## **Online supplement**

## **Poverty-attributable chronic airflow obstruction in the multinational Burden of Obstructive Lung Disease study**

Jaymini Patel<sup>1</sup>, Andre F.S. Amaral<sup>1</sup>, Cosetta Minelli<sup>1</sup>, Fadlalla G.Elfadaly<sup>2</sup>, Kevin Mortimer<sup>3</sup>, Asma El Sony<sup>4</sup>, El Rhazi K<sup>5</sup>, Terence Seemungal<sup>6</sup>, Mahesh PA<sup>7</sup>, Daniel Obaseki<sup>8</sup>, Meriam Denguezli<sup>9</sup>, Rana Ahmed<sup>10</sup>, Hamid Hacene Cherkaski<sup>11</sup>, Parvaiz A Koul<sup>12</sup>, Abdul Rashid<sup>13</sup>, Li Cher Loh<sup>13</sup> Hervé Lawin<sup>14</sup>, Mohamed Al Ghobain<sup>15</sup>, Asaad Ahmed Nafees<sup>16</sup>, Althea Aquart-Stewart<sup>17</sup>, Imed Harrabi<sup>18</sup>, A Sonia Buist<sup>19</sup>, Peter Burney<sup>1</sup> on behalf of the Burden of Obstructive Lung Disease (BOLD) Collaborative Research Group

1 National Heart and Lung Institute, Imperial College London, London, UK

2 School of Mathematics and Statistics, Faculty of Science, Technology, Engineering & Mathematics, The Open University, Milton Keynes, United Kingdom

3 Liverpool School of Tropical Medicine, Liverpool, UK

4 The Epidemiological Laboratory, Khartoum, Sudan

5 Department of Epidemiology and Public Health, Faculty of Medicine of Fez, Morocco

6 Faculty of Medical Sciences, University of West Indies, St Augustine, Trinidad and Tobago

- 7 Department of Respiratory Medicine, JSS Medical College, JSSAHER, Mysore, India
- 8 Obafemi Awolowo University, Ile-Ife, Nigeria
- 9 Université de Sousse, Faculté de Médecine de Sousse, Sousse, Tunisie
- 10 The Epidemiological Laboratory, Khartoum, Sudan
- 11 Department of Pneumology, Faculty of Medicine, Annaba (Algeria)
- 12 Sher-i-Kashmir Institute of Medical Sciences, Srinagar, J&K, India
- 13 Department of Medicine, RCSI & UCD Malaysia Campus, Penang, Malaysia
- 14 Occupational and Environmental Health, University of Abomey Calavi, Cotonou, Benin

15 Department of Medicine, King Saud bin Abdulaziz University for Health Sciences & King Abdullah International Medical Research Centre, Riyadh, Saudi Arabia

- 16 Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan
- 17 University of the West Indies, Kingston, Jamaica
- 18 Faculté de Médecine, Université de Sousse, Sousse, Tunisia

19 Oregon Health & Science University, Portland, OR, USA

**Corresponding author:** Jaymini Patel National Heart and Lung Institute, Imperial College London London SW3 6LR, UK [j.patel@imperial.ac.uk](mailto:j.patel@imperial.ac.uk)

# **Table of contents**



# **Figures and Tables**





## **Supplementary Methods**

#### **BOLD study details**

The BOLD study protocol has been previously published.<sup>1</sup> In short, representative samples of the noninstitutionalised population aged 40 years and above were recruited from a sample of sites from Africa, Asia, Australia, Caribbean, Europe and North America. All sites had approval from their local ethics committee, and participants gave informed consent. All study staff were trained, supervised and monitored during the survey. Spirometry was conducted using an EasyOne spirometer (ndd Medizintechnik AG, Zurich, Switzerland) to take a pre-bronchodilator reading. Post bronchodilator readings were also collected after inhaling Salbutamol (200 µg) via a spacer. All individual spirometry data were reviewed for quality control during the study. Information from participants on several risk factors was collected using a standardised questionnaire.

#### **Calculation of the wealth score**

It is difficult to measure wealth at an individual level, but Howe et al. have shown that household assetbased measures of wealth can be used to collect information about the wealth of an individual, that is stable over short-term economic fluctuations<sup>2</sup>. In the BOLD study, we have developed a wealth score using a Mokken scale approach applied to the data collected on household assets that could be used to compare individuals within and between countries<sup>3</sup>. We have shown that the mean number of assets owned in each site is highly correlated with the Gross National Income (GNI) per capita of the country<sup>4</sup>, which is a strong indicator of the standard of living of an average citizen. We have also shown the high correlation of assets owned with educational level and other variables likely to be related to the socioeconomic position, which suggests that our Mokken scale is broadly indicative of an individual's wealth.

We calculated the wealth score based on questions which were part of the BOLD study core questionnaire<sup>4</sup>. This part of the questionnaire asked whether the participant or any person in the household owned any of the following 15 assets. Only 10 of these assets (highlighted in green below) were used to calculate the wealth score.

Please tell me whether this household, or any person who lives in the household, has/owns the following items [*Possible answers: Yes/No/Don't know*]:

- a. Electricity
- b. Flush toilet
- c. Fixed telephone
- d. Cell telephone
- e. Television
- f. Radio
- g. Refrigerator
- h. Car
- i. Moped/scooter/motorcycle
- j. Washing machine
- k. Own their own home
- l. Indoor bath or shower
- m. Indoor tap
- n. Outdoor tap of their own

The response is then coded as 1 for "YES", 0 for "NO", and missing for "DON'T KNOW".

## **Definition of poverty**

# We used the World Bank data available at

[https://datahelpdesk.worldbank.org/knowledgebase/articles/906519,](https://datahelpdesk.worldbank.org/knowledgebase/articles/906519) by the year of survey for each BOLD site, to classify countries as low-, middle- or high-income. In the sites in low- and middle-income countries, most participants had a wealth score of 7 or less, whereas in the sites in high-income countries it was the opposite, with most participants having a wealth score above 7 (Figure S2). Hence, we used a wealth score lower than or equal to 7 to define poverty.





## **Bayesian hierarchical model to estimate PAR**

For each site, PAR of CAO for the exposure of interest (poverty or wealth score) can be calculated using the formula below:

$$
PAR = \frac{Pe (RR - 1)}{RR} \; X \; Pd
$$

where  $Pe$  is the proportion of cases exposed to poverty and  $Pd$  is the proportion of the disease in the population of interest, for each site.

The RR for CAO of poverty, adjusted for age and sex (and additional covariates in the secondary analyses), was estimated for each site using a Bayesian log-binomial hierarchical model. The hierarchical model allowed us to address the issue of difficult RR estimation (and hence difficult estimation of PAR) in sites with very low numbers. By assuming that all local estimates of RR come from a single distribution of values with a common mean and variance, we could "borrow" information from the other sites to increase the precision of the estimate in each single site. More detail about the model can be found at <https://spiral.imperial.ac.uk/handle/10044/1/89220> (Chapter 2, page 63).

The specification of the Bayesian log-binomial hierarchical model to estimate the RR is as follows.

Let  $\mathit{CAO}_{ij}\,$  be the observed binary outcome (yes/no) for j<sup>th</sup> subject i<sup>th</sup> site,

 $CAO<sub>ij</sub>$  = 1 if subject has a disease

= 0 otherwise,

The model is specified as:

 $CAO_{ij} \sim Bernoulli(p_{ij})$ , i=1,2,.........,21 & j=1,2,.....n<sub>i</sub>  $log(P_{ij}) = \alpha_i + \sum_{s} \beta_{is} * X_{ijs}$ 

where  $p_{ij}$  is the probability of the outcome  $\;CAO_{ij}$ ,  $\alpha_i\;$  is the intercept for site i,  $\;X_{ij}$ s=1 if j<sup>th</sup> subject in the site i<sup>th</sup> is exposed to s<sup>th</sup> exposure of interest and 0 otherwise, and n<sub>i</sub> is the sample size for i<sup>th</sup> site . The RR of the exposure of interest or confounder "s" is given by exp ( $\beta_{is}$ ).

For the hierarchical specification of the model, we assign the following priors for the intercept  $\alpha_i$  and  $\beta_{is}$ :

For (i in 1: 21) {  $\alpha_i$ ~Normal(µ. alpha,  $\sigma^2$ . alpha)  $\beta_{is}$  ~Normal (µ. betas,  $\sigma^2$ . betas)

```
}
```
with hyper priors:

 $\mu$ . alpha~Normal $(0, 10^4)$ 

 $\sigma^2$ . alpha~Gamma(1,0.005)

 $\mu. \, betas \sim Normal(0, 10^2)$ 

 $\sigma^2$ . beta~Gamma $(1,0.005)$ 

The  $\beta_{is}$  and  $\alpha_i$  are assumed to come from a normal distribution, but there is a distinct normal distribution for each  $\beta_{is}$  and  $\alpha_i.$ 

All Bayesian analyses were conducted in Stata 14 with the use of user-written program to call Open Bugs<sup>5,6</sup>. We used two MCMC chains, where 50,000 samples were drawn from the posterior distribution of each parameter, after discarding 50,000 burn-in iterations. For the point estimate of the PAR we used the posterior mean, together with the 95% credibility interval (95%CrI) corresponding to the 2.5% and 97.5% percentiles.

## **Investigation of heterogeneity in the RR estimates across sites**

We investigated the heterogeneity of RR estimates across sites using the  $I^2$  statistic of their meta-analysis<sup>7</sup>. For conveniency, the meta-analysis was performed in a frequentist framework using the 'metan' command in the Stata package<sup>8</sup>.

## **Directed acyclic graphs (DAGs)**

To help us decide which possible confounders we should adjust the analyses for, we drew directed acyclic graphs (DAGs) based on available knowledge. In particular, we used DAGitty<sup>9</sup>, implemented through the *dagitty* command in R [\(www.rdocumentation.org/packages/dagitty/versions/0.3-1/topics/dagitty\)](http://www.rdocumentation.org/packages/dagitty/versions/0.3-1/topics/dagitty), with poverty as the exposure (and similar reasoning for wealth score as a continuous variable) and CAO as the outcome (and again similar reasoning for  $FEV<sub>1</sub>/FVC$ ).

We considered the following factors: age, sex, education, smoking pack-years, exposure to a dusty job, passive smoking, body mass index, childhood hospitalisation due to respiratory infection, family history of respiratory disease, and history of tuberculosis.

The three DAGs in Figure S1 (a to c) illustrate three different scenarios defining possible confounders (for which we should adjust the model) versus possible mediators of the effect of poverty on CAO (for which we should not adjust).

In our main analysis (Figure S1a), we only adjusted for age and sex as, as we hypothesised that all other factors could well act as mediators. In the analysis of  $FEV<sub>1</sub>/FVC$ , age and sex were added to the model, while in the analysis of CAO, age and sex are already accounted for as they are used to calculate the LLN (CAO defined as  $FEV_1/FVC <$  Lower Limit of Normal).

However, to assess the impact on the estimate of the association between poverty and CAO when further adjusting for the other factors, we performed two secondary analyses: a) further adjusting for education (Figure S1b); and b) adjusting for all factors considered (Figure S1c).

## **Figure S2**



# **a) Model 1 (main model)** - Adjusted only for age and sex

**b) Model 2** - Adjusted for age, sex and education



**c) Model 3 –** Adjusted for all variables





#### **Figure S3: Association between wealth score and FEV1/FVC (%) across the 21 sites**

#### **Figure S4: Association between wealth score and FEV1/FVC<LLN across the 21 sites**





#### **Figure S5: Association between poverty and FEV1/FVC (%) across the 21 sites**

**Figure S6: Association between poverty and FEV1/FVC<LLN across the 21 sites**



**Figure S7: Comparison of site-specific estimates of PAR for CAO due to poverty from three different models: (1) main analysis - model adjusted only for age and sex; (2) secondary analysis - model adjusted for age, sex and education; (3) secondary analysis - model adjusted for age, sex, education, BMI, smoking pack years, passive smoking, exposure to dusty job, family history of respiratory infection, hospitalisation before the age of 10, and history of tuberculosis**



# **Table S1**: **Data collection period of BOLD sites**





#### **Table S2: Characteristics of the population by site**

N: Number of people with usable post-bronchodilator spirometry and questionnaire data; FEV<sub>1</sub>: forced expiratory volume in one second; FVC: forced vital capacity; LLN: lower limit of normal; Poverty: defined as a wealth score lower than or equal to 7; BMI: body mass index; smoking pack-years: 1 pack-year equals an average of 20 cigarettes per day for 1 year or the equivalent amount of other types of tobacco smoking; passive smoking: somebody else in the household smoked during the past 2 weeks; post-primary education: the highest level of education completed was above primary school; family history of COPD: a close sibling or parent had been diagnosed as having emphysema, chronic bronchitis or COPD by a health professional; childhood hospitalisation: hospitalised for breathing problems before the age of 10 years; History of tuberculosis: ever diagnosed with tuberculosis



#### **Table S3: Association of wealth score and poverty with FEV1/FVC (%) and CAO, adjusted for possible confounding factors (Model 3)**

FEV<sub>1</sub>: forced expiratory volume in one second; FVC: forced vital capacity; LLN: lower limit of normal; Poverty: defined as a wealth score lower than or equal to 7; BMI: body mass index; smoking pack-years: 1 pack-year equals an average of 20 cigarettes per day for 1 year or the equivalent amount of other types of tobacco smoking; passive smoking: somebody else in the household smoked during the past 2 weeks; Education: the highest level of education completed; family history of COPD: a close sibling or parent had been diagnosed as having emphysema, chronic bronchitis or COPD by a health professional; childhood hospitalisation: hospitalised for breathing problems before the age of 10 years; history of tuberculosis: ever diagnosed with tuberculosis



# **Table S4: PAR for CAO due to poverty, expressed as percent of total population aged ≥40 years, with 95% credible intervals**

#### **References**

1. Buist AS, Vollmer WM, Sullivan SD, et al. The Burden of Obstructive Lung Disease Initiative (BOLD): rationale and design. COPD 2005; 2(2): 277-83.

2. Howe LD, Galobardes B, Matijasevich A, et al. Measuring socio-economic position for epidemiological studies in low- and middle-income countries: a method of measurement in epidemiology paper. Int J Epidemiology 2012; 41, 871-886.

3. Townend J, Minelli C, Harrabi I, et al. Development of an international scale of socio-economic position based on household assets. Emerg Themes Epidemiol 2015; 12: 13.

4. Townend J, Minelli C, Mortimer K, Obaseki DO, et al. The association between chronic airflow obstruction and poverty in 12 sites of the multinational BOLD study. Eur Resp J 2017; 49: 1601880.

5. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. Stat Med 2009; 28: 3049-3067.

6. Palmer T, Thompson J, Moreno S. Performing Bayesian analysis in Stata using WinBUGS. Stata Users Group, United Kingdom Stata Users' Group Meeting 2007 8.

7. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327(7414): 557-60.

8. Palmer TM, Sterne JAC, ed. Meta-Analysis in Stata: An Updated Collection from the Stata Journal. 2016; 2nd ed. College Station, TX: Stata Press.

9. Textor J, Hardt J, Knüppel S. DAGitty: a graphical tool for analyzing causal diagrams. Epidemiology 2011; 22(5): 745.