The Proportional Venn Diagram of Obstructive Lung Disease*

Two Approximations From the United States and the United Kingdom

Joan B. Soriano, MD; Kourtney J. Davis, PhD; Bobbie Coleman, BSc; George Visick, PhD; David Mannino, PhD; and Neil B. Pride, MD

Study objectives: The nonproportional Venn diagram of obstructive lung disease (OLD) produced for the 1995 American Thoracic Society guidelines has not been quantified. We aim to quantify the proportion of the general population with OLD and the intersections of physician-diagnosed asthma, chronic bronchitis, and emphysema in the United States and the United Kingdom, and to examine the relationship to obstructive spirometry.

Design and participants: We analyzed data from the US National Health and Nutrition Examination (NHANES) III survey (1988 to 1994) and the UK General Practice Research Database for the year 1998.

Results: The areas of intersection among the three OLD conditions produced seven mutually exclusive disease groups. The asthma-only group was the largest proportion of OLD patients, accounting for 50.3% and 79.4% of all OLD patients in the United States and the United Kingdom, respectively, and decreased with increasing age. Overall, 17% and 19% of OLD patients in the United States and in the United Kingdom, respectively, reported more than one OLD condition, and this percentage increased with age. According to the spirometry data from NHANES III, only 37.4% of emphysema-only patients had objective airflow obstruction. The prevalence of airflow obstruction was significantly higher among participants with combinations of emphysema and chronic bronchitis (57.7%), with emphysema and asthma (51.9%), and with all three OLD diseases concomitantly (52.0%).

Conclusion: Concomitant diagnosis of asthma, chronic bronchitis, or emphysema is common among OLD patients from the general population, particularly in adults aged ≥ 50 years.

(CHEST 2003; 124:474–481)

Key words: asthma; COPD; epidemiology; obstructive lung disease; Venn diagram

Abbreviations: ATS = American Thoracic Society; GOLD = Global Initiative for Chronic Obstructive Lung Disease; GP = general practitioner; GPRD = UK General Practice Research Database; MEC = mobile examination center; NHANES = National Health and Nutrition Examination Survey; OLD = obstructive lung disease; OXMIS = Oxford Medical Information System; RCT = randomized controlled trials

Asthma, chronic bronchitis, and emphysema are conditions associated with obstructive lung disease (OLD), a leading cause of morbidity and mortality worldwide. Asthma is the most common chronic condition in children but can be diagnosed at any age. The World Health Report of 1998 estimated that 2.9 million adults die each year of COPD, a term that typically includes chronic bronchitis and emphysema, and ranked COPD as the fifth leading cause of mortality worldwide (after ischemic heart disease, cerebrovascular disease, acute lower respiratory infection, and tuberculosis) and as the fifth most prevalent disease (after iron deficiency anemia, neck and back disorders, goiter,

*From Worldwide Epidemiology (Drs. Soriano, Davis, and Visick, and Ms. Coleman), GlaxoSmithKline Research and Development, Research Triangle Park, NC; Air Pollution and Respiratory Health Branch (Dr. Mannino), Division of Environmental Hazards and Health Effects, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA; and Thoracic Medicine (Dr. Pride), National Heart and Lung Institute, Imperial College, London, UK.

This study was supported by GlaxoSmithKline Research and Development.

Manuscript received August 5, 2002; revision accepted February 18, 2003.

Reproduction of this article is prohibited without written permission from the American College of Chest Physicians (e-mail: permissions@chestnet.org).

Correspondence to: Joan B Soriano, MD, Worldwide Epidemiology, GlaxoSmithKline R&D; Greenford Rd, Greenford, Middlesex UB6 0HE, UK; e-mail: joan.b.soriano@gsk.com
and hypertensive disease) for 1997. International consensus definitions of asthma and COPD, including chronic bronchitis and emphysema, are available, but the diagnosis of OLDs is a medical challenge for respiratory specialists and primary care physicians, even with the availability of spirometry or a patient history of tobacco use.

This differential diagnosis is of interest, because asthma, chronic bronchitis, and emphysema are disorders with a different pathogenesis and prognosis, and current respiratory guidelines recommend very different management and drug treatment strategies for patients with asthma vs those with chronic bronchitis and emphysema. However, it is well-established that individual patients commonly share the traits of two or even three of these conditions. A graphic representation of this relationship was first presented as the nonproportional Venn diagram of OLD (Fig 1), which was included in the 1995 American Thoracic Society (ATS) COPD guidelines.6 To date, we are not aware of efforts to quantify this nonproportional Venn diagram. We aim to quantify the subpopulations of patients with OLD and the intersections of the diagnoses of asthma, chronic bronchitis, and emphysema in nationally representative databases in the United States and the United Kingdom.

**Materials and Methods**

**Data Sources**

We analyzed data from the US National Health and Nutrition Examination (NHANES) III survey, conducted from 1988 to 1994, and the UK General Practice Research Database (GPRD) for the year 1998.

The NHANES III was a cross-sectional, multistage probability sample (n = 33,994) that was representative of the total noninstitutionalized civilian population of the United States, and surveys were conducted from 1988 to 1994.7,8 Information about the medical history of respiratory symptoms and diagnoses of respiratory conditions were obtained by questionnaire interview for all participants aged ≥17 years and through a proxy interview with a household adult for those aged <17 years. Spirometry data were obtained for participants aged ≥8 years following the recommendations and standards of the ATS (n = 22,431). Analysis of NHANES III data incorporated sampling weights, in accordance with a probability sampling design, in order to obtain estimates of disease prevalence that reflect the total noninstitutionalized, civilian US population.

The GPRD is an automated medical record database of primary care covering a total population of >3 million inhabitants (approximately 5%) of the population in England and Wales.9,10 Several studies have demonstrated that the study base is a representative sample of the UK general population, with an almost identical gender and age structure as the one provided by the Office of Population Census and Statistics. General practitioners (GPs) from the participating surgeries enter all significant morbidity for each individual patient into the computer record. All diagnoses and procedures communicated to the GP as a result of a hospital or other specialist visit (ie, inpatient, outpatient, or in an accident and emergency unit) must be recorded when the GP is informed. The diagnosis, symptoms, procedure or investigation, referrals, and their outcome must be entered into the relevant sections of the medical record. In this report, estimates are presented for patients who attended their GP, were given a diagnosis of asthma or COPD, and were registered throughout 1998. Pulmonary function data are not systematically or routinely available.

**Case Definitions**

Definitions of OLD in each database differ. NHANES III was a survey, and the questionnaire part of the survey asked participants about self-reported physician-diagnosis of three conditions (ie, asthma, chronic bronchitis, and emphysema). A positive response to the single question “Has a doctor ever told you that you have emphysema?” was sufficient to define emphysema. But, for asthma and chronic bronchitis, positive responses to the following two questions were necessary to define a current

![Figure 1. Nonproportional Venn diagram of COPD showing subsets of patients with chronic bronchitis, emphysema, and asthma. The subsets comprising COPD are shaded. Subset areas are not proportional to the actual relative subset sizes. Asthma is by definition associated with reversible airflow obstruction, although in variant asthma special maneuvers may be necessary to make the obstruction evident. Patients with asthma whose airflow obstruction is completely reversible (ie, subset 9) are not considered to have COPD. Because in many cases it is virtually impossible to differentiate patients with asthma whose airflow obstruction does not remit completely from persons with chronic bronchitis and emphysema who have partially reversible airflow obstruction with airway hyperreactivity, patients with unremitting asthma are classified as having COPD (ie, subsets 6, 7, and 8). Chronic bronchitis and emphysema with airflow obstruction usually occur together (subset 5), and some patients may have asthma associated with these two disorders (ie, subset 8). Individuals with asthma who have been exposed to chronic irritation, as from cigarette smoke, may develop chronic productive cough, which is a feature of chronic bronchitis (ie, subset 6). Such patients often are referred to in the United States as having asthmatic bronchitis or the asthmatic form of COPD. Persons with chronic bronchitis and/or emphysema without airflow obstruction (ie, subsets 1, 2, and 11) are not classified as having COPD. Patients with airflow obstruction due to diseases with known etiology or specific pathology, such as cystic fibrosis or obliterative bronchiolitis (subset 10), are not included in this definition. Reprinted with permission from the American Thoracic Society.]

---

www.chestjournal.org

CHEST / 124 / 2 / AUGUST, 2003 475
condition: “Has the doctor ever told you that you have [asthma or chronic bronchitis]?” and “Do you still have it?” A participant could be classified as having all three conditions.

The GPRD system uses the Oxford Medical Information System (OXMIS) and READ coding system (the first three digits of the OXMIS number corresponds, in most cases, to the first three digits of the International Classification of Diseases, eighth revision, codes). Physician-diagnosed asthma, chronic bronchitis, and emphysema were defined as any individual who was labeled with one or more of the OXMIS/READ codes that were compatible with their respective diagnoses during calendar year 1998. This permitted the use of terms like COPD without reference to chronic bronchitis or emphysema. Acute bronchitis was excluded from the algorithm. The choice of GPRD OXMIS/READ codes was recorded by each individual GP without instructions or guidelines. Direct codes for COPD were defined as overlapping chronic bronchitis and emphysema in the Venn diagrams.

**Spirometry in NHANES III**

Lung function testing was conducted on participants aged ≥ 8 years by a trained technician in a mobile examination center (MEC). Testing also was conducted at the home of participants aged ≥ 60 years who were unwilling or unable to come to the MEC. Before testing, screening questions were asked to determine medical safety exclusions (ie, those who had undergone chest or abdominal surgery within 3 weeks or had experienced heart problems [myocardial infarction or heart attack, angina or chest pain, and congestive heart failure]) within 6 weeks before attending the MEC. Airflow obstruction was defined as stage 1 according to the following Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines: FEV1, < 80% of predicted (based on gender, age, and height); FEV1/FVC ratio, < 70%. As children aged ≤ 7 years did not have spirometry measurements recorded in NHANES III, estimates of airflow obstruction were weighted to the sample of participants aged ≥ 8 years for whom spirometry had been measured.

**Statistical Analysis**

The prevalence for each condition per gender and age group is presented for the United States and the United Kingdom separately. Values have been extrapolated to the total population in each country using the NHANES III sample weights for the United States and using the Office of National Statistics data for the United Kingdom. Each graphic representation in Figures 2 to 5 was presented for the United States and the United Kingdom separately. The prevalence of asthma, chronic bronchitis, and emphysema was associated with younger age.

**Results**

The descriptive characteristics of NHANES III and GPRD participants with OLD are presented in Table 1. Patients were stratified into seven mutually exclusive disease groups by age and sex for each study population. Patients with asthma, in whom diagnoses had been made with or without other OLD conditions, were younger than COPD patients. The combination of asthma with chronic bronchitis, but not with emphysema, was associated with younger age.

In the US NHANES III total population, the prevalence of current asthma was 5.5%, the prevalence of current chronic bronchitis was 3.2%, and the prevalence of ever having had emphysema was 1.5%. In the UK GPRD total population, the prevalence of current asthma was 2.3%, the prevalence of current chronic bronchitis was 0.5%, and the prevalence of current emphysema was 0.5% (Fig 2). The asthma-only group was the largest group of OLD patients, accounting for 4.3% and 2.2%, respectively, of the United States and United Kingdom general populations; that is, 50.3% and 79.4%, respectively, of all OLD patients. Seventeen percent of OLD patients in the United States had more than one OLD condition, and 2.8% had all three conditions simultaneously. The UK figures were 19.1% and 3.6%, respectively. The areas of intersection among the three OLD conditions differed by data source and country (Fig 2).

In the subpopulation of participants aged ≥ 50 years, the prevalence of current asthma was 5.1%, the prevalence of current chronic bronchitis was 5.8%, and the prevalence of ever having had emphysema was 1.5%.
sema was 5.0% in the US NHANES III survey. In the UK GPRD subpopulation of participants aged \( \geq 50 \) years, the prevalence of current asthma was 1.6%, the prevalence of current chronic bronchitis was 1.1%, and the prevalence of current emphysema was 1.1% (data not shown). The seven mutually exclusive disease groups of this proportional Venn diagram also can be displayed as stacked bars, for comparison by gender and age (Fig 3). The relative size of the asthma-only group decreased with increasing age, both in women and men in the United States and the United Kingdom. In the NHANES III data, probably because acute bronchitis could have been misclassified as chronic bronchitis in some cases via the self-reported questionnaire, combinations of chronic bronchitis with asthma or chronic bronchitis alone appeared at very young ages. Emphysema was reported consistently from age 50 years onward. Combinations of two of the three OLD conditions occurred among 21.2%, 31.4%, and 14.4%, in those patients in the age groups of 60 to 69 years, 70 to 79 years, and \( \geq 80 \) years, respectively.

By contrast, in the UK GPRD, emphysema and chronic bronchitis are virtually nonexistent diagnoses before 50 years of age. After age 50 years, the frequent diagnoses of chronic bronchitis and emphysema together as COPD gave little room for chronic bronchitis-only or emphysema-only diagnoses. Therefore, combinations of two of the three OLD conditions occurred among 41.5%, 58.8%, and 65.5% of patients in the age groups of 60 to 69 years, 70 to 79 years, and \( \geq 80 \) years, respectively.

Finally, according to the spirometry data of NHANES III, US OLD participants with chronic bronchitis or emphysema, with or without a concomitant diagnosis of asthma, differed widely regarding the prevalence of airflow obstruction among patients with a diagnosis of emphysema only, with only 37.4% of patients having airflow obstruction confirmed by spirometry. The prevalence of airflow obstruction was significantly higher among participants with combinations of emphysema and chronic bronchitis (57.7%), with emphysema and asthma (51.9%), and with all three OLD diseases concomitantly (52.0%). Among all NHANES III participants with airflow obstruction, accounting for 4.8% of the general population, 58.3% reported no diagnosis of any of the three OLD conditions (Fig 4). The patterns of airflow obstruction prevalence were confirmed when the analysis was restricted to participants aged \( \geq 50 \) years (Fig 5). The prevalence of airflow obstruction in the seven mutually exclusive areas were as follows: asthma only, 26.5%; chronic bronchitis only, 29.6%; emphysema only, 45.5%; asthma plus chronic bronchitis, 55.8%; asthma plus emphysema, 48.7%; chronic bronchitis plus emphysema, 59.7%; and asthma plus chronic bronchitis plus emphysema, 49.0%. Still, 9.3% of the NHANES III population who were \( \geq 50 \) years of age had objective airflow obstruction without any respiratory diagnoses.

**DISCUSSION**

Our analysis highlights the problem of the differential diagnosis among OLDs, particularly in older
adults and the elderly. By analyzing large samples of patients from the general population in the United States and the United Kingdom, we demonstrated that a substantial number of patients are diagnosed with two or even all three OLD conditions concomitantly. By analyzing the NHANES III spirometry data, we observed that diagnoses of asthma, chronic bronchitis, and emphysema are present with and without demonstrable airflow obstruction.

Some limitations of our research deserve discussion. The GPRD does not include reliable information on respiratory function, as pulmonary function is not routinely assessed at the primary care level in the United Kingdom or elsewhere, and the information on tobacco use is not complete. The NHANES III self-reported measurement of current chronic bronchitis by interview questionnaire may have included some misclassified cases of acute bronchitis, which would lead to the overestimation of this disease.

Some strengths of our results are the large study sample sizes, the demonstrated quality control of NHANES III spirometry, and the analyses in two nationally representative populations. As explained previously, the comparison of results from NHANES III and GPRD had to be done indirectly. NHANES III participants were asked about "ever" and "current" physician-diagnosed conditions, while the GPRD analysis was based on OLD diagnoses recorded directly by GPs, with a wide range of possible terms, over a 12-month period. A patient revisiting the GP might not have a diagnosis recorded again that year. Therefore, one potential reason for the lower absolute UK rates relative to the US rates might be a technical artifact. The apparently more than double frequency of OLD conditions that was found in NHANES III compared to the GPRD has to be interpreted cautiously. These estimates should not be considered in absolute terms but in relative terms, as the percentages of each population sample. The most likely major explanation of this difference is the self-reported nature of physician-diagnosed conditions in NHANES III and the potential for misclassification of bronchitis as chronic bronchitis.

However, the GPRD results are no less relevant than those from NHANES III. The GPRD is a real-life source of information, and, as it directly records many thousands of physician diagnoses rather than relying on the self-reporting of these diagnoses as in NHANES III, so it should provide a more accurate estimate of disease burden. Therefore, the GPRD likely produced smaller estimates of disease burden with greater specificity.

We can speculate on the reasons for why the differential diagnosis of asthma vs COPD is still
difficult, including the following: (1) the consideration of conditions as part of a continuum; (2) the consideration of different conditions with strong overlapping features (similar to diabetes mellitus being insulin-dependent or non-insulin-dependent); (3) no incentive to differentiate whether treatment and prognosis are the same; (4) lack of clear guidelines from “experts” as to how the distinction can be
made in clinical practice; or (5) uncertain criteria used by the physician to classify the patient as having asthma, chronic bronchitis, or emphysema.

The controversy on how to label OLDs has been around for decades, debating whether it is generalized OLD, chronic nonspecific lung disease, COPD, or another term that fits the best. The so-called Dutch hypothesis on the interrelationship of OLD, initially described by Orie in 1961, stated that asthma and bronchial hyperreactivity predispose patients to develop COPD later in life, and this theory has not yet been ruled out. Practical implications in this debate are relevant because the management guidelines for asthma and COPD differ. However, a growing number of researchers and clinicians consider that respiratory disease is a continuum from childhood to adulthood, and that asthma leads to chronic OLD becoming COPD in the elderly. Indeed, many physicians treat adult chronic asthma patients and COPD patients with similar drugs. In our study, it was not possible to establish any difference between the GPs and pulmonologists as to how the diagnoses were made. All entries in the GPRD were from GPs only, as the GPRD is a database of primary care doctors only, while NHANES III is a survey of self-reported diagnoses and conditions by the participants. Regrettably, in these two countries and elsewhere, OLD is diagnosed too often without spirometry findings. It may be timely to remind primary care physicians that they should accurately assess the airway status of their patients. It takes only a few minutes to inquire about medical history, to examine patients’ lungs, to confirm their impression with spirometry followed by a course of appropriate therapy, and finally to educate patients on the correct use of inhalers. Overall, only about half of the patients who reported a COPD-compatible diagnosis in the NHANES III survey had abnormal spirometry findings corresponding to grade 1 (or higher) of the GOLD guidelines. If the diagnosis was based on spirometry as well as clinical features, the distribution of the various COPD diagnoses would alter considerably (Fig 4). Our finding that US NHANES III participants with an emphysema-only diagnosis had a lower prevalence of objective airflow obstruction than did participants with concomitant emphysema and asthma or chronic bronchitis might seem to be counterintuitive. If these results are confirmed in other studies, the current practice in COPD randomized controlled trials (RCTs) of excluding patients who have been diagnosed with concomitant COPD and asthma, or COPD patients who have some asthmatic component (i.e., positive results of a bronchodilator test or methacholine challenge, or atopy) might be artificial and would not represent the spectrum of OLD patients represented in the community. Asthma RCTs usually recruit patients who are extremely young, do not smoke, and have mild disease, and whose condition are completely reversible. Conversely, COPD RCTs tend to recruit patients who are old, smoke, and have severe disease, and whose conditions are largely irreversible and include emphysema. Therefore, “mixed” patients are not recruited into RCTs. Physicians often have difficulties in labeling patients as having COPD or asthma among those in the large group of aging, chronic OLD patients with a history of cigarette smoking and an asthmatic component. Indeed, evidence-based medicine cannot be conducted in patients with combined asthma and COPD, who, according to our research, account for as much as half of the OLD population aged ≥ 50 years.

In contrast to investigators and respiratory researchers, GPs often have applied Occam’s razor, treating all OLD patients the same by administering inhaled corticosteroids and bronchodilators, but with some difference in the usage of anticholinergic agents, theophylline, cromones, and, of course, oxygen. This practical approach has made the subdivision of OLDs irrelevant for most doctors, except for those investigators studying etiology, prognosis, or treatment response in detail.

Further research to better characterize the subgroup of patients with objective airflow obstruction but without any respiratory diagnosis (3.1%) [Fig 4] is needed. At least in North America, primary care physicians are still reluctant to consider the diagnosis of COPD, even when confronted by a middle-aged former smoker with chronic cough and dyspnea, especially among women. The newly established Burden of Lung Disease initiative within GOLD aimed to obtain population data on respiratory symptoms, diagnosed conditions, and lung function in patients of different ages and multiple geographic regions will help to quantify the magnitude of these subpopulations.

Finally, we consider our presentation of proportional Venn diagrams to be original. Our diagram allows us to represent the percentage of patients with abnormal spirometry data in all seven groups, which is not possible with the original ATS version. The Venn diagram first appeared in a book written by John Venn at Cambridge in 1866. The idea of a universe box was added later by Charles Dodgson, alias Lewis Carroll, at Oxford in 1891. Usually, these diagrams consist of circles of the same size, so that they were only symbolic of the groups they represented and did not indicate in any way their relative magnitudes. Proportional Venn diagrams are always possible using triangles instead of circles but they lack the clarity and elegance of circles.
We conclude that the concomitant diagnosis of asthma, chronic bronchitis, or emphysema is common among OLD patients from the general population, accounting for as much as half of the OLD population who are aged ≥50 years.

REFERENCES
1 Sears M. Descriptive epidemiology of asthma. Lancet 1997; 350(suppl):1s–27s.
8 National Center for Health Statistics. NHANES III reference manuals and reports. Hyattsville, MD: National Center for Health Statistics, Data Dissemination Branch, 1996; CD-ROM No. 6–1078 (1096)
15 Orie NG. The Dutch hypothesis. Chest 2000; 117(suppl): 298S