Pragmatic research and outcomes in asthma and COPD

Gene L Colice

The George Washington University School of Medicine, Pulmonary, Critical Care and Respiratory Services, Washington Hospital Center, Washington DC, USA

Abstract: Asthma and chronic obstructive pulmonary disease (COPD) are common diseases which cause patients and society considerable difficulties. These are costly diseases which cause substantial morbidity and death. Health care policy makers have made improving outcomes in asthma and COPD a priority. Application of guideline recommended approaches to asthma and COPD care in the real-life setting has been emphasized but outcomes have not improved. Failure to improve outcomes may not be because of inconsistent applications of guideline recommendations, but rather because there are difficulties implementing the Expert Panel Report III (EPR 3) method for categorizing asthma severity and the Global Initiative for Obstructive Lung Disease (GOLD) method for diagnosing COPD. As these serve as the foundation for treatment recommendations for these diseases, alternative approaches should be considered for categorizing asthma severity and identifying COPD patients. Claims-based algorithms provide an intriguing option for identifying persistent asthma patients and symptomatic COPD patients in administrative databases. These methods could be used as the basis for pragmatic research, both retrospective and prospective, on assessing outcomes of guideline recommended treatment approaches in asthma and COPD. Important questions urgently need to be answered about how guideline recommended approaches regarding use of long-acting inhaled β-agonist/inhaled corticosteroid (LABA/ICS) in asthma and long-acting inhaled anti-muscarinic agent (LAMA) and LABA/ICS in COPD affect outcomes in real-life situations.

Keywords: asthma, chronic obstructive pulmonary disease, pragmatic research

Introduction

Asthma affects more than 25 million Americans, about 10% of the childhood population and 8% of adults. In the US more than 5% of adults have chronic obstructive pulmonary disease (COPD). About 12 million Americans have been diagnosed with COPD and an additional 12 million Americans probably have undiagnosed COPD. Both asthma and COPD generate huge direct and indirect health care costs. Although the mortality rate associated with asthma is fortunately low, the death rate attributed to COPD is high. These sobering statistics have led to concerted efforts by health care policy makers to improve overall care for asthma and COPD. The most important initiative intended to advance care in asthma and COPD has been the development of formally structured clinical practice guidelines for aiding in the diagnosis and management of these diseases. Despite the widespread dissemination of clinical practice guidelines for asthma and COPD, there has been little evidence of improved outcomes. Population-based surveys have shown no change in the need for asthma-related acute care interventions, such as emergency department visits and hospitalizations, over the
past decade. Deaths due to COPD are increasing. Health care policy makers have responded to these disappointing findings by suggesting that health care practitioners might not be adhering to treatment approaches recommended in clinical practice guidelines (also known as “best practices”). To more effectively align actual clinical practice with “best practices”, there has been a call in the US for the Centers for Medicare and Medicaid Services to develop performance measures in COPD. This would be a first step towards developing financial incentives, such as pay-for-performance, to reward health care providers (ie, those certified or licensed to practice medicine) who provide care in accordance with guidelines. Pay-for-performance metrics have already been implemented as a method to improve quality of care in asthma.

Improving outcomes in asthma and COPD through more consistent application of recommendations in clinical practice guidelines, even through the use of financial incentives, rests on the critical assumption that following these recommendations will improve outcomes. Unfortunately, the available clinical research in asthma and COPD might not generally apply to assessing outcomes in real-life because this literature is heavily weighted toward either mechanistic studies or clinical trials supporting pharmaceutical products. Furthermore, more than 90% of asthma and COPD patients probably would not qualify for inclusion in typical asthma and COPD clinical trials. An alternative approach to addressing how recommendations in clinical practice guidelines might affect outcomes in typical asthma and COPD patients being managed in real-life would be to use pragmatic study designs. Pragmatic trials evaluate how interventions directly pertinent to patient care affect clinically relevant outcomes in real-world practice. They use broad eligibility criteria to ensure that patients entered into these trials reflect the full spectrum of disease. Pragmatic study methods merge seamlessly into usual clinical care rather than becoming artificial constructs. To best understand how pragmatic research can be used to improve outcomes in these diseases, key aspects of management recommendations in asthma and COPD guidelines should be critically examined.

### Asthma severity as the basis for initiating pharmacotherapy

Guidelines for asthma care rely on accurately categorizing severity prior to beginning treatment. In the stepped-care approach to treating asthma, more aggressive treatments are reserved for more severe disease to appropriately match the risks from drug treatment with the potential benefits. Methods for categorizing asthma severity have evolved from clinically intuitive approaches based on symptoms, short-acting inhaled β-agonist (SABA) use and lung function, as in the Expert Panel Report II (EPR 2), to the more complex approach oriented towards considering the domains of impairment and risk in the Expert Panel Report III (EPR 3). There are difficulties, though, with relying on methods in the EPR 2 and 3 for categorizing asthma severity.

### Awareness and understanding of guidelines

An unconsidered, but limiting, factor for categorizing asthma severity is how well health care providers are aware of and understand guideline methods. In the mid-to late 1990s, there was generally poor adherence of asthma treatment with earlier versions of guidelines, possibly because health care providers were simply unaware of guideline recommendations. Over time, though, health care providers reported basing their asthma management more reliably on guideline recommendations. Although physicians might report that their care is adherent to guidelines, three examples demonstrate that health care providers do not understand guideline approaches to asthma severity categorization. The EPR 2 and 3 guidelines recommend that asthma severity categorization should only be applied to patients not receiving long-term controllers (Figure 1), but the Global Initiative for Asthma (GINA) guidelines point out that asthma severity classification is “often erroneously applied to patients already on treatment.” There are many examples of publications presenting results of asthma severity categorization which have been incorrectly reported in patients already on long-term controllers. Doerschug et al developed a 31-question, multiple-choice test to assess physician understanding of the EPR 2 recommendations for asthma diagnosis and care. This test was administered to asthma specialists, general internists, family physicians, and house staff. Asthma specialists, as expected, scored higher on this test than others. Overall, though, only about 60% of questions were answered correctly. Physicians had particular difficulty answering questions related to severity categorization, answering fewer than 50% of these correctly. Baker et al presented eight case summaries based on actual patients with childhood asthma to pediatric asthma specialists and asked them to categorize asthma severity using the EPR 2 approach. Agreement on asthma severity categorization for the 14 specialists who completed the survey questionnaire was poor. The EPR 2, published in 1997, contains a simpler approach to asthma severity categorization than the EPR 3.
Undoubtedly, if health care providers were questioned about their understanding of asthma severity categorization using the more complicated EPR 3 methods, the results would have been worse.

Physician misunderstanding of the methods for asthma severity categorization is apparent in real life situations.

Physician assessment of asthma severity often disagrees with categorization based on symptom reporting by patients and frequently is incorrect when compared to guideline methods. In a survey of 3468 asthma patients in a managed care organization, patient-reported asthma symptoms were used to categorize their asthma severity using EPR 2 criteria.

Key: FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; ICU, intensive care unit

Notes:
- The stepwise approach is meant to assist, not replace, the clinical decision making required to meet individual patient needs.
- Level of severity is determined by assessment of both impairment and risk. Assess impairment domain by patient’s/caregiver’s recall of previous 2–4 weeks and spirometry. Assign severity to the most severe category in which any feature occurs.
- At present, there are inadequate data to correspond frequencies of exacerbations with different levels of asthma severity. In general, more frequent and intense exacerbations (e.g., requiring urgent, unscheduled care, hospitalization, or ICU admission) indicate greater underlying disease severity. For treatment purposes, patients who had ≥2 exacerbations requiring oral systemic corticosteroids in the past year may be considered the same as patients who have persistent asthma, even in the absence of impairment levels consistent with persistent asthma.

Figure 1 The EPR 3 recommended approach to asthma severity categorization for asthma patients 12 years of age and older is complex.

Notes: Just below the title is the reminder, often not considered, that this severity categorization method should only be applied to patients not currently taking long-term controllers. The categorization method includes two domains. The impairment domain includes five variables which are both subjective and objective. The risk domain includes exacerbations. The worst ranking in any individual impairment and risk domain determines overall severity. Abbreviation: EPR 3, Expert Panel Report III.
The patients’ primary care physicians were separately asked to estimate asthma severity, also using EPR 2 criteria. There was only 31% concordance between asthma severity categorization based on patient symptom reporting and physician estimates. The discordance was greatest in patients categorized by their physician as having mild disease. Over 80% of these patients had symptoms which should have placed them in the moderate or severe category. Incorrect categorization of asthma severity by physicians results in inappropriate treatment. Wolfenden et al determined asthma severity in 4005 asthma patients based on their symptoms. The patients’ physicians were also asked to categorize asthma severity using a method similar to the EPR 2 approach. Physicians both overestimated and underestimated asthma severity. In the 1565 patients categorized as having moderate asthma severity by their symptoms, most were categorized incorrectly by their physicians; physicians categorized 112 (7.2%) as having severe asthma and 824 (52.7%) as mild. Asthma treatment in these patients was based on the (often incorrect) physician asthma severity estimate. Consequently, patients were frequently overtreated and undertreated. These findings again demonstrate practical limitations in using the asthma severity categorization method proposed in the EPR 2 and 3 in real-life situations.

Airway inflammation

The severity categorization methods in EPR 2 and 3 make an important distinction between intermittent and persistent asthma. Regular use of a controller is recommended only in persistent asthma. The preferred controller for persistent asthma is an inhaled corticosteroid (ICS) because this is the most effective class of drugs for controlling the airway inflammation typically found in asthma. The implication of this approach is that severity categorization will identify patients with airway inflammation treatable with an ICS. Vignola et al performed bronchoalveolar lavage and bronchial biopsies on 12 normal subjects, 24 patients with mild intermittent asthma, and 18 patients with persistent asthma. Evidence of airway inflammation, eg, lavage fluid eosinophilia, airway epithelial shedding, and increased basement membrane thickness, was found in the patients with mild intermittent asthma. Van den Toorn et al performed bronchial biopsies in 17 healthy control subjects, 18 patients with a history of asthma but in complete clinical remission and on no asthma medications, and 19 patients with active asthma. Surprisingly, airway inflammation, eg, presence of epithelial and subepithelial inflammatory mediators such as tryptase, chymase and major basic protein, airway epithelial shedding, and increased basement membrane thickness, was seen in the patients with asthma in remission (Figure 2). Although it is not clear that treatment of the airway inflammation in these patients would have been clinically justified, these studies show that patients with intermittent asthma, as determined by clinical severity categorization methods, may have airway inflammation.

Figure 2 (A) A normal bronchial biopsy from a patient without asthma compared with (B) a bronchial biopsy specimen from a patient with a history of asthma but in complete remission demonstrates epithelial shedding and extensive presence of c-major basic protein.

Impairment
There are five variables in the impairment domain of the EPR 3 asthma severity categorization method (Figure 1). These variables include subjective, patient-reported symptoms and objective lung function measures. The scaling of the subjective variables is of uncertain clinical relevance. For example, a minor change in symptom frequency from symptoms two or fewer times per week to more than two times per week is the threshold differentiating intermittent from persistent. No allowance is made for possible confounding factors when assessing two of the subjective variables, daytime symptoms and interference with normal activity. In patients who have asthma and who are also obese, obesity has been recognized to have an independent effect on symptoms relevant to asthma control. Also unclear is how the weighting of a subjective variable against an objective one was determined, eg, daily symptoms are considered the equivalent of a forced expiratory volume in 1 second (FEV1) of 60%–80% predicted.

There is little information available on the relative roles the variables in the impairment domain play in ultimately categorizing asthma severity. In one of the few studies to address this issue, Colice et al categorized asthma severity using EPR 2 methods in 744 asthma patients not on ICS controllers who were about to enter clinical trials with a new pharmaceutical product. They found that 68.3% were categorized as having severe persistent asthma. Remarkably, they found that nocturnal symptoms were the most common determining factor for overall asthma severity categorization. Lung function determined final asthma severity categorization less often than SABA use. There was poor agreement among the variables in categorizing asthma severity for individual patients.

Impairment versus risk
The EPR 3 is a novel approach to asthma severity categorization because it includes two domains for assessment, impairment and risk for exacerbations. Although both domains include highly relevant clinical information for health care providers to consider in caring for asthma patients, it is not clear how health care providers should weigh the impairment and risk domains in deciding on categorization and ultimately treatment. For instance, a long-acting inhaled β-agonist (LABA) in combination with an ICS will improve variables in the impairment domain faster than increasing the dose of an ICS because of the acute bronchodilator effects of the LABA. However, increasing the dose of an ICS might have a greater long-term effect on reducing the risk of exacerbations than adding a LABA to a lower dose of ICS. Health care providers might choose a LABA/ICS specifically because they are more concerned about early symptom control than preventing later exacerbations.

Asthma control
The GINA guidelines suggest basing severity on the intensity of treatment required to obtain control as an alternative approach to asthma severity categorization, but the variables used to assess asthma control are similar to those used for severity categorization. A more fundamental problem with this approach though, is understanding how well determining therapy based on the symptom-based methods recommended in guidelines will affect outcomes. Although this is the approach that clinicians use in practice for managing asthma, clinical trials addressing this issue are limited. A large, prospective, randomized, parallel group trial used EPR 2-based guideline methods for determining asthma control as an endpoint for adjusting pharmacotherapy over time. This study was designed to compare the effects of two pharmaceutical products and did not employ a non-guideline-based method as a comparison arm.

Guideline-based methods for determining asthma control have been compared against methods using measures of airway inflammation in conjunction with symptoms in guiding treatment. Green et al tested an approach to asthma management using eosinophilia in induced sputum together with a guideline method for assessing asthma control against using guideline methods only. Sont et al compared adding methacholine testing to determine bronchial hyperreactivity in conjunction with a guideline method for determining asthma treatment to simply relying on the guideline method. In both studies, adding a measure of airway inflammation resulted in significantly fewer exacerbations than simply relying on guideline methods (Figure 3). These studies suggest that tailoring treatment to asthma control, determined by guideline recommended methods, may not provide the best outcomes. However, the techniques used in those studies are time consuming and require considerable technical expertise to perform correctly. An easier to perform measure of airway inflammation, monitoring exhaled nitric oxide levels, has also been tested as a way to further improve asthma management over using guideline methods alone, but has not been shown to provide consistent benefit in either reducing exacerbations or improving symptom control.

Pragmatic research approaches to identifying asthma patients and categorizing severity
The current guideline-recommended approach to categorizing asthma severity is flawed. Health care providers are
Pragmatic and Observational Research 2012:3

Figure 3 (A) Cumulative severe exacerbations were significantly reduced when asthma treatment was determined by induced sputum eosinophilia used in conjunction with guideline methods (sputum management group) than guideline methods alone (BTS management group). (B) The cumulative incidence of mild first exacerbations was significantly lower when asthma therapy was adjusted based on methacholine testing used along with guideline methods (AHR-strategy) than guideline methods alone (Reference-strategy).34

Note: 3B is reprinted from Lancet, Vol 360, Issue 9347, Ruth H Green, Christopher E Brightling, Susan McKenna, Beverley Hargadon, Debbie Parker, Peter Bradding, Andrew J Wardlaw, Ian D Pavord, Asthma exacerbations and sputum eosinophil counts, Pages 1715–1721, Copyright 2002, with permission from Elsevier.

Abbreviations: AHR, airway hyperresponsiveness; BTS, British Thoracic Society.

Pragmatic research may prove to be an important tool to use in developing methods for categorizing asthma severity which are accurate and easy to use for health care providers. Any approach for asthma severity categorization developed through pragmatic research should meet three criteria: it should be easily understandable for health care providers, appropriate for use in patients who are either on long-term controllers or only short-acting relievers, and applicable in both the clinic setting for the individual patient and the epidemiologic research arena for large populations.

One approach to asthma severity categorization which has generated considerable interest is based on the Health Plan Employer Data and Information Set (HEDIS). The HEDIS was developed by the National Committee for Quality Assurance as a way to assess health plan quality of care performance. Use of the HEDIS approach for evaluating health care quality has been widely accepted by health plans and employers, as well as regulators, consumers, and public purchasers of health care.38

The HEDIS includes a component for identifying persistent asthma. Patients with persistent asthma are defined in HEDIS as those having four or more asthma medication dispensings, one or more acute inpatient or emergency department discharge(s) with a primary diagnosis of asthma, or four or more outpatient visits with asthma listed as one of the diagnoses, and two or more asthma medication dispensings.39

Initial work with the HEDIS method for identifying persistent asthma identified several potential problems with this approach. First, the HEDIS method misclassifies patients with intermittent asthma as having persistent asthma.38–40

To address this problem, Colice et al modified the HEDIS approach by adding another step which identifies persistent asthma based on SABA and oral corticosteroid use over a 1-year period.41,42 This approach has been useful in evaluating asthma costs in database analyses of large populations. As a further modification to the HEDIS method, asthma severity can be categorized into mild, moderate, and severe persistent groups based on patient pharmacotherapy. Second, claims-based algorithms for asthma severity have been criticized because they do not incorporate physiologic measures of lung function. Birnbaum et al have shown that adding spirometry results to the HEDIS claims-based algorithm modified by Colice et al did not appreciably affect asthma severity categorization.43 Third, because asthma is a variable disease over time, patients with persistent asthma identified during 1 year might not have persistent asthma the next year.44 Schatz and Zeiger have suggested that applying the HEDIS method over 2 years, rather than one, will adequately address this issue.45

Other approaches to assessing asthma control and outcomes involve administrative database review of medication use. Studies in health maintenance organizations have shown that records of SABA dispensing can provide an insight into risk for future acute asthma health care use, such as asthma-related emergency department visits and hospitalizations.46,47
The number of SABA canisters dispensed to a patient over time has also been found to be a useful indicator of asthma symptom control. An argument against relying on SABA dispensing records obtained from claims databases has been that some SABAs have been available over the counter, but these products will soon be removed from the market. An approach based on simply counting SABA dispensing has been refined to include dispensing records for long-acting controllers. If SABA use is considered along with controller use, a ratio of medication dispensing for long-term controllers and SABAs can be calculated. A ratio of at least 0.5, indicating that patients are preferentially filling their prescriptions for long-term controllers rather than relying on SABA use to relieve symptoms, has been associated with a lower likelihood for asthma exacerbations.

Schatz and Zeiger suggested that, combined with information obtained from administrative databases on SABA and long-term controller use, and with results from asthma control questionnaires administered by telephone, the HEDIS method is a reasonable approach to evaluating quality of care in asthma management in large populations. There are two intriguing aspects of using claims data for asthma severity categorization and for predicting outcomes. Although these methods were developed for use in large populations, Schatz and Zeiger suggest it should also be applicable to group practices and individual patient care. Studies using Canadian administrative databases suggest that this approach can also be useful outside the US.

**Pragmatic research, asthma treatment, and asthma outcomes**

Development of an approach that health care providers can use to accurately and easily identify patients with persistent asthma and categorize persistent asthma severity is an essential step towards addressing contentious issues regarding treatment recommendations in asthma guidelines. Busse has recently described a number of areas in which there are gaps in our understanding how to best treat asthma. However, it is clear that for the practicing clinician, the most contentious issue confronting them in treating asthma is understanding how combination LABA/ICS therapy affects outcomes in persistent asthma. The EPR 3 and GINA guidelines recommend use of LABA/ICS combination therapy in patients with moderate and severe persistent asthma because adding the LABA will allow use of lower doses of an ICS. These recommendations are consistent with findings from a recent Cochrane review showing that LABA/ICS therapy improves lung function and reduces symptoms to a significantly greater extent than ICS therapy alone. Long-term trials have shown that use of LABA/ICS therapy is associated with a greater likelihood of patients achieving asthma control than treatment with ICS alone.

However, there are concerns about the use of LABA/ICS therapy. Retrospective studies of various insurance claim databases suggest that use of LABA/ICS combination products might be associated with an increased rate of serious asthma exacerbations and asthma exacerbations resulting in death compared to use of ICS alone. Concerns about the safety of LABAs has led the FDA to change the labeling for LABA/ICS products and to restrict their long-term use. Industry sponsored reviews of safety databases compiled from clinical trials involving different LABA/ICS combination products have not shown evidence of a significant increase in asthma-related deaths with use of LABA/ICS. These analyses, though, might not have had sufficient sample size to adequately address this risk issue. Independent work has supported the FDA’s findings of a possible small but finite increased risk of asthma death with use of a LABA/ICS. The recent Cochrane review found that use of an increased dose of ICS was more effective at preventing exacerbations than LABA/ICS combination therapy.

Recent pragmatic research has partly addressed the issue of outcomes with different asthma treatment approaches. Price et al compared the effect of choosing a leukotriene receptor antagonist rather than an ICS as initial therapy and adding a leukotriene receptor antagonist to an ICS rather than a LABA as add-on therapy. In contrast to guideline recommendations, patients treated with a leukotriene receptor antagonist had outcomes similar to those treated with an ICS initially or with a LABA as additional therapy. Important aspects of the approach used in this study were that patients were managed by their primary care physician in a real-life situation, adherence to medication use was considered, clinically relevant outcomes (symptoms and exacerbations) were measured, variations in drug treatment were allowed (as would be expected in real-life practice), and the trials lasted long enough (2 years) to understand effects over time. A limitation
of this study was its relatively small size, which precluded a full understanding of safety issues for each treatment option. However, this type of pragmatic research is a useful model for larger studies which could be designed to compare increasing the dose of an ICS against using a lower dose of an ICS with a LABA. To ensure a large enough sample size to adequately address the risk component of asthma treatment, these studies could preferentially use administrative databases, either prospectively or retrospectively. Claims algorithms based on the modified HEDIS approach could be used to identify patients with persistent asthma. Asthma severity could also be approximated with this approach. Patient care would be determined by the primary care provider without constraints of a clinical protocol. Symptom control could be inferred from SABA dispensings. Adherence to use of long-term controllers could be measured by medication claims. Exacerbations could be defined by oral corticosteroid dispensings, acute care visits, and hospitalizations. Most important, deaths could also be tracked.

**Diagnosing COPD as the basis for initiating pharmacotherapy**

Similar to asthma, guidelines for COPD also depend upon categorizing severity as a basis for determining pharmacotherapy. However, unlike asthma, COPD guidelines include strict criteria for establishing the diagnosis of COPD before proceeding with severity categorization. The clinical presentations of the major COPD subtypes, chronic bronchitis and emphysema, are well recognized. COPD is eloquently defined in the Global Initiative for Obstructive Lung Disease (GOLD) guideline. This definition, though, is not used to diagnose COPD. The GOLD guideline, and others, have an operational definition of COPD, based solely on a spirometric finding of an FEV1/forced vital capacity (FVC) ratio below 0.7 following administration of a bronchodilator, which must be met for the diagnosis of COPD to be confirmed. There are difficulties, though, with relying on a single post-bronchodilator spirometry finding to establish the diagnosis of COPD.

**Performing spirometry in the office**

An obvious practical problem is that primary care physicians often do not have access to spirometers in their clinic, do not train their staff on how to perform spirometry, do not perform quality control on spirometer performance, and do not understand how to interpret results. It is extremely unlikely that pre- and post-bronchodilator testing as part of spirometry is being routinely performed in general practice. It should also be recognized that not all patients can perform spirometry. Hardie et al asked 95 participants in a screening study of elderly asymptomatic non-smokers to perform spirometry. Only 71 (75%) could perform this test. The 25% who could not perform spirometry were more likely to be female and older.

Before interpreting spirometry results, there are a series of technical issues that the health care provider should consider to ensure that the test was performed correctly and that the results were reproducible. These issues represent serious obstacles to widespread use of spirometry for diagnosing COPD.

**Interpreting spirometry results**

If patients can perform technically acceptable spirometry, the health care practitioner must then interpret the results. The current approach to interpreting spirometry results, though, has limitations. The absolute values for FEV1, FVC, and FEV1/FVC are not interpreted directly. Instead the actual values are compared to predicted normal values using regression equations. The regression equations for these predicted values are typically obtained from studying large numbers of asymptomatic, non-smoking subjects of different ages, ethnicity, and physical characteristics, such as in the National Health and Nutrition Examination Survey III (NHANES III). Although predicted values have to be adjusted for ethnic differences, the highly selected populations used for developing reference values would presumably not have underlying COPD. However, other factors besides smoking probably contribute to the development of COPD, because many COPD patients never smoked. In a recent international study, 28% of subjects identified as having COPD by spirometry were never smokers. In this study, a history of asthma and, in women, lower education levels were associated with a COPD diagnosis. Interestingly, there was a suggestion in this analysis that lower educational levels in women was actually a surrogate marker of chronic exposure to biomass fuels used for cooking and heating, another recognized cause of COPD. Van Sickle et al also found that socioeconomic factors, such as high school completion, significantly affected NHANES III predicted values. Wagner suggested that the effect of socioeconomic status on predicted values may have been missed in the past because exclusion of smokers might have preferentially led to fewer subjects in the presumed normal group with lower socioeconomic status. These observations suggest that populations used to develop predicted normal regression equations might have included patients with early COPD.

Spirometry results, even when performed under rigorously controlled settings, may be variable. In two long-term
trials in stable COPD patients, baseline spirometry was performed on two separate occasions between 3 and 12 weeks apart and showed substantial variability in FEV₁ and FVC between the two tests. Spirometry results may be ambiguous when a reduced FVC, suggesting a restrictive defect, is found. Currently, little guidance is available on interpreting a reduced FEV₁ in the context of a concomitant reduction in FVC. Spirometry results for 1831 consecutive patients showed that 470 (25.7%) had a low FVC. Although a low FVC should indicate a restrictive ventilatory defect, only a minority of patients with a low FVC were actually confirmed to have a reduced total lung capacity. Hyatt et al have described a nonspecific pattern of abnormal spirometry, characterized by a low FEV₁, and FVC with a normal FEV₁/FVC ratio, occurring in 9.5% of routine spirometries. In most patients with this nonspecific spirometry pattern bronchodilators did not change the results. Although this nonspecific pattern would not meet the GOLD criteria for COPD, on follow-up over 3 years, 191 (15%) of 1284 patients were confirmed to have airway obstruction. Consequently, a reduced FVC found on spirometry might represent COPD, but current strategies for interpreting spirometry results would not suggest this diagnosis.

Reversibility testing

There is little information available on predicted normal FEV₁, FVC, and FEV₁/FVC values pre- and post-bronchodilator for the general population. The NHANES III data used for predicting normal regression equations relied on pre-bronchodilator spirometry. A study from Norway reported on pre- and post-bronchodilator values for normal subjects aged between 26 and 82 years, but only 515 participants were included. Given the small overall sample size, it is not surprising to note that there were only 39 men and 82 women included in this study over the age of 60, the relevant age for COPD. As expected, they found that the post-bronchodilator values were significantly greater than the pre-bronchodilator values. They also compared the post-bronchodilator FEV₁, FVC, and FEV₁/FVC values to pre-bronchodilator values predicted from standard reference equations and again found significant differences. The authors speculated that relying on only the post-bronchodilator FEV₁/FVC ratio might reduce the detection of obstructive airway disease. In earlier work, this same group showed that this effect did occur. In random screening of adults living in Hordaland County, Norway, 3% of the subjects tested had an FEV₁/FVC ratio below 0.70 pre-bronchodilator but above 0.70 post-bronchodilator. The prevalence of COPD was 27% lower in this screening study using the post-bronchodilator FEV₁/FVC ratio rather than the pre-bronchodilator value. Interestingly, 0.5% of the screened population had a normal FEV₁/FVC ratio pre-bronchodilator but a low ratio post-bronchodilator. This effect probably occurred because there may be a more robust increase in FVC post-bronchodilator than FEV₁. It is often not appreciated that reversibility can be based on changes in either FEV₁ or FVC. Failure to understand how a COPD patient might respond to a bronchodilator could result in confusing situations for the general practitioner. A patient with a low pre-bronchodilator FEV₁/FVC but a normal post-bronchodilator ratio due to a vigorous FEV₁ response would not be considered to have COPD. Conversely, a patient with a normal pre-bronchodilator FEV₁/FVC ratio but a low post-bronchodilator ratio due to a vigorous FVC response would be diagnosed as having COPD by the GOLD guidelines.

There is uncertainty in how to perform and interpret reversibility testing. One approach is to administer four puffs of albuterol by metered dose inhaler with a spacer, but there is no consensus on the type of bronchodilator, the number of puffs to be administered, and the inhalation device to be used when performing reversibility testing. Administering both a SABA and a short-acting inhaled antimuscarinic agent (SAMA) results in a greater bronchodilator effect than found with a SABA alone. Just as there is variability in the FEV₁ over time in stable patients, there is also variability in the FEV₁ response to inhaled bronchodilators in COPD over time. Calverley et al reported on bronchodilator reversibility in 660 COPD patients studied on three separate occasions at 4-week intervals and found that 52.1% of the patients changed responder status between visits. At some visits, patients were classified as having reversible airway obstruction, but at other visits the response to bronchodilators was not as substantial. Han et al observed a similar variability in bronchodilator responsiveness over time. Again, failing to understand how COPD patients respond to bronchodilators could result in odd situations which would be difficult for the general practitioner to interpret. A patient identified as having COPD because of a minimal bronchodilator effect on one visit might not have the diagnosis confirmed on another visit, because of a more robust FEV₁ response to inhaled albuterol. A patient not diagnosed with COPD on one visit could have COPD diagnosed at a later visit if there were a vigorous FVC response. An important caveat to the issue of bronchodilator reversibility, often not recognized in clinical practice, is to ensure that patients do not use either their regularly scheduled long-acting bronchodilators or as-needed...
short-acting bronchodilators in the hours before spirometry, as use of these medications will minimize bronchodilator responsiveness in the laboratory. 

Fixed FEV\textsubscript{1}/FVC ratio

Using a fixed FEV\textsubscript{1}/FVC ratio to diagnose COPD fails to take into account the expected and normal effects of aging on FEV\textsubscript{1} and FVC. With aging both the FEV\textsubscript{1} and FVC decrease, but the FEV\textsubscript{1} tends to fall to a greater extent. Consequently, the FEV\textsubscript{1}/FVC ratio normally decreases with age. This physiologic phenomenon can lead to two different types of diagnostic inaccuracies when relying on the FEV\textsubscript{1}/FVC ratio to establish the diagnosis of COPD. A younger patient with a history consistent with COPD might have an FEV\textsubscript{1}/FVC ratio above 0.70 but a ratio well below expected for their age. This patient might be incorrectly classified as not having COPD (false negative). Conversely, an older patient without a history suggesting COPD might have an FEV\textsubscript{1}/FVC ratio below 0.70 consistent with their age predicted value. This patient might be falsely identified as having COPD (false positive). Miller et al found that false negative findings, ie, missing the diagnosis of COPD, were more likely to occur in women younger than 45. Schermer et al found that false positive findings, ie, overdiagnosis of COPD, occurred more often in middle age and elderly patients. Robberts and Schermer showed that 16% of 3473 men referred for spirometry by their primary care physicians would have been incorrectly diagnosed with COPD if the FEV\textsubscript{1}/FVC ratio <0.70 had been the sole criteria (Figure 4). Although it is argued that the FEV\textsubscript{1}/FVC ratio <0.70 is a simple and practical operational definition of COPD, a more appropriate physiologic approach would be to use the lower limit of predicted normal for the FEV\textsubscript{1}/FVC ratio. Until this issue is resolved, over-diagnosis and under-diagnosis of COPD will occur.

Pragmatic research approaches to identifying COPD patients and categorizing severity

The current operational definition of COPD, which rests solely on the post-bronchodilator FEV\textsubscript{1}/FVC ratio, does not seem either workable or reasonable. There are substantial difficulties performing and interpreting the results of spirometry. These difficulties undoubtedly lead to infrequent use of spirometry by many primary care physicians, which accounts for the overall under-diagnosis and under-detection of COPD in primary care. Predicted normal reference equations are limited by the population studied and provide little information on the predicted post-bronchodilator response. Reversibility testing is not well standardized and is variable over time. Relying on a strict FEV\textsubscript{1}/FVC threshold of 0.70 ignores basic physiologic principles regarding the effect of aging on the lung. Clinical guidelines recommending treatment approaches based on a dysfunctional definition of COPD are fundamentally flawed. Curiously, recent commentaries on controversies in COPD do not mention this issue, and persist in relying on the GOLD definition of an FEV\textsubscript{1}/FVC ratio <0.70 for COPD.

Pragmatic research is needed to develop a definition of COPD that is both accurate and readily usable by the primary care health practitioner. This definition should satisfy two important points. First, it should not be based on spirometry. Second, it should focus on identifying patients with more severe COPD. Recognizing the mild COPD patient, ie, a patient who is either asymptomatic or minimally symptomatic with an FEV\textsubscript{1}/FVC ratio <0.70 but an FEV\textsubscript{1} > 80% predicted, may not be helpful. Office spirometry can significantly improve early detection of COPD, but, as the US Preventive Services Task Force has pointed out, screening asymptomatic patients for COPD provides no net benefit. There is no effective treatment for preventing the accelerated decline in lung function seen in COPD other than smoking cessation. All smokers, whether they have COPD or not, should be advised to stop smoking. Excluding spirometry as a diagnostic standard and developing a definition oriented
towards more severe patients should actually make the process of recognizing COPD simpler.

There are useful non-spirometry-based methods for identifying patients with COPD. A recent systematic review provided helpful insights into the diagnostic value of the history and physical examination for COPD. Features in the history, such as age ≥45 years, current and heavy smoking, female sex, complaints of wheeze and dyspnea, and a self-reported history of COPD, had independent diagnostic value for identifying COPD. Findings on physical examination, like wheeze and prolonged expiration, also had independent diagnostic value. Unfortunately, this systematic review included only a small number of studies and few of the studies reviewed used the FEV₁/FVC ratio <0.70 as the independent determinant of COPD. Others have found that a patient reported smoking history of more than 55 pack years, wheezing heard on auscultation and patient self-reported wheezing almost assures the presence of airflow obstruction. Price et al developed a symptom-based questionnaire for identifying COPD in smokers which they validated by comparison to the standard COPD definition of an FEV₁/FVC ratio <0.70. The final questionnaire contained only eight items and could be easily completed by the patient. Questions were related to age, smoking history, cough, wheeze, and allergies. The questionnaire had reasonably good performance characteristics, a specificity of 80.4%, and sensitivity of 72.0% for identifying COPD.

Approaches to identifying COPD patients using smoking history and symptoms could be improved by incorporating reports of exacerbations, especially those requiring hospitalization, ambulatory claims for care related to COPD, and use of bronchodilators. Recent work has shown that patients with GOLD Stage 2 COPD (an FEV₁ of 50%–79% predicted) had frequent exacerbations (Figure 5). The first year exacerbation rate in this study for Stage 2 patients was 0.85 per person; 22% had frequent exacerbations and 7% had been hospitalized for an exacerbation. Gershon et al found that information available in an administrative database on hospitalizations and ambulatory care visits attributed to COPD by a primary care physician was a reasonably accurate way to identify COPD patients. In a large population study based on administrative data, these authors showed that a physician diagnosis of COPD based on either an ambulatory care claim or hospitalization was a useful approach in identifying patients with COPD over time. Incorporating use of medications typically used for managing COPD, such as inhaled anticholinergic and long-acting bronchodilators, can be used as part of a non-spirometry-based method for determining COPD severity.

Future pragmatic research should concentrate on developing a simple, practical, and accurate method for identifying COPD patients who need treatment. Previous work suggests that a viable approach could be developed based on three components: a questionnaire regarding smoking history, exposure to other noxious dusts, and symptoms of wheeze and cough; findings on physical examination of wheeze; and an administrative database review of exacerbation history and medication use. A further refinement of this approach will be to use elements of this method, such as the number of medications being used, frequency

Figure 5 As COPD severity stage increases, the frequency of exacerbations requiring hospitalization increases. The percent of patients with frequent exacerbations (ie, two or more exacerbations per year) also increased with COPD severity stage. In GOLD Stage 2 COPD the exacerbation rates were 0.85 per year. The high exacerbation rate suggests that an algorithm based on claims for COPD exacerbations could be a useful approach to identifying COPD patients in GOLD Stage 2 and above.


Abbreviations: COPD, chronic obstructive pulmonary disease; GOLD, Global Initiative for Obstructive Lung Disease.
of exacerbations, use of oxygen, etc, to then categorize COPD severity.

Pragmatic research, COPD treatment, and COPD outcomes

With a validated non-spirometric-based method for diagnosing symptomatic COPD, especially one which can also be used to categorize COPD severity, difficult questions about how treatment recommendations in the GOLD guidelines actually impact outcomes can be addressed with pragmatic research. An example of a particularly vexing question is whether currently recommended maintenance pharmacotherapy for COPD in the GOLD guidelines affects survival. For COPD patients in GOLD Stages 2, 3, and 4 severity, the guidelines recommend regular treatment with some combination of a LABA, a long-acting inhaled anti-muscarinic agent (LAMA), and a LABA with an ICS. Each of these products has proven benefits in improving lung function and reducing exacerbations. However, there are contradictory reports regarding how these products affect survival. In a 3-year industry sponsored trial, there was a nonsignificant trend for regular treatment with both a LABA and a LABA/ICS combination to reduce mortality.\(^{102}\) In a 4-year industry sponsored trial there was a nonsignificant trend for regular treatment with a LAMA to reduce mortality.\(^{103}\) A pooled safety analysis compiled from the randomized, double-blind, placebo-controlled clinical trials performed in COPD with a LAMA during its clinical development also suggested a trend for improved survival.\(^{104}\) Retrospective analyses, though, suggested that use of inhaled anti-muscarinic agents, both short-acting and long-acting, was associated with an increased risk of cardiovascular deaths.\(^{105,106}\)

The FDA has publicly commented that the safety data on LAMA use from prospective trials was more meaningful than those from the retrospective analyses.\(^{107}\) Unfortunately, recent data from prospective trials with a LAMA in a novel inhalation device have also suggested an increased risk of death with this drug.\(^{108,109}\) Complicating this situation are results from two other studies comparing outcomes between use of a LAMA and a combination LABA/ICS product and use of LAMA and a LABA. In a 2-year, double-blind, randomized, parallel, industry-sponsored study, patients treated with a LABA/ICS were significantly less likely to die than those receiving the LAMA.\(^{110}\) In a retrospective database analysis, moderate COPD patients initially treated with a LABA had a lower mortality rate over more than 5 years of follow-up than those begun on a LAMA.\(^{111}\) These findings suggest that both LABA and LABA/ICS might provide a survival benefit in COPD over LAMA. Interestingly, a systematic review compared results with LABA/ICS and LABA treatment and found that combination therapy with LABA/ICS did not reduce the risk of either death or severe exacerbations compared to LABA therapy alone.\(^{112}\) Although LABA/ICS therapy did provide significant improvements in lung function and quality of life compared to LABA treatment, these benefits were small and probably clinically unimportant. Use of a LABA/ICS was associated with significantly more side effects, such as pneumonia, than LABA treatment.

On balance, the results of these studies are concerning. They suggest serious possible safety concerns with both LAMA and combination LABA/ICS therapy in COPD. Surprisingly, monotherapy with a LABA might provide benefits with fewer risks than either LAMA monotherapy or combination LABA/ICS treatment. Given the widespread use of LAMA and LABA/ICS products in COPD, there is a clear need to address this issue and pragmatic research is ideally suited for this purpose. As with asthma, large studies would be required to address these safety concerns. Prospective and/or retrospective administrative database analyses would be appropriate research models. Patients with COPD could be identified through a claims-based algorithm. Patient care decisions would be made by the local health care provider. Severity categorization could be inferred from types of therapy being used by patients. Multiple outcomes could be monitored. Exacerbations could be measured by oral corticosteroid and antibiotic dispensings along with acute care visits and hospitalizations as indicators of exacerbations. Deaths would be the ultimate outcome of concern.

Conclusion

Asthma and COPD are common diseases which cause patients and society considerable difficulties. Asthma patients suffer from limitations in their daily lives due to uncontrolled symptoms and intermittent exacerbations. Symptoms and exacerbations in COPD also disrupt daily life activities. Although asthma is associated with a low mortality rate, COPD is a deadly disease with an increasing mortality rate. Both asthma and COPD result in substantial total direct and indirect health care costs for society. Clinical practice guidelines provide comprehensive recommendations for the care of asthma and COPD patients. Health care policy makers have reasonably recognized that improving outcomes in asthma and COPD should be a priority and have emphasized consistent use of guideline recommendations to achieve this end. Unfortunately, failure to improve outcomes may not be because of inconsistent applications of guideline recommendations, but rather because these
recommendations might be based on flawed assumptions. There are difficulties implementing the EPR 3 method for categorizing asthma severity and the GOLD method for diagnosing COPD. As these serve as the foundation for treatment recommendations for these diseases, simpler methods, using readily available information at both the individual practice and the large population level, should be developed to categorize asthma severity and to diagnose COPD. Once these methods are developed and validated, pragmatic research will be of great value in answering important questions about how guideline-recommended approaches, specifically regarding use of LABA/ICS in asthma and LAMA and LABA/ICS in COPD, affect outcomes in real-life situations. Ideally pragmatic research methods developed for use in North America and Europe could be applied worldwide.

Disclosure
The author has served as a speaker/advisory board member/consultant to Teva, MedImmune, Boehringer-Ingelheim, Merck, Schering-Plough, GlaxoSmithKline, Genentech, Pfizer, and Eli Lilly.

References


74. Wagner PD. FEV1 in the suburbs: choose your research subjects wisely. *Am J Respir Crit Care Med*. 2011;184(5):495–496.