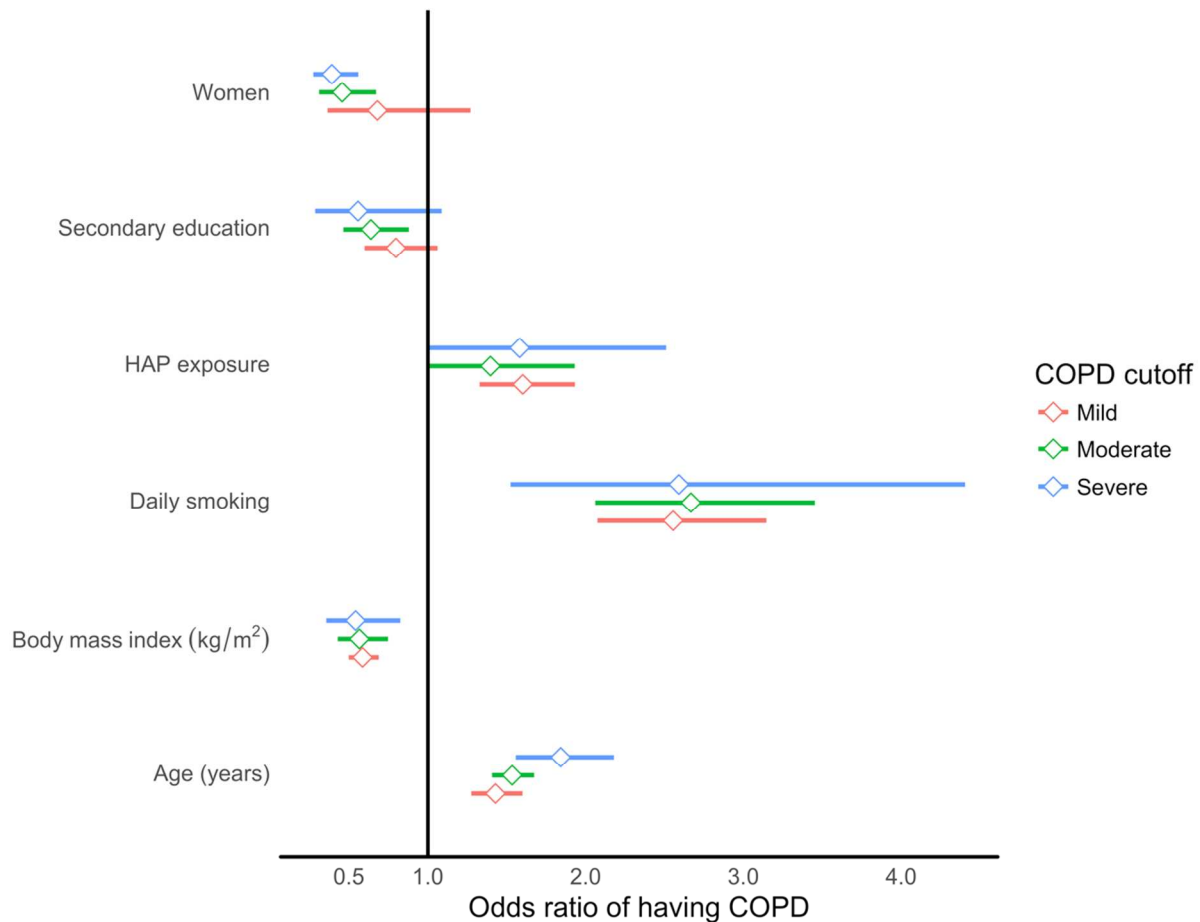
Figure E4: Evaluation of the proportionality of odds assumption across categories of COPD severity by symptoms. Odds ratios and the corresponding 95% confidence intervals are represented by diamonds and lines, respectively. These odds ratios were obtained from single variable logistic regression models using alternating logistic regressions and calculated at each ordinal cutoff (symptomatic and asymptomatic COPD vs. none as reference). Sex, educational attainment, household air pollution (HAP) exposure and daily cigarette smoking were included as dichotomous variables, whereas body mass index and age were included as continuous variable (and interquartile odds ratios were calculated).