A cross-sectional study on prevalence of chronic obstructive pulmonary disease (COPD) in India: rationale and methods

Prabu Rajkumar,1 Kamaraj Pattabi,1 Selvaraj Vadivoo,1 Arvind Bhome,2 Bill Brashier,3 Prashanta Bhattacharya,4 Sanjay M Mehendale5

ABSTRACT

Background Chronic obstructive pulmonary disease (COPD) is a common preventable and treatable chronic respiratory disease, which affects 210 million people globally. Global and national guidelines exist for the management of COPD. Although evidence-based, they are inadequate to address the phenotypic and genotypic heterogeneity in India. Co-existence of other chronic respiratory diseases can adversely influence the diagnosis of COPD. India has a huge burden of COPD with various risk factors and comorbid conditions. However, valid prevalence estimates employing spirometry as the diagnostic tool and data on important comorbid conditions are not available. This study protocol is designed to address this knowledge gap and eventually to build a database to undertake long-term cohort studies to describe the phenotypic and genotypic heterogeneity among COPD patients in India.

Objectives The primary objective is to estimate the prevalence of COPD among adults aged ≥25 years for each gender in India. The secondary objective is to identify the risk factors for COPD and important comorbid conditions such as asthma and post-tuberculosis sequelae. It is also proposed to validate the currently available definitions for COPD diagnosis in India.

Methods and analysis A cross-sectional study will be undertaken among the populations of sub-urban areas of Chennai and Shillong cities, which represent the Southern and Northeastern regions of India. We will collect data on sociodemographic variables, economic characteristics, risk factors of COPD and comorbidities. The Global Initiative for Obstructive Lung Disease (GOLD) and Global Initiative for Asthma (GINA) definitions will be used for the diagnosis of COPD and asthma. Data will be analysed for estimation of the prevalence of COPD, asthma and associated factors.

Ethics and dissemination This study proposal was approved by the respective institutional ethics committees of participating institutions. The results will be disseminated through publications in the peer-reviewed journals and a report will be submitted to the concerned public health authorities in India for developing appropriate research and management policies.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is one of the major preventable chronic respiratory diseases (CRD).

The Global Initiative for Obstructive Lung Disease (GOLD) describes COPD as a common preventable and treatable disease, characterised by persistent airflow limitation that is usually progressive and associated with an enhanced chronic inflammatory response in the airways and the lung to noxious particles or gases.1

COPD is reported to have an estimated disease burden of 210 million people worldwide.2 Globally COPD was the fourth leading cause of death (5.1%) in 2004 and is projected to occupy the third position (8.6%) in 2030.3 Also COPD is a major cause of chronic morbidity; it was ranked 11th in 2002 and is projected to rise to seventh place in 2030.4 The prevalence of COPD in adults ranges between 0.2% in Japan and 37% in USA.5 The Burden of Obstructive Lung Disease (BOLD) group recently reported an average global COPD prevalence of 10.1% with wide variations across the participating countries.5

Additionally, COPD contributes to the economic burden faced by patients as well
as the healthcare infrastructure in the country, incurring 2–4 fold higher costs compared with asthma and ischaemic heart disease (IHD). The prevalence studies from Southeast Asia are scarce and many of the available estimates have been derived using non-standardised methodology. However, even these studies have reported a wide variation in the prevalence of COPD. It is widely accepted that ethnicity is responsible for the observed variability in COPD prevalence. A study conducted in the UK reported the effect of ethnicity in the prevalence and severity of COPD. WHO estimates suggest that 90% of COPD-related deaths occur in low and middle income countries. India and China constitute 33% of the total human population and account for 66% of the global COPD mortality.

Further, it has been estimated that COPD associated mortality is likely to grow by 160% in the Southeast Asian region in the coming decades. Globally the increase in the burden of COPD has been attributed to cigarette smoking among men and women, longer survival of populations, and high levels of air pollution, particularly in developing countries.

India is a large country comprising of people with varying sociodemographic profiles, cultural practices and ethnicities. Hence the risk factors for COPD are also likely to be different across various Indian states and regions. Together COPD, asthma and other respiratory diseases are the second (10.2%) leading cause of death in the population aged 25–69 years in India, as reported in 2001–2003, and they account for 3% of disability adjusted life-years (DALYs) lost. Of the CRD, COPD accounts for about 500,000 deaths in India, which is more than four times the number of people who die due to COPD in USA and Europe. A recently completed nationwide questionnaire-based study estimated the prevalence of COPD at 3.49% in India (ranging from 1.1% in Mumbai to 10% in Thiruvananthapuram). The spirometry test was not employed for the diagnosis of COPD in this study, and it is therefore possible that the reported COPD burden could be underestimated. Recently, the BOLD study conducted in Pune, Mumbai and Srinagar reported overall COPD prevalence estimates of 6.25%, 6.8% and 16.05%, respectively. Though the study adopted standardised procedures, it did not have adequate power to generate dependable prevalence estimates apart from the wide variations of prevalence.

Recent literature shows the co-existence of other important CRD such as asthma and post-tuberculosis (TB) sequelae with COPD. These CRD are considered to be important predisposing conditions for the development of COPD and can seriously influence the course of the disease. No COPD study has estimated the burden of these important comorbid conditions in India.

Without having valid baseline prevalence estimates for COPD from various regions in India, uniform national guidelines cannot be developed. As of now, valid, large-scale spirometry-based and community-level prevalence estimates are not available from India as a whole.

Since prevalence estimates were available from the BOLD study conducted in two urban centres from Western India and one centre from Northern India, we propose to undertake the present study to estimate the prevalence of COPD in Chennai in South India and Shillong in North Eastern India. The current study has been suitably designed and adequately powered to generate robust prevalence estimates of COPD and other comorbidities (asthma and post-TB sequelae) using internationally accepted standardised methodology. This study will also identify associated risk factors for COPD.

METHODS AND ANALYSIS

Objectives

Primary objective
► To estimate the prevalence of COPD among adults aged ≥25 years for each gender separately.

Secondary objectives
► To identify the risk factors for COPD.
► To estimate the prevalence of asthma and post-TB sequelae among adults aged ≥25 years for each gender separately.
► To validate the lower limit of normal (LLN) prediction formula for the diagnosis of COPD.

Study design

The proposed cross-sectional study will be conducted in adult populations living in sub-urban areas of Chennai and Shillong cities, located in South and Northeast regions of India, respectively.

Study sites

Chennai
The National Institute of Epidemiology (NIE) will carry out the study in Chennai. Adjacent to the Institute, NIE has a sub-urban cohort setting in the Ayapakkam area of Chennai. Approximately 10,000 households have been mapped and enumerated covering an approximate population of 45,000. The sample for the study will be drawn from this cohort setting.

Shillong
The North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences (NEIGRIHMS) will carry out the study in the sub-urban area of Shillong City in the Meghalaya state of India. The area has been used to conduct community-based research projects by NEIGRIHMS.

Study population

All residents of the study areas in Chennai and Shillong cities aged ≥25 years and who are able to provide written consent will be eligible.
informed consent will be eligible to participate in this study.

Rationale for including individuals aged 25–40 years

Generally, an age above 40 years is considered as the COPD target age group. This is primarily based on the assumption that tobacco smoking, which is the primary risk factor for COPD, begins in adolescence, and it would take 20–25 years of exposure to tobacco smoke to induce characteristic pathophysiologic changes of COPD in human lungs. However, in India domestic exposure to indoor-air pollution resulting from burning solid biomass, other health-adverse fuels and mosquito coil use has emerged as another important risk factor for COPD. As the exposure to indoor-air pollution may begin from infancy or childhood in homes where biomass fuel is a traditional fuel for cooking, young adults in the Indian subcontinent are likely to develop COPD at an early age. In humans, the lung function keeps improving until early adulthood and subsequently undergoes a natural physiological decline. Therefore, we wish to estimate the prevalence of COPD in the population between 25 and 40 years of age and identify associated risk factors.

The following conditions that could either affect the safety of the study participants during spirometry testing or influence the outcome of spirometry are considered as the exclusion criteria:

1. Any surgery in the abdomen, chest or eye in the last 3 months.
2. Women in the last trimester of pregnancy.
3. Myocardial infarction in the last 3 months.
4. Hospitalisation due to any heart problem within the past month.
5. Currently on treatment for TB.
6. Resting pulse more than 120 per minute.
7. Respiratory infection including common cold in the last 3 weeks.
8. Use of bronchodilators in the last 6 hours.

The following definitions are adopted for the study:

A. COPD: Any one of the following two standard spirometry definitions:
   1. Post-bronchodilator forced expiratory volume in 1 second (FEV₁/FVC) <70%.<sup>1</sup>
   2. Post-bronchodilator FEV₁/FVC <lower limit of normal (LLN)).<sup>22 23</sup>

The GOLD Committee published a consensus statement in 2001 for the use of a fixed FEV₁/FVC <70% value and fixed FEV₁ values to classify severity of COPD. Later, it has been realised that the prevalence of spirometry-based COPD is higher when using fixed values of FEV₁/FVC in comparison to using the LLN. A longitudinal study reported that the in-between group appeared to have a higher risk of hospitalisation and mortality attributable to some lung pathology. Therefore it is believed that using the LLN of FEV₁/FVC underestimates COPD.<sup>26</sup>

In the absence of clear evidence in India in favour of either of the two above-mentioned definitions, we decided to diagnose COPD based on both definitions. This will enable us to determine which criterion is better and more clinically relevant for COPD diagnosis in the Indian setting.

B. Asthma: The diagnosis of asthma will be done using Global Initiative for Asthma (GINA)<sup>27</sup> guidelines:
   - History of variable respiratory symptoms
   - Confirmed variable expiratory airflow limitation with pre- and post-bronchodilator spirometry.

Positive bronchodilator reversibility test (more likely to be positive if bronchodilator medication is withheld before test: short-acting β-agonist (SABA) withheld for ≥4 hours, long-acting β-agonist (LABA) withheld for ≥15 hours); increase in FEV₁ of >12% and 200 mL from baseline, 10–15 min after 400 μg of salbutamol or equivalent.

C. Post-TB sequelae: Information on past history of pulmonary TB will be collected using a validated questionnaire. A chest X-ray will be undertaken to diagnose the radiological sequelae of pulmonary TB. The lung function effects of pulmonary TB on COPD will be measured using spirometry.

Data collection

The data will be collected using tablet-based computers. All the study procedures will be done at convenient access points for the participants. The procedures will be regularly evaluated by the NIE team for quality assurance. The study will include the following study tools and domains of inquiry:

**Questionnaires**
- Core questionnaire.
- Occupational questionnaire.
- Stages of change questionnaire.
- Biomass questionnaire.
- Miscellaneous questionnaire.
- Spirometry questionnaire.
- Minimal data/refusal questionnaire.

The study questionnaires will be adapted from the BOLD study. The permission to utilise the adapted questionnaires in India has already been given by the BOLD Committee. Every participant will answer an interviewer-administered questionnaire. The original questionnaires are in English and will be appropriately translated into Tamil and Khari for the Chennai and Shillong sites, respectively, since the primary spoken language is not English at both these sites. The translations will be pretested in a smaller group for validation and back-translated by a different group.

- Anthropometry: Standing height and weight will be measured as directed by the National Health and Nutrition Examination Survey (NHANES) III Anthropometry Procedures Manual.<sup>28</sup>
- Spirometry testing: We will use the ndd Easyone spirometer for carrying out spirometry testing in this study. Pre and post short-acting bronchodilator inhaled spirometry testing will be done according to American Thoracic Society/European Respiratory...
Sample size and sampling plan

For COPD prevalence estimation, we will consider the following strata of populations:

- Stratum 1: 25–40-year-old males.
- Stratum 2: 25–40-year-old females.
- Stratum 3: 41+ year-old males.
- Stratum 4: 41+ year-old females.

Sample size calculation for each site

Sample size required for 41+ year age group is 3600 individuals (1800 each in male and female strata) which was calculated based on the assumption that the lowest reported prevalence of COPD in this age group in India is about 5% with an absolute precision of 1%, 95% CI and design effect as 1.15.

Sample size required for 25–40 years age group is 6300 individuals (3150 each in male and female strata) which was calculated based on the assumptions of assumed prevalence of COPD in this age group is 2.5% (50% of the prevalence of COPD in 41+ age group) with an absolute precision of 0.5%, 95% CI and design effect as 1.15.

Sample size for 25–40 years age group is 3000 individuals (1500 each in male and female strata) which was calculated based on the assumptions of assumed prevalence of COPD in this age group is 2.5% (50% of the prevalence of COPD in 41+ age group) with an absolute precision of 0.5%, 95% CI and design effect as 1.15.

Sample size formula

\[ n = \frac{[\text{DEFF} \times N \times p \times q]}{[d^2/\hat{Z}_1^2 \times (N-1)+p \times q]} \]

Sampling strategy

The entire study area will be divided into 60 equal portions consisting of 175 adjacent households, which will form a cluster. A single stage cluster sampling method will be adopted for recruiting the required number of participants in each of the 60 clusters. From each cluster 83 males (for 41+ years=30 and for 25–40 years=53) and 83 females (for 41+ years=30 and for 25–40 years=53) will be assessed for COPD. Random selection of a household in a cluster will be the starting point of the survey. All the eligible members in that selected household will be included in the study. The next nearest neighbourhood concept will be adopted until the required sample size in each stratum is achieved.

Data management and analysis

The questionnaire and anthropometry measurement data from the tablet computers will be transferred to the data management server at NIE on a daily basis from both the Chennai and Shillong sites. Necessary steps will be taken while designing the data collection tools and actually collecting the data to minimise missing data.

The spirometry data will be transferred to the designated computers in the study sites on a daily basis by the technicians. In the next step, the collated spirometry data will be sent to the central server at NIE on a weekly basis in an Access database using a secured and encrypted internet platform. The spirometry reading team (SRT) at NIE will grade and assign scores to each test. The SRT will also review the quality of the tests by each spirometry technician and recommend corrective measures, if necessary. Duplicate data will be retained at the study site. The three spirometric measurements, namely FEV1, FVC and FEV1/FVC, will be considered for analysis. We will use Global Lung Initiative (GLI) 2012 prediction equations for interpreting the observed lung volumes.

Descriptive statistics including proportions, mean and SD, median and IQR will be computed as applicable to cluster sampling. The prevalence of COPD and asthma for each age group and gender will be estimated based on cluster sampling. Odds ratio (adjusted and unadjusted) will be estimated for the factors associated with COPD and asthma through logistic regression analysis. Odds ratios (adjusted and unadjusted) will also be estimated for the comorbidities including TB with COPD and asthma through logistic regression analysis. The burden of COPD on life will be estimated in terms of impact on quality of life, activity limitation, respiratory symptoms and use of healthcare services. The distribution of various known community level risk factors will be described. The sensitivity and specificity of selected clinical symptoms for COPD will be assessed using spirometry as the gold standard.

Quality control

The manual of procedures is in place for all components of the study. The site investigators will serve as master trainers and be trained and certified by the central team. Study staff will be trained and certified in their respective activities and functions by the master trainers before initiation of the study. The training and certification activity will be supervised by the central team for assuring the quality of the process.

All questions in the questionnaires are coded for ease of data analysis and to ensure comparability between study sites. The study sites will submit the forward-translated
version of the questionnaires to the QC team at NIE. The translated versions will be back-translated to English and checked to assure the questionnaires are comparable with the original after accounting for the cultural norms of the respective areas.

The study tools will be pilot tested for operational feasibility and finalised before initiation of the survey.

**Discussion**

There are many unanswered questions about the epidemiology of COPD in India and there is a paucity of systematically collected prevalence data using well-standardised protocols from India. Most of the available prevalence estimates are not based on spirometry testing or adopted non-standard methods. However, spirometry is the internationally accepted gold standard for the diagnosis of COPD. Hence, the available data cannot be interpreted in the global context. A recent study on COPD prevalence from three cities in India using standardised methodology did not have an adequate sample size to reflect the true burden of the disease. Further, the same study did not represent Southern and Eastern parts of India.

Also, no study in India has concurrently estimated the other CRD such as asthma and post-TB sequelae, which are important comorbidities of COPD and are reported as important comorbidities of COPD. However, spirometry is the internationally accepted gold standard for the diagnosis of COPD. Hence, the available data cannot be interpreted in the global context. A recent study on COPD prevalence from three cities in India using standardised methodology did not have an adequate sample size to reflect the true burden of the disease. Further, the same study did not represent Southern and Eastern parts of India.

In this age group the information on COPD burden is not readily available. On successful completion of the project, we will be able to decide whether more centres will be required to obtain more dependable, uniform and geographically more representative data that could provide a clear and holistic picture of the prevalence of COPD across the country. Knowledge of the prevalence of COPD from multiple sites would also help in ‘modelling studies’ for COPD disease burden estimation at the national and sub-national levels. Further, the prevalence study sites will act as future focal points for initiating long-term cohort studies to define phenotypes and genotypes of obstructive lung diseases in India.

**Limitations**

Non-availability of valid spirometric reference equations for the Indian population will be a limitation, for which we propose to use GLI 2012 equations for data analysis.

**Outcome**

This study will be the first of its kind in India and will specifically contribute to developing a long-term cohort study to characterise the heterogeneity of COPD in various parts of India and in various ethnic groups. It will also contribute to clearly define the phenotypes and genotypes of COPD in subsequent studies. This may provide a replicable model to study the prevalence and heterogeneity of COPD in India and other developing countries.

**ETHICS AND DISSEMINATION**

The final study protocol, including the final version of the other essential documents, have been reviewed and approved in writing by the Institutional Human Ethics Committees of NIE (ID # NIE/IHEC/201/401–01) and NEIGRIHMS (ID #NEIGR/IEC/2015/80). Written informed consents will be obtained from all the participants before enrolment in the study. All the study procedures will be done free of cost at the convenient access points for the participants. The participant will be referred to the nearest public health facility for appropriate treatment, if COPD or any other ailment/abnormality gets detected during the examination.

This study will provide population-based COPD prevalence estimates, based on internationally accepted standardised methods. We will disseminate the study findings by publishing manuscripts of high impact in peer-reviewed scientific journals. Due to the interdisciplinary nature of the study findings, we will also present the study findings at national and international conferences, which will attract other researchers in the field for collaboration. The study findings will be submitted to the concerned public health authorities in India for developing appropriate research and management policies.

**Acknowledgements** We thank Dr MV Murhekar and Dr Tarun Bhatnagar for reviewing and providing helpful comments for improving the manuscript. We thank Dr Peter Burney, BOLD Executive Committee, for providing access to the BOLD questionnaires and permission to use the questionnaires for our study.

**Contributors** RP drafted the article. RP, PK, VS, AB, BB, PKB and SMM contributed to the development of study design and participant recruitment plan. PK and VS contributed in calculation of sample size, sampling design and development of data analysis plan. All authors provided feedback and approved the final protocol.

**Funding** Funding has been approved for the study by Shillong site from Science and Engineering Research Board, Department of Science and Technology, Government of India through grant # EMR/2015/000737. The data collection for the study will be initiated in March 2017 and the site is currently in the process of finalizing pre-initiation arrangements. We have approached other agencies to seek funding support for the Chennai site.

**Competing interests** None declared.

**Ethics approval** Institutional Human Ethics Committees of National Institute of Epidemiology (ICMR), Chennai, India and Northeastern Indira Gandhi Regional Institute of Health and Medical Sciences, Shillong, India.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) licence, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

**REFERENCES**

1. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis management and prevention of Chronic