Prevalence of Respiratory Symptoms and Chronic Obstructive Pulmonary Disease, and Reference Values for Lung Function Testing in Kinondoni District, Dar Es Salaam, Tanzania

G. Toril Mørkve Knudsen

Centre for International Health and Institute for Internal Medicine

Faculty of Medicine and Dentistry

University of Bergen, Norway

Muhimbili Medical Research Centre, National Institute for Medical Research, Tanzania

2010

Prevalence of Respiratory Symptoms and Chronic Obstructive Pulmonary Disease, and Reference Values for Lung Function Testing in Kinondoni district, Dar Es Salaam, Tanzania

G. Toril Mørkve Knudsen

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Philosophy in International Health at the University of Bergen.

Centre for International Health and Institute of Internal Medicine

Faculty of Medicine and Dentistry

University of Bergen, Norway

Muhimbili Medical Research Centre, National Institute for Medical Research, Tanzania

2010

Acknowledgements

I want to express my gratitude to Dr. Jon Andrew Hardie who has been my main supervisor throughout my period as master-student. His knowledge of respiratory physiology, as well as his scientific skills and capacity, have provided me with both increased understanding and attention to analyses and statistics as well as for pulmonary function. He has given me the best scientific training a master-student could wish for, and both motivated, inspired and challenged me to perform better.

I also want to extend my gratitude to my co-supervisor, Dr. Sayoki Mfinanga, for support, guidance and valuable advice prior to and during my fieldwork. I would never have accomplished the data collection without help and support from him and the kind collaboration of all the employees at Muhimbili Research Centre, National Institute for Medical Research, NIMR. Also, I thank Dr. Esther Ngadaya and Dr. Veneranda Bwanza at NIMR for their practical help and guidance during the initial stages of the field work.

I give a special and warm thanks to the fieldworkers; Michael Stanislaus, Peter Mbelele, and Salala Fortunatus, for their high quality work, endurance and good spirit during the field work. Also thanks to Azizi Juma for driving us safely back and forth from the field every day.

I also wish to thank dr. Svein Gudmund Hinderaker, at the Centre for International Health, for his patience and valuable supervising in the initial stages of cleaning and coding of the data. Also, I am deeply grateful for his help and expertise with some of the analyses.

Last, but not by any means, least I would like to thank my father, Odd Mørkve, for his reassurance, support and advice throughout the process, for inspiring scientific discussions as well as indispensible feedback at every stage of the work. It has been a privilege to share this period with you, and I am blessed to have you in my life.

Table of content

Executive Summary

In the wake of socioeconomic development, changing demographics towards ageing populations and the expatiation of non-communicable diseases put additional strains and new challenges unto health services and disease management in low-income countries. Chronic Obstructive Pulmonary Disease (COPD) is by the WHO Global Burden of Disease Study projected to rank as third leading cause of death worldwide in 2020, though studies concerning prevalence of COPD and presence of risk factors in low-income settings are yet rare. The expected frequency and burden of COPD is however anticipated to be of same magnitude in these countries, where exposures to occupational and domestic hazards are considerable, as for infections and morbidities, influences known to be detrimental to respiratory health. Assessing prevalence and diagnosis is an important step towards clinical management and awareness regarding implementation of prevention initiatives. A necessity in diagnosis of COPD and measuring disease severity is the use of reference values for normal lung function. Adhering to American Thoracic Society (ATS) recommendations, such prediction equations should be based on healthy people with the same anthropometric characteristics and ethnic origin as the subjects being tested. This study aims at deriving reference values of a healthy adult suburban population in Dar Es Salaam, Tanzania, and to estimate prevalence of COPD from a random sample of the population.

The study is based on a descriptive cross-sectional design. The source population comprised all adults above the age of 15 in Kinondoni district. Estimating prevalence to 15 % and accounting for reduced precision due to cluster sampling procedure, the sample size required was set to 300 participants. Approval was obtained from the Ethical Committee of Western Norway (REK-Vest) and the Medical Research Coordinating Committee of the National Institute for Medical Research (MRCC). Subjects within the age-span examined and who were willing to participate, were eligible for participation. A total of 365 subjects were enrolled in the study. Sub-selection to the reference sample and for generation of prediction equations were based on ATS recommendations, where subjects with negative responses to core questions from ATS-DLD regarding respiratory symptoms and doctor diagnosed heart/chest illnesses were selected, providing their spirometric data met ATS criteria.

Lung volumes was tested using the ndd EasyOne spirometer, and the following parameters were measured; peak expiratory flow (PEF), forced expiratory volume in one second (FEV1), forced expiratory volume in six seconds (FEV6) and the ratio of these two measurements

(FEV1/FEV6). All spirograms were reviewed by an experienced chest physician, adhering to ATS acceptability and reproducibility criteria for selection of best pulmonary function curve. Anthropometric data, including weight, standing height and sitting height were collected, and age was recorded as birth-date or as mid year in the year of birth. Respiratory symptoms were recorded using the ATS-DLD validated questionnaire, and questions regarding socioeconomic conditions were derived from a culture specific questionnaire used in the Tanzania Demographic Health Survey from 1999 and 2002.

Prevalence of COPD was estimated according to the Global initiative for chronic Obstructive Lung Disease (GOLD) and ATS criteria, and severity was determined by GOLD disease stages.

Based on the inclusion criteria, a total of 150 subjects, 52 men (32,7 %) and 98 women (47,6 %) were selected to the reference value group. Median age for men and women in the reference sample was 34 years. The spirometric parameters FEV1, FEV6, FEV1/FEV6 and PEF were regressed against sample mean age, height and weight, and the models were assessed in terms of whether a linear or curvilinear prediction produced best fit for the present data. In the male strata, an exponential model was selected as regression equation for all of the spirometric variables, with the exception of prediction of PEF when sitting height was independent variable, where a linear model was applied. For women, a linear model was chosen in further analyses, except for the prediction equations for PEF, where an exponential equation best fitted the data. For women, weight was retained in the final model when standing height was part of the independent variables and regressed against PEF. For the other spirometric variables, weight was non-significant, and was removed from the equations in further analyses. Height, and height square in the exponential equations for men and women, entered all the regression models with the exception of the model predicting FEV1/FEV6, as it did not provide a significant contribution to the variance of the dependent variables. All the spirometric parameters were negatively related to age, and all increased with height, with exception of FEV1/FEV6. The lower limit of normal (LLN) is presented as – 1,645 x SEE which is the age and height specific, estimated 5th percentile for the reference sample. The reference equations derived from our study do not allow direct comparison with previously published predictions due to differences in source populations and for the effects of altitude, however, on average, our equations generated lower reference values. When comparing our predictions to those published by Mustafa in 1977, no secular changes in FEV1 values could be traced. .

Both presence of respiratory symptoms and illnesses were more frequently reported among the female participants, though considerable in both sexes.

Regarding questions on smoking exposure, a proportion of 14,0 % of the study population responded that they were current smokers, and 4,7 % was ex-smokers. Patterns of cigarette smoking differed substantially between the sexes, where the proportion of current smokers was considerably higher for men (30,2 %) than for women (1,5 %).

Prevalence of COPD when applying a fixed cut off ratio of $FEV1/FEV6 < 0,73$, was 12,6 %, and it was equally distributed among men (13,9 %) and women (11,5 %). Prevalence of COPD when classified as below the lower limit of normal was 7,9 %, also equally distributed with 6,9 % and 8,7 % of men and women respectively. In our study, prevalence of stage I and II COPD was 6,1 % and 5,8 % respectively.

Direct logistic regression was conducted to assess the impact of certain determinants on COPD. Separate analyses were carried out for GOLD and ATS/ERS defined COPD. The multivariate models contained the following independent variables; sex, age, domestic exposure, occupational exposure, smoking status and socioeconomic position. When GOLD defined COPD was dependent variable, the following covariates made a statistically significant contribution to the model; being ex-smoker, $(p = 0,01)$, and age $(p < 0,0005)$. The strongest predictor was ex-smoker, with an odds ratio (OR) of 5,37. In the model using the ATS defined COPD prevalence, the independent variable, ex-smoker was still significant $(p = 0,004)$ with an OR of 7,92. In addition, domestic exposure made a statistically significant contribution to the variance of the model $(p = 0,013)$. However, the OR was less than 1 (0,19), indicating a negative association.

Introduction

"Chronic obstructive pulmonary disease (COPD) is a preventable and treatable disease with some significant extrapulmonary effects that may contribute to the severity in individual patients. Its pulmonary component is characterized by airflow limitation that is not fully reversible. The airflow limitation is usually progressive and associated with an abnormal inflammatory response of the lung to noxious particles or gases".

This is the clinical definition of COPD recommended by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) (1). According to the same guidelines, COPD is determined by use of a fixed percentage of spirometric predicted values, with airflow limitation defined as a forced expiratory volume in one second/forced vital capacity (FEV1/FVC) ratio less than 70 % (1, 2, 3). When applying ATS[1]/ERS[2] guidelines, ventilatory obstructivity is defined as a reduced FEV1/FVC ratio below the statistically derived fifth percentile of the predicted value, and this lower limit of normal is used as a cut off to determine proportion of responders with COPD (3, 4, 5). The WHO Global Burden of Disease Study estimated COPD to rank as the fifth leading cause of death in 2001, and predicts the disease to be the third leading cause of death worldwide in 2020 (6). The expected increase in mortality is brought forth by both changing demographics, with ageing populations reaching the stage of developing chronic diseases, as well as an anticipated increase of cigarette smoking in many African and Asian countries (1, 6).

Besides cigarette smoking, occupational pollutants as vapours, irritants and fumes are found to be important contributing factors in development of COPD (1, 2, 4). Several studies within occupational medicine have examined the association between environmental exposures and respiratory impairment (7, 8, 9). Some researchers have also shown an additive effect of smoking and occupational pollutants on mortality from COPD (1, 7). Additional risk factors, like indoor pollution from solid fuels for cooking and heating, have been thought to increase the prevalence of COPD, especially among women in developing countries (1, 6, 10, 11). Three billion people in the world, mainly in the Middle East, Africa and Asia, use biomass fuels and coal as their main source of energy. Many of them live in poorly ventilated houses

---

[1] American Thoracic Society [2] European Respiratory Society

(1). This illustrates the interconnection of respiratory symptoms and socioeconomic factors that besides living conditions and increased levels of indoor pollution, also influence exposure to potential environmental hazards and episodes of respiratory illness (4, 6, 8). Socioeconomic status is thus thought to be an important determinant of lung function (4, 6). HIV infection has been shown to accelerate the onset of smoking-related emphysema, and might be an important cofactor in the development and burden of COPD (12). This reinforces the thought of COPD as both underreported and under-diagnosed in low-income countries, none the less causing high morbidity and mortality in poorly resourced populations, with less access to health care and preventive actions.

The fact that COPD among other non communicable diseases is growing worldwide, putting strains on already heavily pressured populations, justifies the need to investigate the burden of obstructive lung disease in low-income countries, and to increase awareness, implement prevention strategies and improve management of this disease. Measurement of lung function and prevalence of obstructive lung diseases in the catchment area and age group for the present study is important, as there is no recent published data on respiratory function in the general population in Dar Es Salaam, Tanzania.

This study has been conducted under the framework of collaborations between National Institute for Medical Research (NIMR), Tanzania, and the University of Bergen. The two partners have a Memorandum Of Understanding (MOU) whereby a number of Tanzanian scientists have conducted several projects for their Master and PhD degrees. The study presented here is one in the series of training projects that are conducted under this framework and the project has investigated an important health problem which is defined in the National Health Research Agenda of Tanzania. In addition, the project will contribute and complement to NIMR, supported by the International Association of National Public Institute (IANPHI), with an overall goal of developing an evidence-based approach to establishing national surveillance for Non Communicable Diseases and creating the foundation for a national programme through training and exchange.

Assessing COPD

Quantifying prevalence and assessing diagnosis and severity of COPD is a challenge in low-income countries and in epidemiological field work, where resources are scarce and there is less availability of appropriate equipment and tools needed to fully apply standardized and recommended guidelines (1, 2, 3, 5). GOLD acknowledges this in its report, emphasizing that in areas lacking access to state-of-the-art diagnostic tools, diagnosis should be made with the equipment available (1[3]). The Burden of Obstructive Lung Disease (BOLD) Initiative addresses the scarce documentation of prevalence, true burden and key risk factors across countries (10). In one of their studies, involving 12 different countries, Cape Town, South Africa had by far the highest prevalence of clinically manifested COPD at 12 %. By comparison Bergen, Norway, had an estimated prevalence of 7 % (13). The South African study site revealed high numbers of participants with previous tuberculosis in addition to extensive exposures of occupational pollutants and cigarette smoking (13). This pattern of co-morbidity and presence of risk factors fits well in other sub-Saharan countries, and indicates that the expected frequencies of COPD in other African countries are of similar magnitude, though this has not yet been established.

Lung function measurements

In accordance with GOLD and ATS guidelines, the following basic parameters are frequently used in lung function testing and in assessing respiratory diseases; forced expiratory volume in one second (FEV1), forced vital capacity (FEV6) and the ratio of these two measurements (FEV1/FEV6). Due to increased risk of false-positive diagnosis in the interpretation, it is recommended not to include too many indices of lung function in the testing (1, 4, 5). The selected parameters are important in identifying obstructive airflow limitation, and FEV1 and FEV1/FEV6 are also shown to be independent predictors of mortality from respiratory diseases (2, 4, 5, 13, 14, 15, 16).

A necessity in assessment of COPD and disease severity is the use of reference values for normal lung function. According to ATS recommendations, selection of such prediction equations should be based on healthy populations sharing the same anthropometric characteristics (sex, age and height) and ethnic origin as the subject being tested (11). When

---

[3] GOLD report, page 89

compared to Caucasians of European descent, other ethnic groups have shown to have smaller static and dynamic lung volumes (4). When measured at same standing height, spirometric values derived from White populations tend to over-predict values in Black subjects by 12 % for total lung capacity (TLC), forced expiratory volume in one second (FEV1) and forced vital capacity (FVC) (4, 11). The practice of using fixed adjustment factors of published reference values to control for such ethnic differences is however questioned, and when possible, it is recommended to use race specific predictions (11, 17, 18, 19, 20).

The causality of ethnic differences is not fully accounted for, though it is thought to be partly due to differences in body build, and that Blacks have smaller trunk-to-leg ratio than Whites (4, 11). Studies have tried to adjust for this by using sitting height as anthropometric measure in lung function testing, which reduces, but not eliminates the differences between Blacks and Whites (11, 17, 18, 20, 21). Ethnic differences have also been thought to concur with environmental and socioeconomic factors, rather than being merely genetic inherent, a connection which makes the selection and use of appropriate reference values even more complex (4, 17, 18, 19, 20, 21). Another critique of using ethnic scaling- factors, is the inability of such a method to face socio-environmental factors as nutrition, growth and health status as influences of lung function , and how these determinants are subject to development and change throughout the decades, a phenomenon known as the cohort-effect (4, 11, 21, 22, 23, 24). Several studies of normal lung function on the African continent have pointed out such a secular trend, where higher spirometric values have been measured in more recent studies (19, 20, 24, 25, 26). This has made researchers aware of the need of reference values to be updated at regular intervals, preferably every decade (4, 11, 26). Even though there are several published reference values derived from healthy African populations (19, 20, 22, 25, 26, 27), they all acknowledge their limitations in terms on generalisability across countries. Important geographic determinants like altitude has been shown to explain some of the differences in lung function within populations on the African continent (19). Spirometric values of forced expiratory flows are increased at high altitudes (above 1,500 m), and it has been shown that people living at higher altitudes have larger lung volumes (4). Selection of study population, whether randomly selected from a community, or through a workforce has further been demonstrated to influence the prediction equations derived. Occupation based samples of men and women have had higher lung volumes than those derived from community source, a bias known as the healthy worker effect (19, 20). The different determinants and factors contributing to lung function, and the possible bias in the process of deriving reference values, enhance the importance of appropriate prediction equations in the

diagnosis of COPD and disease severity. In Tanzania, the work of Mustafa (28) has contributed to generation of proper prediction equations in assessment of lung function and respiratory diseases. The present study seeks to update reference values of the general adult population of today.

Aims and objectives

The present study has a twofold aim: To derive reference values of a healthy adult suburban population in the district of Kinondoni, Dar Es Salaam, Tanzania and to estimate prevalence of COPD from a random sample of the population

Specific objectives:

1 To measure spirometric values of FEV1, FEV6, FEV1/FEV6 ratio and PEF in all study
   participants

2 To derive reference values from spirometric measurements of non-symptomatic
   study subjects

3 To estimate prevalence of respiratory symptoms based on ATS-DLD[4] (American Thoracic
   Society- Division of Lung Disease) and culture specific questionnaire

4 To determine prevalence of COPD based on spirometric values according to GOLD and
ATS diagnostic criteria

---

[4] ATS-DLD: validated respiratory questionnaire

Methods

Study design and participants

The study is based on a descriptive cross-sectional design. The source population comprised all adults above the age of 15 in Kinondoni district, a suburban area of Dar Es Salaam, Tanzania, roughly estimated to hold a population of 350,000. Exact population registries do not exist. Participants were recruited by multistage cluster random sampling, and the procedure comprised three stages of selection. In the first stage, two out of 27 wards in Kinondoni district were selected. At next stage, two of a total of five areas within each ward were chosen. In the last stage, four ten-cell leaders were selected from each area, adding up to a total of 16 clusters. The term 'Ten-cell leader' refers to the smallest administrative unit in Tanzania, and one ten-cell leader has the responsibility for approximately 10-12 households (29). The number of ten-cell leaders varied in the selected areas; one area held a total of 40 ten-cell leaders, whereas another comprised 10 ten-cell leaders and the last two areas consisted of 8 ten-cell leaders each. The primary sampling unit comprised all households under the selected ten-cell leaders.

Sample size was calculated on the basis of the objective to determine prevalence of COPD, at a 95 % confidence interval, and with precision set at 0,05. Assumptions of expected frequency of COPD were based on a study in South Africa, showing a prevalence of COPD around 12 % (13). Estimating prevalence of COPD to 15% and accounting for reduced precision due to cluster sampling through use of a design effect at 1.5[5], the minimum sample size required was calculated to 300 participants. The presented formula was applied in the sample size calculation (30)

$$n = \frac{Z^2 P(1-P)}{d^2}$$

Where n = sample size

Z = level of confidence

P = expected prevalence

d = precision, here set at 0.05

---

[5] Inflation factor of 1,5 obtained on basis of mean cluster size of 25 individuals and an intra correlation coefficient of 0,02.

Recruiting study subjects was initiated by first informing the randomly selected ten-cell leaders about the aim of the study, and seeking their approval to collect the necessary data. Next, a mobile testing team carried out a knock-on-door approach to the households under their selected ten-cell leaders. Subjects within the age-span examined and who were willing to participate, were included in the study. Selection of subjects eligible for generating reference values followed ATS recommendations, where subjects with negative responses to core questions from ATS-DLD regarding respiratory symptoms and doctor diagnosed heart/chest illnesses were selected, providing their spirometric data met ATS criteria (4, 31). Questions regarding self reported occupational exposures were excluded from the stratification because they yielded unreliable information. However, duration of symptoms like cough, phlegm, wheezing and breathlessness were accounted for, as were past or present history of smoking and respiratory illnesses like chronic bronchitis, emphysema, asthma, tuberculosis, doctor diagnosed heart problem and high blood pressure confirmed by a doctor.

Data collection methods: definitions and measurement

Spirometry

The spirometric testing was performed using the *ndd EasyOne* spirometer (ndd Medizintechnik AG), which has proven to be suitable in field work as it operates on batteries and requires no calibration, yet achieving a high degree of accuracy and reliability (10). The following parameters were measured; peak expiratory flow (PEF), forced expiratory volume in one second (FEV1), and forced expiratory volume in six seconds (FEV6). In the present study FEV6 was chosen over FVC, as it has been shown to be advantageous due to less exhalation time, implying both shorter coaching time as well as reduced physical discomfort for the participants. It has also been demonstrated to display less test variability when compared to FVC (10, 14, 15, 16). All study subjects performed a minimum of three and a maximum of eight tests, and the best FEV1, FEV6 and PEF were selected and used in further analysis. The spirometric testing was performed by trained assistants following a standardized procedure adherent to ATS/ERS guidelines (4, 31). The test subjects performed spirometry without nose clips, and sitting position were allowed if participants found the maneuver in standing position to be exhausting. All spirograms were reviewed by an experienced chest physician, where ATS acceptability and reproducibility criteria were followed for selection of best pulmonary function curve; forced exhalation time should exceed a minimum of 6 seconds or show a plateau of the volume-time graph. Further each participant should display at least

two acceptable reproducible maneuvers with both FEV1 and FEV6 within 200 ml and with absence of cough during the first second of the maneuvers (31).

Anthropometry

Anthropometric data, including weight, standing height and sitting height were obtained using robust equipment easy to transport in the field. The anthropometric measurements were collected by field personnel trained by the author, and the standardized procedure was monitored by the principal investigator throughout the data collection.

The participants' weights were obtained using a weight scale. Standing height was measured using a metal tape-measure and with heels, shoulders and occiput positioned against the tape. Sitting height was measured with a tape-measure affixed to a vertical wooden plank with the subjects sitting on a firm wooden chair with buttock, shoulder-blades and occiput touching the plank. Participants performed the anthropometric measuring barefoot.

Age was recorded as birth-date or as mid year in the year of birth. The data were transformed to age in years in further analysis.

Questionnaire data

Respiratory symptoms were recorded using the ATS-DLD validated questionnaire, which was translated into Swahili, then back-translated to English prior to data collection, to ensure that questions and the phrasing used were adherent to the original format (see appendix 1). Questions regarding socioeconomic status were derived from a culture specific questionnaire used in the Tanzania Demographic Health Survey from 1999 and 2002 (see appendix 2). The questionnaire was administered face-to-face by field personnel, trained to reduce information bias regarding interview technique, where phrasing of questions, and appropriate wording, was enhanced to obtain acceptable intra and inter- observer variability.

Ethical considerations

Approval was obtained from the Ethical Committee of Western Norway (REK-Vest), and the Medical Research Coordinating Committee of the National Institute for Medical Research (MRCC). After subjects eligible for participation were informed of the nature and purpose of the study, and had given their informed consent, they were enrolled providing inclusion criteria were met. Informed consent form is attached in appendix 3.

Statistical analysis

Data regarding demography, respiratory symptoms and socio-environmental conditions included all responders. Estimation of normal spirometric values and COPD prevalence included study participants with acceptable spirometry only. The data used to generate reference values were analyzed using standardized and hierarchical multiple regression techniques. Both linear and curvilinear formulae were obtained, and the model best fitted for the present data were chosen. The following dependent variables were applied; FEV1, FEV6, PEF and FEV1/FEV6, and regressed against the following independent variables; age, height and weight. The explanatory variables were assessed in terms of highest contributing R statistic, variable significance at $p<0.05$, partial correlation coefficients and lowest SEE values. Outliers were defined as cases that had standardised residual values above 3,0 or below -3,0, and were excluded from further analyses. Preliminary analyses were conducted to ensure no violation of the assumption of normality, linearity and multicollinearity. Prediction equations were estimated separately for men and women. The lower limit of normal (LLN) was defined as the 5th percentile, estimated as $1,645 \times SEE$, for each of the dependent variables. The FEV1/FEV6 ratio did not perfectly match a Gaussian distribution, and logarithm and square root transformation were conducted to assess whether the proportion of skewness decreased. As the distribution still showed some deviation from the normal curve and the standard error of estimate was nearly identical to the non log transformed, the non log transformed FEV1/FEV6 was selected in further analyses.

Reference values generated from the present study were compared to previously published predictions by use of paired t-tests, and the magnitude of mean differences was obtained by Eta squared calculations. The Independent-Sample T-test was used to compare continuous variables, and Pearson Chi square and Wald statistic tests were used to compare group differences for categorical variables. Definition and classification of COPD was based on ATS and adjusted GOLD criteria, as the present study used FEV1/FEV6 ratio over FEV1/FVC in diagnosing (15, 16, 17). Severity was determined by GOLD disease stages (1, 4, 5). Diagnoses were made on the basis of simple spirometry. The study further defined the lower 5 % of the non-symptomatic responders as below the normal limit (LLN) and used this as a cut off to determine proportion of responders with COPD (4, 5). Descriptive statistics were applied, using frequencies and cross tabulation. Logistic regression analyses were

carried out to assess for potential determinants of COPD. The models were adjusted for the following variables; sex, age, smoking status, pack years, presence of occupational and domestic exposures and level of socioeconomic position. The smoking exposure variable, pack years, was computed by dividing the number of cigarettes smoked per day by 20, multiplied by number of years smoked. Pack years were computed for both current and ex-smokers. Smoking status contained the following variables; never smoked ex-smoker and current smoker. Occupational exposure was computed by the following variables; worked ≥ 1 year in dusty work or ever exposed to gas/chemical in work. Domestic exposure was constructed by positive responses to question regarding use of biogas and charcoal at home. Measures regarding socioeconomic information were collected at individual level, and a socioeconomic position variable (SEP index) was constructed according to both economic and social dimensions represented by questions regarding level of education, present employment, recent work for pay and available money resources derived from the culture specific questionnaire used.

Logistic regression analyses were generated to estimate odds ratios (OR) with 95 % confidence intervals (CI) of COPD. Preliminary analyses were conducted to ensure no violation of the assumption of normality and multicollinearity. Both univariate and multivariate analyses were conducted. The models were assessed by goodness of fit tests. Prevalence estimates were adjusted for multistage cluster sampling during analyses. Each participant was given a sampling weight constructed by use of SPSS complex samples analysis, where inclusion probabilities at each of the three stages of the sampling procedure were added. Variables and differences were assessed in terms of significance at p< 0,05 and their 95 % confidence intervals. Descriptive statistics are reported as means and standard deviations (SD) when normally distributed, skewed data are expressed in medians and percentiles. Calculations were done using the Statistical Package for the Social Sciences version 15 (SPSS Inc., Chicago, IL, USA).

Questionnaire data were double entered, cleaned and coded using Epi data version 3.1 (Centre for Disease Control and Prevention, Atlanta, GA, USA), and were exported to SPSS for the statistical analyses. The spirometric data were automatically stored in the spirometres used, and were electronically exported to SPSS.

Results

Of the 16 selected clusters, 14 were visited, and a total of 365 subjects were enrolled in the study, 159 (43,6 %) men and 206 (56,4 %) women. As 135 subjects in the visited households were absent at the time of testing, the response rate attained was 74%. However, the denominator in this fraction included subjects with unknown eligibility status. Further, 41 participants were excluded from the study due to conditions contraindicating spirometry testing (10): 19 were due to known heart disorder, 15 had undergone chest/abdominal surgery, 5 were in last trimester of pregnancy, and 2 had undergone eye surgery. Eleven eligible subjects refused to participate, giving a cooperation rate of 97 %. After evaluating the spirometric measurements, 39 study participants (10,7 %) did not accomplish tests results in compliance to ATS criteria, and were excluded from analysis regarding spirometric lung indices.

Synopsis of paper 1

On the basis of the selection criteria, a total of 150 subjects, 52 men (32,7 %) and 98 women (47,6 %) were selected to the reference value group. Median age for the selected men and women were 34 years. Bivariate analyses were conducted to evaluate correlations for FEV1, FEV6, FEV1/FEV6 and PEF according to sample mean age, height and weight. Age, height and sitting height correlated significantly with the dependent variables FEV1 and FEV6 for both sexes. For women this was also the case with PEF. FEV1/FEV6 showed no significant correlation with height, and neither of the dependent variables correlated with weight in the female strata, whereas for men, weight showed a correlation with the FEV1/FEV6 ratio. The regression analyses were performed separately for males and females. Assessment whether a linear or curvilinear model produced best fit for the present data were done, according to R square change and p values. In the male strata, an exponential model was selected as regression equation for all of the spirometric variables, with the exception of prediction of PEF using sitting height as independent variable, were a linear model was applied. For women, a curvilinear model was selected for the prediction of PEF, whereas for the other spirometric parameters, a linear model was chosen. For women, weight was retained in the final model when standing height was part of the independent variables and regressed against PEF. For the dependent variables FEV1, FEV6 and FEV1/FEV6, weight was non-significant,

and was removed from the equations in further analyses. Height, and height square in the exponential equations for men and women, entered all the regression models with the exception of the model predicting FEV1/FEV6, as it did not provide a significant contribution to the variance in the present sample and only showed a marginally improvement of the R statistics and the corresponding SEE values. Tables 4 and 5, see attached paper 1, summarize the prediction equations generated. All the spirometric parameters were negatively related to age, and all increased with height, with exception of FEV1/FEV6. The lower limit of normal (LLN) is presented as – 1,645 x SEE which is the age and height specific, estimated 5th percentile for the reference sample.

The prediction equations derived from our sample are compared to those of Mokoetle, reference values generated from a healthy South African University workforce (21), and with Hankinson's equation (NHANES iii),of African-American above 20 years of age (32). Reference values for the male strata are also compared to Louw equations of Black and White South African men (20). Although the reference equations derived from our study do not allow direct comparison with previously published predictions due to differences in selection of source populations and effects of altitude, on average, our equations generated lower reference values. We did not trace any secular changes in FEV1 values when comparing our predictions to those of Mustafa from 1977 (28).

Synopsis of paper 2

Median age was 37 and 34 years for the male and female strata respectively. Age ranged from 16-79 for men and 15-90 for women. Tables 2 and 3, see paper 2, summarize frequencies and proportions of respiratory symptoms and conditions, based on the participants' responses to the ATS-DLD questionnaire. Both presence of symptoms and illnesses were more frequently reported among the female participants, though considerable in both sexes.

In tables 4 and 5, exposures to various risk factors are reported; table 4 presents history and magnitude of current/previous smoking among study participants, and table 5 summarizes self judged exposure to occupational and domestic pollutants. Men reported more frequently exposure of dust and gas/chemicals in work, whereas a higher proportion among women reported presence of domestic exposure as biogas and charcoal.

A total of 14,0 % of the study population responded that they were current smokers, and 4,7 % was ex-smokers. However, patterns of cigarette smoking differed substantially between the sexes, where the proportion of current smokers was considerably higher for men (30,2 %) than for women (1,5 %). Also 8,8 % of men, in contrast to only 1,5 % of women, reported a

previously history of smoking. Mean pack-years for current smokers were 8,7 and 6,6 years for men and women respectively, whereas mean pack-years for ex-smokers were 7,5 for men, and 1,47 for women.

Prevalence of COPD when applying a fixed cut off ratio of FEV1/FEV6 < 0,73, was 12,6 %, and it was equally distributed among men (13,9 %) and women (11,5 %). Prevalence of COPD when classified as below the lower limit of normal was 7,9 %, also equally distributed with 6,9 % and 8,7 % of men and women respectively. This is presented in table 6, together with frequency and proportion of cases of COPD according to GOLD's four disease severity stages. Among the study participants with COPD, only a small fraction were classified in stage 3 or higher (0,3 %). The age distribution of COPD cases according to GOLD and ATS/ERS are also summarized in table 7.

Direct logistic regression was performed to assess the impact of certain determinants on COPD. Separate analyses were carried out for GOLD and ATS/ERS defined COPD. The multivariate models contained the following independent variables; sex, age, domestic exposure, occupational exposure, smoking status and socioeconomic position. The following independent variables made a statistically significant contribution to the GOLD defined COPD model; being ex-smoker, (p = 0,01), and age (p < 0,0005). The strongest predictor was ex-smoker, with an OR of 5,37. When the ATS defined COPD was dependent variable, ex-smoker was still significant (p = 0,004) with an OR of 7,92. In addition, domestic exposure made a statistically significant contribution (p = 0,013). However, the OR was less than 1 (0,19), indicating a negative association. The logistic regression was not adjusted for complex sampling, and the results should be interpreted with caution as the precision of the estimates presupposes simple random sampling. Table 8, summarizes the odds ratios generated and present them with their corresponding p-values and 95 % confidence intervals.

General discussion

Discussion of methods

Selection of source population

Our study was conducted in Dar Es Salaam, which is situated at coastal level in Tanzania. The source population comprised adults above the age of 15 in Kinondoni district, a suburban area of Dar Es Salaam, thus the sample was derived from a community source. Both altitude and type of source population are found to explain some of the variability in spirometric measurements across studies. By using a workforce to generate a study sample, a bias, known

as the healthy-worker effect, is thought to generate higher spirometric values. Both Mokoetle and Louw used a workforce as source population to generate their reference values (20, 21). Both studies were conducted in Johannesburg, South Africa, which is situated 2000 m above sea level. According to White et al, this can account for up to 400 ml differences in spirometric measurements (19). Our study generated lower spirometric values when compared to those of Mokoetle and Louw. When applying Mustafa prediction equations from 1977, derived from a northern male Sudanese population and also tested to fit a Tanzanian population, this generated similar predicted values for FEV1 as in our study (28). The Sudanese site was situated at 400 m altitude, and the sample was derived from an urban community source. This raise questions to what extent secular changes, and the bias known as the cohort effect influence on the spirometric measurements, though possible differences in socio-environmental conditions for the study populations generating the reference values have been thought to influence lung function. On the other hand, the present sample size might not been sufficiently large to detect differences in lung volumes according to changes in determinants as nutrition, health status and environmental factors. The study also failed to point out any interconnectedness between lung volumes and socioeconomic factors.

In accordance with the Barker Fetal Origins of Disease hypothesis, *in utero* under-nutrition is thought to result in permanent detrimental changes associated with an increased susceptibility of developing chronic diseases later in life (33). Normal lung growth is found to be related to exposures during gestation and childhood, and a reduced attained lung function is associated with an increased risk of developing COPD. Researchers have also found a relation between birth weight and FEV1 in adulthood. Whether this applies to our study population is unknown.

Sample size

The sample size was calculated at a 95 % confidence level, with precision set at 0.05, and assumptions of expected frequency of COPD to 15%. In order to allow for reduced precision due to a cluster sampling – technique, and assurance of not making erroneous conclusions caused by variation or error in results, a design effect at 1.5 was applied. This inflation factor was obtained anticipating a mean cluster size of 25 individuals and an intra correlation coefficient of 0,02. On the basis of these calculations, the minimum sample size required was 300 participants. The study sample was selected randomly, through a multistage cluster technique. As the prevalence of COPD in our study was 12,7 % using the fixed ratio, and

7,9 % when applying the lower limit of normal, the study had sufficient precision, as the calculated frequency was 15 %. However, the available sample in the non-symptomatic group was small, particularly in the male strata. Thus the power to detect differences between the subgroups and sex strata might not have been sufficient. Also, correlations between the selected pulmonary parameters and explanatory variables might have been weakened and in some cases falsely negative due to a small sample size. In retrospect, it would have been preferable to calculate sample size requirements according to sufficiently large non-symptomatic observations within age and sex specific cells, over expected frequencies of COPD, which was the case in the present study. During analysis, adjustments due to complex sampling were made, and all study subjects were given a sampling weight calculated by adding the inclusion probabilities at each of the three selection stages. This reduced the precision of the confidence intervals computed. However, difficulties in obtaining a correct sampling frame prior to the selection, and thus ensuring selection probabilities proportionate to size at the different stages, may limit the extent to which the study results can be generalised. It would have been preferable to ensure a self weighted sample prior to the selection, which also would not violate the rule of selection proportionate to size.

Due to time and capacity restraints, the two last clusters were not visited. This could also have implications on the generalisability of the study outcomes and increase the overall sampling error.

Test procedures

Spirometry

One of the strengths of the present study is that the spirometry was carried out by use of a high- quality test device, and that strict adherence to ATS/ERS acceptability criteria were followed when assessing the manoeuvres. All study subjects performed a minimum of three and a maximum of eight tests, and the best FEV1, FEV6 and PEF were selected and used in further analysis. The spirometric testing was performed by trained assistants following a standardized procedure adherent to ATS/ERS guidelines (4, 5, 31). All spirograms were reviewed by an experienced chest physician, where ATS acceptability and reproducibility criteria were followed for selection of best pulmonary function curve; forced exhalation time should exceed a minimum of 6 seconds or show a plateau of the volume-time graph. Further each participant should display at least two acceptable reproducible maneuvers with both FEV1 and FEV6 within 200 ml and with absence of cough during the first second of the maneuvers (31).

Anthropometric data

As the testing was conducted out-door in the field, the test devices were carried from house to house, and robust equipment, suitable for this setting was used. This might have influenced on the accuracy of the data obtained. Though difficult to carry through in the present study, it would have been preferable to have a stationary test lab, where adherence to gold standards within test procedures and equipment is made easier as the testing is less prone to environmental influences. Also, the anthropometric test equipment would not have been selected at the expense of being robust and easy to transport.

However, training and supervision of field assistants and use of a standardized test procedure throughout the data collection have been ensured, to reduce the influence of measurement errors.

Questionnaires

The feasibility of the ATS-DLD questionnaire to obtain information regarding respiratory symptoms in this population must be considered, where both meaning and phrasing of words and illnesses are rooted in a different cultural context and might not capture the true picture of the study participants interviewed. The process of translating and back translating of the questionnaire was conducted with the purpose of reducing such measurements errors. However, developing culture specific questionnaires assessing respiratory symptoms and illness should be prioritised in future research.

Also, the possible presence of response bias must be considered. As the questionnaires were both elaborate and time consuming, the bias known as satisficing, and the way the respondents are administrating the questionnaire, might reduce the validity of the recorded data. However, as the questionnaires were administered face to face by research assistants, there is an immediate opportunity to clear up misconceptions on the part of the respondent and this might reduce the risk of invalid responses.

Discussion of findings

The following pulmonary parameters were measured in all study participants; peak expiratory flow (PEF), forced expiratory volume in one second (FEV1), forced expiratory volume in six seconds (FEV6) and the FEV1/FEV6 ratio. In the present study FEV6 was chosen over FVC, as it has been shown to be less demanding for both the participant as well as the technician, as

it requires less exhalation time. It has also been demonstrated to display less test variability when compared to FVC (14, 15, 16).

After evaluating the spirometric measurements, 39 study participants (10,7 %) did not accomplish tests results in compliance to ATS criteria, and were excluded from analysis regarding spirometric lung indices.

We have generated prediction equations for the following pulmonary parameters; FEV1, FEV6, FEV1/FEV6 and PEF from a non-symptomatic subsample of the study participants. Due to increased risk of false-positive diagnosis in the interpretation, it is recommended not to include too many indices of lung function in the testing, and the selected spirometric parameters are frequently used in assessment of respiratory diseases (1, 4, 5, 31). FEV1 is also shown to be an independent predictor of mortality from respiratory diseases (2, 4, 3, 5). In our study, we chose predicted values of FEV6 over FVC, both for its advantages in epidemiological fieldwork, and that it has shown to be an acceptable surrogate for FVC in diagnosing airflow obstruction (14, 15, 16). However, this makes a direct comparison to previously published reference values of FVC not applicable. We also derived prediction values of PEF from the study participants. Peak flow meters are relatively inexpensive compared to spirometric devices, and can be a more feasible investment for health services in low-income countries, yet produce useful information in clinical assessment of pulmonary illnesses, providing appropriate reference values are applied.

Frequencies and proportions of respiratory symptoms and conditions are calculated and listed in table 2 and 3 in paper 2, based on the participants' responses to the ATS-DLD respiratory questionnaire. Both presence of symptoms and illnesses were more frequently reported among the female participants, though considerable in both sexes. However, one must question the validity of answers regarding presence of respiratory illnesses as emphysema, which requires diagnostic tools as x-ray or whole body pletysmography to verify. Also, presence of asthma should be determined based on spirometric testing. Thus, reporting of prevalence of emphysema (1,9 % and 1,0 % in men and women respectively), and cases of asthma (recorded in 10,7 % of men and 10,2 % of women) should be interpreted with caution as it might not reflect the true picture in the population. There are no previous studies on respiratory symptoms in the general population from Sub-Saharan Africa, but in Mokoetle's respiratory survey of a workforce in Johannesburg, the presence of respiratory symptoms among the study participants were also considerable and of similar magnitude (20).

Prevalence of COPD when applying a fixed cut off ratio of FEV1/FEV6 < 0,73, was 12,6 %, and it was equally distributed among men (13,9 %) and women (11,5 %). Prevalence of COPD when classified as below the lower limit of normal was 7,9 %, also equally distributed with 6,9 % and 8,7 % of men and women respectively. The latter method pays attention to the relation between age and lung volumes, and the bias of over-diagnosing COPD in the elderly and *vice versa* among younger adults is reduced, which may explain the lower prevalence calculated (3, 34). A limitation of our study is that COPD is diagnosed on basis of simple spirometry, though responses to a post-bronchodilator spirometric test and assessment of reversibility of airway obstruction are considered a necessity in diagnosing COPD (1, 2). However, due to limited resources and the nature of the data collection, it was not feasible to include post bronchodilator spirometry as part of the testing. Besides increased costs, transportation and storage demands, this would also imply a more elaborate training of assistants as well as a more time-consuming test procedure for the study participants.

In any case, one can question whether the prevalence of COPD in our study consists of truly positive COPD cases, as the diagnostic method used fails to discriminate between subjects who are responsive and non responsive to a bronchodilator and thus possibly insufficient in differentiating COPD from asthma. According to a Norwegian study by Lehmann et al, a reversibility response of FEV1 increase ≥ 12 % and ≥ 200 ml after administration of a bronchodilator, was found in 2 % and 4 % of middle-aged (47-48 years) and elderly (71-73 years) participants respectively (35). Thus only a small fraction of the study population had a clinical relevant bronchodilatation after inhalation of salbutamol. This indicates that the proportion of possibly false positive COPD cases in our study might not be considerable. However, Lehmann did not investigate the reversibility response among those of younger age. In any case, the subjects classified as having COPD in our study, had a clear obstructive ventilatory pattern, and all of the subjects were in their habitual condition at the time of testing, which we consider relevant for the diagnosing, despite the lack of a post-bronchodilator test. GOLD acknowledges in its report that epidemiological field studies face challenges in quantifying prevalence and assessing diagnosis and severity of COPD, and emphasizes that in areas lacking access to state-of-the-art diagnostic tools, diagnosis should be made with the equipment available (1[6]). This helps justify the diagnostic approach used in the present study. Prevalence of stage I and II COPD was 6,1 % and 5,8 % respectively.

---

[6] GOLD report, page 89

Among the study participants with COPD, only a small fraction were classified in stage 3 or higher (0,3 %).

Conclusions

- The following pulmonary parameters have been measured in all study participants: peak expiratory flow (PEF), forced expiratory volume in one second (FEV1), forced expiratory volume in six seconds (FEV6) and the FEV1/FEV6 ratio.

- We have generated prediction equations for the following pulmonary parameters: FEV1, FEV6, FEV1/FEV6 and PEF from a non-symptomatic subsample of the study participants.

- Frequencies and proportions of respiratory symptoms among the study participants have been assessed based on responses to the ATS-DLD respiratory questionnaire.

- Prevalence of COPD when applying a fixed cut off ratio of FEV1/FEV6 < 0,73, was 12,6 %, and it was equally distributed among men (13,9 %) and women (11,5 %). Prevalence of COPD when classified as below the lower limit of normal was 7,9 %, also equally distributed with 6,9 % and 8,7 % of men and women respectively. Prevalence of stage I and II COPD was 6,1 % and 5,8 % respectively. Among the study participants with COPD, only a small fraction were classified in stage III or higher (0,3 %).

Recommendations

- The differences in predicted lung volumes derived from our study compared to values obtained by using previously published equations, emphasise the importance of using appropriate reference values in clinical assessment of respiratory illness, based on the same population as those being tested. However, in a longitudinal perspective, our predictions face limitations as they can only be considered reference values of today. Prediction equations should be updated at regular intervals, preferably every decade (4, 5).

- Increased awareness, and more attention to causal relations of COPD are important steps towards developing strategies for both prevention and treatment of COPD in accordance with the local epidemiological context, and should be prioritized in future research.

- The prevalence of current smoking among the participants in our study must be considered disturbingly high among men (30 %). This should be given more attention in future research, as it has important implications concerning the health status and the expatiation of non-communicable diseases in the population in Dar Es salaam, Tanzania.

References

1 GOLD. Global strategy for the diagnosis, management, and prevention of Chronic Obstructive Pulmonary Disease. Updated 2008

2 Mannino DM. Chronic Obstructive Pulmonary Disease: Epidemiology and Evaluation *Hospital Physician* 2001; 22-31

3 Hansen JE, Sun XG, Wasserman K. Spirometric criteria for Airway Obstruction. *Chest* 2007; 131; 349-355

4 American Thoracic Society. Lung Function Testing: Selection of Reference Values and Interpretative Strategies. *Am Rev Respir Dis* 1991; 144; 1202-1218

5 Brusasco V, Crapo R, Viegi G. ATS/ERS task force. Interpretative strategies for lung Function tests. *Eur Respir J*; 2005; 26; 948-968

6 World Health Organization. Available from; http://www.who.int (Accessed May 01, 2010).

7 Hnizdo E. Combined effect of silica dust and tobacco smoking on mortality from chronic obstructive lung disease in gold miners. *British Journal of Industrial Medicine* 1990; 47; 656-664

8 Mamuya SHD, Bråtveit M, Mashalla Y, Moen BE. High prevalence of respiratory symptoms among workers in the development section of a manually operated coal mine in a developing country: A cross-sectional study. *BMC Public Health,* 2007. Available from: http://www.biomedcentral.com/1471-2458/7/17

9 Shamssain MH. Pulmonary function and symptoms in workers exposed to wood dust. *Thorax* 1992; 47; 84-87

10 Buist SA et al. The Burden of Obstructive Lung Disease Initiative (BOLD): Rationale and Design. *Journal of Chronic Obstructive Pulmonary Disease* 2005; 2; 277-283

11 Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *Lancet* 2009; 374; 733-43

12 Diaz PT et al. Increased susceptibility to Pulmonary Emphysema among HIV-Seropositive Smokers. *Am Intern Med* 2000; 132; 369-372

13 Buist SA et al. International variation in the prevalence of COPD (The BOLD Study): a population-based prevalence study. *Lancet* 2007; 370; 741-750

14 Swanney MP, Jensen RL, Chrichton DA, Beckert LE, Cardno LA, Crapo RO. FEV6 is an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and restriction. *Am J Respir Crit Care Med* 2000; 162; 917-919

15   Jing J, Huang T, Cui W, Xu F, Shen H. Should FEV1/FEV6 Replace FEV1/FVC Ratio To Detect Airway Obstruction? A Metaanalysis. *Chest* 2009; 135; 991-998

16   Vandevoorde J, Verbanck S, Schuermans D, Kartounian J, Vincken W. Obstructive and Restricitve spirometric patterns: fixed cut-offs for FEV1/FEV6 and FEV6. *Eur Respir J* 2006; 27; 378-383

17   Myers JE. Differential ethnic standards for lung functions, or one standard for all? *S Afr Med J* 1984; 65; 768-772

18   Goldin JG, Louw SJ, Joubert G. Spirometry of healthy adult South African men: Part II. Interrelationships between socioenvironmental factors and 'race' as determinants of spirometry. *S Afr Med J* 1996; 86; 820-826

19   White NW, Hanley JH, Lalloo UG, Becklake MR. Review and Analysis of Variation between Spirometric Values Reported in 29 Studies of Healthy African Adults. *Am J Respir Crit Care Med* 1994; 150; 348-355

20   Louw SJ, Goldin JG, Joubert G. Spirometry of healthy adult South African men: Part I. Normative values. *S Afr Med J* 1996; 86; 814-819

21   Mokoetle KE, de Beer M, Becklake MR. A respiratory survey in a black Johannesburg workforce. *Thorax* 1994; 49; 340-346

22   Shamssain MH. Pulmonary function in normal non-smoking black Southern African adults. *Respiratory Medicine* 1994; 88; 287-291

23   Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the Normal Maximal Expiratory Flow-Volume Curve with Growth and Aging. *Am Rev Respir Dis* 1983; 127; 726-734

24   Hnizdo E, Churchyard G, Dowdeswel R. Lung function prediction equations derived from healthy South African gold miners. *Occup Environ Med* 2000; 57; 698-705

25   Njoku CH, Anah CO. Reference values for peak expiratory flow rate in adults of African descent. *TROPICAL DOCTOR* 2004; 34; 135-140

26   Cookson JB, Pilane CN, Tembo PM. Normal pulmonary function in Botswana. *C Afr J Med* 1992; 38; 238-239

27   Ele PU. Reference values for FEV1 and FVC in male adolescents and young adults of Ibo origin. *E Afr Med J* 1992; 69; 105-109

28   Mustafa KY. Spirometric lung function tests in normal men of African ethnic group. *Am Rev Respir Dis* 1977; 16; 209-13

29   Mfinanga SG et al. The role of livestock keeping in tuberculosis trends in Arusha, Tanzania. *Int J Tuberc Lung Dis* 2003; 7 (7); 695-704

30 Naing L, Winn T, Rusli BN. Practical Issues in Calculating the Sample Size for Prevalence Studies. *Archives of orofacial Sciences* 2006; 1; 9-14

31 M.R Miller et al. Standardisation of spirometry. Series "ATS/ERS Task Force: Standardisation of Lung Function Testing" *Eur Respir J* 2005; 26; 319-338

32 Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; 159; 179-187.

33 Intrauterine Growth and Disease in Later Life: The Barker Hypothesis. Available from http://www.medscape.com/viewarticle/453242_2

34 Hardie JA, Buist AS, Vollmer WM, Ellingsen I, BAkke PS, Morkve O. Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. *Eur Respir J* 2002; 20: 1117-1122.

35 Lehmann S, Bakke PS, Eide GE, Gulsvik A. Bronchodilator response to adrenergic β2- agonists: Relationships to symptoms in an adult community. *Respiratory Medicine* 2007; 101; 1183-1190

Paper I

Predictive equations for spirometric reference values in a healthy adult suburban population in Tanzania.

Toril Mørkve Knudsen[1], Odd Mørkve[1], Sayoki Mfinanga[2], Jon A Hardie[3, 4]

[1]*Centre for International Health,, University of Bergen, Norway,* [2]*Muhimbili Medical Research Centre, National Institute for Medical Research, Tanzania*
[3]*Institute for Internal Medicine, University of Bergen, Norway,* [4]*Department of Thoracic Diseases, Haukeland University Hospital, Norway,*

Introduction

Availability of appropriate reference values is a necessity for interpretation of pulmonary function tests and in assessment of lung function and respiratory diseases. According to American Thoracic Society (ATS) recommendations, such reference values should be based on healthy people with the same anthropometric characteristics and ethnic origin as the subjects being tested (1). Nevertheless, many low–income countries are faced with scarce documentation regarding national prediction equations and must rely on foreign reference values adjusted with estimated ethnic correction factors (2, 3, 4, 5). The use of ethnic adjustment factors assumes that the relationship between spirometry outcome and sex, age, and anthropometric predictors are the same for the local population as for the foreign reference population. This assumption is rarely well founded. Also, this method does not account for secular changes nor addresses underlying causes of ethnic differences in lung function, implicates artificially low ethnic standards, and may possibly result in a high false negative rate and missed diagnosis of respiratory diseases (2, 3, 4, 5).

The present study aims to derive reference values from a healthy adult suburban population in Dar Es Salaam, Tanzania. This paper presents prediction equations generated from a non-symptomatic subsample of the subjects being tested, and compares these equations to already published reference values. Both sitting and standing height have been measured and included as part of presumptive determinants of lung function, to allow for comparison of study results across countries and ethnicity (1, 2, 3, 5, 6, 7).

Methods

Study design and participants

The study is based on a descriptive cross-sectional design. The source population comprised all adults above the age of 15 in Kinondoni district, a suburban area of Dar Es Salaam, Tanzania, roughly estimated to hold a population of 350,000. Exact population registries do not exist. Participants were recruited by multistage cluster random sampling. In the first stage, two of the 27 wards of Kinondoni district were selected. At next stage, two areas in each ward were chosen. Each ward consisted of five areas. In the last stage, four ten-cell leaders were selected from each area, conducing to a total of 16 clusters. The term 'Ten-cell leader' refers to the smallest administrative unit in Tanzania, and one ten-cell leader has the responsibility for approximately 10-12 households (8). The number of ten-cell leaders varied in the selected areas; one area constituted a total of 40 ten-cell leaders, another consisted of 10 ten-cell

leaders and two areas had 8 ten-cell leaders. The primary sampling unit comprised all households under the selected ten-cell leaders.

Another primary objective for the study,dealt with in a parallel paper, was to determine the prevalence of chronic obstructive pulmonary disease (COPD) in Tanzania. As such, the sample size was calculated at a 95 % confidence level, and with precision set at 0.05, and assumptions of expected frequency of COPD were based on a study from South Africa, showing a prevalence of COPD around 12 % (9). Estimating prevalence of COPD to 15% and accounting for reduced precision due to cluster sampling through use of a design effect at 1.5, the minimum sample size required was calculated to 300 participants. Recruiting study subjects was accomplished by informing the randomly selected ten-cell leaders about the aim of the study, seeking their approval to collect the necessary data. Next, a mobile testing team carried out a knock-on-door approach to the households under their selected ten-cell leaders. Subjects within the age-span examined and who were willing to participate, were included in the study. Selection of subjects eligible for generating reference values followed ATS recommendations, where subjects with negative responses to core questions from ATS-DLD regarding respiratory symptoms and doctor diagnosed heart/chest illnesses were selected, providing their spirometric data met ATS criteria (6, 10). Questions regarding self reported occupational exposures were excluded from the stratification because they yielded unreliable information. However, duration of symptoms like cough, phlegm, wheezing and breathlessness were accounted for, as were past or present history of smoking and respiratory illnesses like chronic bronchitis, emphysema, asthma, tuberculosis, doctor diagnosed heart problem and high blood pressure confirmed by a doctor.

Data collection methods: definitions and measurement
Spirometry

Spirometry was performed using the *ndd EasyOne* spirometer (ndd Medizintechnik AG), which has proven to be suitable in field work as it operates on batteries and requires no calibration while achieving a high degree of accuracy and reliability (11). The following parameters were measured; peak expiratory flow (PEF), forced expiratory volume in one second (FEV1), and forced expiratory volume in six seconds (FEV6). In the present study FEV6 was chosen over FVC, as it has been shown to be advantageous due to less exhalation time, implying both shorter coaching time as well as reduced physical discomfort for the participants. It has also been demonstrated to display less test variability when compared to FVC (11, 12). All study subjects performed a minimum of three and a maximum of eight

tests, and the best FEV1, FEV6 and PEF were selected and used in further analysis. The spirometric testing was performed by trained assistants following a standardized procedure adherent to ATS/ERS guidelines (6, 10). The test subjects performed spirometry without nose clips, and sitting position was allowed if participants found the maneuver in standing position to be exhausting. All spirograms were reviewed by an experienced chest physician, where ATS acceptability and reproducibility criteria were followed for selection of best pulmonary function curve; forced exhalation time should exceed a minimum of 6 seconds or show a plateau of the volume-time graph. Further, each participant should display at least two acceptable reproducible maneuvers with both FEV1 and FEV6 within 200 ml and with absence of cough during the first second of the maneuvers (10).

Anthropometry

Anthropometric data, including weight, standing height and sitting height were obtained using robust equipment easy to transport in the field. The anthropometric measurements were collected by field personnel trained by the author, and the standardized procedure was monitored by the principal investigator throughout the data collection.

The participants' weights were obtained using a weight scale. Standing height was measured using a metal tape-measure and with heels, shoulders and occiput positioned against the tape. Sitting height was measured with a tape-measure affixed to a vertical wooden plank with the subjects sitting on a firm wooden chair with buttock, shoulder-blades and occiput touching the plank. Participants performed the anthropometric measuring barefoot.

Age was recorded as birth-date or as mid year in the year of birth. The data were transformed to age in years in further analysis.

Questionnaire data

Respiratory symptoms were recorded using the ATS-DLD validated questionnaire, which was translated into Swahili, then back-translated to English prior to data collection, to ensure that questions and the phrasing used were adherent to the original format.

Questions regarding socioeconomic status were derived from a culture specific questionnaire used in the Tanzania Demographic Health Survey from 1999 and 2002. The questionnaire was administered face-to-face by field personnel, trained to reduce information bias regarding interview technique, where phrasing of questions, and appropriate wording, was enhanced to obtain acceptable intra and inter- observer variability.

Ethical considerations

Approval was obtained from the Ethical Committee of Western Norway (REK-Vest) and the Medical Research Coordinating Committee of the National Institute for Medical Research (MRCC). After subjects eligible for participation were informed of the nature and purpose of the study, and had given their informed consent, they were enrolled providing inclusion criteria were met. If disease that needed medical attention was found in any study participant, the person was aided with a referral to proper medical care.

Statistical analysis

Data were analyzed using standardized and hierarchical multiple regression techniques. Both linear and curvilinear formulae were obtained, and the model best fitted for the present data were chosen. The following dependent variables were applied; FEV1, FEV6, PEF and FEV1/FEV6, and regressed against the following independent variables; age, height and weight. The explanatory variables were assessed in terms of highest contributing R statistic, variable significance at $p<0.05$, partial correlation coefficients and lowest SEE values. Outliers were defined as cases that had standardized residual values above 3,0 or below -3,0, and were excluded from further analyses. Preliminary analyses were conducted to ensure no violation of the assumption of normality, linearity and multicollinearity. Prediction equations were estimated separately for men and women. The lower limit of normal (LLN) was defined as the 5th percentile, estimated as $1,645 \times SEE$, for each of the dependent variables. The FEV1/FEV6 ratio did not perfectly match a Gaussian distribution, and logarithm and square root transformation were conducted to assess whether the proportion of skewness decreased. As the distribution still showed some deviation from the normal curve and the standard error of estimate was nearly identical to the non log transformed, the non log transformed FEV1/FEV6 was selected in further analyses.

Reference values generated from the present study were compared to previously published predictions by use of paired t-tests, and the magnitude of mean differences was obtained by Eta squared calculations. The Independent-Sample T-test was used to compare continuous variables, and Pearson Chi square and Wald statistic tests were used to compare group differences for categorical variables. Descriptive statistics are reported as means and standard deviations (SD) when normally distributed, skewed data are expressed in medians and percentiles. Calculations were done using the Statistical Package for the Social Sciences version 15 (SPSS Inc., Chicago, IL, USA).

Questionnaire data were double entered, cleaned and coded using Epi data version 3.1 (Centre for Disease Control and Prevention, Atlanta, GA, USA), and were exported to SPSS for the statistical analyses. The spirometric data were automatically stored in the spirometres used, and were electronically exported to SPSS.

Results

Of the 16 selected clusters, 14 were visited, and a total of 365 subjects were enrolled in the study, 159 (43,6 %) men and 206 (56,4 %) women. As 135 subjects in the visited households were absent at the time of testing, the response rate attained was 74%. However, the denominator in this fraction included subjects with unknown eligibility status. Further, 41 participants were excluded from the study due to conditions contraindicating spirometry testing (10): 19 were due to known heart disorder, 15 had undergone chest/abdominal surgery, 5 were in last trimester of pregnancy, and 2 had undergone eye surgery. Eleven eligible subjects refused to participate, giving a cooperation rate of 97 %. After evaluating the spirometric measurements, 39 study participants (10,7 %) did not accomplish tests results in compliance to ATS criteria, and were excluded from analysis regarding spirometric lung indices.

On the basis of the selection criteria, a total of 150 subjects, 52 men (32,7 %) and 98 women (47,6 %) were selected to the reference value group. Median age for the selected men and women were 34 years. Tables 1 and 2 summarize socio-demographic characteristics in the reference sample; table 1 presents the age distribution and descriptive statistics regarding anthropometry and spirometry parameters, whereas table 2 gives a summary of occupations. Prior to the regression modeling, bivariate analyses were conducted to evaluate correlations for FEV1, FEV6, FEV1/FEV6 and PEF according to sample mean age, height and weight. This is summarized in table 3. Age, height and sitting height correlated significantly with the dependent variables FEV1 and FEV6 for both sexes. For women this was also the case with PEF. FEV1/FEV6 showed no significant correlation with height, and neither of the dependent variables correlated with weight in the female strata, whereas for men, weight showed a correlation with the FEV1/FEV6 ratio.

Further, group differences for categorical variables regarding exposure to occupational pollutants and other socioeconomic conditions were investigated in relation to the above mentioned pulmonary parameters. Based on these preliminary analyses, no significant group differences could be traced, and they were not taken into further consideration in the regression analysis.

The regression analyses were performed separately for males and females. The exponential variable height square was entered into the equation together with the age, height and weight, to assess whether a linear or curvilinear model produced best fit for the present data. In the male strata, an exponential model was selected as regression equation, as height square explained an additional 3,2 % and 3,6 % of the variance in FEV1 and FEV6 respectively, and contributed 1,8 % to the total R square for PEF. For FEV1 and FEV6, R statistics increased from 0,462 to 0,494 (p = 0,09) and from 0,445 to 0,481 (p = 0,08) respectively, when the effect of the other independent variables had been statistically controlled for. For PEF, R square change was more modest, increasing from 0,223 to 0,241 (p = 0,3), after controlling for age and height , but was retained in the final model. When assessing the model fit for men using sitting height as explanatory variable, a curvilinear formulae was also considered most suitable for predicting FEV1 and FEV6 where sitting height square explained an additional 3,1 % and 3,5 % of the variance in these variables and R square values changed from 0,419 – 0,450 (p = 0,1) and from 0,473 to 0,508 (p = 0,07) respectively. For the dependent variable PEF, a linear model was applied, as the exponential variable did not contribute much to the total R square (0,4 %) or for the R square values (0,271 – 0,274, p = 0,6). For women, the exponential variable height square, explained an additional 2,0 % of the total variance in PEF, with R statistics increasing from 0,219 – 0,239 (p = 0,1). Similarly, sitting height square, contributed 1,9 % to the total R square change for PEF, with R square increasing from 0,237 – 0,256 (p = 0,1), and a curvilinear model were selected as equation. For the dependent variables FEV1 and FEV6, a linear model was chosen in further analysis, as neither of the exponential variables height square or sitting height square contributed to the total R square. Weight turned out non-significant in the male strata, and showed only a minor contribution to the variability of the regression models, and was excluded from the formulas. For women weight had a part correlation coefficient of 0,178, when regressed against PEF, explaining an additional 3,17 % of the total R square (p = 0,05), and was retained in the final model when standing height was part of the independent variables. For the dependent variables FEV1, FEV6 and FEV1/FEV6, weight was non-significant, and was removed from the equations in further analyses. Height, and height square in the exponential equations for men and women, entered all the regression models with the exception of the model predicting FEV1/FEV6, as it did not provide a significant contribution to the variance in the present sample and only showed a marginally improvement of the R statistics and the corresponding SEE values. Tables 4 and 5 summarize the prediction equations generated. All the spirometric parameters were negatively related to age, and all increased with height, with exception of FEV1/FEV6.

The lower limit of normal (LLN) is presented as – 1,645 x SEE which is the age and height specific, estimated 5th percentile for the reference sample.

The prediction equations derived from our sample are compared to those of Mokoetle, reference values generated from a healthy South African University workforce (7), and with Hankinson's equation (NHANES iii),of African-American above 20 years of age (13). Reference values for the male strata are also compared to Louw's equations of Black and White South African men and to the predictions of Mustafa of Northern Sudanese but also tested to fit Tanzanian men (5, 14). As Mustafa's predictions were generated in 1977, they are considered old. However, they are included in the comparison of published equations to assess for secular changes in predicted lung volumes. Figure 1-4 present scatter plots of predicted FEV1 for men and women applying the equations of Mokoetle, Hankinson and Louw. Table 6 summarizes mean difference of FEV1 values between our prediction equations to those in comparison.

Discussion

We have generated prediction equations for the following pulmonary parameters; FEV1, FEV6, FEV1/FEV6 and PEF from a non-symptomatic subsample of the study participants. In accordance with GOLD and ATS guidelines the selected lung function indices are frequently used in pulmonary testing and in assessment of respiratory diseases (1, 10). When applying Mokoetle's and Louw's predictive equations of Black South African men to our male strata, this generated larger values of FEV1 compared to those derived from our study, thus overestimating normal values of FEV1 in the present sample population. This was also evident when using Hankinson's equation of African-American adults above 20 years of age, though the mean difference to our predictive equations was smaller. Louw's reference values of White South African men, showed an even larger mean difference in FEV1 at 1,2 litre when compared to the present study values. Applying equations using sitting height as a proxy estimate of height reduced this difference to 1,0 litre. This is in accordance with previous publications, were ethnic differences in lung volumes is thought to be partly due to difference in body build and that Blacks have smaller trunk-to-leg ratio than Whites (10). An adjustment using sitting height as an anthropometric measure in lung function testing reduces these ethnic differences (1, 2, 3, 5, 10). The predicted FEV1 values for the women in our study were also significantly lower compared to those of Mokoetle and Hankinson, with a mean difference of 0,30 and 0,23 litres respectively. Louw did not generate reference values for women in his study. A factor that might have contributed to the differences in predicted

FEV1 values, is the study population used for generation of reference values. Both Mokoetle and Louw used occupation based samples in their studies, which have proven to generate higher lung volumes than studies deriving reference values from a community source, which was the case in the present study (4). This bias, known as the healthy-worker effect, might explain some of the variation in mean FEV1 values, but also possible differences in socio-environmental conditions for the study populations generating the reference values must be considered, since it have been thought to influence lung function (2, 3, 5, 10, 11, 15). In addition, the altitude in Johannesburg is 2000 m above sea level, which can account for up to 400 ml according to the study of White et al, published in 1994 (4). Also, methodological differences between the studies, like the equipment used for spirometric testing and variations in cut-off criteria for acceptability assessment, might play a role in deriving different predicted values. Such possible biases in the process of generating reference values emphasise the importance of using appropriate prediction equations to prevent a possible risk of under-diagnosing or over-diagnosing respiratory illnesses and disease severity. When comparing our equations to those of Mustafa in 1977, this generated almost similar predicted values for FEV1 (14). The Sudanese site was situated at 400 m altitude, and the sample was derived from an urban community source. This raise questions to what extent secular changes, and the bias known as the cohort effect influence on the spirometric measurements, though possible differences in socio-environmental conditions for the study populations generating the reference values has been thought to influence lung function. On the other hand, the present sample size might not been sufficiently large to detect differences in lung volumes according to changes in determinants as nutrition, health status and environmental factors. The study also failed to point out any interconnectedness between lung volumes and socioeconomic factors.

In our study, we have chosen predicted values of FEV6 over FVC, both for its advantages in epidemiological fieldwork, and that it has been shown to be an acceptable surrogate for FVC in diagnosing airflow obstruction (16, 17). However, this makes a direct comparison to previously published reference values of FVC not applicable.

We also derived prediction values of PEF from the study participants. Peak flow meters are relatively inexpensive compared to spirometric devices, and can be a more feasible investment for health services in low-income countries, yet produce useful information in clinical assessment of pulmonary illnesses, providing appropriate reference values are applied.

One of the strengths of the present study is that the spirometric testing was carried out by use of high- quality apparatus in a realistic "in-the-field" setting, and that strict adherence to ATS/ERS acceptability criteria was followed when assessing the manoeuvres. As the testing was conducted out-door in the field, the test devices were carried from house to house, and robust anthropometric measures, suitable for this setting were used. This might have influenced on the accuracy of the anthropometric data obtained. However, training and supervision of field assistants and use of a standardized test procedure throughout the data collection have been ensured, to reduce the influence of measurement errors.

The feasibility of the ATS-DLD questionnaire to obtain information regarding respiratory symptoms in this population must be considered, where both meaning and phrasing of words and illnesses are rooted in a different cultural context and might not capture the true picture of the study participants interviewed. Developing culture specific questionnaires assessing respiratory symptoms and illness should be prioritised in future research.

Some limitations of the present study need to be considered. The study sample was selected randomly, through a multistage cluster technique. However, difficulties in obtaining a correct sampling frame prior to the selection, and thus ensuring selection probabilities proportionate to size at the different stages, may limit the extent to which the study results can be generalised. Due to time and capacity restraints, the two last clusters were not visited. This also have implications on the generalisability of the study outcomes and increase the overall sampling error. The available sample in the non-symptomatic group was small, particularly in the male strata. Thus the power to detect differences between the subgroups and sex strata might not have been sufficient. Also, correlations between the selected pulmonary parameters and explanatory variables might have been weakened and in some cases falsely negative due to a small sample size. In retrospect, it would have been preferable to calculate sample size requirements according to sufficiently large non-symptomatic observations within age and sex specific cells, over expected frequencies of COPD, which was the case in the present study. However, the study results still present valuable information. The differences in predicted lung volumes derived from our study compared to values obtained by using previously published equations, emphasise the importance of using appropriate reference values in clinical assessment of respiratory illness, based on the same population as those being tested. Measurement of lung function in the proposed catchment area and age group is also important, as there is no recent published data on respiratory function in the general population in Dar Es Salaam, Tanzania.

References

1   Brusasco V, Crapo R, Viegi G. ATS/ERS task force. Interpretative strategies for lung Function tests. *Eur Respir J* 2005; 26; 948-968

2   Myers JE. Differential ethnic standards for lung functions, or one standard for all? *S Afr Med J* 1984; 65; 768-772

3   Goldin JG, Louw SJ, Joubert G. Spirometry of healthy adult South African men: Part II. Interrelationship between socioenvironmental factors and 'race' as determinants of spirometry. *S Afr Med J* 1996; 86; 820-826

4   White NW, Hanley JH, Lalloo UG, Becklake MR. Review and Analysis of Variation between Spirometric Values Reported in 29 studies of Healthy African Adults. *Am J Respir Crit care Med* 1994; 150; 348-355

5   Louw SJ, Goldin JG, Joubert G. Spirometry of healthy adult South African men: Part I. Normative values. *S Afr Med J* 1996; 86; 814-819

6   American Thoracic Society. Lung Function Testing: Selection of Reference Values and Interpretative Strategies. *Am Rev Respir Dis* 1991; 144; 1202-1218

7   Mokoetle KE, de Beer M, Becklake MR. A respiratory survey in a black Johannesburg workforce. *Thorax* 1994; 49; 340-346

8   Mfinanga SG et al. The role of livestock keeping in tuberculosis trends in Arusha, Tanzania. *Int J Tuberc Lung Dis* 2003; 7 (7); 695-704

9   Buist SA et al.International variation in the prevalence of COPD (The Bold Study): a population-based prevalence study. *Lancet* 2007; 370; 741-750

10   M.R Miller et al. Standardisation of spirometry. Series "ATS/ERS Task Force: Standardisation of Lung Function Testing" *Eur Respir J* 2005; 26; 319-338

11   Buist SA et al. The burden of Obstructive Lung Disesase Initiative (BOLD): Rationale and Design. *Journal of Chronic Obstructive Pulmonary Disease* 2005; 2; 277-283

12   Swanney MP, Jensen RL, Chrichton DA, Beckert LE, Cardno LA, Crapo RO. FEV6 is an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and restriction. *Am J Respir Crit Care Med* 2000; 162; 917-919

13   Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; 159; 179-187.

14   Mustafa KY. Spirometric lung function tests in normal men of African ethnic group. *Am Rev Respir Dis* 1977; 16; 209-13

15 GOLD. Global strategy for the diagnosis, management and prevention of Chronic Obstructive Pulmonary Disease. Updated 2008

16 Jing J, Huang T, Cui W, Xu F, Shen H. Should FEV1/FEV6 Replace FEV1/FVC Ratio To Detect Airway Obstruction? A Metaanalysis. *Chest* 2009; 135; 991-998

17 Vandevoorde J, Verbanck S, Schuermans D, Kartounian J, Vincken W. Obstructive and Restricitve spirometric patterns: fixed cut-offs for FEV1/FEV6 and FEV6. *Eur Respir J* 2006; 27; 378-383

Table 1

Table 1; age distribution and descriptive statistics reference sample

	Men (n = 52)	Women (n = 98)
Age categories, no (%)		
< 20	7 (13,5)	12 (12,2)
20-29	16 (30,8)	34 (34,7)
30-39	11 (21,2)	25 (25,5)
40-49	10 (19,2)	13 (13,3)
50-59	4 (7,7)	10 (10,2)
60-69	3 (5,8)	3 (3,1)
70+	1 (1,9)	1 (1,0)
Height, mean ± SD, cm	166,38 ± 7,86	157,16 ± 5,96
Sitting height mean ± SD, cm	81,45 ± 3,61	78,08 ± 3,42
Weight, mean ± SD, kg	65,58 ± 11,50	61,48 ± 14,99
FEV1, mean ± SD, L	3,02± 0,65	2,24 ± 0,51
FEV6, mean ± SD, L	3,59 ± 0,68	2,71 ± 0,56
FEV1/FEV6, mean ± SD	0,840 ± 0,079	0,825 ± 0,070
PEF, mean ± SD, L	7,98 ± 2,32	5,48 ± 1,38

N = number of responders, % = proportion of responders, SD = standard deviation

Table 2

Table 2; distribution of job categories reference sample

Occupation*, N (%)	Men (n = 41)	Women (n = 66)
White-collar worker/professional	7 (13,5)	5 (5,1)
Agriculture	12 (23,1)	23 (23,5)
Blue-collar worker	13 (25,0)	4 (4,1)
Sales/shopkeeping	5 (9,6)	15 (15,3)
Student	4 (7,7)	7 (7,1)
Housewife/housegirl		12 (12,2)

N = number of responders, % = proportion of responders

* Total missing rate of 28,7 % when listing occupation

Table 3

Table, bivariate correlation				
Men (n = 52)	FEV1, L	FEV6, L	FEV1/FEV6	PEF, L
Age, years	-0,461**	-0,333*	-0,415*	-0,205
Height, cm	0,536**	0,603**	0,026	0,416**
Sitting height, cm	0,475**	0,621**	-0,109	0,482**
Weight, kg	-0,007	0,136	-0,294*	0,202
Women (n = 98)				
Age, years	-0,738**	-0,628**	-0,557**	-0,397**
Height, cm	0,405**	0,467**	-0,044	0,258*
Sitting height, cm	0,383**	0,426**	-0,008	0,314**
Weight, kg	0,142	0,153	0,009	0,195

** correlation is significant at the 0,01 level (2-tailed)
* correlation is significant at the 0,05 level (2-tailed)

Table 4

Table 4; prediction equations reference sample			
Model	R^2	SEE	5th percentile
Men			
$FEV1 = 33{,}173 - 0{,}0203 \times A - 0{,}394 \times H + 0{,}00130 \times H^2$	0,491	0,477	0,785
$FEV6 = 37{,}052 - 0{,}0149 \times A - 0{,}444 \times H + 0{,}00148 \times H^2$	0,479	0,503	0,827
$PEF = 94{,}661 - 0{,}0314 \times A - 1{,}143 \times H + 0{,}00377 \times H^2$	0,222	2,108	3,468
$FEV1/FEV6 = 0{,}920 - 0{,}00231 \times A$	0,172	0,073	0,120
Women			
$FEV1 = -1{,}133 - 0{,}0261 \times A + 0{,}0270 \times H$	0,658	0,317	0,521
$FEV6 = -2{,}031 - 0{,}0288 \times A + 0{,}0350 \times H$	0,553	0,385	0,633
$PEF = 104{,}066 - 0{,}0386 \times A - 1{,}286 \times H + 0{,}00419 \times H^2 + 0{,}0179 \times W$	0,239	1,229	2,022
$FEV1/FEV6 = 0{,}920 - 0{,}00282 \times A$	0,310	0,058	0,095

A = age in years, H = height (cm), H² = height square, W = weight

Table 5

Table 5; prediction equations using sitting height reference sample

Model	R^2	SEE	5th percentile
Men			
FEV1 = 29,530 – 0,0222 × A – 0,713 × SH + 0,00486 × SH²	0,446	0,498	0,819
FEV6 = 30,233 – 0,0168 × A – 0,753 × SH + 0,00531 × SH²	0,508	0,489	0,804
PEF = – 15,682 – 0,0289 × A + 0,303 × SH	0,264	2,030	3,339
Women			
FEV1 = – 0,0196 – 0,0253 × A + 0,0398 × SH	0,614	0,319	0,525
FEV6 = – 0,762 – 0,0232 × A + 0,0544 × SH	0,502	0,401	0,660
PEF = 51,653 – 0,0359 × A – 1,267 × SH + 0,00883 × SH²	0,236	1,225	2,015

A, age in years; SH, sitting height (cm); SH², sitting height square

Table 6

Table 6; Mean difference in FEV1 values; Our predictive compared to Mokoetle, Louw and Hankinson.

Mean difference in FEV1 to Our predictions (95 % CI)

	Men (n = 52)	Women (n =98)	
Mokoetle predicitive Black S Afr	0,45 (0,38 - 0,53)	0,30	(0,28 - 0,31)
Louw predictive Black S Afr	0,72 (0,66 - 0,78)		
Louw predictive White S Afr	1,2 (1,17 – 1,33)		
Hankinson predictive Afr Am	0,22 (0,19 - 0,25)	0,23	(0,21 - 0,25)
Mustafa predictive Sudanese	0,063 (0,046 – 0,080)		

Figure 1: Scatter plot of predicitive FEV1 values of male strata applying equations of Louw, Hankinson, Mokoetle and those generated in the present study
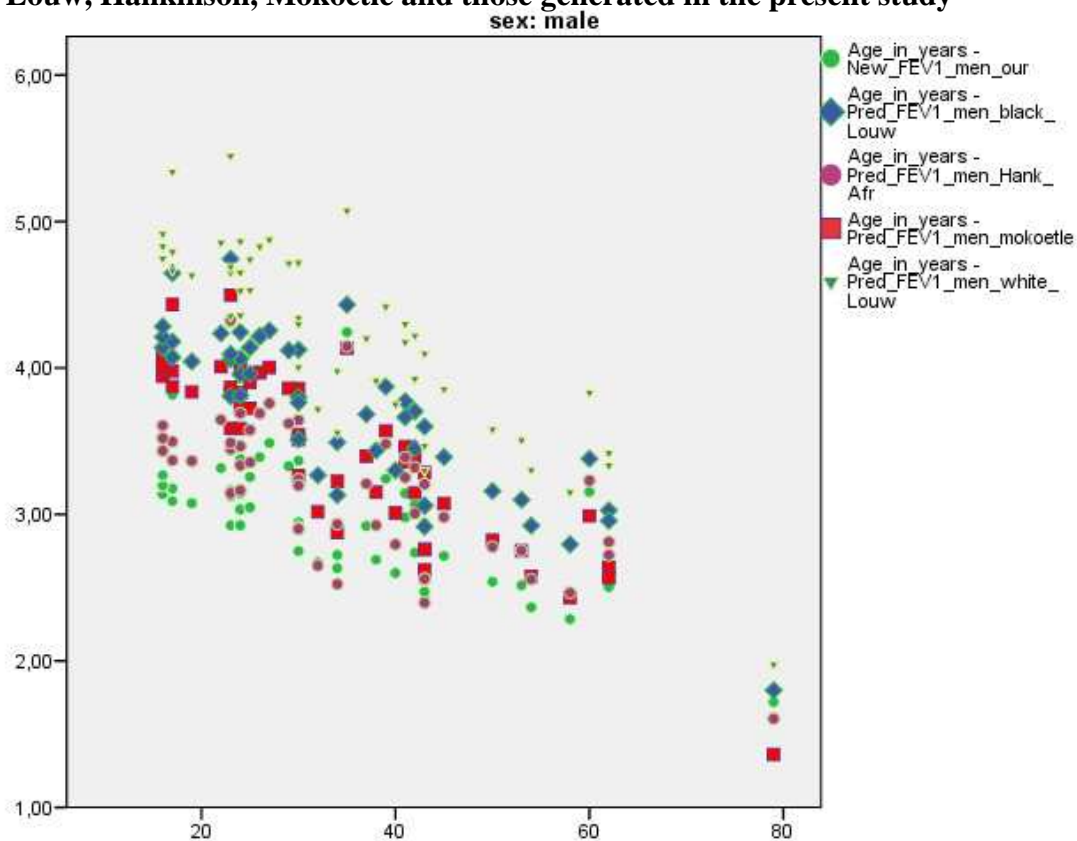
Figure 2; scatter plot of predictive FEV1 values of male strata applying equations of Louw, Mokoetle and those generated in the present study. Use of sitting height in equations
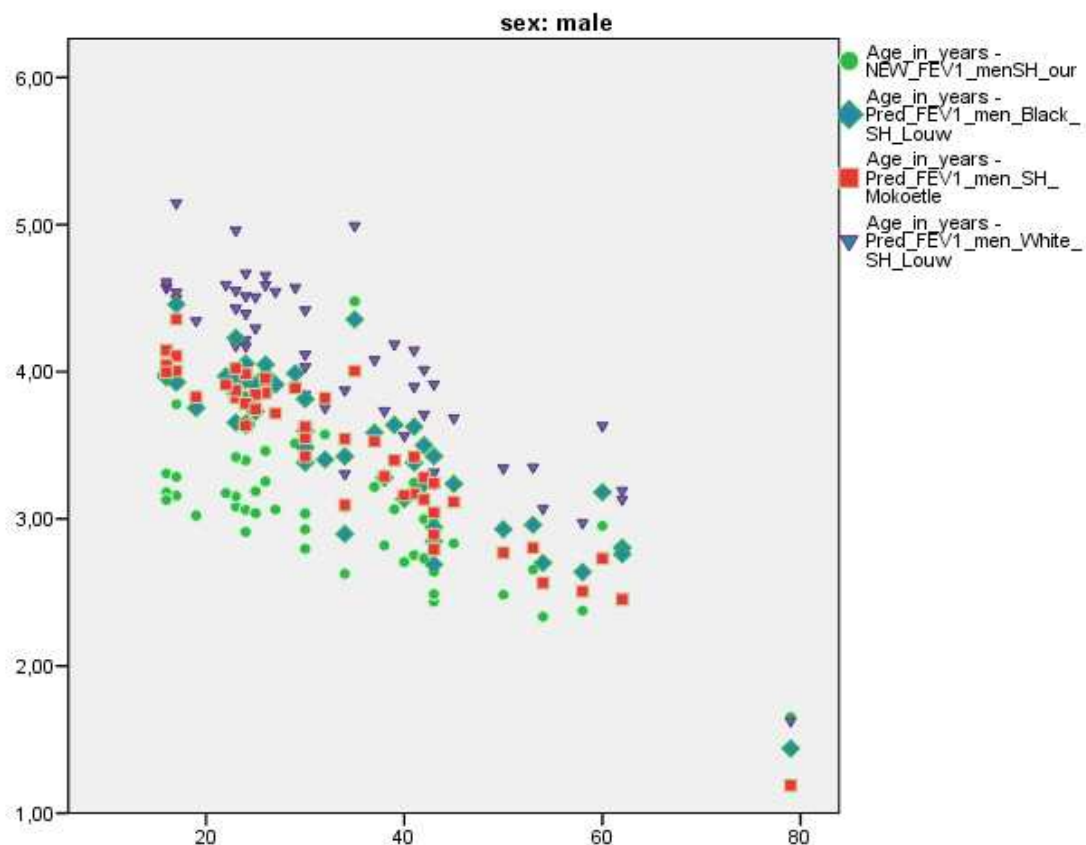
Figure 3; scatter plot of predictive FEV1 values in female strata applying equations of Mokoetle, Hankinson and those generated in the present study
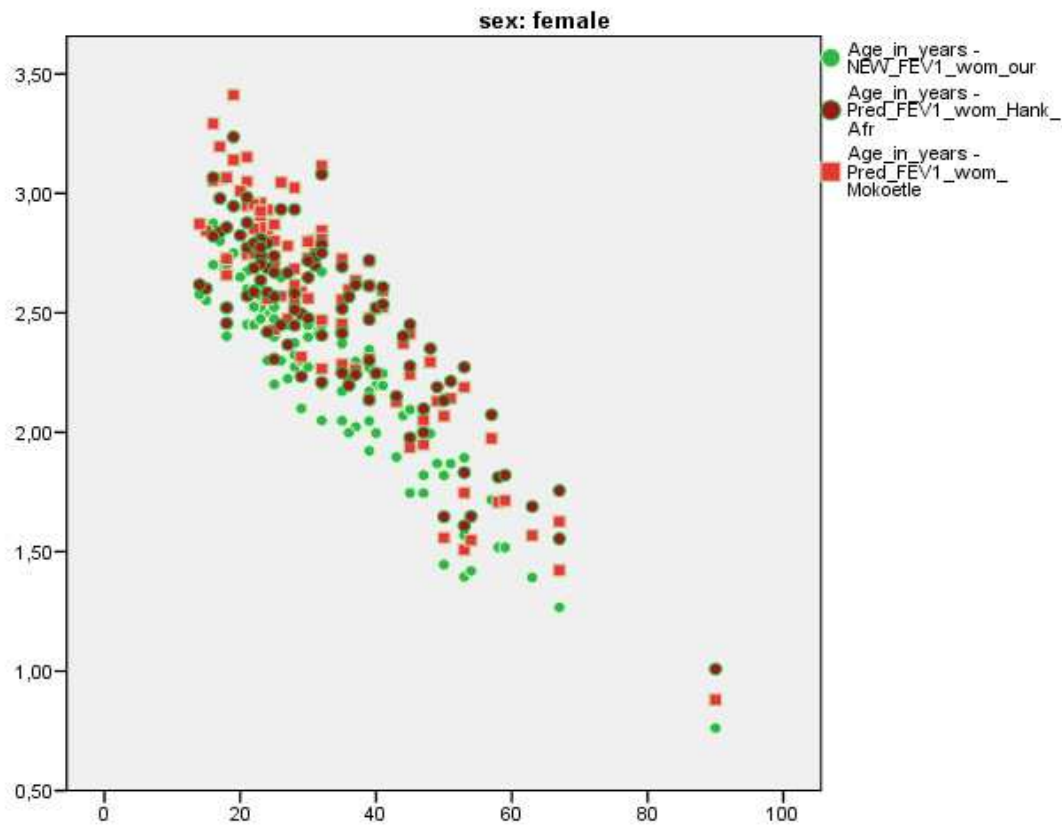
Figure 4; scatter plot of predictive FEV1 values applying equations of Mokoetle and those in present study. Use of sitting height in equations
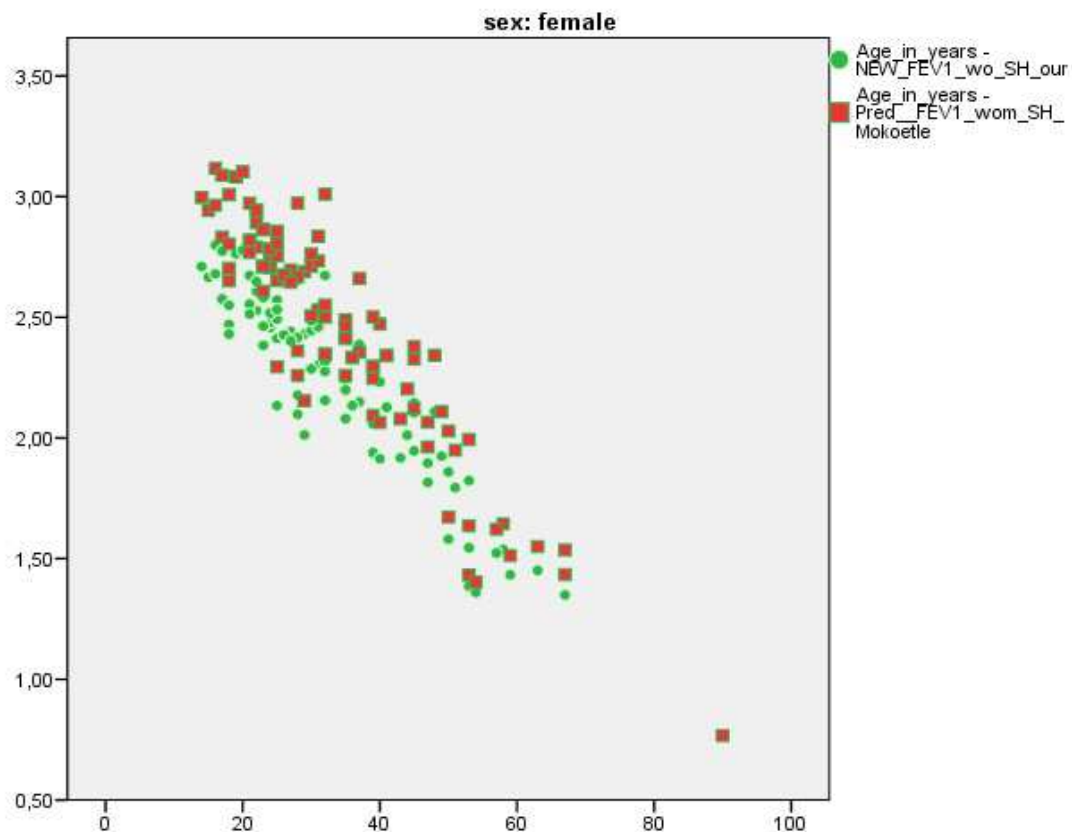
Paper II

Prevalence of respiratory symptoms and Chronic Obstructive Pulmonary Disease (COPD) in a suburban population in Dar Es Salaam, Tanzania.

Toril Mørkve Knudsen[1], Odd Mørkve[1], Sayoki Mfinanga[2] Jon A Hardie[3,4]

[1]*Centre for International Health,, University of Bergen, Norway,* [2]*Muhimbili Medical Research Centre, National Institute for Medical Research, Tanzania,* [3]*Institute for Internal Medicine, University of Bergen, Norway,* [4]*Department of Thoracic Diseases, Haukeland University Hospital, Norway,*

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is by the WHO Global Burden of Disease Study projected to increase in magnitude and rank as third leading cause of death worldwide in 2020, an expansion brought forth by both an ongoing health transition characterizing many of today's low and middle-income populations, as well as an anticipated increase in cigarette smoking in several African and Asian countries (1, 2). Studies concerning respiratory symptoms and prevalence of obstructive lung disease in low-income setting are yet rare. The expected frequency and burden of respiratory impairment is however anticipated to be considerable in these countries, both for the presence of potential hazards as occupational and domestic exposures, as for infections and morbidities, like tuberculosis and HIV, all influences known to be detrimental to respiratory health (2, 3, 4, 5). This reinforces the concept of COPD as both underreported and under-diagnosed in low-income countries, none the less likely causing high morbidity and mortality in poorly resourced populations, with less access to health care and preventive actions.

This paper aims to estimate prevalence of respiratory symptoms from a random sample of the population. Further, prevalence of COPD will be determined based on spirometric values according to adjusted GOLD[7] and ATS[8]/ERS[9] diagnostic criteria (2, 6, 7, 8, 9).

Methods

Study design and participants

The study is based on a descriptive cross-sectional design. The source population comprised all adults above the age of 15 in Kinondoni district, a suburban area of Dar Es Salaam, Tanzania, roughly estimated to hold a population of 350,000. Exact population registries do not exist. Participants were recruited by multistage cluster random sampling. In the first stage, two of the 27 wards of Kinondoni district were selected. At next stage, two areas in each ward were chosen. Each ward consisted of five areas. In the last stage, four ten-cell leaders were selected from each area, adding up to a total of 16 clusters. The term 'Ten-cell leader' refers to the smallest administrative unit in Tanzania, and one ten-cell leader has the responsibility for approximately 10-12 households (10). The number of ten-cell leaders varied in the selected areas; one area constituted a total of 40 ten-cell leaders, another consisted of 10 ten-

---

[7] Global initiative for Chronic Obstructive Lung Disease
[8] American Thoracic Society
[9] European Respiratory Society

cell leaders and two areas had 8 ten-cell leaders. The primary sampling unit comprised all households under the selected ten-cell leaders. Sample size was calculated at a 95 % confidence level, and with precision set at 0.05. Assumptions of expected frequency of COPD were based on a study in South Africa, showing a prevalence of COPD around 12 % (11). Estimating prevalence of COPD to 15% and accounting for reduced precision due to cluster sampling through use of a design effect at 1.5, the minimum sample size required was calculated to 300 participants. Recruiting study subjects was accomplished by informing the randomly selected ten-cell leaders about the aim of the study, and seeking their approval to collect the necessary data. Next, a mobile testing team carried out a knock-on-door approach to the households under their selected ten-cell leaders. Subjects within the age-span examined and who were willing to participate, were included in the study.

Data collection methods: definitions and measurement
Spirometry
Spirometry was performed using the *ndd EasyOne* spirometer (ndd Medizintechnik AG), which has proven to be suitable in field work as it operates on batteries and requires no calibration while achieving a high degree of accuracy and reliability (11). The following parameters were measured; peak expiratory flow (PEF), forced expiratory volume in one second (FEV1), and forced expiratory volume in six seconds (FEV6). In the present study FEV6 was chosen over FVC, as it has been shown to be advantageous due to less exhalation time, implying both shorter coaching time as well as reduced physical discomfort for the participants. It has also been demonstrated to display less test variability when compared to FVC (11, 12, 13, 14). All study subjects performed a minimum of three and a maximum of eight tests, and the best FEV1, FEV6 and PEF were selected and used in further analysis. The spirometric testing was performed by trained assistants following a standardized procedure adherent to ATS/ERS guidelines (6, 8). The test subjects performed spirometry without nose clips, and sitting position was allowed if participants found the maneuver in standing position to be exhausting. All spirograms were reviewed by an experienced chest physician, where ATS/ERS acceptability and reproducibility criteria were followed for selection of best pulmonary function curve; forced exhalation time should exceed a minimum of 6 seconds or show a plateau of the volume-time graph. Further each participant should display at least two acceptable reproducible maneuvers with both FEV1 and FEV6 within 200 ml and with absence of cough during the first second of the maneuvers (8).

Age was recorded as birth-date or as mid year in the year of birth. The data were transformed to age in years in further analysis.

Questionnaire data

Respiratory symptoms were recorded using the ATS-DLD[10] validated questionnaire, which was translated into Swahili, then back-translated to English prior to data collection, to ensure that questions and the phrasing used were adherent to the original format.

Questions regarding socioeconomic status were derived from a culture specific questionnaire used in the Tanzania Demographic Health Survey from 1999 and 2002. The questionnaire was administered face-to-face by field personnel, trained to reduce information bias regarding interview technique, where phrasing of questions, and appropriate wording, was enhanced to obtain acceptable intra and inter- observer variability.

Ethical considerations

Approval was obtained from the Ethical Committee of Western Norway (REK-Vest) and the Medical Research Coordinating Committee of the National Institute for Medical Research (MRCC). After subjects eligible for participation were informed of the nature and purpose of the study, and had given their informed consent, they were enrolled providing inclusion criteria were met. If disease that needed medical attention was found in any study participant, the person was aided with a referral to proper medical care.

Statistical analysis

Data regarding demography, respiratory symptoms and socio-environmental conditions included all responders. Estimation of COPD prevalence included study participants with acceptable spirometry only. Definition and classification of airway obstruction was based on, both adjusted GOLD criteria and ATS/ERS criteria (2, 3, 9, 12, 14). According to GOLD, COPD is determined by use of a fixed percentage of spirometric predicted values, with airflow limitation defined as an FEV1/FVC[11] ratio less than 70 % (2, 3, 9). In our study we used an adjusted GOLD definition, as FEV6 was selected over FVC as spirometric parameter, and a FEV1/FEV6 ratio less than 73% was applied as cut off point (14). When applying ATS/ERS guidelines, ventilatory obstructivity is defined as a reduced FEV1/FEV6 ratio below the statistically derived fifth percentile of the predicted value, and we also used this

---

[10] American Thoracic Society- Division of Lung Disease
[11] Forced Vital Capacity

lower limit of normal as a cut off to determine proportion of responders with COPD (6, 7, 9). Disease severity was based on GOLD disease stages (2, 6, 7). Diagnoses were made on the basis of simple spirometry without any attempt at bronchodilation. Descriptive statistics were applied, using frequencies and cross tabulation. The Independent-Sample T-test was used to compare continuous variables, and Pearson Chi square and Wald statistic tests were used to compare group differences for categorical variables. Logistic regression analyses were carried out to assess and adjust for potential confounders for determinants of COPD. The models were adjusted for the following variables; sex, age, smoking status, pack years, presence of occupational and domestic exposures and level of socioeconomic position. The smoking exposure variable, pack years, was computed by dividing the number of cigarettes smoked per day by 20, multiplied by number of years smoked. Pack years were computed for both current and ex-smokers. Smoking status contained the following variables; never smoked, ex-smoker and current smoker. Occupational exposure was computed by the following variables; worked ≥ 1 year in dusty work or ever exposed to gas/chemical in work. Domestic exposure was constructed by positive responses to question regarding use of biomass and charcoal at home. Measures regarding socioeconomic information were collected at individual level, and a socioeconomic position variable (SEP index) was constructed according to both economic and social dimensions represented by questions regarding level of education, present employment, recent work for pay and available money resources derived from the culture specific questionnaire used.

Logistic regression analyses were generated to estimate odds ratios (OR) with 95 % confidence intervals (CI) of COPD. Preliminary analyses were conducted to ensure no violation of the assumption of normality and multicollinearity. Both univariate and multivariate analyses were conducted. The models were assessed by goodness of fit tests. Prevalence estimates were adjusted for multistage cluster sampling during analyses. Each participant was given a sampling weight constructed by use of SPSS complex samples analysis, where inclusion probabilities at each of the three stages of the sampling procedure were added. Variables and differences were assessed in terms of significance at p< 0,05 level and their 95 % confidence intervals. Descriptive statistics are reported as means and standard deviations (SD) when normally distributed, skewed data are expressed in medians and percentiles. Calculations were done using the Statistical Package for the Social Sciences version 15 (SPSS Inc., Chicago, IL, USA).

Questionnaire data were double entered, cleaned and coded using Epi data version 3.1 (Centre for Disease Control and Prevention, Atlanta, GA, USA), and were exported to SPSS for the

statistical analyses. The spirometric data were automatically stored in the spirometres used, and were electronically exported to SPSS.

Results

Of the 16 selected clusters, 14 were visited, and a total of 365 subjects were enrolled in the study, 159 (43,6 %) men and 206 (56,4 %) women. As 135 subjects in the visited households were absent at the time of testing, the response rate attained was 74%. However, the denominator in this fraction included subjects with unknown eligibility status. Further, 41 participants were excluded from the study due to conditions contraindicating spirometry testing (8): 19 were due to known heart disorder, 15 had undergone chest/abdominal surgery, 5 were in last trimester of pregnancy, and 2 had undergone eye surgery. Eleven eligible subjects refused to participate, giving a cooperation rate of 97 %. After evaluating the spirometric measurements, 39 study participants (10,7 %) did not accomplish tests results compliant with ATS criteria, and were excluded from analysis regarding spirometric lung indices. Median age was 37 and 34 years for the male and female strata respectively. Age ranged from 16-79 for men and 15-90 for women. Table 1 presents sociodemographic characteristics of the study participants, as age categories, job categories, level of education and proportion of subjects with low socioeconomic position defined from a set of constructed core variables from the socio environmental questionnaire.

Tables 2 and 3 summarize frequencies and proportions of respiratory symptoms and conditions, based on the participants' responses to the ATS-DLD questionnaire. Both presence of symptoms and illnesses were more frequently reported among the female participants, though considerable in both sexes, with 26,4 % of men and 35,0 % of women reporting that they usually cough. Similarly, usually phlegm was recorded in 25,2 % and 32,0 % in men and women respectively. Also 22,0 % of the male strata, and 19,9 % of women responded that they have had problems with wheezing. Positive responses to questions regarding lung trouble before the age of 16, was recorded in 8,8 % in men and 12,6 % in women, and a total of 19,5 % and 17,5 % of men and women respectively, responded that they have had problems with chest illnesses during the past three years.

In tables 4 and 5 exposures to various risk factors are reported; table 4 presents history and magnitude of current/previous smoking among study participants, and table 5 summarizes self judged exposure to occupational and domestic pollutants. Men reported more frequently

exposure of dust and gas/chemicals in work, whereas a higher proportion among women reported presence of domestic exposure as biogas and charcoal. A total of 14,0 % of the study population responded that they were current smokers, and 4,7 % was ex-smokers. However, patterns of cigarette smoking differed substantially between the sexes, where the proportion of current smokers was considerably higher for men (30,2 %) than for women (1,5 %). Also 8,8 % of men, in contrast to only 1,5 % of women, reported a previously history of smoking. Mean pack-years for current smokers were 8,7 and 6,6 years for men and women respectively, whereas mean pack-years for ex-smokers were 7,5 for men, and 1,47 for women.

Prevalence of adjusted GOLD defined COPD when applying a fixed cut off ratio of $FEV1/FEV6 < 0,73$, was 12,6 %, and it was equally distributed among men (13,9 %) and women (11,5 %). Prevalence of ATS/ERS defined COPD when classified as below the lower limit of normal was 7,9 %, also equally distributed with 6,9 % and 8,7 % of men and women respectively. This is presented in table 6, together with frequency and proportion of cases of COPD according to GOLD's four disease severity stages. Among the study participants with COPD, only a small fraction were classified in stage 3 or higher (0,3 %). Figures 1-4 presents scatter plots of FEV1/FEV6 values *versus* age in both sexes. The proportions of observations defined as COPD cases according to both a fixed $FEV1/FEV6$ ratio < 0.73 and below the LLN are marked. The age distribution of COPD cases according to GOLD and ATS/ERS are also summarized in table 7.

Direct logistic regression was performed to assess the impact of certain determinants on COPD. Separate analyses were carried out for GOLD and ATS/ERS defined COPD. None of the categories in pack years made a significant contribution to the models in univariate analyses, and were therefore not included in further analyses, since smoking status was selected as predictor. The models contained the following covariates; sex, age, domestic exposure, occupational exposure, smoking status and socioeconomic position. When containing all predictors, both models were statistically significant, with χ^2 (7, N=326) = 43,93, p < 0,0005 when GOLD defined COPD was dependent variable, and χ^2 (7, N = 326) = 21,17, p = 0,004 when using the ATS/ERS defined COPD prevalence. The models correctly classified 88,7 % (GOLD defined COPD) and 92,0 % (ATS/ERS defined COPD) of the cases. The following independent variables made a statistically significant contribution to the GOLD defined COPD model; being ex-smoker, (p = 0,01), and age (p < 0,0005). The strongest predictor was ex-smoker, with an OR of 5,37. When the ATS defined COPD was dependent

variable, ex-smoker was still significant (p = 0,004) with an OR of 7,92. In addition, domestic exposure made a statistically significant contribution (p = 0,013). However, the OR was less than 1 (0,19), indicating a negative association. As the logistic regression was not adjusted for complex sampling, the precision of the OR estimates assumes simple random sampling. The odds ratios with their corresponding p-values and 95 % confidence intervals are presented in table 8.

Discussion

This paper has aimed to estimate prevalence of respiratory symptoms and COPD, defined according to both adjusted GOLD and ATS criteria, from a sample of the general adult population in Dar Es Salaam, Tanzania.

The prevalence of respiratory symptoms is high among the study participants, with 26,4 % of men and 35,0 % of women reporting that they usually cough. Similarly, usually phlegm is recorded in 25,2 % and 32,0 % in men and women respectively. Also 22,0 % of men and 19,9 % of women responded that they have had/ have problems with wheezing. There are no previous studies on respiratory symptoms in the general population from Sub-Saharan Africa, but when compared to Mokoetle's respiratory survey of a workforce in Johannesburg, the presence of respiratory symptoms among the study participants were also considerable and of similar magnitude (15). So far, mainly studies within occupational medicine have investigated the prevalence of respiratory symptoms, and how they are related to certain occupational pollutants and exposures (16, 17, 18). The burden of respiratory symptoms in the general population, beyond a specific workforce, should be paid more attention to in future research.

A limitation of our study is that COPD is diagnosed on basis of simple spirometry, though responses to a post-bronchodilator spirometric test and assessment of reversibility of airway obstruction are considered a necessity in diagnosing COPD (2, 3). However, due to limited resources and the nature of the data collection, it was not feasible to include post bronchodilator spirometry as part of the testing. Besides increased costs, transportation and storage demands, this would also imply a more elaborate training of assistants as well as a more time-consuming test procedure for the study participants.

In any case, one can question whether the prevalence of COPD in our study consists of truly positive COPD cases, as the diagnostic method used fails to discriminate between subjects who are responsive and non responsive to a bronchodilator and thus possibly insufficient in

differentiating COPD from asthma. According to a Norwegian study by Lehmann et al, a reversibility response of FEV1 increase ≥ 12 % and ≥ 200 ml after administration of a bronchodilator, was found in 2 % and 4 % of middle-aged (47-48 years) and elderly (71-73 years) participants respectively (19). Thus only a small fraction of the study population had a clinical relevant bronchodilatation after inhalation of salbutamol. This indicates that the proportion of possibly false positive COPD cases in our study might not be considerable. However, Lehmann did not investigate the reversibility response among those of younger age. In any case, the subjects classified as having COPD in our study, had a clear obstructive ventilatory pattern, and all of the subjects were in their habitual condition at the time of testing, which we consider relevant for the diagnosing, despite the lack of a post-bronchodilator test. GOLD acknowledges in its report that epidemiological field works face challenges in quantifying prevalence and assessing diagnosis and severity of COPD, and emphasizes that in areas lacking access to state-of-the-art diagnostic tools, diagnosis should be made with the equipment available (2[12]). This helps justify the diagnostic approach used in the present study.

One of the strengths of our study is that the spirometric testing was carried out by use of high-quality apparatus, and that strict adherence to ATS/ERS acceptability criteria were followed when assessing the manoeuvres. As the testing was conducted out-door in the field, the test devices were carried from house to house, and robust equipment, suitable for this setting was used. This might have influenced on the accuracy of the data obtained. In future research, it would be preferable to have a stationary test lab, where adherence to gold standards within test procedures is made easier as the testing is less prone to environmental influences. However, training and supervision of field assistants and use of a standardized test procedure throughout the data collection have been ensured, hopefully reducing the influence of measurement errors.

The prevalence of COPD when applying the fixed cut off ratio of 0.73, was 12,6 %, a thought-provoking number as the study population is young, with median age of 37 and 34 for men and women respectively. When using the lower limit of normal, prevalence of COPD was 7,9 %. The latter method pays attention to the relation between age and lung volumes, and the bias of over-diagnosing COPD in the elderly and *vice versa* among younger adults is reduced (9, 20). In our study, the consequences of using GOLD versus ATS criteria when

---

[12] GOLD report, page 89

defining COPD cases were not too evident, though the prevalence calculated when applying the lower limit of normal was lower compared to when GOLD's fixed FEV1/FEV6 ratio was used. We would anticipate that the age effect and the possibly bias of misclassifying COPD in both elderly and younger participants when using GOLD criteria would have been more evident if our study sample had been larger.

In any case, irrespective of the method used, the prevalence of COPD is surprisingly high. There are no previous studies in East Africa on prevalence of COPD, but according to the BOLD study, their site in Cape town, South Africa revealed an overall prevalence rate of stage I and stage II COPD at 4,7 % and 12,4 % respectively (11). This was the highest numbers of COPD reported in that study. The BOLD study used post-bronchodilator values when assessing COPD and in classification of disease severity (4). The authors also found that risk factors like smoking and occupational exposures were significant, and that prevalence of morbidities like prior tuberculosis was high. In our study, prevalence of stage I and II COPD was 6,1 % and 5,8 % respectively.

Our study showed a statistically significant between being ex-smoker and having COPD. No such association could be traced between current smoking and COPD. This finding might reflect that former smokers could have stopped smoking because of respiratory symptoms and airflow obstruction. Current smokers might not yet have developed obstructive lung disease, maybe due to a "healthy worker effect". The study failed to show a significant relationship between occupational exposure and COPD. There might be several explanations to this finding. Firstly, the reliability of the information collected on occupational exposure is poor. There was a high missing rate (26 %) on questions regarding these issues. People who are actively employed tend to have a more favorable health than the population at large, and those chronically ill and disabled are naturally excluded from being part of a workforce. This is known as the healthy worker effect, and could partly explain the non-significant association between occupational exposure and COPD.

We have no obvious explanation to our finding that domestic exposure seems to be negatively associated with COPD. The term "domestic exposure" was calculated based on answers to different questions, and the construct might not have captured the true exposures. The reliability of the predictor "domestic exposure" could therefore be questionable. In addition, the sample size was calculated with the purpose of determining prevalence of COPD, and the study might therefore not have enough power to detect any differences between sub-groups. Neither of the regression models was adjusted for complex sampling, thus the precision of the

estimates presupposes simple random sampling. The odds ratios with their corresponding p-values and 95 % confidence intervals should therefore be interpreted wit caution.

The prevalence of current smoking among the study participants must be considered disturbingly high among men (30,2 %). This should be given more attention in future research, as it has important implications concerning the health status and the expatiation of non-communicable diseases in the population in Dar Es salaam, Tanzania.

The diagnostic tools used to assess respiratory symptoms and conditions among the study participants, might have limitations. The feasibility of the ATS-DLD questionnaire to obtain information regarding respiratory symptoms in this population must be considered, where both meaning and phrasing of words and illnesses are rooted in a different cultural context and might not capture the true picture of the study participants interviewed. Developing culture specific questionnaires assessing respiratory symptoms and illness should be prioritised in future research. Also, the possible presence of response bias must be considered. As the questionnaires were both elaborate and time consuming, the bias known as satisficing, and the way the respondents are administrating the questionnaire, might reduce the validity of the recorded data. However, as the questionnaires were administered face to face by research assistants, there is an immediate opportunity to clear up misconceptions on the part of the respondent and this might reduce the risk of invalid responses. One must also question the validity of questions regarding presence of respiratory illnesses as emphysema, which requires diagnostic tools as x-ray or whole body pletysmography to verify. Also presence of asthma should be determined based on spirometric testing. Thus reporting of prevalence of emphysema (1,9 % and 1,0 % in men and women respectively), and cases of asthma (recorded in 10,7 % of men and 10,2 % of women) should be interpreted with caution as it might not reflect the true picture in the population.

Some further limitations of the present study need to be considered. The study sample was selected randomly, through a multistage cluster technique. However, difficulties in obtaining a correct sampling frame prior to the selection, and thus ensuring selection probabilities proportionate to size at the different stages, may limit the extent to which the study results can be generalised. Due to time and capacity restraints, the two last clusters were not visited. This could also have implications on the generalisability of the study outcomes and increase the overall sampling error though unlikely to be of significant magnitude.

Our study did not seem to identify expected determinants for development of COPD. Still the study provides valuable information regarding pulmonary symptoms and obstructive pulmonary disease in a population sample of Dar Es Salaam, and it puts respiratory health on the agenda. Increased awareness, and more attention to causal relations of COPD are important steps towards developing strategies for both prevention and treatment of COPD in accordance with the local epidemiological context, and should be prioritized in future research.

References

1   World Health Organization. Available from; http://www.who.int
        (accessed May 01, 2010

2   GOLD. Global strategy for the diagnosis, management, and prevention of Chronic
        Obstructive Pulmonary Disease. Updated 2008

3   Mannino DM. Chronic Obstructive Pulmonary Disease: Epidemiology and Evaluation.
        Hospital Physician 2001; 22-31

4   Buist SA et al. The Burden of Obstructive Lung Disease Initiative (BOLD): Rationale
        and Design. *Journal of Chronic Obstructive Pulmonary Disease.* 2005; 2; 277-283

5   Diaz PT et al. Increased susceptibility to Pulmonary Emphysema among HIV-
        seropositive smokers. *Ann Intern Med* 2000; 132; 369-372

6   American Thoracic Society. Lung function Testing: Selection of Reference Values and
        Interpretative Strategies. *Am Rev Respir Dis* 1991; 144; 1202-1218

7   Brusasco V, Crapo R, Viegi G. ATS/ERS task force. Interpretative strategies for Lung
        Function tests. *Eur Respir J* 2005; 26; 948 – 968

8   M.R Miller et al. Standardisation of spirometry. Series "ATS/ERS Task Force:
        Standardisation of Lung Function Testing" *Eur Respir J* 2005; 26; 319-338

9   Hansen JE, Sun XG, Wasserman K. Spirometric Criteria for Airway Obstruction.
        Chest 2007; 131; 349-355

10   Mfinanga SG et al. The role of livestock keeping in tuberculosis trends in Arusha,
        Tanzania. *Int J Tuberc Lung Dis* 2003; 7 (7); 695-704

11   Buist SA et al.International variation in the prevalence of COPD (The Bold Study): a
        population-based prevalence study. *Lancet* 2007; 370; 741-750

12   Swanney MP, Jensen RL, Chrichton DA, Beckert LE, Cardno LA, Crapo RO. FEV6 is
        an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and
        restriction. *Am J Respir Crit Care Med* 2000; 162; 917-919

13   Jing J, Huang T, Cui W, Xu F, Shen H. Should FEV1/FEV6 Replace FEV1/FVC
        Ratio To Detect Airway Obstruction? A Metaanalysis. *Chest* 2009; 135; 991-998

14   Vandevoorde J, Verbanck S, Schuermans D, Kartounian J, Vincken W. Obstructive
        and Restricitve spirometric patterns: fixed cut-offs for FEV1/FEV6 and FEV6. *Eur
        Respir J* 2006; 27; 378-383

15   Mokoetle KE, de Beer M, Becklake R. A respiratory survey in a black Johannesburg
        workforce. *Thorax* 1994; 49; 340-346

16   Hnizdo E. Combined effect of silica dust and tobacco smoking on mortality from chronic obstructive lung disease in gold miners. *British Journal of Industrial Medicine* 1990; 47; 656-664

17   Mamuya SHD, Bråtveit M, Mashalla Y, Moen BE. High prevalence of respiratory symptoms among workers in the development section of a manually operated coal mine in a developing country: A cross-sectional study. *BMC Public Health,* 2007. Available from: http://www.biomedcentral.com/1471-2458/7/17

18   Shamssain MH. Pulmonary function and symptoms in workers exposed to wood dust. *Thorax* 1992; 47; 84-87

19   Lehmann S, Bakke PS, Eide GE, Gulsvik A. Bronchodilator response to adrenergic β2- agonists: Relationships to symptoms in an adult community. *Respiratory Medicine* 2007; 101; 1183-1190

20   Hardie JA, Buist AS, Vollmer WM, Ellingsen I, BAkke PS, Morkve O. Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. *Eur Respir J* 2002; 20: 1117-1122.

21   Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *Lancet* 2009; 374; 733-43

Table 1

Table sociodemographic characteristics. Presented in frequencies and percentages						
Age categories, years	Total (n = 365)		Men (n = 159)		Women (n = 206)	
	N	(%)	N	(%)	N	(%)
<20	47	(12,9)	14	(8,8)	33	(16,0)
20-29	121	(33,2)	49	(30,8)	72	(35,0)
30-39	81	(22,2)	40	(25,2)	41	(19,9)
40-49	48	(13,2)	25	(15,7)	23	(11,2)
50-59	40	(11,0)	16	(10,1)	24	(11,7)
60-69	23	(6,3)	12	(7,5)	11	(5,3)
70-79	4	(1,1)	3	(1,9)	1	(0,5)
80+	1	(0,3)			1	(0,5 %)
Job categories *	Total (n = 269)		Men (n = 121)		Women (n = 148)	
White collar/professional	25	(9,3)	16	(13,2)	9	(6,1)
Agriculture	78	(29,0)	31	(25,6)	47	(31,8)
Sales/shop keeping	51	(19,0)	21	(17,4)	30	(20,3)
Blue collar	60	(22,3)	42	(34,7)	18	(12,2)
Student	25	(9,3)	11	(9,1)	14	(9,5)
Housewife/house girl					29	(19,6)
Level of education,	Total (n = 365)		Men (n = 159)		Women (n = 206)	
Never school	39	(10,7)	7	(4,4)	32	(15,5)
Primary level	210	(57,5)	91	(57,2)	119	(57,8)
Secondary level	96	(26,3)	47	(29,6)	49	(23,8)
Advanced secondary	7	(1,9)	4	(2,5)	3	(1,5)
Tertiary	13	(3,6)	10	(6,3)	3	(1,5)
	Total (n = 365)		Men (n = 159)		Women (n = 206)	
Low SEP index	90	(24,7)	29	(18,2)	61	(29,6)

N = number of respondents. % = proportion of respondents. SEP = socioeconomic position
* non response rate 26 %

Table 2

Table; frequency (%) of respiratory symptoms. Presented with 95 % confidence intervals*						
Symptoms	Men (n = 159)			Women (n = 206)		
	N	(%)	95 % CI	N	(%)	95 % CI
Usually cough	42	(26,4)	20,6 % – 33,2 %	72	(35,0)	28,8 % - 41,6 %
Cough more than 5 months during year in last 2 years	8	(5,0)	2,4 % - 10,3 %	7	(3,4)	1,7 % - 6,5 %
Usually phlegm	40	(25,2)	18, 8 % - 32,8 %	66	(32,0)	26,9 % - 37,7 %
Phlegm more than 3 months during year in last 2 years	8	(5,0)	2,4 % - 10,4 %	8	(3,9)	1,7 % - 6,5 %
Ever wheeze	35	(22,0)	14,2 % - 32,5 %	41	(19,9)	15,8 % - 24,7 %
Wheezing attack	19	(11,9)	7,5 % - 18,6 %	26	(12,6)	8,6 % - 18,2 %
Breathless walking	30	(18,9)	12,5 % - 27,5 %	65	(31,6)	24,6 % - 39,4 %
Need to stop for breath at own pace	19	(11,9)	7,8 % - 18,0 %	40	(19,4)	14,1 % - 26,1 %
Need to stop for breath after few minutes walk	13	(8,2)	4,9 % - 13,4 %	21	(10,2)	6,6 % - 15,4 %

N =, number of respondents; % = proportion of respondents; CI = confidence interval
*Adjusted for complex sampling

Table 3

Condition	Men (n = 159)			Women (n = 206)		
	N	(%)	95 % CI	N	(%)	95 % CI
Lung trouble before the age of 16	14	(8,8)	5,0 % - 14,9 %	26	(12,6)	9,3 % - 16,9 %
Usually chest colds	32	(20,1)	12,6 % - 30,6 %	55	(26,7)	20,9 % - 33,4 %
Chest illnesses past 3 years	31	(19,5)	13,1 % - 28,0 %	36	(17,5)	11,1 % - 26,4 %
Bronchitis confirmed by a doctor	9	(5,7)	2,2 % - 13,9 %	25	(12,1)	7,2 % - 19,8 %
Pneumonia confirmed by a doctor	17	(10,7)	6,1 % - 18,1 %	30	(14,6)	9,0 % - 22,7 %
Hayfever confirmed by a doctor	9	(5,7)	2,9 % - 10,9 %	21	(10,2)	5,7 % - 17,5 %
Ever chronic bronchitis	2	(1,3)	0,2 % - 9,0 %	7	(3,4)	1,6 % - 7,3 %
Ever emphysema	3	(1,9)	0,7 % - 5,0 %	2	(1,0)	0,1 % - 7,3 %
Ever asthma	17	(10,7)	5,9 % - 18,5 %	21	(10,2)	6,8 % - 15,0 %
Ever tuberculosis	6	(3,8)	1,3 % - 10,4 %	3	(1,5)	0,5 % - 4,5 %

Table; frequency (%) of respiratory illnesses. Presented with 95 % confidence intervals*

N = number of respondents; % = proportion of respondents; CI = confidence interval
*Adjusted for complex sampling

Table 4

Table; Smoking history and burden. Presented with 95 % confidence interval*			
Smoking status, N (%)	Total (n = 365)	Men (n = 159)	Women (n = 206)
Never smoker,	297 (81,4)	97 (61,0)	200 (97,1)
Ex-smoker	17 (4,7)	14 (8,8)	3 (1,5)
Current smoker	51 (14,0)	48 (30,2)	3 (1,5)
Pack-year ex-smoker	Total (n = 17**)	Men (n = 14**)	Women (n = 3**)
M, 95 % CI	6,44 -2,59 – 15,47	7,51 -2,76 – 17,77	1,47 -0,46 – 3,39
Pack-year categories ex-s N (%)			
<1	5 (29,4)	3 (21,4)	2 (66,7)
1-9,99	10 (58,8)	9 (64,3)	1 (33,3)
10 and above	2 (4,8)	2 (14,3)	
Pack-year current smoker	Total (n = 51)	Men (n = 48)	Women (n = 3**)
M, 95 % CI	8,61 5,67 – 11,56	8,74 5,61 – 11,88	6,55 0,27 – 12,83
Pack-year categories cu-s N (%)			
<1	8 (15,7)	8 (16,7)	
1-9,99	26 (51,0)	24 (50,0)	2 (66,7)
10 and above	17 (33,0)	16 (33,3)	1 (33,3)

N = number of respondents. % = proportion of respondents. M = mean. CI = confidence interval.
* Adjusted for complex sampling ** Inference difficult due to few observations

Table 5

Table; Exposure to occupational or domestic pollutants			
	Total (n = 365)	Men (n = 159)	Women (n = 206)
Occupational exposure of dust, gas or chemical pollutants N (%)	24 (6,6)	22 (13,8)	2 (1,0)
Domestic exposure of biogas and charcoal N (%)	108 (29,6	48 (30,2)	60 (29,1)

No, number of respondents; %, proportion of respondents; M, mean; SE, standard error.

Table 6

Table; frequency and proportion of COPD, defined by ATS and GOLD severity stages Presented with 95 % CI*						
	Total (n = 326[13])		Men (n=144)		Women (n = 182)	
COPD FEV1/FEV6<0.73 N (%), 95 % CI	41 (12,6)	8,8 % – 17,6 %	20 (13,9)	9,0 % - 20,9 %	21 (11,5)	8,0 % - 16,4 %
COPD LLN N (%), 95 % CI GOLD, N (%) 95 % CI	29 (7,9)	5,2 % – 11,9 %	11 (6,9)	3,6 % - 12,9 %	18 (8,7)	5,6 % - 13,3 %
GOLD stage 1	20 (6,1)	3,8 % - 9,7 %	8 (5,6)	2,8 % - 10,8 %	12 (6,6)	3,7 % - 11,4 %
GOLD stage 2	19 (5,8)	3,8 % – 8,9 %	11 (7,6)	4,6 % - 12,4 %	8 (4,4)	2,3 % - 8,1 %
GOLD stage 3	1 (0,3)	0,0 % - 2,4 %			1 (0,5)	0,1 % - 4,2 %
GOLD stage 4	1 (0,3)	0,0 % - 2,4 %	1 (0,7)	0,1 % - 5,4 %		

N = number of respondents, % = proportion of respondents, CI = Confidence Interval, LLN=Lower Limit of Normal
* Adjusted for complex sampling

Table 7

Table; age distribution of frequency and proportion of GOLD defined COPD and ATS defined COPD				
Age categories	COPD (GOLD) (n = 41)		COPD (ATS) (n = 29)	
	N (%)	95 % CI*	N (%)	95 % CI
<20			1 (2,4)	0,3 % - 16,8 %
20-29	9 (8,6)	5,0 % - 14,2 %	11 (10,5)	5,6 % - 18,7 %
30-39	4 (5,5)	1,5 % - 17,7 %	2 (2,7)	0,3 % – 20,6 %
40-49	6 (13,0)	8,1 % - 20,4 %	4 (8,7)	4,4 % - 16,6 %
50-59	11 (31,4)	16,6 % - 51,4 %	6 (17,1)	8,3 % - 32,1 %
60-69	9 (40,9)	25,1 % - 58,9 %	5 (22,7)	8,3 % - 32,1 %
70 +	2 (50,0)	10,4 % - 89,6 %		

N = number of respondents, % = proportion of respondents, CI = confidence interval
*Adjusted for complex sampling

---

[13] Denominator includes only subjects with acceptable spirometry

Table 8

Table; Logistic regression. Determinants of COPD (defined by GOLD and ATS).
Presented with OR and 95 % CI*

Determinants	COPD (GOLD)			COPD (ATS)		
	p	OR	95 % CI for OR	p	OR	95 % CI for OR
Sex	0,55	1,30	0,55 – 3,09	0,28	1,71	0,64 – 4,54
Age	<0,0005	1,06	1,04 – 1,09	0,09	1,02	1,00 – 1,05
Ex-smoker	0,011	5,37	1,46 – 19,71	0,004	7,92	1,91 – 32,76
Current smoker	0,24	1,85	0,66 – 5,15	0,85	0,87	0,21 – 3,59
Domestic exposure	0,11	0,48	0,20 – 1,17	0,013	0,19	0,05 – 0,71
Occupational exposure	0,56	1,45	0,42 – 4,98	0,57	1,57	0,33 – 7,41
SEP	0,82	1,10	0,48 – 2,54	0,52	0,75	0,31 – 1,82

OR= Odds Ratio. CI= Confidence Interval. p = p-value. SEP= Socioeconomic position
* Not adjusted for complex sampling

Figure 1; scatter plots of COPD cases *versus* healthy among men defined by GOLD plotted against age
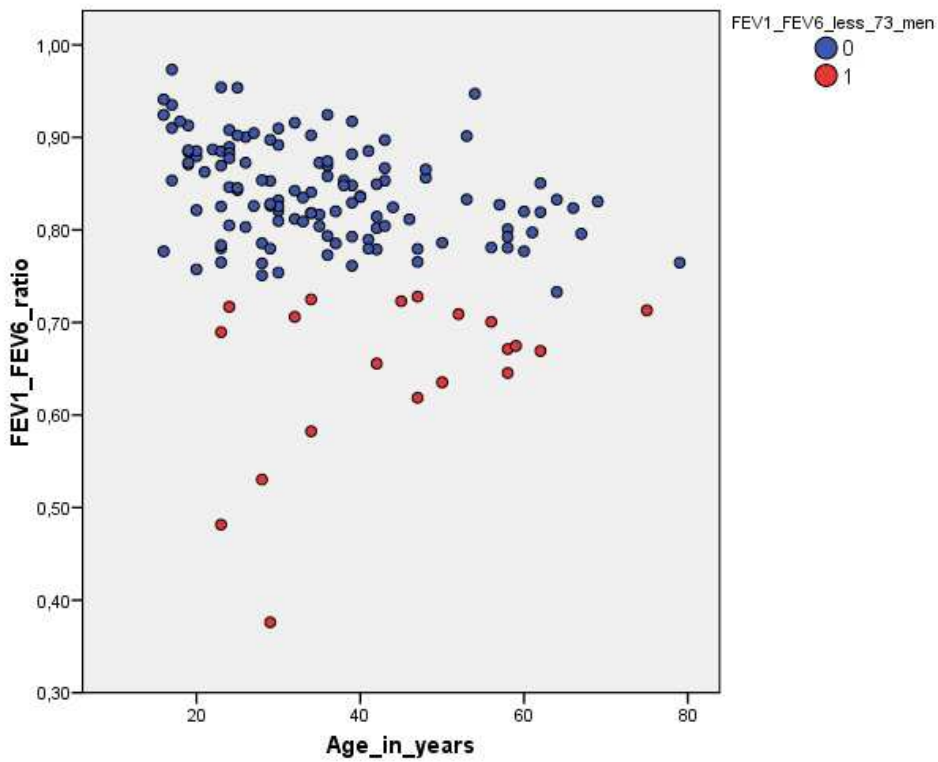
Figure 2; scatter plots of COPD cases *versus* healthy among men defined by ATS plotted against age. Lower limit of normal highlighted
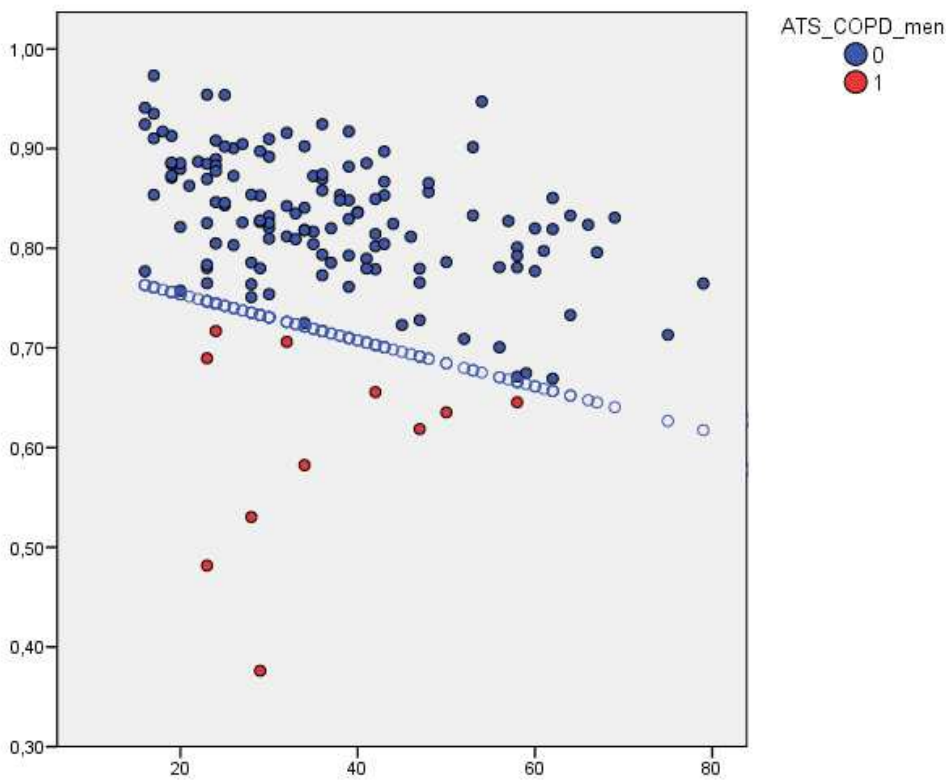
Figure 3; scatter plots of COPD cases *versus* healthy among women defined by GOLD Plotted against age
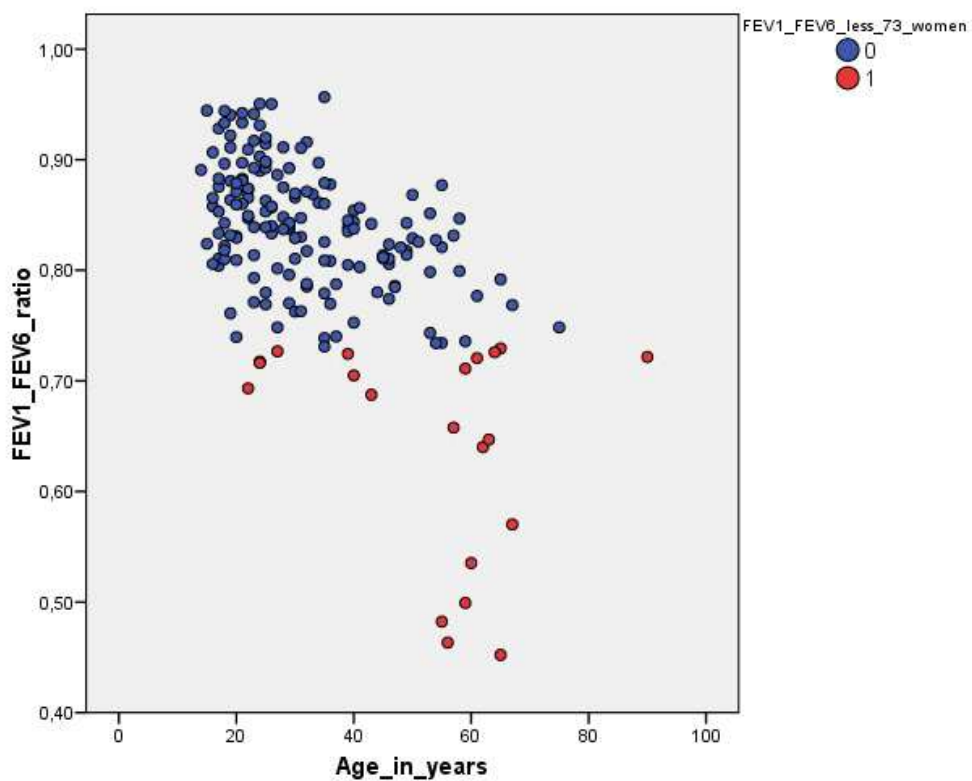
Figure 4; scatter plots of COPD cases *versus* healthy among women defined by ATS against age. Lower limit of normal highlighted
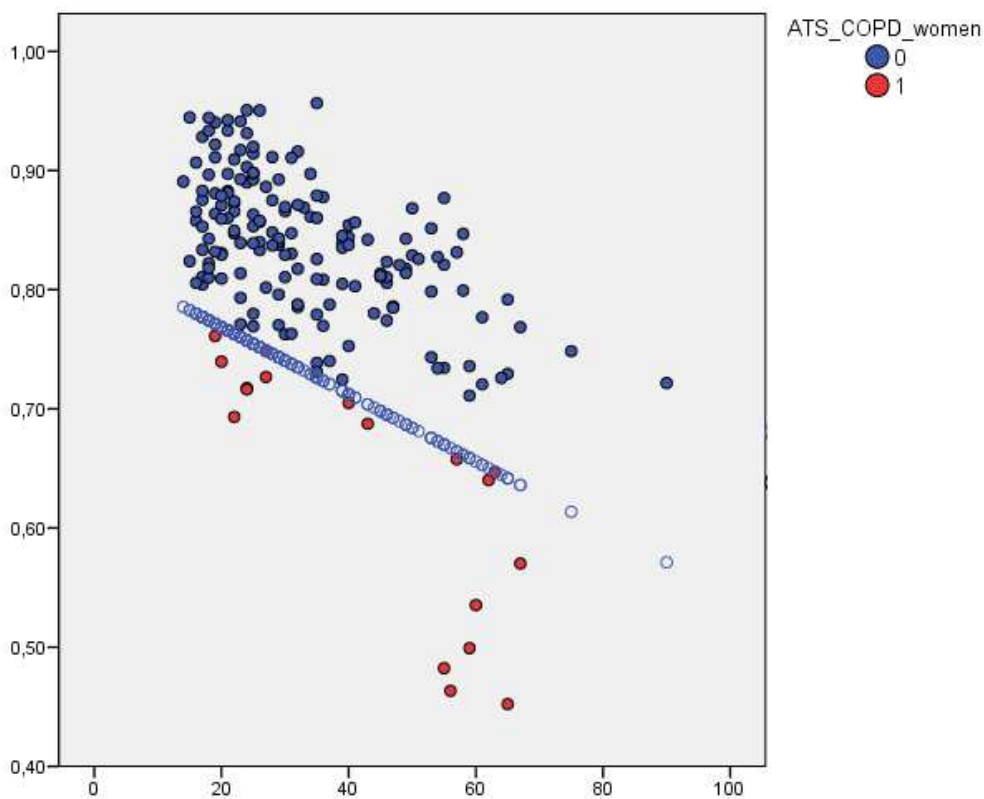
Appendix 1

ATS-DLD questionnaire

ATS-DLD-78-A |

| ADULT QUESTIONNAIRE - SELF COMPLETION |

| (for those 13 years of age and older) |

Thank you for your willingness to participate. You were selected by a scientific sampling procedure, and your cooperation is very important to the success of this study.   This is a questionnaire you are asked to fill out. Please answer the questions as frankly and accurately as possible.

ALL INFORMATION OBTAINED IN THE STUDY WILL BE KEPT CONFIDENTIAL AND USED FOR MEDICAL RESEARCH ONLY. Your personal physician will be informed about the test results if you desire.

|

|---|

                IDENTIFICATION

        IDENTIFICATION NUMBER: #####

        NAME:_____ _____ ___
                (Last)   (First)   (MI)

        STREET _____

        CITY _____ STATE ____ ZIP _____

        PHONE NUMBER: (   ) _____-_____

        INTERVIEWER: ###

        DATE: _____
            MO   DAY   YR

1. BIRTHDATE: _____ ____ _____
            Month   Day   Year

2. Place of Birth: _____

3. Sex: 1. Male ____
                        2. Female ____

4. What is your marital status?   1. Single ____

2. Married ____

3. Widowed ____

4. Separated/Divorced ____

5. Race: 1. White ____

2. Black ____

3. Oriental ____

4. Other ____

6. What is the highest grade completed in school? _____

(For example: 12 years is completion of high school)

SYMPTOMS

These questions pertain mainly to your chest. Please answer yes or no if possible. If a question does not appear to be applicable to you, check the "does not apply" space. If you are in doubt about whether your answer is yes or no, record no.

COUGH

7A. Do you usually have a cough?   1. Yes ___ 2. No ___

(Count a cough with first smoke or on first going

out-of-doors. Exclude clearing of throat.)[If no,

skip to question 7C.]

B. Do you usually cough as much as 4 to 6 times a   1. Yes ___ 2. No ___

day, 4 or more days out of the week?

C. Do you usually cough at all on getting up, or   1. Yes ___ 2. No ___

first thing in the morning?

D. Do you usually cough at all during the rest   1. Yes ___ 2. No ___

of the day or at night?

IF YES TO ANY OF THE ABOVE(7A,7B,7C, OR 7D), ANSWER THE FOLLOWING:

IF NO TO ALL, CHECK DOES NOT APPLY AND SKIP TO 8A.

E. Do you usually cough like this on most days for   1. Yes ___ 2. No ___

5 consecutive months or more during the year?

8. Does not apply __

F. For how many years have you had this cough?   _____

<div align="center">Number of years</div>
<div align="center">88. Does not apply __</div>

PHLEGM

8A. Do you usually bring up phlegm from your chest?   1. Yes ___ 2. No ___
    (Count phlegm with the first smoke or on first
    going out-of-doors. Exclude phlegm from the
    nose. Count swallowed phlegm)
    [If no, skip to 8C.]

B. Do you usually bring up phlegm like this as   1. Yes ___ 2. No ___
   much as twice a day, 4 or more days out of the week?

C. Do you usually bring up phlegm at all on get-   1. Yes ___ 2. No ___
   ting up or first thing in the morning?

D. Do you usually bring up phlegm at all during   1. Yes ___ 2. No ___
   the rest of the day or at night?

IF YES TO ANY OF THE ABOVE (8A, B, C, OR D),
ANSWER THE FOLLOWING:
IF NO TO ALL, CHECK DOES NOT APPLY AND SKIP TO 9A.

E. Do you bring up phlegm like this on most days   1. Yes ___ 2. No ___
   for 3 consecutive months or more during the
   year?   8. Does not apply __

F. For how many years have you had trouble with   _____
   phlegm?   Number of years
<div align="center">88. Does not apply __</div>

EPISODES OF COUGH AND PHLEGM

9A. Have you had periods or episodes of (in-   1. Yes ___ 2. No ___
   creased*) cough and phelgm lasting for 3
   weeks or more each year?

*(For individuals who usually have cough and/or
phlegm)

IF YES TO 9A:

B. For how long have you had at least 1 such episode per year?   _____(Number of years)

88. Does not apply __

WHEEZING

10A. Does your chest ever sound wheezy or whis-
tling:

    1. When you have a cold?   1. Yes ___ 2. No ___

    2. Occaisonally apart from colds?   1. Yes ___ 2. No ___

    3. Most days or nights?   1. Yes ___ 2. No ___

IF YES TO 1, 2, OR 3 IN 10A:

B. For how many years has this been present?   _____

Number of years

88. Does not apply __

11A. Have you ever had an ATTACK of wheezing that   1. Yes ___ 2. No ___
has made you feel short of breath?

IF YES TO 11A:

B. How old were you when you had your first   _____ Age in years
such attack?   88. Does not apply __

C. Have you had 2 or more such episodes?   1. Yes ___ 2. No ___
8. Does not apply __

D. Have you ever required medicine or treatment   1. Yes ___ 2. No ___
for the(se) attack(s)?   8. Does not apply __

BREATHLESSNESS

12. If disabled from walking by any condition
    other than heart or lung disease, please
    describe and proceed to Question 14A.

    Nature of condition(s):_____

13A. Are you troubled by shortness of breath when
     hurrying on the level or walking up a slight hill?   1. Yes ___ 2. No ___

     IF YES TO 13A:

  B. Do you have to walk slower than people of your   1. Yes ___ 2. No ___
     age on level because of breathlessness?   8. Does not apply __

  C. Do you ever have to stop for breath when walk-   1. Yes ___ 2. No ___
     ing at your own pace on the level?   8. Does not apply __

  D. Do you ever have to stop for breath after walk   1. Yes ___ 2. No ___
     ing about 100 yards(or after a few minutes) on   8. Does not apply __
     the level?

  E. Are you too breathless to leave the house or   1. Yes ___ 2. No ___
     breathless on dressing or undressing?   8. Does not apply __

   CHEST COLDS AND CHEST ILLNESSES

14A. If you get a cold, does it usually go to your   1. Yes ___ 2. No ___
     chest? (Usually means more than 1/2 the time)   8. Don't get colds__

15A. During the past 3 years, have you had any   1. Yes ___ 2. No ___
     chest illnesses that have kept you off work, in-
     doors at home, or in bed?

     IF YES TO 15A:

  B. Did you produce phlegm with any of these   1. Yes ___ 2. No ___
     chest illnesses?   8. Does not apply __

C. In the last 3 years, how many such illnesses, _____Number of illnesses
with (increased) phlegm, did you have which _____No such illnesses
lasted a week or more? _____Does not apply

PAST ILLNESSES

16. Did you have any lung trouble before the age   1. Yes ___ 2. No ___
of 16?

17. Have you ever had any of the following:
1A. Attacks of Bronchitis?   1. Yes ___ 2. No ___

  IF YES TO 1A:
B. Was it confirmed by a doctor?   1. Yes ___ 2. No ___
         8. Does not apply __

C. At what age was your first attack?   _____ Age in years
        88. Does not apply __

2A. Pneumonia (include bronchopneumonia)?   1. Yes ___ 2. No ___

IF YES TO 2A:
B. Was it confirmed by a doctor?   1. Yes ___ 2. No ___
         8. Does not apply __

C. At what age did you first have it?   _____ Age in years
        88. Does not apply __

3A. Hayfever?   1. Yes ___ 2. No ___

  IF YES TO 3A:
B. Was it confirmed by a doctor?   1. Yes ___ 2. No ___
         8. Does not apply __

C. At what age did it start?   _____ Age in years
        88. Does not apply __

18A. Have you ever had chronic bronchitis?   1. Yes ___ 2. No ___

IF YES TO 18A:

B. Do you still have it?   1. Yes ___ 2. No ___

                                      8. Does not apply __

C. Was it confirmed by a doctor?   1. Yes ___ 2. No ___

                                      8. Does not apply __

D. At what age did it start?   _____ Age in years

                                  88. Does not apply __

19A. Have you ever had emphysema?   1. Yes ___ 2. No ___

IF YES TO 19A:

B. Do you still have it?   1. Yes ___ 2. No ___

                                      8. Does not apply __

C. Was it confirmed by a doctor?   1. Yes ___ 2. No ___

                                      8. Does not apply __

D. At what age did it start?   _____ Age in years

                                  88. Does not apply __

20A. Have you ever had asthma?   1. Yes ___ 2. No ___

IF YES TO 20A:

B. Do you still have it?   1. Yes ___ 2. No ___

                                      8. Does not apply __

C. Was it confirmed by a doctor?   1. Yes ___ 2. No ___

                                      8. Does not apply __

D. At what age did it start?   _____ Age in years

                                  88. Does not apply __

E. If you no longer have it, at what age did it   _____ Age stopped
    stop?   88. Does not apply __

21. Have you ever had:

    A. Any other chest illnesses?   1. Yes ___ 2. No ___

       If yes, please specify _____

    B. Any chest operations?   1. Yes ___ 2. No ___

       If yes, please specify _____

    C. Any chest injuries?   1. Yes ___ 2. No ___

       If yes, please specify _____

22A. Has doctor ever told you that you had heart   1. Yes ___ 2. No ___
      trouble?

     IF YES to 22A:

    B. Have you ever had treatment for heart trouble   1. Yes ___ 2. No ___
      in the past 10 years?   8. Does not apply __

23A. Has a doctor ever told you that you have high   1. Yes ___ 2. No ___
      blood pressure?

     IF YES to 23A:

    B. Have you had any treatment for high blood   1. Yes ___ 2. No ___
      pressure (hypertension) in the past 10 years?   8. Does not apply __

OCCUPATIONAL HISTORY

24A. Have you ever worked full time (30 hours per   1. Yes ___ 2. No ___
      week or more) for 6 months or more?

     IF YES to 24A:

    B. Have you ever worked for a year or more in   1. Yes ___ 2. No ___
      any dusty job?   8. Does not apply __

     Specify job/industry: _____ Total years worked __
     Was dust exposure 1. Mild ___   2. Moderate ___   3. Severe ___ ?

C. Have you ever been exposed to gas or chemical   1. Yes ___ 2. No ___
   fumes in your work?   8. Does not apply __

   Specify job/industry: _____ Total years worked __
   Was dust exposure 1. Mild ___   2. Moderate ___   3. Severe ___ ?

D. What has been your usual occupation or job -- the one you have
   worked at the longest?

   1. Job-occupation: _____
   2. Number of years employed in this occupation:_____
   3. Position-job title: _____
   4. Business, field, or industry: _____

TOBACCO SMOKING

25A. Have you ever smoked cigarettes? (NO means   1. Yes ___ 2. No ___
    less than 20 packs of cigarettes or 12 oz. of
    tobacco in a lifetime or less than 1 cigarette a
    day for 1 year.

   IF YES to 25A:

B. Do you now smoke cigarettes (as of 1 month   1. Yes ___ 2. No ___
   ago)?   8. Does not apply __

C. How old were you when you first started reg-   ____ Age in Years
   cigarette smoking?   88.Does not apply __

D. If you have stopped smoking cigarettes com-   ____ Age stopped
   pletely, how old were you when you stopped?   Check if
                                                   still smoking ___
                                                   88.Does not apply __
E. How many cigarettes do you smoke per day now?   ___ Cigarettes/day
                                                   88.Does not apply __

F. On the average of the entire time you smoked,   ___ Cigarettes/day
   how many cigarettes did you smoke per day?   88.Does not apply __

G. Do or did you inhale the cigarette smoke?   1. Does not apply __

                                      2. Not at all _____

                                        3. Slightly _____

                                        4. Moderately _____

                                        5. Deeply _____

26A. Have you ever smoked a pipe regularly?   1. Yes ___ 2. No ___
(YES means more than 12 oz tobacco in a
lifetime.)

    IF YES to 26A:

B1. How old were you when you started to   ____ Age
   smoke a pipe regularly?

 2. If you have stopped smoking a pipe com-   ____ Age stopped
   pletely, how old were you when you stopped?   Check if still
                                  smoking pipe ____
                                88.Does not apply __

C. On the average over the entire time you   ____ oz per week (a stan-
   smoked a pipe, how much pipe tobacco did dard pouch of tobacco con-
   you smoke per week ?   tains 1 1/2 oz)
                                88.Does not apply __

D. How much pipe tobacco are you smoking now?   ___ oz per week
                   88. Not currently smoking a pipe ___

E. Do or did you inhale the pipe smoke?   1. Never smoked ____

                                         2. Not at all _____

                                         3. Slightly _____

                                         4. Moderately _____

                                         5. Deeply _____

27A. Have you ever smoked cigars regularly?   1. Yes ___ 2. No ___
(Yes means more than 1 cigar a week for a
year).

IF YES to 27A:

B1. How old were you when you started smok-   ____ Age
    ing cigars regularly?

 2. If you have stopped smoking cigars com-   ____ Age stopped
    pletely, how old were you when you stopped?   Check if still
                        smoking cigars___
                     88.Does not apply __

C. On the average over the entire time you   ___ Cigars per week
    smoked cigars, how many cigars did you smoke   88.Does not apply __
    per week ?

D. How many cigars are you smoking per week   ___ Cigars per week
    now?   88. Check if not smoking cigars currently __

E. Do or did you inhale the cigar smoke?   1. Never smoked ____
                        2. Not at all _____
                        3. Slightly _____
                        4. Moderately _____
                        5. Deeply _____

FAMILY HISTORY

28. Were either of your natural parents ever told by a doctor that they
    had a chronic lung condition such as:

          FATHER   MOTHER

     1. YES 2. NO 3. DON'T   1. YES 2. NO 3. DON'T
            KNOW   KNOW

A. Chronic
   brochitis? _____ _____ _____   _____ _____ _____

B. Emphysema? _____ _____ _____   _____ _____ _____

C. Asthma?   _____   _____ _____   _____   _____ _____

D. Lung cancer? _____   _____ _____   _____   _____ _____

E. Other chest
  conditions? _____   _____ _____   _____   _____ _____

29A. Is parent currently alive?

                _____   _____ _____   _____   _____ _____

B. Please Specify:

        _____ Age if living   _____ Age if living

        _____ Age at death   _____ Age at death

      8. Don't know _____   8. Don't know _____

C. Please specify cause of death.

    _____   _____

Appendix 2

Questionnaire, socioeconomic conditions

1 **Level of Education**: *Please tick appropriate box*

Never attended school   1

Primary level education   2

Secondary level education   3

Advanced secondary school (High Level)   4

Tertiary education, including undergraduate and diploma level   5

Higher advanced learning = masters and PHD programmes   6

2. Question

a) Are you employed?

Yes	1	Go to questions 8b and 8c
No	2	Go to question 9

b) Where do you work? _____

c) What is your position there? _____

3. In the past 7 days have you had any work for pay?

Yes	1	Go to question 10
No	2	Skip to question 11

4 What kind of work was it? _____

5. When was the last time that you worked for money?

Within the last month	1

Within the last 3 months	2
Over six months ago	3

6. In the past 10 days have you had money available from other sources than work?

Yes	1	Go to question 11
No	2	Skip to question 12

7. What are the sources_____?

8. What is the number of people living in the household (excluding visitors)? _____ _____

9. Who is the head the household? Male or female _____

10. What is the main source of drinking water for members of your household?

Piped	1		In residence	1	Outside residence	2	Public tap	3
Open well	2		In yard/plot	1	Outside yard	2		
Protected well	3		In yard/plot	1	Outside yard	2		
Borehole	4		In yard/plot	1	Outside yard	2		
Spring water	5							
Rain water	6							
Tanker truck water	7							
River, canal or surface water	8							
Bottled water	9							
Water from gravity flow scheme	10							

11. What Type of toilet does your household have?

Flush toilet (own)	1
Flush toilet (shared with other household)	2
Traditional pit toilet/ latrine	3
Ventilated improved pit latrine	4
No facility/ Bush/ Field	5
Other : Please specify_____	6

12. Please indicate by ticking if your household has:

Electricity	1
Radio	2
Television	3
Telephone (fixed) /(mobile)	4
Refrigerator	5
Lantern	6
Cupboard	7

13. What is the principal flooring material in your house?

Earth/ dirt/sand/dung	1
Cement	2
Vinyl or asphalt tile	3
Ceramic tiles	4

Parquet or polished wood	5
Other, please specify_____	6

14. What is the main material in the walls?

Thatched	1
Mud and pole	2
Un burnt bricks	3
Burnt bricks with mud	4
Burnt bricks with cement	5
Timber	6
Cement blocks, bricks, concrete walls	7
Stone	8
Other, please specify_____	9

15. What is the principal roofing material in your house?

Thatched	1
Iron sheets	2
Asbestos	3
Tiles	4
Tin	5
Cement	6
Other, please specify_____	7

16. Does any member of your household own………?

A bicycle	1
A motorcycle or motor scooter	2
A car or truck	3
A boat or canoe	4
A donkey	5

17. What do you use for lighting?

Electricity	1
Biogas	2
Kerosene	3
Charcoal	4
Dung	5
Others please specify _____	

18. Would you say that people in your household often, sometimes, seldom or never go hungry?

Often	1
Sometimes	2
Seldom	3
Never	4

Appendix 3

Informed consent form

Informed Consent
You are being asked to participate in a study assessing lung function for an adult population in Kinondoni district, Dar Es Salaam, Tanzania. The aim of the study is partly to establish values for normal lung function, and partly to determine the frequency of respiratory symptoms and a condition called COPD (Chronic Obstructive Pulmonary Disease), which is an astma-like disease. Since there is poor knowledge on how frequent this condition is in Tanzania, it is important to collect more information to enable health policy makers to plan for intervention strategies.
We ask you to participate in this study because you are a member of this community, and are above the age of 15.
If you are willing to participate in the proposed study, the information you provide will be kept confidential. Your name will not be required, and all tests and answers will be handled anonymously. We will request you to blow 3 times in a spirometer(an apparatus that measures lung volumes), and we will register your age, and measure your height and weight. Finally we want you to answer some questions regarding respiratory symptoms and living conditions. If we detect disease that needs medical attention, you will be aided with referral to proper medical care.
Participation in the study is voluntary and free of charge, and you can withdraw at any time, also after consenting to participate, without giving any reason. Your answers and test results will then be deleted.
Thank you for your cooperation

Correspondence:
Toril Morkve Knudsen- Principal investigator
National Institute for Medical Research
Muhimbili Research Centre
Tel 0744 307624

Study participant:
I have been informed about the nature and aim of the study, and I give my consent to participate
Signature of study participant:……………………………………………….
Thumb print of study participant:
Date:………/………/………………………