ORIGINAL ARTICLE

A New Approach for Identifying Patients with Undiagnosed Chronic Obstructive Pulmonary Disease

Fernando J. Martinez[1], David Mannino[2], Nancy Kline Leidy[3], Karen G. Malley[3], Elizabeth D. Bacci[4], R. Graham Barr[5,6], Russ P. Bowler[7], MeiLan K. Han[8], Julia F. Houfek[9], Barry Make[7], Catherine A. Meldrum[8], Stephen Rennard[10,11], Byron Thomashow[12], John Walsh[13], and Barbara P. Yawn[14]; on behalf of the High-Risk-COPD Screening Study Group*

[1]Joan and Sanford I. Weill Department of Medicine, Weill Cornell Medical College, New York, New York; [2]Department of Preventive Medicine and Environmental Health, University of Kentucky, Lexington, Kentucky; [3]Outcomes Research, Evidera, Bethesda, Maryland; [4]Outcomes Research, Evidera, Seattle, Washington; [5]Department of Medicine and [6]Department of Epidemiology, Columbia University Medical Center, New York, New York; [7]Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, National Jewish Health, Denver, Colorado; [8]Division of Pulmonary and Critical Care, University of Michigan, Ann Arbor, Michigan; [9]College of Nursing and [10]Pulmonary, Critical Care, Allergy and Sleep Medicine Division, Department of Internal Medicine, University of Nebraska Medical Center, Omaha, Nebraska; [11]Clinical Discovery Unit, Early Clinical Discovery, AstraZeneca, Cambridge, United Kingdom; [12]Division of Pulmonary, Allergy and Critical Care Medicine, Columbia University, New York, New York; [13]COPD Foundation, Washington, District of Columbia; and [14]Department of Family and Community Health, University of Minnesota, Minneapolis, Minnesota

Abstract

Rationale: Chronic obstructive pulmonary disease (COPD) is often unrecognized and untreated.

Objectives: To develop a method for identifying undiagnosed COPD requiring treatment with currently available therapies (FEV_1 <60% predicted and/or exacerbation risk).

Methods: We conducted a multisite, cross-sectional, case-control study in U.S. pulmonary and primary care clinics that recruited subjects from primary care settings. Cases were patients with COPD and at least one exacerbation in the past year or FEV_1 less than 60% of predicted without exacerbation in the past year. Control subjects were persons with no COPD or with mild COPD (FEV_1 ⩾60% predicted, no exacerbation in the past year). In random forests analyses, we identified the smallest set of questions plus peak expiratory flow (PEF) with optimal sensitivity (SN) and specificity (SP).

Measurements and Main Results: PEF and spirometry were recorded in 186 cases and 160 control subjects. The mean (SD) age of the sample population was 62.7 (10.1) years; 55% were female; 86% were white; and 16% had never smoked. The mean FEV_1 percent predicted for cases was 42.5% (14.2%); for control subjects, it was 82.5% (15.7%). A five-item questionnaire, CAPTURE (COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk), was used to assess exposure, breathing problems, tiring easily, and acute respiratory illnesses. CAPTURE exhibited an SN of 95.7% and an SP of 44.4% for differentiating cases from all control subjects, and an SN of 95.7% and an SP of 67.8% for differentiating cases from no-COPD control subjects. The PEF (males, <350 L/min; females, <250 L/min) SN and SP were 88.0% and 77.5%, respectively, for differentiating cases from all control subjects, and they were 88.0% and 90.8%, respectively, for distinguishing cases from no-COPD control subjects. The CAPTURE plus PEF exhibited improved SN and SP for all cases versus all control subjects (89.7% and 78.1%, respectively) and for all cases versus no-COPD control subjects (89.7% and 93.1%, respectively).

Conclusions: CAPTURE with PEF can identify patients with COPD who would benefit from currently available therapy and require further diagnostic evaluation.

Clinical trial registered with clinicaltrials.gov (NCT01880177).

Keywords: screening; chronic obstructive pulmonary disease; primary care; questionnaire; random forests

(Received in original form March 25, 2016; accepted in final form October 7, 2016)

*A complete list of members may be found before the beginning of the References.

Funding for this work was provided by the NHLBI under grant R01 HL114055 (F.J.M.). A portion of B.M.'s efforts was also supported by NHLBI grants R01 HL089856 and R01 HL089897.

Author Contributions: F.J.M., D.M., and N.K.L.: served as principal investigators on the project; R.G.B., R.P.B., M.K.H., J.F.H., B.M., C.A.M., S.R., B.T., B.P.Y., and J.W.: served as coinvestigators; E.D.B. and K.G.M.: contributed to analytical design, statistical analyses, and writing; and all authors: contributed to design, implementation, and analyses, as well as to the writing, review, and approval of the paper prior to submission.

Correspondence and requests for reprints should be addressed to Fernando J. Martinez, M.D., M.S., Joan and Sanford I. Weill Department of Medicine, Weill Cornell Medical College, New York-Presbyterian Hospital/Weill Cornell Medical Center, 525 East 68th Street, Room M-522, Box 130, New York, NY 10065. E-mail: fjm2003@med.cornell.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 195, Iss 6, pp 748–756, Mar 15, 2017
Copyright © 2017 by the American Thoracic Society
Originally Published in Press as DOI: 10.1164/rccm.201603-0622OC on October 26, 2016
Internet address: www.atsjournals.org

At A Glance Commentary

Scientific Knowledge on the Subject: The challenges associated with efficiently identifying people with undiagnosed chronic obstructive pulmonary disease (COPD) in primary care settings are well known. Identifying symptomatic patients with more severe airflow obstruction or at risk for exacerbation who will benefit from currently available therapeutic interventions has immediate clinical importance for these individuals. To date, questionnaires have been designed to identify people with COPD through population- or clinic-based screening programs without reference to disease severity or exacerbation risk, resulting in the identification of a high proportion of patients with mild disease. The use of peak expiratory flow using various methods for gathering and interpreting data has been proposed.

What This Study Adds to the Field: In this study, we used a novel, multimethod approach to develop a process for identifying undiagnosed cases of COPD requiring treatment. CAPTURE (COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk), a simple, five-item, patient-completed questionnaire, plus peak expiratory flow, with an inexpensive, easy-to-use mechanical device and interpretive thresholds, is able to differentiate cases and control subjects with remarkable precision, suggesting this is a viable approach for patient screening and COPD case identification in primary care settings. Further study of this approach is warranted.

Evidence suggests that chronic obstructive pulmonary disease (COPD) is underdiagnosed in primary care settings, with most cases being identified during an exacerbation or after significant loss of lung function (1). Undiagnosed patients have been suggested to have impaired health status (2) and outcomes (3, 4). Therapies are available to improve lung function, reduce exacerbations, and improve health status in patients with COPD, with evidence of therapeutic benefit clearly demonstrated and strongly recommended by the American College of Physicians, the American College of Chest Physicians, the American Thoracic Society, and the European Respiratory Society in symptomatic people with FEV_1 less than 60% of predicted or who are at risk for acute exacerbations (5, 6). The efficient identification of this group of unrecognized patients with COPD, whom we arbitrarily label as having "clinically significant COPD," is therefore important clinically.

To date, spirometry has served as the "gold standard" for COPD diagnosis (7), but, as recently noted in the U.S. Preventive Services Task Force (USPSTF) report, it is not recommended for routine, general population, or practice-based screening in asymptomatic patients (8). Questionnaires offer a practical triage or case-finding method for identifying symptomatic people in practice settings who may have unrecognized COPD and would benefit from treatment. The USPSTF noted that few data exist to support the widespread use of case-finding approaches and that improved clinical outcomes and the limitations of overdiagnosis have not been established (8). Existing questionnaires were generally designed to identify people with COPD without reference to disease severity or exacerbation risk, resulting in the identification of a high proportion of patients with mild disease (9–17). To date, no methodology has been designed explicitly for the identification of people with undiagnosed COPD who are most likely to benefit from currently available therapies (18). Several studies have tested the accuracy of handheld flow meters (FEV_1, FEV_6, peak expiratory flow [PEF]) for case identification, with varying sensitivity and specificity being reported (17). Although informative in terms of airflow obstruction, flow meters are unable to identify patients at risk of exacerbation, nor can they identify symptomatic patients. To overcome these weaknesses, one investigative group suggested that a three-stage approach (risk factor questionnaire, PEF, and spirometry) for identifying moderate to severe COPD (FEV_1 <60% predicted) might improve sensitivity and specificity (19).

Although identification of patients with mild COPD is important for research into COPD natural history and disease prevention (6, 8, 20), identification of patients with symptomatic disease, with more severe airflow obstruction (20), or at risk for exacerbation has immediate clinical importance for individual patients (5, 6, 21). An NHLBI task force reviewed the available literature and suggested that the identification of these patients may prove to be an ideal initial stage in systematically evaluating the potential impact of COPD case finding in primary care (22). We hypothesized that use of a combination of a questionnaire and PEF would optimally identify patients who would benefit from further diagnostic evaluation, such as those with an FEV_1 less than 60% predicted and/or at risk of an exacerbation in primary care settings (17, 18). This paper describes the empirical methods used to develop a new COPD case-finding methodology prior to testing its performance properties in a large, prospective study in primary care settings across the United States.

Methods

Design

We conducted a prospective, cross-sectional, multisite, case-control study (clinicaltrials.gov identifier NCT01880177) to select the best, smallest set of questions capable of differentiating cases and control subjects, with and without PEF. COPD was defined by medical diagnosis with prescribed pharmacological maintenance therapy and an FEV_1/FVC ratio less than 0.70. To address our primary goal, cases included subjects with COPD and a history of at least one exacerbation in the prior 12 months (group 1) or COPD with moderate to severe airflow obstruction (FEV_1 <60% predicted) and being exacerbation-free longer than 12 months (group 2). Control subjects included those with no known diagnosis or treatment for COPD (group 3) and those with mild COPD (FEV_1 >60% predicted and no exacerbation in the prior 12 mo) (group 4). The mild COPD group was included in the control group to focus the item selection process on identifying patients with COPD most likely to benefit from currently available therapies. Identifying milder, symptomatic patients with undiagnosed disease would be an added benefit, but it was not the intent of this study. An adjudication step was included to ensure unequivocal group assignment (*see* online supplement).

Procedures

The investigators engaged primary care clinicians to identify men and women at least 40 years of age in six diverse geographical locations in the United States. The protocol was reviewed and approved by a central institutional review board and institutional review boards at each investigation site. After providing informed consent, each subject participated in one study visit, completing a questionnaire booklet with candidate items and sociodemographics- and health-related questions, PEF, and spirometry (*see* online supplement). To evaluate questionnaire test-retest reliability, a subset of subjects (n = 111) completed the questionnaire booklet a second time, with additional questions to identify stable patients, defined as little or no self-reported change in breathing-related health during the past week. This booklet was completed at home 7–14 days following the clinic visit and was returned by mail.

Measures

Questionnaire candidate item pool. Results of earlier work, including a comprehensive literature review (18), qualitative interviews with patients from the target population (23), and analyses of existing datasets (24), were used to create 44 candidate items covering 6 content areas: exposure (6 items), family and personal health history (7 items), respiratory events during the prior 12 months (6 items), respiratory symptoms (12 items), other symptoms (5 items), and impact or effect of breathing-related issues on daily life (8 items). For ease of use, all items were dichotomous (yes or no), with the exception of frequency of respiratory events (scored on a 3-point scale; none, 1, 2 or more).

COPD Assessment Test and modified Medical Research Council dyspnea scale. The COPD Assessment Test (CAT) (25) and the modified Medical Research Council dyspnea scale (mMRC) (26) were used to assess the presence and magnitude of respiratory symptoms in the sample and to test the final questionnaire.

Analyses

A model-free data-mining approach using random forests (27) [RF; R package randomForest (28)] analysis was used to derive the best, smallest set of questions from the pool of 44 candidate questions. Additional information on these analytical methods is provided in the online supplement.

The following predictive precision estimates were used to test the questionnaire and PEF: receiver operating characteristic (ROC) curve, area under the receiver operator characteristic curve (AUC) (29), sensitivity, specificity, and overall misclassification error estimates. Questionnaire scores were also tested using traditional validation methods, including test-retest (intraclass correlation coefficient) reliability and validity, including relationship with pulmonary function, CAT and mMRC scores, and patient self-assessment of breathing-related health during the past week (Pearson product-moment correlation coefficient). Analysis of covariance was used to examine scores by Global Initiative for Chronic Obstructive Lung Disease (GOLD) and COPD Foundation (COPDF) airflow limitation categories (6, 7), controlling for sex, smoking status, age, and group-by-sex interaction. Nonsignificant control variables were removed, and the model was retested with the final variable set.

Performance properties of PEF alone were evaluated, including relationship to spirometry (Pearson product-moment correlation coefficient), as were GOLD and COPDF airflow limitation categories (6, 7). Predictive precision estimates were systematically tested using 50-ml increments stratified by sex to determine the optimal cutoff for differentiating cases from control subjects. Results were used to develop guidelines for using the questionnaire and PEF to refer patients for further diagnostic workup for COPD.

Results

Sample

Three hundred ninety-three English-speaking subjects were enrolled in the study, and 380 subjects provided spirometry data for confirmation of case versus control status (196 cases and 184 control subjects). Of these, 47 subjects exhibited spirometric values and clinical characteristics inconsistent with group assignment and were excluded from the analyses, yielding an analytical sample of 346 subjects (186 cases, 184 with peak flow; 160 control subjects, all with peak flow).

The demographic and clinical characteristics of the analytical sample, cases, and control subjects are shown in Table 1. Sample characteristics for groups comprising cases (group 1 [n = 97] and group 2 [n = 89]) and control subjects (group 3 [n = 87] and group 4 [n = 73]) are provided in Tables E1 and E2 in the online supplement.

Questionnaire Item Reduction

Using RF, the 44 candidate items were reduced to 34-item, 21-item, and finally to one 8-item and two different 5-item sets. Throughout the reduction process, the item sets maintained good performance, consistently misclassifying less than 27% of cases and control subjects, and with sensitivity greater than 80% and specificity greater than 70% (Figure E1A). Segregating cases from control subjects with no COPD was more precise, misclassifying less than 14% of subjects and with sensitivity greater than 85% and specificity greater than 88% (Figure E1B). All estimates improved with PEF in the model. Content coverage for the final three candidate sets is shown in Table E3.

Final Questionnaire

The final questionnaire selected for further testing, CAPTURE (COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk), is shown in Figure 1. Various scoring algorithms were tested, including weighted and unweighted summation, with clinical use in mind (efficient and precise). The selected algorithm is a simple summation of patient responses to each of the five items, yielding a questionnaire score ranging from 0 ("no" to all 5 questions) to 6 ("yes" to all questions and at least two respiratory events during the past year). Score distributions for cases and control subjects and by group are provided in Table E4. Precision estimates for two scoring thresholds when using the questionnaire alone, without PEF, are shown in Table E5, for clinicians who wish to use only the questionnaire to predict the need for spirometry.

The receiver operating characteristic curve and AUC for CAPTURE scores are shown in Figure 2. The performance of

Table 1. Demographic and Clinical Characteristics

Characteristic	Analytic Sample* (*n* = *346*)	Cases (*n* = *186*)	Control Subjects (*n* = *160*)
Age, yr			
Mean (SD)	62.7 (10.1)	64.0 (9.7)	61 (10.5)
Range	40–88	42–88	40–88
Male sex, n (%)	154 (45)	88 (47)	66 (41)
Ethnic background, n (%)†			
Hispanic or Latino‡	7 (2)	5 (3)	2 (1)
Not Hispanic or Latino	325 (94)	173 (93)	152 (95)
Racial background, n (%)†			
White	299 (86)	160 (86)	139 (87)
Black	34 (10)	18 (10)	16 (10)
American Indian, Alaska Native, Asian, or other	13 (4)	8 (4)	5 (3)
Employment, n (%)			
Employed full- or part-time	118 (34)	48 (26)	70 (44)
Retired	137 (40)	78 (42)	59 (37)
Disabled	69 (19)	48 (26)	21 (13)
Other§	22 (6)	12 (6)	10 (6)
Education level, n (%)			
High school or less	143 (41)	89 (48)	54 (34)
Some college, vocational training	76 (22)	40 (22)	36 (23)
College degree or higher	127 (37)	57 (31)	70 (44)
Smoking history, n (%)			
Never or <100 cigarettes	60 (18)	7 (4)	53 (33)
Former	196 (57)	120 (65)	76 (48)
Current	90 (26)	59 (32)	31 (19)
Spirometry, mean (SD)			
FEV_1	1.7 (0.82)	1.2 (0.47)	2.3 (0.69)
FEV_1 % predicted	61.0 (24.90)	42.5 (14.20)	82.5 (15.67)
FEV_1/FVC	0.6 (0.17)	0.5 (0.13)	0.7 (0.11)
GOLD classification, airflow limitation, n (%)			
No COPD	87 (25)	0 (0)	87 (54)
GOLD 1/2, mild/moderate	131 (38)	58 (31)	73 (46)
GOLD 3, severe	90 (26)	90 (48)	0 (0)
GOLD 4, very severe	38 (11)	38 (20)	0 (0)
COPD Foundation classification, n (%)			
SG0: normal	68 (20)	0 (0)	68 (43)
SG1: mild	86 (25)	13 (7)	73 (46)
SG2: moderate	135 (39)	135 (73)	0 (0)
SG3: severe	38 (11)	38 (20)	0 (0)
SGU: undefined	19 (6)	0 (0)	19 (12)
CAT, mean (SD)	15.2 (9.6)	19.7 (8.4)	10.1 (8.4)
mMRC, mode (%)	1 (34)	1 (39)	0 (56)
Comorbid health conditions (any), n (%)	317 (92)	168 (90)	149 (93)
Self-report activity on most days, n %			
Sit or lie down most of the day	65 (19)	49 (26)	16 (10)
Very active or exercise	153 (44)	63 (34)	90 (56)

Definition of abbreviations: CAT = COPD Assessment Test; COPD = chronic obstructive pulmonary disease; GOLD = Global Initiative for Chronic Obstructive Lung Disease; mMRC = modified Medical Research Council dyspnea scale; SG = spirometry grade.
*English-speaking, groups 1–4, with informed consent and spirometry.
†Subject self-identified.
‡Excludes Spanish language (n = 31), analyzed separately.
§Other: homemaker, unemployed, not specified.

the final, recommended threshold for CAPTURE alone is shown in Table 2.

Test-retest reliability (intraclass correlation coefficient) was 0.85 (n = 111). CAPTURE scores were significantly related to spirometry, based on the following data: FEV_1 ($r = 0.47$), FEV_1 percent predicted ($r = 0.53$), and FEV_1/FVC ratio ($r = 0.50$) (all $P < 0.0001$; n = 344); and CAT ($r = 0.74$), mMRC rating ($r = 0.58$), and self-assessment of breathing-related health ($r = 0.65$) (all $P < 0.0001$; n = 346). Differentiation of GOLD ($F = 28.67$) and COPDF ($F = 29.59$) categories was also significant (both $P < 0.0001$). The Flesch-Kincaid grade level (United States), based on a combination of words, sentences, and syllables comprising the questionnaire, was determined to be 6.4 years, indicating that this new questionnaire should be comprehensible to adults with a sixth-grade education level or above (30).

PEF

PEF was significantly correlated with spirometric values ($P < 0.0001$; n = 344), including FEV_1 ($r = 0.82$), FEV_1 percent predicted ($r = 0.70$), and FEV_1/FVC ratio ($r = 0.64$), and with differentiated GOLD

For each question, place an X in the box with the answer that is best for you.
There are no right or wrong answers, only answers which are right for you.

Please answer each question	No	Yes
1. Have you ever lived or worked in a place with dirty or polluted air, smoke, second-hand smoke, or dust?	☐	☐
2. Does your breathing change with seasons, weather, or air quality?	☐	☐
3. Does your breathing make it difficult to do things such as carry heavy loads, shovel dirt or snow, jog, play tennis, or swim?	☐	☐
4. Compared to others your age, do you tire easily?	☐	☐

	0	1	2 or more
5. In the past 12 months, how many times did you miss work, school, or other activities due to a cold, bronchitis, or pneumonia?	☐	☐	☐

*COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease & Exacerbation Risk

Figure 1. The CAPTURE questionnaire (Chronic Obstructive Pulmonary Disease Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk).

(7) and COPDF categories (6) (Figure E2). PEF values distinguished cases from control subjects ($P < 0.0001$), as well as groups 1–4 ($P < 0.0001$) (Figure E3), but they were unable to differentiate patients with COPD with previous exacerbation (group 1) from COPD cases with an FEV_1 less than 60% of predicted (Group 2).

Using sensitivity and specificity data, the following cutoff scores were selected for identifying cases of clinically significant COPD using PEF alone: males, less than 350 L/min; females, less than 250 L/min (Figure E4). Sample sensitivity, specificity, and overall prediction error for these thresholds are shown in Table 2. Estimates by sex are provided in Table E6.

Questionnaire with PEF

The best method for predicting cases versus control subjects or group membership was a combination of questionnaire and PEF, where PEF is used only for midrange scores, as explained below. The performance of the questionnaire with selective use of PEF is shown in Table 2. The AUC for CAPTURE alone was inferior to that of CAPTURE with selective use of PEF ($P < 0.0001$). Similarly, the AUC for PEF alone was inferior to that of CAPTURE with selective use of PEF ($P = 0.0065$). Under our scoring scenario, patients with scores of 0 or 1 are not considered at risk of exacerbation or COPD with an FEV_1 less than 60% of predicted; they would not require further evaluation. Those with a score of 5 or 6 ("yes" answer to all items) are considered to have a high likelihood of symptomatic respiratory disease and/or exacerbation risk and should be referred for further evaluation, including spirometry. Thus, for low scores (0 or 1) or high scores (5 or 6), PEF testing is not required. Patients scoring in the middle range (2–4) undergo PEF testing, applying the 350/250 interpretation thresholds. In our sample, with four roughly equal-sized subject groups of exacerbation risk, severity risk, no COPD, and mild COPD, 52% of the subjects required PEF to determine if further diagnostic evaluation was indicated. The other 48% needed only the questionnaire.

Discussion

COPD leads to substantial morbidity and mortality worldwide, appears to be greatly underdiagnosed, and is frequently first diagnosed after significant loss of lung function or at the time of an exacerbation. Earlier detection of COPD in patients most likely to benefit from current therapies could lead to improvement in short- and long-term patient outcomes (31). Although spirometry is the diagnostic gold standard (7), it is often perceived as time-consuming and difficult to implement in primary care settings (32–34). Even the availability of less expensive and easily used spirometers, such as those used in the Burden of Obstructive Lung Disease Study (35), has not resulted in increased use of testing in primary care settings (9). Using PEF to screen all patients in primary care is an unrealistic and expensive expectation, requiring supplies, staff time, and sufficiently careful execution to yield reliable results. Furthermore, neither spirometry nor PEF assesses clinical manifestations of disease, such as symptoms, impact, or exacerbation history. The USPSTF noted that few data exist to support the widespread use of COPD case finding and that improved clinical outcomes have not been established through their use (8). Existing questionnaires have generally identified a high proportion of patients with mild disease (9–17). No methodology has been designed explicitly for the identification of patients with undiagnosed COPD who are most likely to benefit from currently available therapies (18). Identifying these patients was recommended by an NHLBI task force as an ideal, initial stage in systematically confirming the positive impact of COPD case finding in primary care (22). We used an innovative, multimethod approach to develop a case-finding methodology that uses a brief, patient self-administered questionnaire as an initial screen, with PEF performed on a subset of patients with positive questionnaire results, to determine which patients should be referred for further diagnostic evaluation for COPD.

Our questionnaire development method used data mining to select the best, smallest set of items from among a comprehensive list of candidate items derived from the literature (18), analyses of existing data (24), and qualitative research (23). This approach was unique for several reasons. First, it included focus groups and interviews with people from the target population to inform the content and wording of the candidate questions. Second, we generated a comprehensive item pool based on the literature, existing data, and patient insight, with all questions treated as viable candidates for the final instrument.

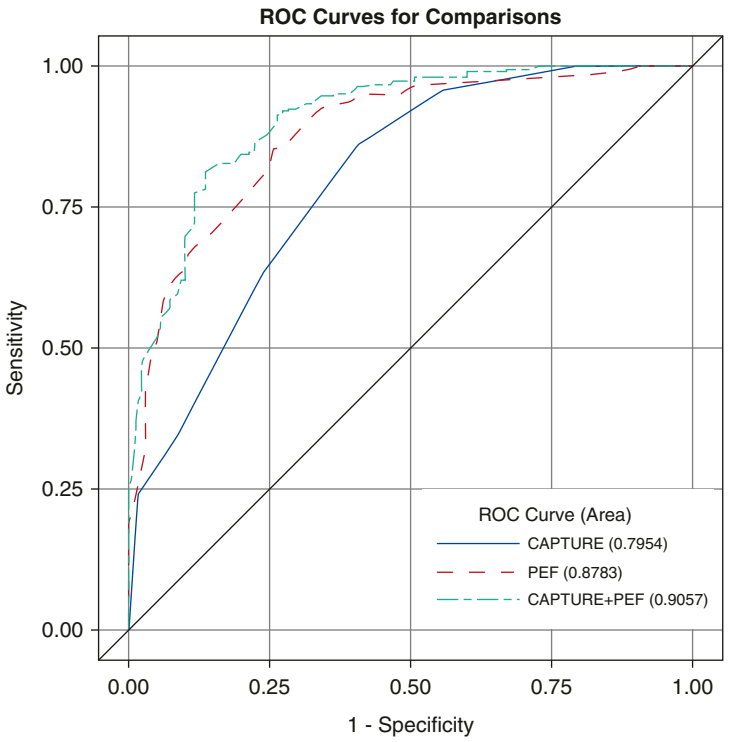

Figure 2. Receiver operating characteristic (ROC) curve and area under the curve statistics for differentiating cases from control subjects using CAPTURE alone, peak expiratory flow (PEF) alone, and PEF + CAPTURE. CAPTURE = Chronic Obstructive Pulmonary Disease Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk.

Third, we used RF analysis, rather than bivariate or multivariate statistical models, for item selection. RF is completely model-free; it does not presume any distribution for predictor variables, linear or otherwise, and takes into account the entanglement, or hidden interactions, of predictor variables, which would otherwise have to be poorly specified or missed completely. This method uses the full complexity of the data file, enabling patterns to emerge that traditional techniques would overlook.

CAPTURE is a short, five-item questionnaire that can be easily completed by patients in primary care settings prior to or during a clinic visit. Simple yes-or-no questions and a summated scoring algorithm are used to identify individuals who may have undiagnosed, clinically significant COPD, with PEF furthering the accuracy of case identification. Importantly, the patient-centered item reduction process we used resulted in a questionnaire with content that differs from previous screening or case-finding tools (17, 18). Specifically, it does *not* include an explicit question on smoking history. Rather, it asks about exposure, extending the risk assessment beyond smoking to occupational and environmental history that can increase the risk of COPD. This does not preclude clinicians from asking smokers to complete the questionnaire, but rather assesses risk beyond smoking, which is likely to be particularly useful for those in high-risk settings. Seasonal or daily variation in breathing, the impact of shortness of breath on activity, easy fatigability, and the number of activities missed because of a respiratory event in the previous year complete the assessment. These items are understandable and meaningful to patients (23) and, taken together, yield important information for clinicians on risk factors for and health effects of lung disease that could be allayed through education and treatment.

We propose a case-finding methodology that integrates a simple, self-administered questionnaire with the selective use of PEF measurement to optimize sensitivity, specificity, and efficiency. Individuals with midrange CAPTURE scores (2–4) undergo PEF measurement using a familiar, inexpensive mechanical device for a quick clinical assessment of airflow obstruction using thresholds for easy interpretation. We chose PEF measurement based on previous research (19) and known difficulties with establishing spirometry in primary care settings (36). Previous investigators have used PEF in a broader fashion to identify airflow obstruction (37, 38). In our study, PEF was remarkably sensitive and specific for differentiating cases from control subjects, with precision improving when control subjects were limited to subjects with no COPD. Our estimates with a simple mechanical meter were as good as or better than those in previous studies of handheld flow meters (17) and simpler than results obtained with diagnostic-quality spirometers (38). It is unrealistic to propose or expect PEF to be used as a screening tool in primary care settings for all patients. More importantly, PEF does not address exacerbation risk or symptomatic manifestations of disease, unless the risk or symptoms coincide with airway obstruction. The latter has not proven to fully be the case (39). Our approach begins with a simple questionnaire patients can complete independently, at home or in the office, with the results easily scored and interpreted by the clinician. PEF adds precision to the case identification process, but it is performed only as needed. The combination of our questionnaire and PEF exhibited improved operating characteristics than either alone. The sensitivity of CAPTURE scores will permit fewer missed cases of clinically significant COPD, and the higher levels of specificity provided by PEF will result in fewer false positives and lower overall screening costs. The operating characteristics of our approach, which should minimize overdiagnosis, are improved over the majority of previous strategies (17, 18, 40); the others with similar or slightly better characteristics were developed in higher-risk populations (41, 42).

We elected to develop an approach optimized for the identification of undiagnosed patients with significant airflow obstruction and/or exacerbation risk (i.e., those most likely to benefit from currently available therapies and included in recent therapeutic algorithms) (6, 7, 21). This flows from the recommendations of an NHLBI task force that suggested the identification of these patients as an initial

Table 2. Predictive Performance of Peak Expiratory Flow Alone, Questionnaire Alone, and Questionnaire plus Selective Use of Peak Expiratory Flow

Measure	Cases (Groups 1 and 2) vs. Control Subjects (Groups 3 and 4) (*n* = *346*)	Cases (Groups 1 and 2) vs. No COPD (Group *3*) (*n* = *273*)	COPD with Exacerbation (Group 1) vs. Control Subjects (Groups 3 and 4) (*n* = *257*)	COPD with FEV_1 <60% Predicted (Group 2) vs. Control Subjects (Groups 3 and 4) (*n* = *249*)
PEF alone; threshold (males, <350 L/min; females, <250 L/min)				
Sensitivity	88.0%	88.0%	91.7%	84.1%
Specificity	77.5%	90.8%	77.5%	77.5%
Overall error	16.9%	11.1%	17.2%	20.2%
CAPTURE alone; score ⩾2				
Sensitivity	95.7%	95.7%	96.9%	94.4%
Specificity	44.4%	67.8%	44.4%	44.4%
Overall error	28.0%	13.2%	35.8%	37.8%
CAPTURE + PEF; score 0–1 = control; score 5–6 = case; score 2, 3, or 4 = PEF used for group assignment				
Sensitivity	89.7%	89.7%	93.8%	85.2%
Specificity	78.1%	93.1%	78.1%	78.1%
Overall error	15.7%	9.2%	16.0%	19.4%

Definition of abbreviations: CAPTURE = COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk; COPD = chronic obstructive pulmonary disease; PEF = peak expiratory flow.

stage in confirming the positive impact of COPD case finding in primary care (22). We did not specifically attempt to separate these two groups of patients with COPD who would benefit from current therapies, nor did we attempt to generate a severity measure. Importantly, our approach should be viewed not as a diagnostic test but as a case-finding approach to identify patients who should undergo additional, definitive diagnostic testing (7).

We did not attempt to develop a case-finding approach to identify all patients with COPD. A group of smokers with symptoms and adverse clinical outcomes but without airflow obstruction has been identified (43). The role of therapy in these patients remains unclear. Treatments that could prevent COPD progression would be a major advance and would support identification of individuals with early disease (20). The extent to which CAPTURE and PEF would be useful for this purpose remains to be determined.

Several limitations of this study should be noted. First, sites included pulmonary clinics in addition to primary care settings, although primary clinicians were engaged at all specialty centers in identifying appropriate study subjects. Further study is needed to assure generalizability to patients in a broader number of primary care practices. Second, experienced clinical research personnel administered PEF and spirometry. The feasibility and precision of administering these tests as a complement to a simple, patient self-administered questionnaire should be evaluated in a variety of primary care clinical settings. Third, we enrolled a limited number of patients with mild airflow obstruction and exacerbation risk. Researchers in future studies should adequately sample this population because cohort studies have suggested that patients with COPD with lesser airflow obstruction (38) or nonobstructed but symptomatic smokers (43) are at risk of exacerbations. Fourth, our approach was focused on the identification of obstructive airway disease; patients with other cardiorespiratory diseases were not the target population. Fifth, RF analysis is one of many learning machines, with new ones emerging regularly. It is possible that another learning machine applied to the same data could perform better, although in our experience RF is competitive across a wide range of datasets. Finally, it is not known whether the identification of previously undiagnosed but symptomatic patients meeting our criteria would lead to earlier treatment and improved outcomes. Prospective studies testing the effects of case identification and treatment on patient outcomes are urgently needed.

Conclusions

We developed a case-finding methodology that uses five simple patient-reported questions and selective use of PEF for identifying patients in need of further diagnostic evaluation for COPD, initially focusing on those most likely to benefit from treatment. Results of the development work suggest that this method is sensitive and specific and may offer an efficient case-finding approach for primary care. Future study of the performance properties of this method in primary care settings is warranted. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The COPD Foundation contributed peak expiratory flow meters used in this study, Christine Thompson assisted with statistical programming, and Kathryn Miller performed manuscript text editing and formatting.

High-Risk-COPD Screening Study Group members: Rebecca Copeland, B.S., University of Kentucky; Tim Dorius, M.D., University of Nebraska Medical Center; David Hengerer, B.A., Evidera; Karen Ishitani, R.N., M.S.N., Olmsted Medical Center; Patricia Jellen, R.N., M.S.N.,

New York Presbyterian Hospital; Katherine Kim, M.P.H., Evidera; Marge Kurland, R.N., Olmsted Medical Center; James Melson, R.N., B.S.N., University of Nebraska Medical Center; Randel Plant, B.A., COPD Foundation; Wilson Quezada, M.D., Columbia University; Christina Schnell, B.A., C.C.R.C., National Jewish Health; Jason Shiffermiller, M.D., M.P.H., University of Nebraska Medical Center; Sonja Stringer, M.P.H., Evidera; Deb Sumnick, P.B.T., University of Nebraska; Kyle Textor, B.A., Olmsted Medical Center; Jennifer Underwood, C.C.R.C., National Jewish Health; and Beth Whippo, R.N.-B.C., M.S.N., New York Presbyterian Hospital.

References

1. Ford ES, Mannino DM, Wheaton AG, Giles WH, Presley-Cantrell L, Croft JB. Trends in the prevalence of obstructive and restrictive lung function among adults in the United States: findings from the National Health and Nutrition Examination surveys from 1988-1994 to 2007-2010. *Chest* 2013;143:1395–1406.
2. Miravitlles M, Soriano JB, García-Río F, Muñoz L, Duran-Tauleria E, Sanchez G, Sobradillo V, Ancochea J. Prevalence of COPD in Spain: impact of undiagnosed COPD on quality of life and daily life activities. *Thorax* 2009;64:863–868.
3. Martinez CH, Mannino DM, Jaimes FA, Curtis JL, Han MK, Hansel NN, Diaz AA. Undiagnosed obstructive lung disease in the United States: associated factors and long-term mortality. *Ann Am Thorac Soc* 2015;12:1788–1795.
4. Labonté LE, Tan WC, Li PZ, Mancino P, Aaron SD, Benedetti A, Chapman KR, Cowie R, FitzGerald JM, Hernandez P, *et al*.; Canadian Respiratory Research Network, CanCOLD Collaborative Research Group. Undiagnosed chronic obstructive pulmonary disease contributes to the burden of health care use: data from the CanCOLD Study. *Am J Respir Crit Care Med* 2016;194:285–298.
5. Qaseem A, Snow V, Shekelle P, Sherif K, Wilt TJ, Weinberger S, Owens DK; Clinical Efficacy Assessment Subcommittee of the American College of Physicians. Diagnosis and management of stable chronic obstructive pulmonary disease: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2007;147:633–638.
6. Rennard S, Thomashow B, Crapo J, Yawn B, McIvor A, Cerreta S, Walsh J, Mannino D. Introducing the COPD Foundation Guide for Diagnosis and Management of COPD, recommendations of the COPD Foundation. *COPD* 2013;10:378–389.
7. Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, Barnes PJ, Fabbri LM, Martinez FJ, Nishimura M, *et al*. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013;187:347–365.
8. Guirguis-Blake JM, Senger CA, Webber EM, Mularski RA, Whitlock EP. Screening for chronic obstructive pulmonary disease: evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 2016;315:1378–1393.
9. Freeman D, Nordyke RJ, Isonaka S, Nonikov DV, Maroni JM, Price D, Halbert RJ. Questions for COPD diagnostic screening in a primary care setting. *Respir Med* 2005;99:1311–1318.
10. Hanania NA, Mannino DM, Yawn BP, Mapel DW, Martinez FJ, Donohue JF, Kosinski M, Rendas-Baum R, Mintz M, Samuels S, *et al*. Predicting risk of airflow obstruction in primary care: validation of the Lung Function Questionnaire (LFQ). *Respir Med* 2010;104: 1160–1170.
11. Martinez FJ, Raczek AE, Seifer FD, Conoscenti CS, Curtice TG, D'Eletto T, Cote C, Hawkins C, Phillips AL; COPD-PS Clinician Working Group. Development and initial validation of a self-scored COPD Population Screener Questionnaire (COPD-PS). *COPD* 2008;5:85–95.
12. Price DB, Tinkelman DG, Halbert RJ, Nordyke RJ, Isonaka S, Nonikov D, Juniper EF, Freeman D, Hausen T, Levy ML, *et al*. Symptom-based questionnaire for identifying COPD in smokers. *Respiration* 2006;73:285–295.
13. Price DB, Tinkelman DG, Nordyke RJ, Isonaka S, Halbert RJ; COPD Questionnaire Study Group. Scoring system and clinical application of COPD diagnostic questionnaires. *Chest* 2006;129:1531–1539.
14. Raghavan N, Lam YM, Webb KA, Guenette JA, Amornputtisathaporn N, Raghavan R, Tan WC, Bourbeau J, O'Donnell DE. Components of the COPD Assessment Test (CAT) associated with a diagnosis of COPD in a random population sample. *COPD* 2012;9:175–183.
15. Yawn BP, Mapel DW, Mannino DM, Martinez FJ, Donohue JF, Hanania NA, Kosinski M, Rendas-Baum R, Mintz M, Samuels S, *et al*.; Lung Function Questionnaire Working Group. Development of the Lung Function Questionnaire (LFQ) to identify airflow obstruction. *Int J Chron Obstruct Pulmon Dis* 2010;5:1–10.
16. Kotz D, Nelemans P, van Schayck CP, Wesseling GJ. External validation of a COPD diagnostic questionnaire. *Eur Respir J* 2008;31:298–303.
17. Haroon S, Jordan R, Takwoingi Y, Adab P. Diagnostic accuracy of screening tests for COPD: a systematic review and meta-analysis. *BMJ Open* 2015;5:e008133.
18. Han MK, Steenrod AW, Bacci ED, Leidy NK, Mannino DM, Thomashow BM, Barr RG, Make BJ, Bowler RP, Rennard SI, *et al*. Identifying patients with undiagnosed COPD in primary care settings: insight from screening tools and epidemiologic studies. *Chronic Obstr Pulm Dis* 2015;2:103–121.
19. Nelson SB, LaVange LM, Nie Y, Walsh JW, Enright PL, Martinez FJ, Mannino DM, Thomashow BM. Questionnaires and pocket spirometers provide an alternative approach for COPD screening in the general population. *Chest* 2012;142:358–366.
20. Rennard SI, Drummond MB. Early chronic obstructive pulmonary disease: definition, assessment, and prevention. *Lancet* 2015;385:1778–1788.
21. Criner GJ, Bourbeau J, Diekemper RL, Ouellette DR, Goodridge D, Hernandez P, Curren K, Balter MS, Bhutani M, Camp PG, *et al*. Executive summary: Prevention of acute exacerbation of COPD: American College of Chest Physicians and Canadian Thoracic Society guideline. *Chest* 2015;147:883–893.
22. National Heart, Lung, and Blood Institute. A case-finding strategy for moderate-to-severe COPD in the United States; 2009 [accessed 2016 July 4]. Available from: http://www.nhlbi.nih.gov/research/reports/2009-case-finding-exesum
23. Leidy NK, Kim K, Bacci ED, Yawn BP, Mannino DM, Thomashow BM, Barr RG, Rennard SI, Houfek JF, Han MK, *et al*.; High-Risk-COPD Screening Study Group. Identifying cases of undiagnosed, clinically significant COPD in primary care: qualitative insight from patients in the target population. *NPJ Prim Care Respir Med* 2015;25:15024.
24. Leidy NK, Malley KG, Steenrod AW, Mannino DM, Make BJ, Bowler RP, Thomashow BM, Barr RG, Rennard SI, Houfek JF, *et al*. Insight into best variables for COPD case identification: a random forests analysis. *Chronic Obstr Pulm Dis* 2016;3:406–418.
25. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009;34:648–654.
26. Ferris BG. Epidemiology Standardization Project (American Thoracic Society). *Am Rev Respir Dis* 1978;118:1–120.
27. Malley JD, Malley KG, Pajevic S. Statistical learning for biomedical data. Cambridge, UK: Cambridge University Press; 2011.
28. The Comprehensive R Archive Network; 2014 [accessed 2016 Mar 24]. Available from: http://cran.us.r-project.org/
29. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–845.
30. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for Navy enlisted personnel. Chief of Naval Technical Training, Naval Air Station Memphis; 1975. Research Branch Report 8–75.
31. Ramsey SD, Sullivan SD. Chronic obstructive pulmonary disease: is there a case for early intervention? *Am J Med* 2004;117(Suppl 12A):3S–10S.
32. Ferguson GT, Enright PL, Buist AS, Higgins MW. Office spirometry for lung health assessment in adults: a consensus statement from the National Lung Health Education Program. *Chest* 2000;117:1146–1161.
33. Han MK, Kim MG, Mardon R, Renner P, Sullivan S, Diette GB, Martinez FJ. Spirometry utilization for COPD: how do we measure up? *Chest* 2007; 132:403–409.

34. Joo MJ, Lee TA, Weiss KB. Geographic variation of spirometry use in newly diagnosed COPD. *Chest* 2008;134:38–45.
35. Buist AS, McBurnie MA, Vollmer WM, Gillespie S, Burney P, Mannino DM, Menezes AM, Sullivan SD, Lee TA, Weiss KB, *et al*.; BOLD Collaborative Research Group. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet* 2007;370:741–750.
36. Enright PL. Should we keep pushing for a spirometer in every doctor's office? Respir Care 2012;57:146–153.
37. Mahboub B, Alzaabi A, Soriano JB, Salameh L, Mutairi YA, Yusufali AA, Alsheikh-ali A, Almahmeed W, Haughney J. Case-finding of chronic obstructive pulmonary disease with questionnaire, peak flow measurements and spirometry: a cross-sectional study. *BMC Res Notes* 2014;7:241.
38. Perez-Padilla R, Vollmer WM, Vázquez-García JC, Enright PL, Menezes AM, Buist AS; BOLD and PLATINO Study Groups. Can a normal peak expiratory flow exclude severe chronic obstructive pulmonary disease? *Int J Tuberc Lung Dis* 2009;13:387–393.
39. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, Miller B, Lomas DA, Agusti A, Macnee W, *et al*.; Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) Investigators. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;363: 1128–1138.
40. Guirguis-Blake JM, Senger CA, Webber EM, Mularski R, Whitlock EP. Screening for chronic obstructive pulmonary disease: a systematic evidence review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 130. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ); 2015. AHRQ Publication No. 14-05205-EF-1.
41. Zarowitz BJ, O'Shea T, Lefkovitz A, Peterson EL. Development and validation of a screening tool for chronic obstructive pulmonary disease in nursing home residents. *J Am Med Dir Assoc* 2011;12: 668–674.
42. Müllerová H, Wedzicha J, Soriano JB, Vestbo J. Validation of a chronic obstructive pulmonary disease screening questionnaire for population surveys. *Respir Med* 2004;98:78–83.
43. Woodruff PG, Barr RG, Bleecker E, Christenson SA, Couper D, Curtis JL, Gouskova NA, Hansel NN, Hoffman EA, Kanner RE, *et al*.; SPIROMICS Research Group. Clinical significance of symptoms in smokers with preserved pulmonary function. *N Engl J Med* 2016;374: 1811–1821.