## Online Supplement for:

# Influence of Country-Level Differences on COPD Prevalence 

Shawn D. Aaron, Andrea Gershon, Yuan Gao, Jenna Yang, G. A. Whitmore For The Canadian Respiratory Research Network (CRRN)

Section 1 of this online supplement describes the PubMed literature search and data sources. Section 2 presents a table of summary statistics for the data. Section 3 provides information about data analysis and statistical computation. It includes a fragment of data for one qualified study to illustrate the formatting used for data records. Section 4 gives the reference list for the 80 qualified articles of our study. The Appendix contains a fragment of Stata computer code for the stepwise binomial regression analysis used in this study.

## 1. PubMed Literature Search and Data Sources

## PubMed literature search

Table S1 presents a list of our initial keyword searches in PubMed for qualified articles. The table shows both the keyword combinations and the number of cited articles obtained in each search. Observe that the searches, by design, are not overlapping.

Table S1. List of initial keyword searches in PubMed for qualified articles.

| Search | Key Words | Combination | Citations |
| :---: | :--- | :---: | :---: |
| 1 | COPD prevalence spirometry GOLD population | $(11111)$ | 85 |
| 2 | ((COPD prevalence spirometry population) NOT GOLD) | $(11101)$ | 288 |
| 3 | $(($ COPD prevalence GOLD population) NOT spirometry $)$ | $(11011)$ | 83 |
| 4 | $(($ COPD prevalence spirometry GOLD) NOT population) | $(11110)$ | 98 |
| 5 | $((($ COPD prevalence GOLD) NOT spirometry) NOT population) | $(11010)$ | 167 |
| Total |  |  | $\mathbf{7 2 1}$ |

## Study-level data

Table S2 shows the number of countries and reference numbers for the 80 qualified articles, classified by world region. Literature citations for the articles are presented alphabetically at the end of this online supplement because they are too numerous to be listed in the main report.

Table S2. Numbers of countries and articles, and reference numbers for qualified articles, classified by world region.

| World Region | Number of <br> Countries | Number of <br> Articles | Reference Numbers for Qualified Articles |
| :--- | :---: | :---: | :--- |
| Australasia | 2 | 6 | $6,12,45,62,67,74$ |
| Latin America | 5 | 3 | $47,48,56$ |
| Mediterranean | 10 | 18 | $2,3,6,7,10,12,13,18,44,50,51,52,63,64,65,68,73,77$ |
| North America | 2 | 11 | $6,8,9,12,22,27,41,42,49,66,70$ |
| Northern Europe | 14 | 30 | $4,5,6,11,12,14,15,19,20,21,23,26,29,30,31,32,36,37,38$, <br> $40,46,53,54,57,58,59,60,69,71,72$ |
| Southeast Asia | 5 | 17 | $6,16,17,24,25,28,33,34,35,39,55,61,75,76,78,79,80$ |
| Other | 3 | 3 | $1,6,43$ |
| Total | $\mathbf{4 1}$ | $\mathbf{8 8}$ |  |

Note: Some countries were covered by several articles so the number of articles exceeds the number of countries. The total number of articles exceeds 80 because some qualified articles studied multiple countries.

Prevalence data were extracted from tables, figures and texts of the qualified articles. Some articles did not provide sample sizes for prevalence percentages reported in risk classes but did provide standard errors or $95 \%$ confidence intervals. The latter were then used to estimate sample sizes, assuming simple random sampling. In all cases, sums of imputed sample sizes were adjusted to match reported subtotals. Some studies adjusted sample estimates of prevalence for differences between sample and population distributions, resulting in the reporting of population-adjusted estimates. Adjusted prevalence estimates were used only when unadjusted estimates were not reported. The statistical analysis required imputation of representative values for closed- and open-end classes for age groups and smoking amounts. Mid-points were uniformly used for closed-end classes. Judgment was used to choose representative values for open-end classes.

## Country-level data

The following public databases provided country-level statistics on demographic, socioeconomic, environmental, and geographic factors that might be associated with COPD.

1. Smoking prevalence was assessed using the WHO tobacco surveillance database which provides country-level prevalence rates of tobacco and cigarette smoking by men and women in 2012. http://www.who.int/tobacco/surveillance/policy/country_profile/en/.
2. A percentage breakdown of population by age and sex was obtained from the CIA World Fact Book for 2013. https://www.cia.gov/library/publications/the-worldfactbook/geos/xx.html.
3. Education was assessed using UNESCO data. We focused on a single statistic, namely, the percent of the population that is literate, defined as having more than primary education (ISCED level 2 or higher). http://data.uis.unesco.org/.
4. Income was assessed by per-capita gross national income (GNI) by country in current international dollars for 2013. http://data.worldbank.org/country.
5. Overall health system performance levels were assessed using system ranks for 191 member states presented in the World Health Report 2000:
http://www.who.int/whr/2000/en/whr00_en.pdf?ua=1.
6. CO 2 emissions in 2010 were assessed using World Bank data http://data.worldbank.org/country.
7. Motor vehicles per capita in 2011 were assessed using World Bank statistics http://data.worldbank.org/indicator/IS.VEH.NVEH.P3.
8. Elevation and latitude for major cities were obtained from http://dateandtime.info/citycoordinates.php.

## 2. Data Summary Statistics

Table S3 lists variables and their definitions, variable names (used in later computer codes), and descriptive statistics for the master data set, classified by (a) study-level variables and (b) country-level variables.

Study-level variables in Table S3 have 1583 records, except for risk factors sex, age, smoking status and smoking amount which do not appear in every study record. Country-level variables have records for 41 countries and, in the case of latitude and elevation, 112 districts (subnational regions) and countries. For studies that involved a representative national sample over a wide geographical area, we computed population weighted averages for elevation and latitude using the readings for the five largest cities in each country. If multiple sites were chosen, but were not nationally representative, sample weighted averages for sites were used. Preliminary analysis revealed that districts having elevations above a threshold tended to have lower COPD prevalence and that the protective effect increased with elevation above this threshold. The threshold, estimated as a tuning parameter by minimizing model deviance, was found to be 6.16 on the log-elevation scale (about 470 meters above sea level). Therefore, a covariate highelev was created which equals 0 when log-elevation is 6.16 or less and which increases linearly with log-elevation above 6.16. The Stata coding for the variable follows:

```
gen highelev=0;
replace highelev=lnelev-6.16 if lnelev>6.16;
```

Because the country-level socioeconomic variables are in disparate units, we standardized each variable (mean 0 , standard deviation 1 ) to achieve comparable odds ratios in our regression analyses. In the case of the overall health system performance variable, the WHO health system performance ranks for countries were converted into standard normal percentiles or z-scores. Standardized variables have prefix $\mathbf{z}_{\mathbf{-}}$ in the Stata code in the Appendix.

## 3. Data Analysis and Statistical Computation

## Data record format

Table S4 presents a fragment of the data for one illustrative article.

Table S3. Variables and definitions, variable names, and descriptive statistics for the master data set: (a) Studylevel variables; (b) Country-level variables.

| Variable | Variable Name | Records | Mean or percent | Std. Dev. | Min. | Max. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| (a) Study-level variables |  |  |  |  |  |  |
| Prevalence (percent points) | prev | 1583 | 12.4 | 10.6 | 0 | 74.2 |
| COPD GOLD I-IV severity | defn | 1583 | 60.0\% |  |  |  |
| Post-bronchodilation used | pbd | 1583 | 68.9\% |  |  |  |
| Respondents in the study | ntotal | 1583 | 2793.2 | 4091.9 | 106 | 20245 |
| Respondents in the record | n | 1583 | 749.8 | 1571.8 | 2 | 15440 |
| COPD cases in the record | r | 1583 | 81.1 | 205.8 | 0 | 3434 |
| Sex (male) | male | 1008 | 49.7\% |  |  |  |
| Age (years) | age | 796 | 57.0 | 14.1 | 12.5 | 90.0 |
| Smoking status - Never | smkstatus | 454 | 43.4\% |  |  |  |
| Ever (former) |  | 454 | 23.6\% |  |  |  |
| Current |  | 454 | 33.0\% |  |  |  |
| Smoking amount (pack-years) | smk | 422 | 16.9 | 16.9 | 0 | 80.1 |
|  |  |  |  |  |  |  |
| (b) Country-level variables |  |  |  |  |  |  |
| Smoking prevalence (men, \%) | smkpc_m | 41 | 32.6 | 10.0 | 13.0 | 52.9 |
| Smoking prevalence (women, \%) | smkpc_f | 41 | 17.8 | 9.3 | 1.2 | 38.0 |
| Children (male, under 15, \%) | under15pc_m | 41 | 10.3 | 2.9 | 6.7 | 17.2 |
| Children (female, under 15, \%) | under15pc_f | 41 | 9.8 | 2.9 | 6.3 | 16.5 |
| Working-age men (15-64, \%) | age15_64pc_m | 41 | 33.5 | 1.8 | 30.5 | 38.7 |
| Working-age women (15-64, \%) | age15_64pc_f | 41 | 33.0 | 1.4 | 30.5 | 35.8 |
| Seniors (men, over 64, \%) | over64pc_m | 41 | 5.8 | 2.6 | 1.6 | 11.2 |
| Seniors (women, over 64, \%) | over64pc_f | 41 | 7.5 | 3.4 | 1.5 | 14.6 |
| Literacy level (percent points) | literacy | 41 | 77.6 | 19.0 | 34.1 | 100.0 |
| Gross national income (000s/ capita) | gni | 41 | 30.9 | 15.2 | 5.0 | 66.5 |
| Health system performance | WHOrank | 41 |  |  |  |  |
| CO2 emissions (metric tons/ capita) | co2 | 41 | 7.43 | 4.22 | . 7 | 17.6 |
| Motor vehicles per capita | vehicles | 41 | . 423 | . 227 | . 014 | . 786 |
| Latitude (North or South, degrees) | latitude | 112 | 40.6 | 12.6 | 10 | 66 |
| Elevation above sea level (meters) | elevation | 112 | 188.1 | 325.8 | 1 | 2240 |

Table S4. A fragment of the data record for one qualified article. Many record lines and some variables for this study have been removed in order to fit this page. The fragment is for illustrative purposes only.

| Country | district | Table | prev | defn | ntotal | n | r | agegrp | male | smkstatus | elevation | latitude | pbd |
| :--- | :--- | :--- | ---: | :--- | ---: | ---: | ---: | :--- | :--- | :--- | :--- | :--- | :--- |
| Swe | umea | t3t2t1 | 3.4 | g1plus | 1237 | 89 | 3 | $46-47$ | 1 | never | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 6.3 | g1plus | 1237 | 63 | 4 | $61-62$ | 1 | never | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 28 | g1plus | 1237 | 25 | 7 | $76-77$ | 1 | never | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 3.7 | g1plus | 1237 | 107 | 4 | $46-47$ | 1 | ever | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 15.8 | g1plus | 1237 | 101 | 16 | $61-62$ | 1 | ever | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 27.6 | g1plus | 1237 | 76 | 21 | $76-77$ | 1 | ever | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 13 | g1plus | 1237 | 77 | 10 | $46-47$ | 1 | current | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 38.9 | g1plus | 1237 | 54 | 21 | $61-62$ | 1 | current | 22.7 | 64.2 | 1 |
| Swe | umea | t3t2t1 | 66.7 | g1plus | 1237 | 12 | 8 | $76-77$ | 1 | current | 22.7 | 64.2 | 1 |

## Stata Computational Procedures

In each step of the step-wise binomial regression, the effects of variables that entered the regression model in preceding steps were included as a fixed offset in the binomial regression function, that is, the regression function estimated in the preceding step was forced into the regression function at the current step with a regression coefficient of 1 . Thus, the explanatory contribution of each new group of variables was evaluated as an increment to the contribution of groups that preceded it. A fragment of the Stata computer program for the step-wise binomial regression analysis reported in the article appears in the Appendix to this document. Comment lines are inserted to explain analytical steps as the computation progresses.

Table S5. Frequency of records for each combination (profile) of the established risk factors of sex, age, smoking status and smoking amount, represented by indicator variables agegrp, male, smkstatus, and smkexp.

| profile | agegrp | male | smkstatus | smkexp | frequency |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 1 | 1 | 1 | 1 | 52 |
| 2 | 1 | 1 | 1 | 0 | 74 |
| 3 | 1 | 1 | 0 | 1 | 43 |
| 4 | 1 | 1 | 0 | 0 | 370 |
| 5 | 1 | 0 | 1 | 1 | 0 |
| 6 | 1 | 0 | 1 | 0 | 93 |
| 7 | 1 | 0 | 0 | 1 | 0 |
| 8 | 1 | 0 | 0 | 0 | 164 |
| 9 | 0 | 1 | 1 | 1 | 10 |
| 10 | 0 | 1 | 1 | 0 | 69 |
| 11 | 0 | 1 | 0 | 1 | 214 |
| 12 | 0 | 1 | 0 | 0 | 176 |
| 13 | 0 | 0 | 1 | 1 | 9 |
| 14 | 0 | 0 | 1 | 0 | 147 |
| 15 | 0 | 0 | 0 | 1 | 94 |
| 16 | 0 | 0 | 0 | 0 | 68 |
| Total |  |  |  |  | 1583 |

Step 2 of the stepwise binomial regression requires explanation. This step takes account of the studylevel established risk factors for sex, age, smoking status (current, ever, never) and smoking amount (in total pack years). Studies differed in their reporting formats for these risk factors. Frequently one or more of the four risk factors were omitted. Thus, these risk factors appeared singly or jointly in various combinations across the 80 studies. The Stata routine for this second step creates a categorical variable profile that represents every possible combination of the four variables representing age, sex, smoking status and smoking amount. There are 16 possible combinations of the four variables but combinations profile==5 and profile==7 do not occur and, hence, only 14 combinations enter the regression analysis. Table S5 shows the number of prevalence records for each profile. The most common reporting profile for prevalence was a cross-classification by sex and age group (370 among
the 1583 records). A separate binomial regression was run for each of the 14 profiles with the reported established risk factors for that profile included in the regression function. Each of the 14 regression functions also included the fitted regression function for the COPD outcome measure from step 1 as a fixed offset.

## 4. References For Qualified Articles

1. Al Ghobain M, Al-Hajjaj MS, Wali SO. Prevalence of chronic obstructive pulmonary disease among smokers attending primary healthcare clinics in Saudi Arabia. Ann Saudi Med 2011;31(2):129-133.
2. Arslan Z, Ilgazlı A, Etiler N, Hamzaoğlu O. Prevalence of chronic obstructive pulmonary disease in kocaeli: an industrialised city in Turkey. Balkan Med J. 2013; 30(4):387-93.
3. Barbara C, Rodrigues F, Dias H et al. Chronic obstructive pulmonary disease prevalence in Lisbon, Portugal: the burden of obstructive lung disease study. Revista Portuguesa de Pneumologia (English Edition) 2013;19(3):96-105.
4. Bednarek M, Plywaczewski R, Jonczak L, Zielinski J. There is no relationship between chronic obstructive pulmonary disease and obstructive sleep apnea syndrome: a population study. Respiration 2005;72(2):142-149.
5. Bednarek M, Maciejewski J, Wozniak M, Kuca P, Zielinski J. Prevalence, severity and underdiagnosis of COPD in the primary care setting. Thorax 2008;63(5):402-407.
6. Buist AS, McBurnie MA, Vollmer WM et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. Lancet 2007;370(9589):741-750.
7. Cabrera Lopez C, Julia Serda G, Cabrera Lacalzada C et al. Prevalence of COPD in the Canary Islands. Arch Bronconeumol 2014;50(7):272-277.
8. Celli BR, Halbert RJ, Isonaka S, Schau B. Population impact of different definitions of airway obstruction. Eur Respir J 2003;22(2):268-273.
9. Celli BR, Halbert RJ, Nordyke RJ, Schau B. Airway obstruction in never smokers: results from the Third National Health and Nutrition Examination Survey. Am J Med 2005;118(12):13641372.
10. Daldoul H, Denguezli M, Jithoo A et al. Prevalence of COPD and Tobacco Smoking in TunisiaResults from the BOLD Study. International journal of environmental research and public health 2013;10(12):7257-7271.
11. Danielsson P, Ólafsdóttir IS, Benediktsdóttir B, Gíslason T, Janson C. The prevalence of chronic obstructive pulmonary disease in Uppsala, Sweden ГÇô the Burden of Obstructive Lung Disease (BOLD) study: cross-sectional population-based study. The Clinical Respiratory Journal 2012;6(2):120-127.
12. De Marco R, Accordini S, Cerveri I et al. An international survey of chronic obstructive pulmonary disease in young adults according to GOLD stages. Thorax 2004;59(2):120-125.
13. Deveci F, Deveci SE, Turko-lu S et al. The prevalence of chronic obstructive pulmonary disease in Elazig, Eastern Turkey. European journal of internal medicine 2011;22(2):172-176.
14. Eriksson B, Lindberg A, Müllerova H, Rönmark E, Lundbäck B. Association of heart diseases with COPD and restrictive lung function-results from a population survey. Respir Med 2013;107(1):98-106.
15. Fabricius P, Løkke A, Marott JL, Vestbo J, Lange P. Prevalence of COPD in Copenhagen. Respir Med 2011;105(3):410-417.
16. Fukuchi Y, Nishimura M, Ichinose M et al. COPD in Japan: the Nippon COPD Epidemiology study. Respirology 2004;9(4):458-465.
17. Funakoshi Y, Omori H, Mihara S, Marubayashi T, Katoh T. Association between airflow obstruction and the metabolic syndrome or its components in Japanese men. Internal Medicine 2010;49(19):2093-2099.
18. Gunen H, Hacievliyagil SS, Yetkin O, Gulbas G, Mutlu LC, Pehlivan E. Prevalence of COPD: first epidemiological study of a large region in Turkey. European journal of internal medicine 2008;19(7):499-504.
19. Hagstad S, Ekerljung L, Lindberg A, Backman H, Rönmark E, Lundbäck B. COPD among nonsmokers - report from the obstructive lung disease in Northern Sweden (OLIN) studies. Respir Med 2012;106(7):980-988.
20. Hagstad S, Bjerg A, Ekerljung L et al. Passive smoking exposure is associated with increased risk of COPD in never smokers. CHEST Journal 2014;145(6):1298-1304.
21. Hansen JG, Pedersen L, Overvad K, Omland Ø, Jensen HK, Sørensen HT. The prevalence of chronic obstructive pulmonary disease among Danes aged 45ГÇô84 years: population-based study. COPD 2008;5(6):347-352.
22. Hill K, Goldstein RS, Guyatt GH et al. Prevalence and underdiagnosis of chronic obstructive pulmonary disease among patients at risk in primary care. Can Med Assoc J 2010;182(7):673678.
23. Hvidsten SC, Storesund L, Wentzel-Larsen T, Gulsvik A, Lehmann S. Prevalence and predictors of undiagnosed chronic obstructive pulmonary disease in a Norwegian adult general population. The Clinical Respiratory Journal 2010;4(1):13-21.
24. Hwang YI, Kim CH, Kang HR et al. Comparison of the prevalence of chronic obstructive pulmonary disease diagnosed by lower limit of normal and fixed ratio criteria. Journal of Korean medical science 2009;24(4):621-626.
25. Idolor LF, De Guia TS, Francisco NA et al. Burden of obstructive lung disease in a rural setting
in the Philippines. Respirology 2011;16(7):1111-1118.
26. Johannessen A, Omenaas ER, Bakke PS, Gulsvik A. Implications of reversibility testing on prevalence and risk factors for chronic obstructive pulmonary disease: a community study. Thorax 2005;60(10):842-847.
27. Johnston AK, Mannino DM, Hagan GW, Davis KJ, Kiri VA. Relationship between lung function impairment and incidence or recurrence of cardiovascular events in a middle-aged cohort. Thorax 2008; 63(7):599-605.
28. Joo H, Park J, Lee SD, Oh YM. Comorbidities of chronic obstructive pulmonary disease in Koreans: a population-based study. Journal of Korean medical science 2012;27(8):901-906.
29. Jordan RE, Miller MR, Lam KbH, Cheng KK, Marsh J, Adab P. Sex, susceptibility to smoking and chronic obstructive pulmonary disease: the effect of different diagnostic criteria. Analysis of the Health Survey for England. Thorax 2012;67(7):600-605.
30. Jyrki-Tapani K, Sovijärvi A, Lundbäck B. Chronic obstructive pulmonary disease in Finland: prevalence and risk factors. COPD 2005;2(3):331-339.
31. Kainu A, Rouhos A, Sovijärvi A, Lindqvist A, Sarna S, Lundbäck B. COPD in Helsinki, Finland: socioeconomic status based on occupation has an important impact on prevalence. Scandinavian journal of public health 2013;41(6):570-578.
32. Kanervisto M, Vasankari T, Laitinen T, Heliövaara M, Jousilahti P, Saarelainen S. Low socioeconomic status is associated with chronic obstructive airway diseases. Respir Med 2011;105(8):1140-1146.
33. Kim DS, Kim YS, Jung KS et al. Prevalence of chronic obstructive pulmonary disease in Korea: a population-based spirometry survey. Am J Respir Crit Care Med 2005;172(7):842847.
34. Kim SJ, Suk MH, Choi HM et al. The local prevalence of COPD by post-bronchodilator GOLD criteria in Korea. The International Journal of Tuberculosis and Lung Disease 2006;10(12):1393-1398.
35. Lam HT, Ekerljung L, T Ng, Rönmark E, Larsson K, Lundbäck B. Prevalence of COPD by Disease Severity in Men and Women in Northern Vietnam. COPD 2014;11(5):575-581.
36. Lange P, Mogelvang R, Marott JL, Vestbo J, Jensen JS. Cardiovascular morbidity in COPD: A study of the general population. COPD 2011;7(1):5-10.
37. Leivseth L, Brumpton BM, Nilsen TIL, Mai XM, Johnsen R, Langhammer A. GOLD classifications and mortality in chronic obstructive pulmonary disease: the HUNT Study, Norway. Thorax 2013;68(10):914-921.
38. Lindberg A, Jonsson AC, Rönmark E, Lundgren R, Larsson LG, Lundbäck B. Prevalence of chronic obstructive pulmonary disease according to BTS, ERS, GOLD and ATS criteria in
relation to doctor diagnosis, symptoms, age, gender, and smoking habits. Respiration 2005;72(5):471-479.
39. Liu S, Zhou Y, Wang X et al. Biomass fuels are the probable risk factor for chronic obstructive pulmonary disease in rural South China. Thorax 2007;62(10):889-897.
40. Lundbäck B, Lindberg A, Lindstrom M et al. Not 15 but $50 \%$ of smokers develop COPD? ГС̧öreport from the obstructive lung disease in Northern Sweden Studies. Respir Med 2003;97(2):115-122.
41. Mannino DM, Doherty DE, Buist AS. Global Initiative on Obstructive Lung Disease (GOLD) classification of lung disease and mortality: findings from the Atherosclerosis Risk in Communities (ARIC) study. Respir Med 2006;100(1):115-122.
42. Mannino DM, Davis KJ. Lung function decline and outcomes in an elderly population. Thorax 2006;61(6):472-477.
43. Martins P, Rosado-Pinto J, do Céu Teixeira M et al. Under-report and underdiagnosis of chronic respiratory diseases in an African country. Allergy 2009;64(7):1061-1067.
44. Mascarenhas J, Falcão H, Lourenço P et al. Population-based study on the prevalence of spirometric obstructive pattern in Porto, Portugal. Respiratory Care 2011;56(5):619-625.
45. Matheson MC, Benke G, Raven J et al. Biological dust exposure in the workplace is a risk factor for chronic obstructive pulmonary disease. Thorax 2005;60(8):645-651.
46. Melville AM, Pless-Mulloli T, Afolabi OA, Stenton SC. COPD prevalence and its association with occupational exposures in a general population. Eur Respir J 2010;36(3):488-493.
47. Menezes A, Macedo SC, Gigante DP et al. Prevalence and Risk Factors for Chronic Obstructive Pulmonary Disease According to Symptoms and Spirometry\#. COPD 2004;1(2):173-179.
48. Menezes AM, Perez-Padilla R, Jardim JRB et al. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. Lancet 2005;366(9500):18751881.
49. Methvin JN, Mannino DM, Casey BR. COPD prevalence in southeastern Kentucky: the burden of lung disease study. CHEST Journal 2009;135(1):102-107.
50. Minas M, Hatzoglou C, Karetsi E et al. COPD prevalence and the differences between newly and previously diagnosed COPD patients in a spirometry program. Prim Care Respir J 2010;19(4):363-370.
51. Miravitlles M, Soriano JB, Garcia-Rio F et al. Prevalence of copd in spain: impact of undiagnosed copd on quality of life and daily life activities. Thorax 2009; 64(10):863-8.
52. Mohammad Y, Shaaban R, Al-Zahab BA, Khaltaev N, Bousquet J, Dubaybo B. Impact of active and passive smoking as risk factors for asthma and COPD in women presenting to primary care
in Syria: first report by the WHO-GARD survey group. International journal of chronic obstructive pulmonary disease 2013;8:473.
53. Murtagh E, Heaney L, Gingles J et al. The prevalence of obstructive lung disease in a general population sample: The NICECOPD study. Eur J Epidemiol 2005;20(5):443-453.
54. Nizankowska-Mogilnicka E, Mejza F, Buist AS et al. Prevalence of COPD and tobacco smoking in Malopolska region-results from the BOLD study in Poland. Polskie Archiwum Medycyny Wewnetrznej 2007;117(9):402.
55. Omori H, Nonami Y, Mihara S, Marubayashi T, Morimoto Y, Aizawa H. Prevalence of airflow limitation on medical check-up in Japanese subjects. Journal of UOEH 2007;29(3):209-219.
56. Queiroz MC, Moreira MA, Rabahi MF. Underdiagnosis of COPD at primary health care clinics in the city of Aparecida de Goiânia, Brazil. Jornal Brasileiro de Pneumologia 2012;38(6):692699.
57. Runarsdottir SB, Gudmundsson G, Aspelund T et al. Prevalence of airflow obstruction in nonsmoking older individuals using different spirometric criteria: the AGES Reykjavik study. COPD 2013;10(4):493-499.
58. Schirnhofer L, Lamprecht B, Vollmer WM et al. COPD prevalence in Salzburg, Austria: results from the Burden of Obstructive Lung Disease (BOLD) study. CHEST Journal 2007;131(1):2936.
59. Scholes S, Moody A, Mindell JS. Estimating population prevalence of potential airflow obstruction using different spirometric criteria: a pooled cross-sectional analysis of persons aged 40 Г̧̧ô95 years in England and Wales. ВМJ open 2014;4(7): e 005685.
60. Shahab L, Jarvis MJ, Britton J, West R. Prevalence, diagnosis and relation to tobacco dependence of chronic obstructive pulmonary disease in a nationally representative population sample. Thorax 2006;61(12):1043-1047.
61. Shin C, In KH, Shim JJ et al. Prevalence and correlates of airway obstruction in a communitybased sample of adults. CHEST Journal 2003;123(6):1924-1931.
62. Shirtcliffe P, Weatherall M, Marsh S et al. COPD prevalence in a random population survey: a matter of definition. Eur Respir J 2007;30(2):232-239.
63. Sichletidis L, Tsiotsios I, Gavriilidis A et al. Prevalence of chronic obstructive pulmonary disease and rhinitis in northern Greece. Respiration 2005;72(3):270-277.
64. Soriano JB, Miravitlles M, Borderas L et al. Geographical variations in the prevalence of COPD in Spain: relationship to smoking, death rates and other determining factors. Archivos de Bronconeumología (English Edition) 2010;46(10):522-530.
65. Stav D, Raz M. Prevalence of chronic obstructive pulmonary disease among smokers aged 45 and up in Israel. IMAJ-RAMAT GAN- 2007;9(11):800.
66. Tan WC, Bourbeau J, Fitzgerald JM et al. Can age and sex explain the variation in COPD rates across large urban cities? A population study in Canada. The International Journal of Tuberculosis and Lung Disease 2011;15(12):1691-1698.
67. Toelle BG, Xuan W, Bird TE et al. Respiratory symptoms and illness in older Australians: the Burden of Obstructive Lung Disease (BOLD) study. Med J Aust 2013;198(3):144-148.
68. Tzanakis N, Anagnostopoulou U, Filaditaki V, Christaki P, Siafakas N. Prevalence of COPD in Greece. CHEST Journal 2004;125(3):892-900.
69. Vanfleteren LE, Franssen FM, Wesseling G, Wouters EF. The prevalence of chronic obstructive pulmonary disease in Maastricht, the Netherlands. Respir Med 2012;106(6):871-874.
70. Vaz Fragoso CA, Concato J, McAvay G et al. Chronic obstructive pulmonary disease in older persons: a comparison of two spirometric definitions. Respir Med 2010;104(8):1189-1196.
71. Voll-Aanerud M, Eagan TM, Wentzel-Larsen T, Gulsvik A, Bakke PS. Respiratory symptoms, COPD severity, and health related quality of life in a general population sample. Respir Med 2008;102(3):399-406.
72. Waatevik M, Skorge TD, Omenaas E, Bakke PS, Gulsvik A, Johannessen A. Increased prevalence of chronic obstructive pulmonary disease in a general population. Respir Med 2013;107(7):1037-1045.
73. Waked M, Salame J, Khayat G, Salameh P. Correlates of COPD and chronic bronchitis in nonsmokers: data from a cross-sectional study. International journal of chronic obstructive pulmonary disease 2012;7:577.
74. Wilson D, Adams R, Appleton S, Ruffin R. Difficulties identifying and targeting COPD and population-attributable risk of smoking for COPD: a population study. CHEST Journal 2005;128(4):2035-2042.
75. Yin P, Jiang CQ, Cheng KK et al. Passive smoking exposure and risk of COPD among adults in China: the Guangzhou Biobank Cohort Study. Lancet 2007;370(9589):751-757.
76. Yoo KH, Kim YS, Sheen SS et al. Prevalence of chronic obstructive pulmonary disease in Korea: the fourth Korean National Health and Nutrition Examination Survey, 2008. Respirology 2011;16(4):659-665.
77. Zachariades AG, Zachariadou T, Adamide T, Anagnostopoulou U, Georgiou A, Gourgoulianis KI. Prevalence of chronic obstructive pulmonary disease in Cyprus: A population-based study. COPD 2012;9(3):259-267.
78. Zhong N, Wang C, Yao W et al. Prevalence of chronic obstructive pulmonary disease in China: a large, population-based survey. Am J Respir Crit Care Med 2007;176(8):753-760.
79. Zhou Y, Wang C, Yao W et al. COPD in Chinese nonsmokers. Eur Respir J 2009;33(3):509-518.
80. Zhou Y, Wang D, Liu S et al. The association between BMI and COPD: the results of two population-based studies in Guangzhou, China. COPD 2013;10(5):567-572.

## Appendix

The following is a fragment of the Stata computer program for the step-wise binomial regression analysis reported in the article. The Stata procedure binreg offers an option offset (varname) that includes varname in the regression model with a coefficient constrained to 1 . The generic command predict varname, $\mathbf{x b}$ is used to generate the fitted values from the current step in the regression analysis that can be carried forward to the next regression step as an offset with name varname.

```
***
*Remove the effects of the COPD diagnostic definition,
*the post-bronchodilation indicator, and their interaction
***
char defn[omit] "g2plus"
char smkstatus[omit] "never"
#delimit ;
xi: binreg r
i.defn*i.pbd,
or n(n);
predict offset0, xb;
#delimit cr
***
*Variable 'profile' represents all 16 possible
*combinations of risk factors for sex, age, smoking
*status and smoking amount reported in the article
***
***
*Fit binomial regression models for all risk factor
*profiles (profiles 5 and 7 have no cases)
#delimit ;
quietly xi: binreg r
age agesq male i.smkstatus smk offset0 if profile==1,
or n(n);
predict off1 if profile==1, xb;
gen offset1=off1 if off1!=.;
quietly xi: binreg r
age agesq male i.smkstatus offset0 if profile==2,
or n(n);
predict off2 if profile==2, xb;
replace offset1=off2 if off2!=.;
quietly xi: binreg r
age agesq male smk offset0 if profile==3,
or n(n);
```

```
predict off3 if profile==3, xb;
replace offset1=off3 if off3!=.;
quietly xi: binreg r
age agesq male offset0 if profile==4,
or n(n);
predict off4 if profile==4, xb;
replace offset1=off4 if off4!=.;
*xi: binreg r
*age agesq i.smkstatus smk offset0 if profile==5,
*or n(n);
*predict off5 if profile==5, xb;
*replace offset1=off5 if off5!=.;
quietly xi: binreg r
age agesq i.smkstatus offset0 if profile==6,
or n(n);
predict off6 if profile==6, xb;
replace offset1=off6 if off6!=.;
*xi: binreg r
*age agesq smk offset0 if profile==7,
*or n(n);
*predict off7 if profile==7, xb;
*replace offset1=off7 if off7!=.;
quietly xi: binreg r
age agesq offset0 if profile==8,
or n(n);
predict off8 if profile==8, xb;
replace offset1=off8 if off8!=.;
#delimit ;
quietly xi: binreg r
male i.smkstatus smk offset0 if profile==9,
or n(n);
predict off9 if profile==9, xb;
replace offset1=off9 if off9!=.;
quietly xi: binreg r
male i.smkstatus offset0 if profile==10,
or n(n);
predict off10 if profile==10, xb;
replace offset1=off10 if off10!=.;
quietly xi: binreg r
male smk offset0 if profile==11,
or n(n);
predict off11 if profile==11, xb;
replace offset1=off11 if off11!=.;
quietly xi: binreg r
male offset0 if profile==12,
or n(n);
predict off12 if profile==12, xb;
replace offset1=off12 if off12!=.;
quietly xi: binreg r
i.smkstatus smk offset0 if profile==13,
or n(n);
predict off13 if profile==13, xb;
replace offset1=off13 if off13!=.;
```

```
quietly xi: binreg r
i.smkstatus offset0 if profile==14,
or n(n);
predict off14 if profile==14, xb;
replace offset1=off14 if off14!=.;
quietly xi: binreg r
smk offset0 if profile==15,
or n(n);
predict off15 if profile==15, xb;
replace offset1=off15 if off15!=.;
quietly xi: binreg r
offset0 if profile==16,
or n(n);
predict off16 if profile==16, xb;
replace offset1=off16 if off16!=.;
#delimit cr
***
*Remove effects of study-level risk factors for
*sex, age, smoking status and smoking amount
***
xi: binreg r, or n(n) offset(offset1)
***
*Sequentially fit binomial regression models for
*groups of country-level effects
***
char defn[omit] "g2plus"
char smkstatus[omit] "never"
char worldregion[omit] "NEurope"
#delimit ;
xi: binreg r
smkpc_m smkpc_f a15_64pc_m a15_64pc_f over64pc_m over64pc_f,
or n(n) offset(offset1) ;
predict offset2, xb;
xi: binreg r
z_literacy z_gni z_hsys,
or n(n) offset(offset2) ;
predict offset3, xb;
xi: binreg r
z_co2 z_vehicles,
or n(n) offset(offset3) ;
predict offset4, xb;
xi: binreg r
latitude highelev,
or n(n) offset(offset4);
predict offset5, xb;
xi: binreg r
i.worldregion,
or n(n) offset(offset5) ;
```

predict lgtfit, xb;
\#delimit cr

