Comparison of Spirometry Criteria for the Diagnosis of COPD: Results from the BOLD Study

William M. Vollmer,[1] Þórarinn Gíslason,[2] Peter Burney,[3]
Paul L. Enright,[4] Amund Gulsvik,[5] Ali Kocabas,[6] A. Sonia Buist[7]
for the BOLD Collaborative Research Group

[1] Kaiser Permanente, Center for Health Research, Portland, OR, USA
[2] University of Iceland, Medical Faculty, Landspitali University Hospital, Reykjavik, Iceland
[3] National Heart and Lung Institute, Imperial College, London, UK
[4] The University of Arizona, Tucson, AZ, USA
[5] Institute of Medicine, University of Bergen, Bergen, Norway
[6] Cukurova University School of Medicine, Balcali, Adana, Turkey
[7] Oregon Health & Science University, Portland, OR, USA

Corresponding Author
William M. Vollmer, PhD
Center for Health Research
3800 N Interstate Ave
Portland, OR USA
97227-1110
503-335-6755 telephone
503-335-6311 fax
email: william.vollmer@kpchr.org

Key Words: Adult, Chronic Obstructive Pulmonary Disease, Epidemiology

Running Title: Comparing spirometry-based COPD definitions

Abstract

Background: Published guidelines recommend spirometry to accurately diagnose COPD. However, even spirometry-based COPD prevalence estimates can vary widely. We compared properties of several spirometry-based COPD definitions using data from the international Burden of Obstructive Lung Disease (BOLD) study.

Methods: 14 sites recruited population-based samples of adults aged 40 and older. Procedures included standardized questionnaires and post-bronchodilator spirometry. 10,001 individuals provided usable data.

Results: Use of the lower limit of normal (LLN) FEV_1/FVC reduced the age-related increases in COPD prevalence that are seen among healthy never smokers when using the fixed ratio criterion ($FEV_1/FVC < 0.7$) as recommended by the Global Initiative for Obstructive Lung Disease (GOLD). The added requirement of an FEV_1 either $< 80\%$ predicted, or below the LLN, further reduced age-related increases and also led to the least site-to-site variability in prevalence estimates after adjusting for potential confounders. Use of the FEV_1/FEV_6 ratio in place of the FEV_1/FVC yielded similar prevalence estimates.

Conclusion: Use of the $FEV_1/FVC < LLN$ criterion instead of the $FEV_1/FVC < 0.7$ should minimize known age biases and better reflect clinically significant irreversible airflow limitation. Our study also supports the use of the FEV_1/ FEV_6 as a practical substitute for the FEV_1/ FVC.

Introduction

Although COPD is recognized as a major public health problem world-wide, estimates of its prevalence vary widely.[1] Much of this variation likely reflects differences in the populations studied, spirometry methods, and data-quality control, and the rules used to define COPD. For example, self-reported physician diagnosis of COPD typically results in estimated prevalences well below those obtained based on spirometry.[1;2]

Although no gold-standard definition of COPD exists, published guidelines recommend use of spirometry to define it.[3;4]. However, even spirometry-based COPD prevalence estimates can vary by two-fold or more, depending on the definition used to classify mild disease.[5;6] The most widely used definition comes from the Global Initiative for Obstructive Lung Disease (GOLD), which recommends using a post-bronchodilator $FEV_1/FVC<0.7$ to define irreversible airflow limitation, and the FEV_1 to stage disease.[3] This "fixed ratio" approach, while easy to apply, appears to overestimate COPD in older individuals [2;7-10] and to underestimate it in young adults.[9;11]Alternative definitions that account for normal aging can alleviate this bias,[9;12] but in turn raise questions about which reference equations are appropriate for which populations. In addition, if pre- (rather than post-) bronchodilator spirometry is used, COPD prevalence may be overestimated by as much as 30%.[8;10;13;14]

The Burden of Obstructive Lung Disease (BOLD) study is an international effort to collect population-based estimates of the prevalence and economic burden of COPD using standardized methods.[15;16] Using BOLD data, we examine the impact on prevalence estimates of using the fixed-ratio criterion versus various other spirometry-based definitions of COPD. We also

compare the effects of using central versus site-specific prediction equations and of using the FEV_1/FEV_6 in place of the FEV_1/FVC.

Methods and Materials

The design of BOLD is described in detail elsewhere [15;16] and only summarized here. Participating entities in the BOLD Collaborative Research Group are listed as an Addendum.

<u>Population</u>

Participating sites were expected to recruit population-based samples of at least 600 non-institutionalized adults aged ≥ 40. We report data from the first 14 BOLD sites (Table 1), consisting of 10,001 individuals (93% of all responders) with acceptable post bronchodilator spirometry. Each site obtained approval from local ethical committees and written informed consent from each participant.

<u>Questionnaires</u>

Questionnaires were administered by trained and certified staff and covered respiratory symptoms, smoking history, respiratory diagnoses, and comorbidities. We defined pack-years of cigarette smoking exposure as average number of packs smoked per day (20 cigarettes = 1 pack) multiplied by number of years smoked. Never-smoking was defined as fewer than 20 packs of cigarettes in a lifetime.

Site-specific prediction equations were developed using never-smokers who had <u>never</u> been told by a health care provider that they had emphysema, COPD, or tuberculosis, and did not report a <u>current</u> diagnosis of asthma or chronic bronchitis. We were unable to restrict to asymptomatic never-smokers due to the extremely small numbers of (particularly male) never-smokers at some sites.

<u>Height and Weight</u>

We measured height (to the nearest centimeter) with the participant standing on a firm, level surface that was perpendicular to the vertical board of the height measurement device (ideally a wall-mounted stadiometer). Participants were instructed to remove their shoes and stand erect with feet flat on the floor, heels together, and head in the horizontal (Frankfort) plane.

Sites used calibrated scales (preferably balance beam or digital) to measure weight to the nearest 0.1 kg. Participants were instructed to remove shoes, hats, coats, and heavy items in their pockets in order to be weighed in light indoor clothing.

Body mass index (BMI) was computed as weight over height-squared and expressed in units of kg/m^2.

Spirometry

Lung function data were collected using the ndd EasyOne Spirometer (ndd Medical Technologies, Zurich, Switzerland), which was chosen for its portability and level of accuracy.[17] Lung function was measured before and 15 minutes after administration of $200\mu g$ of albuterol/salbutamol. Spirometry measures reported here include the FEV_1, FEV_6, and FVC, as well as the FEV_1/FVC and FEV_1/FEV_6 ratios. Percent predicted FEV_1 (FEV_1,%), though not reported separately, was used to stage COPD.[3]

All spirograms were reviewed by the BOLD Pulmonary Function Reading Center and assigned an overall quality score based on standardized criteria.[18] Local-site spirometry technicians were trained and certified, and received regular quality-control feedback during data collection. Usable spirometry was defined as two or more acceptable blows, with FEV_1 and FVC repeatability within 200ml. Acceptable maneuvers were defined as those with a rapid start (back-extrapolated volume <150ml or <5% of the FVC), lack of a cough during the first second, and a small end-of-test volume (<40ml during the final second). The calibration of all spirometers was

verified to be accurate within 3.0% using a 3.00 liter syringe at the beginning of each day of testing. Biological controls were not used.

Definition of COPD

BOLD uses the GOLD criteria for defining and staging COPD,[3] which are consistent with the 2004 ATS/ERS criteria [4] and define COPD as a post-bronchodilator $FEV_1/FVC < 0.70$. The $FEV_1,\%$ is used to further stage disease ($FEV_1,\% \geq 80$ = stage 1; $50 \leq FEV_1,\% < 80$ = stage 2; $30 \leq FEV_1,\% < 50$ = stage 3; $FEV_1,\% < 30$ = stage 4). BOLD also uses the prediction equations for Caucasian adult men and women derived from the third United States National Health and Nutrition Examination Survey (NHANES III)[19] as its primary reference equations for all participants, although this paper also examines the impact of using equations derived from Norway's Hordaland County Respiratory Health Study [20] as well as site-specific prediction equations in place of the NHANES equations.

In addition, we assess the impact of restricting COPD to GOLD stages 2 or higher, and of using the LLN of the FEV_1/FVC, and the FEV_1 in place of the fixed ratio and the $FEV_1,\% < 80$ criteria, in the GOLD definitions. Finally, we examine the impact of using FEV_1/FEV_6 in place of FEV_1/FVC in our definitions. Table 2 summarizes the various definitions of COPD assessed in this manuscript.

Although the text focuses on post-bronchodilator spirometry, the results of comparable analyses based on pre-bronchodilator data are included in the online supplement.

Analysis

To provide comparability with earlier reports,[16] the site-specific prevalences presented in Figure 1 are population-based estimates reflecting sampling designs used at each site. For all other analyses, data are pooled across sites and presented as unweighted prevalences with

standard errors accounting only for correlations within site and, where applicable, for clustering in the sampling plan. Comparisons of the prevalence estimates in Figures 1-3 and in Table 4 were computed using McNemar's test.

A desired characteristic of any prevalence estimator is that it gives comparable estimates in different populations after adjusting for known confounders. In order to compare the residual site-to-site variability associated with our various prevalence estimators, we report the Wald statistic for the "site" effect as derived from logistic regression models that adjusted for age (40-49, 50-59, 60-69, 70+ years), gender, pack-years of cigarette smoking (never-smokers, 0-9, 10-19, 20+), body mass index (<20, 20-25, 25-30, 30-35, >35 kg/m^2), years worked in a dusty job (0, 1-9, 10+) and interactions of gender with both age and pack-years. We also report Wald statistics for testing the significance of age in selected regression models. Where appropriate, we tested heterogeneity of age effects across strata using appropriate interaction terms. Under the null hypothesis of no effect, the Wald statistic will have an F-distribution with an expected value equal to one, and higher values indicate greater heterogeneity across subgroups. All Wald tests are adjusted for clustering in the sampling plan.

All analyses were conducted using Stata, version 9.2 [Stata Corp., College Station, TX].

Results

Participants exhibited marked differences in smoking patterns across sites and between genders within sites (Table 3). BOLD sites also differ markedly in prevalences of occupational and other potential COPD risk factors.[16]

Use of the fixed ratio criterion (GOLD Stage 1 and higher) produced overall population prevalence estimates that, for each site, were significantly greater than those for each of the other estimators (all but one p<0.0001).   The fixed ratio estimates were generally 5-11 percentage

points higher than those for GOLD stages 2-4 (Figure 1a). The LLN (FEV$_1$/FVC) criterion

produced estimates that tended to be intermediate to these two GOLD-based definitions, though

generally closer to the GOLD stages 2-4 criterion than to the fixed-ratio criterion. The added

requirement of an FEV$_1$,%<80 *and* an FEV1/FVC ratio below the LLN resulted in estimates that

were 1-3 percentage points lower than estimates for GOLD stages 2-4. Finally, use of

FEV$_1$<LLN in place of FEV$_1$,%<80 in this latter definition further reduced estimates (though

generally by less than one percentage point). These patterns were generally consistent across

sites.

Regardless of the definition used, we observed sizable site-to-site variation in prevalence

estimates (Figure 1b). After adjusting for potential confounders, site-to-site variance in COPD

prevalence (as measured by the Wald statistic) ranged from 7.1 to 8.6 and was lowest (7.1 and

7.3 respectively) using the 'LLN(FEV$_1$/FVC) & LLN(FEV$_1$)' and 'LLN(FEV$_1$/FVC) &

FEV$_1$,%<80' criteria, respectively. These Wald statistics all indicated highly statistically

significant (p<.0001) residual site-to-site variability in prevalence estimates.

All prevalences reported in Figure 1 were lower than they would have been had we based them

on *pre-bronchodilator* measurements (see online supplement). For the fixed-ratio criterion,

absolute declines between pre- and post-bronchodilator values ranged from 1-11 percentage

points across centers, while using GOLD stages 2-4 instead of the fixed-ratio criterion led to a

decline in prevalence ranging from 1-6 percentage points across centers. On a *relative* basis,

prevalence estimates declined between 25-29% (depending on the definition used) across the five

measures in going from pre- to post-bronchodilator measurements.

The prevalence of "COPD" per the fixed ratio criterion increased sharply with age even among

healthy never-smokers (Figure 2), a population in which COPD is expected to be rare. By

contrast, for the other measures we observed much more muted increases with age and, except for the LLN(ratio) criterion for the lowest age group (p=0.14), the fixed ratio prevalence estimates were all significantly greater than those for each of the other estimators (p<0.0001). These age-related increases in prevalence were lowest for the 'LLN(FEV_1/FVC) & LLN(FEV_1)' and 'LLN(FEV_1/FVC) & FEV_1,%<80' criteria, for which the age-specific prevalence estimates varied from 2% among forty year-olds, to 4-5% among those aged 70 and older. We observed the same general patterns within each site (data not shown).

The Wald statistic for testing for age effects in Figure 2 dropped from a high of 62.6 for the fixed- ratio criterion, to 24.5 for GOLD stages 2-4, to about 6.6 for the two analogues of these criteria in which FEV_1/FVC<0.7 is replaced by FEV_1/FVC<LLN, and to 3.4 for the 'LLN(FEV_1/FVC) & LLN(FEV_1)' criteria. All were statistically significant. We found modest evidence of a statistically significant gender-age interaction using the fixed ratio criterion (Wald=3.1, p=.027) and no evidence of statistically significant gender-age interactions in these healthy never-smokers using any of the other prevalence estimators.

When we assessed site differences in the group of healthy, never-smoking individuals, we observed smaller site differences for the GOLD stages 2-4 criterion (Wald statistic = 1.6) than for the LLN (FEV_1/FVC) criterion (Wald statistic = 2.9, though once again the smallest site differences were seen for the 'LLN(FEV_1/FVC) & LLN(FEV_1)' and 'LLN(FEV_1/FVC) & FEV_1,%<80' criteria (Wald = 0.9-1.1). Indeed, for both of these latter criteria, the site differences did not come close to reaching statistical significance (p>0.35), whereas for the other three criteria the p-values were all less than 0.07.

Figure 3 illustrates the impact on prevalence of using a single common prediction equation (the U.S. NHANES III Caucasian equations or the Hordaland County Respiratory Health Study

equations) versus site-specific prediction equations. For both men and women, the estimated GOLD stage 2-4 prevalences were higher (by 2-3 percentage points overall, p<0.0001) when using common reference equations for all sites (NHANES and Hordaland County) than when using local prediction equations. The NHANES and Hordaland County prevalence estimates were generally similar, although they differed significantly overall and for the oldest age group. The Wald statistic for site differences computed from the site-specific equations (4.7) was less than the Wald statistic for the NHANES (9.5) and Hordaland County (8.4) equations, though all were highly statistically significant (p<.0001). We observed similar patterns when we replaced the GOLD stage 2-4 criterion with the LLN(ratio) & FEV_1<80% criterion (data not shown), although the Wald statistics were closer (6.5 vs. 7.3 and 8.9).

Finally, the use of the FEV_1/FEV_6 in place of the FEV_1/FVC when using the 'LLN(FEV_1/FVC) and FEV_1,%<80' criterion had little clinically relevant impact on prevalence estimates, whether computed overall, by age or pack-years categories, or by site (Table 4). When we did observe statistically significant differences, the prevalences were generally smaller for the FEV_1/FEV_6 -based criterion.

Discussion

This analysis of data from the BOLD study confirmed previously reported limitations associated with the use of the fixed-ratio criterion to define COPD. Adjusting the FEV_1/FVC for normative aging effects appears to reduce the rate of false-positive diagnoses that has been reported for older individuals,[2;7-10] and the added requirement of a low FEV_1 further reduced the age-related increases in COPD prevalence seen among healthy never-smokers.

A strength of this analysis is that data were gathered using a standardized approach from a wide range of populations, with close attention paid to spirometry quality control. The qualitative

similarity of results across sites (Figure 1a) provides strong evidence of the robustness of our findings. The wide variation in characteristics of BOLD sites enabled us to use site-to-site variation in prevalence (assessed using the Wald statistic) as a convenient metric for comparing alternative measures of COPD prevalence, since a desired characteristic of any prevalence estimate is that it yield comparable estimates in different populations after adjusting for known risk factors.

An obvious limitation of this analysis is the lack of a gold standard against which to assess our putative definitions of COPD (indeed, a more accurate descriptor of what we are measuring may simply be chronic airflow limitation). Nonetheless it *is* possible to evaluate how alternative definitions perform in individuals who have a low *a priori* probability of disease. Our results confirm previous reports that the fixed-ratio criterion lacks specificity and, as age increases, increasingly misclassifies apparently healthy never-smokers as having COPD. [2;7-10;12] This pattern of (apparent) misclassification with increasing age was greatly muted by adding the requirement that the FEV_1,% be below a defined threshold, or by replacing the fixed-ratio criterion with a criterion that the FEV_1/FVC be below the LLN (Figure 2). However, only the method requiring *both* an FEV_1/FVC below the LLN *and* a low FEV_1 (measured as either $FEV_1 < LLN$ or FEV_1,% < 80) largely eliminated this age-related increase.

The upward trend that still persists in Figure 2 even with our "best" definitions of COPD may reflect the fact that our "healthy" never-smokers did include some individuals with symptoms. As noted below, this was a pragmatic decision due to the limited number of never-smokers at some sites. The fact that the NHANES prediction equations were fit to a cohort whose upper age limit was 80 years also may create an upward bias for very old individuals that helps explain the upward drift in Figure 2. However fewer than 4% of the BOLD cohort were aged 80 or older; in

addition, the NHANES prediction equations for FEV_1 include an age-squared term and so allow for accelerated aging effects.

Notably, the recent American Thoracic Society and European Respiratory Society Joint Statement [21] recommends using the LLN of the FEV_1/FVC in place of the fixed-ratio criterion to diagnose airflow obstruction. (A recent paper by Swanney et al [12]., albeit using pre-bronchodilator spirometry, also supports this recommendation.) Use of both an FEV_1/FVC below the LLN and a low FEV_1 was consistently associated with low site-to-site and age-related variation relative to other measures, after adjusting for known risk factors. Assuming that variability about the prediction equations is stable, using the LLN as a threshold for defining low FEV_1 should produce less misclassification, [22] although in practice these two measures performed similarly.

The results of our study also add to the evidence suggesting that, without both a low FEV_1/FVC and a low FEV_1, confidence is low that a true lung function abnormality (or airway disease) exists. The current GOLD stage I classification was based solely on expert opinion, not on evidence of airway disease or subsequent rapid loss of lung function. Patients with GOLD stage 1 do not have reduced exercise capacity [23]. Among Lung Health Study participants, a rapid fall in FEV_1 was not seen when baseline FEV_1 was above 70% predicted [24].

Apart from the fixed-ratio criterion, the competing measures we evaluated all require use of prediction equations. One of the purported benefits of the fixed-ratio criterion is that it does not rely on such equations. However, as Swanney et al [12] note, this easy-to-apply definition is only valid around age 50. In addition, the fixed-ratio criterion is not necessarily easier to use in practice, since even inexpensive pocket spirometers include a microprocessor that calculates the appropriate LLN for FEV_1/FVC, FEV_1/FEV_6, and FEV_1. Lastly, even GOLD relies on prediction

equations to stage disease, so any advantage of the fixed ratio in terms of its simplicity disappears as soon as one looks at clinically relevant impairment (nominally GOLD stage 2 or higher). The only way to overcome the limitations of the current fixed ratio criterion while still avoiding the need for prediction equations would be to establish a series of separate fixed ratio cutpoints for different ages.

The question then arises, what is a suitable prediction equation for any given population, and what if normative prediction equations do not exist for that population? While the documented variability in lung function that exists among "healthy" never-smokers in different racial groups may reflect, at least in part, true genetic differences in these populations, it also may represent the cumulative effect of environmental exposures, including childhood factors. For this reason, BOLD chose to use a single set of gender-specific prediction equations for all subjects in all sites. We chose the U.S. NHANES III equations for Caucasian adults because they were derived from a large study conducted in a diverse population with rigorous attention to quality control. We observed similar prevalence estimates using equations derived from Norway's Hordaland County Respiratory Health Study [20] in place of the NHANES equations.

The PLATINO study, conducted in five Latin American countries using methods similar to BOLD's, used site-specific prediction equations.[25] In BOLD, the use of local prediction equations led to prevalence estimates 2-3 percentage points lower, on average, than those based on a single, common equation. Whether this means that the BOLD prevalence estimates overestimate the "true" estimate, or the local equations underestimate it, we cannot say, but on balance we prefer to maintain the site-to-site variation and see if it can be explained by other risk factors. Because our local equations were fitted to individuals aged 40 and older, while the NHANES equations were fitted to adults aged 18 and older, the former may better describe the

accelerated aging that is known to occur in healthy adults. Also, we included symptomatic individuals in our prediction equations as long as they did not report diagnosed disease, whereas the HNANES equations required individuals to be asymptomatic. Since there can be large discrepancies between prediction equations based on individuals with and without major respiratory symptoms, [26] this also may help explain the somewhat lower prevalence estimates between the two approaches. One final consideration relating to the use of site-specific prediction equations, particularly if reliable normative equations for that population do not exist, is that the resulting estimates may be highly variable due to limited sample sizes. For instance, despite relatively large sample sizes from each site, the number of healthy never-smokers available to build our prediction equations was very limited in some sites due to extremely high rates of ever having smoked.

Considerable attention is now being paid to the use of the FEV_1/FEV_6 as an alternative to the FEV_1/FVC, particularly in older, less healthy populations for whom achievement of a high quality, reproducible FVC may be problematic.[27] Several studies have shown that the FEV_1/FEV_6, for which reliable reference equations exist [19], is a more reproducible measure than is the FEV_1/FVC, [28;29] and predicts subsequent lung function decline about as well as the FEV_1/FVC.[30] Our results (Table 4) show that using the FEV_1/FEV_6 in place of the FEV_1/FVC in our definition of 'LLN(FEV_1/FVC) & $FEV_1,\%<80$' yields very similar prevalence estimates, thus further supporting the use of this alternative measure in future studies of COPD prevalence. Once an obstructive lung disease has been diagnosed, however, changes in FEV_1 should be used to follow disease progression or treatment responses.

Finally, our observation that use of pre-bronchodilator spirometry results in consistently inflated estimates of chronic airflow obstruction, regardless of the definition used, further emphasizes the

need for using post-bronchodilator spirometry to classify COPD.[14] Our findings that prevalence estimates dropped, on average, about 25% when using post-bronchodilator spirometry is generally consistent with other reports.[8;10;13] Although we recognize that well-assessed, normal pre-bronchodilator spirometry has high negative predictive value even in the absence of post-bronchodilator testing, its use is associated with the more serious risk of increased false-positive diagnoses.

In summary, data from the BOLD study confirm previous reports of misclassification using the fixed-ratio criterion to measure COPD. As an alternative, we recommend a definition based on an FEV_1/FVC ratio less than the lower limit of normal, and an FEV_1 either less than 80% of predicted, or below the lower limit of normal. This modification of the current GOLD stage 2 severity threshold appears to better account for known aging effects in healthy never-smokers. While this new definition will likely miss many individuals with mild COPD, it should capture most individuals with clinically significant disease, while minimizing the risk of false positive diagnoses. Finally, substitution of the FEV_1/ FEV_6 in place of the FEV_1/FVC in this definition appears to yield similar prevalence estimates and, based on previous reports, may be a more reproducible and practical measure.

Acknowledgments

Legends for Figures

Figure 1. Comparison of COPD prevalence for various alternative definitions of COPD for participants in the BOLD study. Top panel contrasts prevalences for each definition within site, while bottom panel contrasts site differences within each definition. Significant site-site variability persisted for each of the estimates even after adjusting for age, gender, pack-years of cigarette smoking, BMI, years worked in a dusty job, and interactions of gender with both age and pack-years in logistic regression models (Wald statistics ranging from 7.1 to 8.6, all p-values<0.0001). Prevalences based on the fixed ratio are significantly higher than for all other estimators at each site (all p-values <0.001)
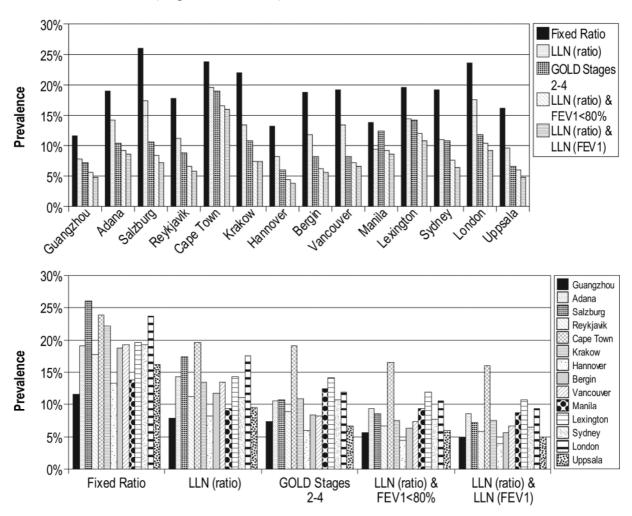
Figure 2. Prevalence of "COPD" among healthy never-smokers in the BOLD study (data from all sites combined). Wald statistics for comparing the four age groups (and adjusted for site, gender, BMI, and years worked in a dusty job) ranged from 62.6 for the fixed-ratio criterion, to 24.5 for GOLD stages 2-4, to about 6.6 for the two analogues of these criteria in which $FEV_1/FVC < 0.7$ is replaced by $FEV_1/FVC < LLN$, to 3.4 for the '$LLN(FEV_1/FVC)$ & $LLN(FEV_1)$' criteria. All were statistically significant. Significant gender-age interactions were observed only for the fixed ratio criterion. Except for the LLN(ratio) criterion for the lowest age group (p=0.14), the fixed ratio prevalence estimates were all significantly greater than those for each of the other estimators (p<0.0001).
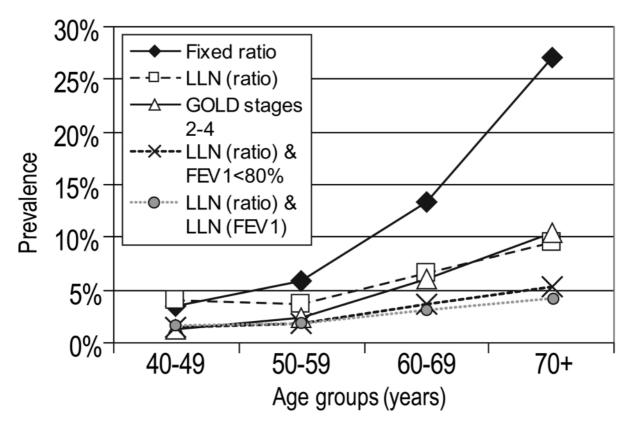

Figure 3. Prevalence of GOLD stages 2-4 COPD in the BOLD study using NHANES III and site-specific prediction equations (data from all sites combined). Site differences were

greater when using the NHANES equations (Wald statistic = 9.5, p<0.0001) or Hordalund county equations (Wald statistic = 8.4, p<0.0001) than when using the site-specific equations (Wald statistic = 4.7, p<0.0001). The estimated prevalences were significantly higher (p<0.0001 for all age groups) when using common reference equations for all sites (NHANES and Hordaland County) than when using local prediction equations. The NHANES and Hordaland County prevalence estimates differed significantly overall and for the oldest age group.
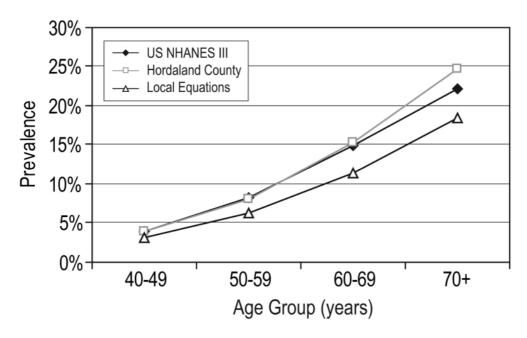
Table 1. Summary of BOLD Sites Included in this Analysis

Site	Description of Target Population	Sampling Design	# Respondents[1]	Response Rate[2]	Cooperation Rate[3]
Guangzhou, China	Residents of the Number Two Community Center of the Liwan district of the city of Guangzhou.	Stratified random sample	602	87%	87%
Adana, Turkey	Residents of all districts (urban and rural) that make up the Adana province.	Stratified cluster sample	875	82%	85%
Salzburg, Austria	Residents of Salzburg County.	Stratified random sample	1349	65%	67%
CapeTown, South Africa	Residents of the towns of Ravensmead and Uitsig, both located in the general Tygerberg area.	Cluster sample[4]	896	63%	68%
Reykjavik, Iceland	All age-eligible Icelandic citizens living in Reykjavik and surrounding suburbs.	Simple random sample	758	81%	84%
Hanover, Germany	Residents of Hannover city and four of the twenty town councils that make up the rest of Hannover Region as their target population.	Stratified random sample	713	59%	61%
Krakow, Poland	Residents of the Chrzanów and Proszowice regions within the Malopolska district of Poland.	Stratified random sample	603	78%	79%
Bergen, Norway	Residents of the city of Bergen, Norway.	Stratified random sample	707	68%	71%
Vancouver, Canada	Residents of the Vancouver Health Service Delivery Area of Vancouver, Canada.	Random digit dialing	856	26%	51%
Lexington, USA	Residents of the 5th congressional district of Kentucky.	Random digit dialing	563	14%	27%
Manila, Philippines	Residents of one of the 6 districts that make up the city of Manila.*	Cluster sample	918	58%	58%
Sydney, Australia	Residents of the federal electorates, Kingsford Smith and Barton, which make up the southern beachside suburbs of Sydney.	Stratified random sample	585	25%	33%
London, England	The patients served by 3 general practitioners from west London, which is broadly representative of the Hammersmith & Fulham neighborhood of west London.	Stratified random sample	697	17%	37%
Uppsala, Sweden	Residents of the city of Uppsala.	Stratified random sample	588	61%	63%

1 Participants with core questionnaire and any post-bronchodilator spirometry.
2 Denominator includes persons of unknown eligibility status who could not be contacted. Only known ineligible participants were excluded.
3 Denominator includes only participants who were contacted and eligible.

Table 2. Definitions of COPD Used in this Analysis

Threshold for COPD	Formula
Fixed Ratio	$FEV_1/FVC < 0.7$ (GOLD stage 1 or higher)
LLN (FEV_1/FVC)	$FEV_1/FVC < 5^{th}$ percentile (predicted $FEV_1/FVC - 1.645*$ SD)
GOLD Stages 2-4	$FEV_1/FVC < 0.7$ <u>and</u> FEV_1,% <80
LLN (FEV_1/FVC) and FEV_1,%<80	$FEV_1/FVC < 5^{th}$ percentile <u>and</u> FEV_1,% <80
LLN (FEV_1/FVC) and LLN (FEV_1)	$FEV_1/FVC < 5^{th}$ percentile <u>and</u> $FEV_1 < 5^{th}$ percentile
LLN (FEV_1/ FEV_6) and FEV_1,%<80	$FEV_1/ FEV_6 < 5^{th}$ percentile <u>and</u> FEV_1,% <80

For all but the fixed ratio, alternative versions may be computed based on choice of prediction equation. This table just describes general formulas.

Table 3. Characteristics[1] of BOLD participants included in analysis

	Guangzhou	Adana	Salzburg	Reykjavik	Cape Town	Krakow	Hannover	Bergen	Vancouver	Manila	Lexington	Sydney	London	Uppsala
	n=473	n=806	n=1258	n=757	n=847	n=526	n=683	n=658	n=827	n=893	n=508	n=541	N=677	N=547
Age (yrs)														
40-49	43.1%	42.1%	29.2%	35.9%	39.6%	37.5%	26.9%	25.1%	35.2%	47.3%	26.6%	29.0%	28.2%	22.3%
50-59	26.4%	30.6%	28.5%	29.0%	31.1%	27.4%	26.8%	28.1%	31.9%	31.0%	37.4%	27.2%	26.1%	34.2%
60-69	20.5%	17.5%	26.0%	18.0%	19.7%	20.3%	29.7%	22.0%	17.3%	14.3%	25.6%	21.4%	27.8%	26.3%
70+	9.9%	9.8%	16.3%	17.1%	9.7%	14.8%	16.5%	24.8%	15.6%	7.4%	10.4%	22.4%	17.9%	17.2%
Male	N=236	N=389	N=685	N=403	N=313	N=266	N=349	N=324	N=344	N=378	N=206	N=265	N=323	N=283
never smoker	18.6%	19.5%	40.2%	38.5%	15.6%	21.1%	30.7%	32.1%	42.2%	16.6%	30.1%	44.5%	32.5%	37.5%
0-10 packyears	11.9%	13.6%	12.3%	16.1%	30.5%	11.3%	16.6%	16.4%	18.3%	24.3%	7.8%	18.9%	18.6%	23.0%
10-20 packyears	21.2%	11.1%	13.1%	16.1%	24.4%	15.8%	13.2%	19.1%	11.0%	23.5%	6.3%	10.2%	11.1%	14.5%
20+ packyears	48.3%	55.8%	34.5%	29.0%	28.9%	51.5%	39.5%	32.4%	28.2%	35.4%	55.8%	26.4%	37.8%	25.1%
Height (cm)	166±6.2	168±6.6	176±6.6	180±6.5	169±7.3	173±6.3	175±7.6	178±7.1	175±7.3	163±6.7	176±7.3	172±7.5	175±7.1	178±7.0
BMI (kg/m^2)	23.3±3.1	27.7±4.6	26.7±3.5	28.3±4.4	24.7±5.7	27.4±3.9	27.7±4.0	26.9±3.8	27.1±4.0	24.4±4.8	30.3±4.9	28.0±4.4	27.5±4.5	27.2±3.8
Female	N=237	N=417	N=573	N=354	N=532	N=260	N=334	N=334	N=483	N=515	N=302	N=276	N=354	N=264
never smoker	93.7%	69.5%	55.8%	39.7%	42.1%	56.5%	50.0%	42.2%	51.8%	69.5%	46.4%	53.6%	43.2%	48.1%
0-10 packyears	2.1%	15.8%	16.8%	25.8%	27.8%	15.0%	13.8%	19.8%	18.2%	19.0%	11.3%	18.1%	21.8%	20.1%
10-20 packyears	1.7%	5.8%	10.1%	13.3%	15.8%	13.1%	12.9%	18.9%	10.6%	5.2%	8.3%	8.7%	9.6%	12.5%
20+ packyears	2.5%	8.9%	17.3%	21.2%	14.3%	14.6%	23.4%	19.2%	19.3%	5.2%	34.1%	19.6%	25.4%	18.9%
Height (cm)	155±6.5	154±6.0	163±6.1	166±6.3	157±6.7	161±5.8	163±7.0	164±6.4	161±7.4	151±5.9	161±5.8	159±6.6	162±7.3	164±6.2
BMI (kg/m^2)	23.4±3.4	31.5±5.3	26.1±4.9	27.5±5.5	29.8±7.7	28.0±5.3	26.8±5.1	26.2±4.8	26.4±5.9	25.3±4.6	31.1±7.8	27.9±5.7	26.7±5.3	26.8±4.9

data represent percent of individuals in each group (age and smoking categories) or mean ± standard deviation

Table 4. Prevalence of COPD[1] Computed Using FEV_1/FEV_6 in place of the FEV_1/FVC

	N	LLN(FEV_1/FEV_6) and FEV_1,%<80	LLN(FEV_1/FVC) and FEV_1,%<80	P-value[2]
		COPD Criterion		
Overall	10001	8.1%	8.3%	.10
Age (years)				
40-49	3381	4.1%	4.1%	.84
50-59	2968	6.5%	7.2%	<.01
60-69	2172	11.2%	11.5%	.57
70+	1480	15.9%	15.7%	.73
Packyears				
never smoker	4291	4.0%	4.0%	1.0
0-10	1777	5.0%	4.8%	.66
10-20	1270	9.6%	9.6%	1.0
20+	2654	16.3%	17.1%	.01
Gender				
Male	4766	9.1%	9.1%	.90
Female	5235	7.2%	7.6%	.01
Site				
Guangzhou	473	6.8%	5.5%	.03
Adana	806	8.3%	9.3%	.04
Salzburg	1258	7.2%	7.5%	.34
CapeTown	847	17.1%	16.3%	.25
Reykjavik	757	5.9%	6.5%	.29
Hanover	683	5.1%	5.0%	1.0
Krakow	526	7.6%	7.4%	1.0
Bergen	658	5.9%	6.8%	.11
Vancouver	827	5.7%	6.4%	.07
Lexington	508	11.2%	12.4%	.15
Manila	893	9.2%	8.4%	.09
Sydney	541	8.1%	7.6%	.38
London	677	9.2%	10.2%	.02
Uppsala	547	4.9%	5.7%	.12

[1] Data are univariate classifications and not adjusted for other terms in the table.

[2] two-tailed exact p-values based on McNemar's test for comparing prevalences within each subgroup.

Reference List

1. Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. *Eur Respir J* 2006;28(3):523-532.

2. Celli BR, Halbert RJ, Isonaka S, Schau B. Population impact of different definitions of airway obstruction. *Eur Respir J* 2003;22(2):268-273.

3. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van WC, Zielinski J. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 9-15-2007;176:532-555.

4. Celli BR, MacNee W. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004;23:932-946.

5. Viegi G, Pedreschi M, Pistelli F, Di Pede F, Baldacci S, Carrozzi L, Giuntini C. Prevalence of airways obstruction in a general population: European Respiratory Society vs American Thoracic Society definition. *Chest* 2000;117:339S-345S.

6. Lindberg A, Eriksson B, Larsson LG, Ronmark E, Sandstrom T, Lundback B. Seven-year cumulative incidence of COPD in an age-stratified general population sample. *Chest* 2006;129:879-885.

7. 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(5):1117-1122.

8. Shirtcliffe P, Weatherall M, Marsh S, Travers J, Hansell A, McNaughton A, Aldington S, Muellerova H, Beasley R. COPD prevalence in a random population survey: a matter of definition. *Eur Respir J* 2007;30:232-239.

9. Hnizdo E, Glindmeyer HW, Petsonk EL, Enright P, Buist AS. Case definitions for chronic obstructive pulmonary disease. *COPD* 2006;3(2):95-100.

10. Perez-Padilla R, Hallal PC, Vazquez-Garcia JC, Muino A, Maquez M, Lopez MV, de Oca M, Talamo C, Valdivia G, Pertuze J, Jardim J, Menezes AMB. Impact of bronchodilator use on the prevalence of COPD in population-based samples. *Journal of Chronic Obstructive Pulmonary Disease* 2007;4:113-120.

11. Cerveri I, Corsico AG, Accordini S, Niniano R, Ansaldo E, Anto JM, Kunzli N, Janson C, Sunyer J, Jarvis D, Svanes C, Gislason T, Heinrich J, Schouten JP, Wjst M, Burney P, de MR. Underestimation of airflow obstruction among young adults using FEV1/FVC<70% as a fixed cut-off: a longitudinal evaluation of clinical and functional outcomes. *Thorax* 5-20-2008.

12. Swanney MP, Ruppel G, Enright PL, Pedersen OF, Crapo RO, Miller MR, Jensen RL, Falaschetti E, Schouten JP, Hankinson JL, Stocks J, Quanjer PH. Using the lower limit of

normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction. *Thorax* 9-11-2008.

13. Johannessen A, Omenaas ER, Bakke PS, Gulsvik A. Implications of reversibility testing on prevalence and risk factors for chronic obstructive pulmonary disease: a community study. *Thorax* 2005;60(10):842-847.

14. Sterk PJ. Let's not forget: the GOLD criteria for COPD are based on post-bronchodilator FEV1. *Eur Respir J* 2004;23(4):497-498.

15. Buist AS, Vollmer WM, Sullivan S, Weiss KB, Lee TA, Menezes AMB, Crapo RO, Jensen RL, Burney PGJ. The Burden of Obstructive Lung Disease Initiative (BOLD): Rationale and design. *Journal of Chronic Obstructive Pulmonary Disease* 2005;2:277-283.

16. Buist AS, McBurnie MA, Vollmer WM, Gillespie S, Burney P, Mannino DM, Menezes AM, Sullivan SD, Lee TA, Weiss KB, Jensen RL, Marks GB, Gulsvik A, Nizankowska-Mogilnicka E. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet* 9-1-2007;370:741-750.

17. Perez-Padilla R, Vazquez-Garcia JC, Marquez MN, Jardim JR, Pertuze J, Lisboa C, Muino A, Lopez MV, Talamo C, de Oca MM, Valdivia G, Menezes AM. The long-term stability of portable spirometers used in a multinational study of the prevalence of chronic obstructive pulmonary disease. *Respir Care* 2006;51(10):1167-1171.

18. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J. Standardisation of spirometry. *Eur Respir J* 2005;26:319-338.

19. 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(1):179-187.

20. Johannessen A, Lehmann S, Omenaas ER, Eide GE, Bakke PS, Gulsvik A. Post-bronchodilator spirometry reference values in adults and implications for disease management. *Am J Respir Crit Care Med* 6-15-2006;173:1316-1325.

21. Pellegrino R, Brusasco V, Viegi G, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J, Jensen R, Johnson DC, MacIntyre N, McKay R, Miller MR, Navajas D, Pedersen OF, Wanger J. Definition of COPD: based on evidence or opinion? *Eur Respir J* 2008;31:681-682.

22. Stanojevic S, Wade A, Stocks J, Hankinson J, Coates AL, Pan H, Rosenthal M, Corey M, Lebecque P, Cole TJ. Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med* 2-1-2008;177:253-260.

23. Pinto-Plata VM, Celli-Cruz RA, Vassaux C, Torre-Bouscoulet L, Mendes A, Rassulo J, Celli BR. Differences in cardiopulmonary exercise test results by American Thoracic

Society/European Respiratory Society-Global Initiative for Chronic Obstructive Lung Disease stage categories and gender. *Chest* 2007;132:1204-1211.

24. Scanlon PD, Connett JE, Waller LA, Altose MD, Bailey WC, Buist AS. Smoking cessation and lung function in mild-to-moderate chronic obstructive pulmonary disease. The Lung Health Study. *Am J Respir Crit Care Med* 2000;161:381-390.

25. Menezes AM, Perez-Padilla R, Jardim JR, Muino A, Lopez MV, Valdivia G, Montes de Oca M, Talamo C, Hallal PC, Victora CG. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. *Lancet* 11-26-2005;366(9500):1875-1881.

26. Johannessen A, Omenaas ER, Eide GE, Bakke PS, Gulsvik A. Feasible and simple exclusion criteria for pulmonary reference populations. *Thorax* 2007;62:792-798.

27. Pezzoli L, Giardini G, Consonni S, Dallera I, Bilotta C, Ferrario G, Cristina SM, Annoni G, Vergani C. Quality of spirometric performance in older people. *Age Ageing* 2003;32:43-46.

28. Swanney MP, Jensen RL, Crichton DA, Beckert LE, Cardno LA, Crapo RO. FEV(6) 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.

29. Rosa FW, Perez-Padilla R, Camelier A, Nascimento OA, Menezes AM, Jardim JR. Efficacy of the FEV1/FEV6 ratio compared to the FEV1/FVC ratio for the diagnosis of airway obstruction in subjects aged 40 years or over. *Braz J Med Biol Res* 2007;40:1615-1621.

30. Enright PL, Connett JE, Bailey WC. The FEV1/FEV6 predicts lung function decline in adult smokers. *Respiratory Medicine* 2002;96:444-449.