EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original article

Population-Based Case-Finding to Identify Subjects with Undiagnosed Asthma or COPD

Matthew Preteroti, G. A. Whitmore, Katherine L. Vandemheen, J. Mark FitzGerald, Catherine Lemière, Louis-Philippe Boulet, Erika Penz, Stephen K. Field, Samir Gupta, R. Andrew McIvor, Irvin Mayers, Paul Hernandez, Diane Lougheed, Martha Ainslie, Christopher Licskai, Tanweer Azher, Ian Fraser, Masoud Mahdavian, Shawn D. Aaron

Please cite this article as: Preteroti M, Whitmore GA, Vandemheen KL, *et al*. Population-Based Case-Finding to Identify Subjects with Undiagnosed Asthma or COPD. *Eur Respir J* 2020; in press (https://doi.org/10.1183/13993003.00024-2020).

Population-Based Case-Finding to Identify Subjects with Undiagnosed Asthma or COPD

Short Title: Undiagnosed COPD and Asthma Population Study

Matthew Preteroti, BScH

G. A. Whitmore, PhD

Katherine L. Vandemheen, MScN

J. Mark FitzGerald, MD

Catherine Lemière, MD

Louis-Philippe Boulet, MD

Erika Penz, MD

Stephen K. Field, MD

Samir Gupta, MD

R. Andrew McIvor, MD

Irvin Mayers, MD

Paul Hernandez, MD

Diane Lougheed, MD

Martha Ainslie, MD

Christopher Licskai, MD

Tanweer Azher, MD

Ian Fraser, MD

Masoud Mahdavian, MD

Shawn D. Aaron, MD

The Ottawa Hospital Research Institute, University of Ottawa, (MP, KLV, SDA); Desautels Faculty of Management, McGill University, (GAW); Department of Medicine, The University of British Columbia

(JMF); Department of Medicine, Université de Montreal (CL); Centre de recherche, Hôpital Laval, Université Laval (LPB); Department of Medicine, University of Saskatchewan (EP); Cumming School of Medicine, University of Calgary (SKF); Department of Medicine and the Li Ka Shing Knowledge Institute of St. Michael's Hospital University of Toronto (SG); Firestone Institute for Respiratory Health, McMaster University (RAM); Department of Medicine, University of Alberta (IM); Department of Medicine, Dalhousie University (PH); Department of Medicine, Queen's University (DL); Department of Medicine, University of Manitoba (MA); Department of Medicine, University of Western Ontario (CL); Department of Medicine, Memorial University (TA), Michael Garron Hospital, Department of Medicine, University of Toronto (IF); Royal Victoria Regional Health Centre (MM); all in Canada.

Correspondence and request for reprints:

Dr. Shawn Aaron

The Ottawa Hospital, General Campus

501 Smyth Road

Ottawa, Ontario, Canada

K1H 8L6

Telephone: (613) 739-6636

Fax: (613) 739-6807

Email: **saaron@ohri.ca**

Conception and design: SDA, GAW.

Analysis and interpretation: MP, GAW, SDA.

Drafting the manuscript for important intellectual content: MP, GAW, SDA, KLV, JMG, CL, LPB, EP, SKF, SG, RAM, IM, PH, DL, MA, CL, TA, IF, MM.

Word count: 3643

Keywords: asthma, COPD, case-finding, undiagnosed, questionnaire, obstructive lung disease.

Funding: Canadian Institutes of Health Research- FDN grant #154322

Background: Approximately 5-10% of adults may have undiagnosed airflow obstruction. The objective of this study was to develop a population-based case-finding strategy to assess the prevalence of undiagnosed airflow obstruction (asthma or COPD) amongst adults with respiratory symptoms in Canada.

Methods: Adults without a previous history of asthma, COPD, or lung disease were recruited by random digit-dialing and asked if they had symptoms of dyspnea, cough, sputum or wheeze within the past 6 months. Those who answered affirmatively completed the Asthma Screening Questionnaire (ASQ), COPD-Diagnostic Questionnaire (COPD-DQ), and COPD Assessment Test (CAT). Those with ASQ score ≥ 6 or COPD-DQ score ≥ 20 underwent pre and post bronchodilator spirometry to diagnose asthma or COPD.

Results: 12,117 individuals were contacted at home and assessed for study eligibility. Of 1260 eligible individuals, 910 (72%) enrolled and underwent spirometry. Ultimately, 184 subjects (20% of those enrolled) had obstructive lung disease (73 asthma and 111 COPD). Individuals found to have undiagnosed asthma or COPD had more severe respiratory symptoms and impaired quality of life compared to those without airflow obstruction. The ASQ, COPD-DQ, and CAT had ROC areas for predicting undiagnosed asthma or COPD of 0.49, 0.64 and 0.56 respectively. Four descriptive variables; age, BMI, sex and pack-years smoked, produced better ROC values than the questionnaires (ROC area= 0.68).

Conclusion: Twenty percent of randomly-selected individuals who report respiratory symptoms in Canada have undiagnosed airflow obstruction due to asthma or COPD. Questionnaires could exclude subjects at low risk but lacked ability to accurately find subjects with undiagnosed disease.

Introduction:

Within the North American population, 8% of adults have been diagnosed with asthma and 6% have been diagnosed with COPD[1-3]. Collectively, these diseases cost North Americans in excess of 90 billion USD in direct and indirect costs every year [4,5]. However, physician-diagnosed asthma and COPD represent only a fraction of individuals living with these diseases. Population-based studies suggest that the prevalence of undiagnosed obstructive lung disease (OLD) may be double that of diagnosed disease[6].

A clinical diagnosis of asthma or COPD requires that a patient exhibit respiratory symptoms and have evidence of physiologic airflow limitation on spirometry. For asthma there should be evidence of variable airflow obstruction or airway hyper-responsiveness, whereas a diagnosis of COPD requires evidence of persistent post-bronchodilator airflow obstruction [7,8]. Spirometry is the most important test used to establish a diagnosis, however, studies have demonstrated that spirometry is often not integrated into clinical practice in North American communities[9,10]. This under-utilization of spirometry leads to under-diagnosis of individuals living with the disease. Under-diagnosis of asthma and COPD can also occur if individuals living with undiagnosed asthma or COPD fail to seek medical attention for their respiratory symptoms, or if they have limited access to medical care[11-13].

The US Preventive Services Task Force (USPSTF) published a systematic review assessing published evidence on the benefits and risks of screening for COPD among asymptomatic adults[14]. The task force recommended "against screening for COPD in asymptomatic adults" [15]." However, the USPSTF noted that this recommendation applies only to asymptomatic adults. Critical to the assessment is the USPSTF's conclusion that current medical treatments for COPD reduce symptoms and respiratory exacerbations. Thus, even though the USPSTF recommended against screening asymptomatic adults, its report pointed out that this recommendation is not applicable to "at-risk persons who present to clinicians with symptoms." The USPSTF Recommendation Statement also "encourages clinicians… to pursue active case-finding for COPD in patients with risk factors, such as exposure to cigarette smoke or respiratory symptoms"[15].

Case-finding studies to identify undiagnosed asthma in the community are relatively sparse relative to studies for COPD. Nevertheless, undiagnosed asthma remains a prevalent, important public health problem. A Danish study of 1149 subjects who reported respiratory symptoms found 493 (43%) with definite asthma, and of these 249 (51%) were undiagnosed[16]. Studies have also shown impaired quality of life and higher healthcare utilization in adult individuals with undiagnosed asthma compared to those without asthma[17].

Current task force recommendations require consideration of the difference between screening, ie, testing large numbers of apparently healthy people to detect unrecognized disease at an earlier stage; and case-finding, ie, evaluating subgroups of people at increased risk of having a disease to make a diagnosis earlier than would occur by waiting for them to present with symptoms or signs. A case-finding approach such as used in this study specifically targets those people who have symptoms they have not reported, perhaps because they have not been asked.

A systematic review was conducted by the US Preventive Services Task Force (USPSTF)[14,15] to assess tools for identification of undiagnosed COPD. The review determined that the COPD-Diagnostic Questionnaire (COPD-DQ) is the most extensively studied questionnaire (5 studies, n=3048). The questionnaire score ranges from 0 to 38. At a threshold score of 20, sensitivity and specificity values are 63.0% and 70.1% respectively [18,19]. Case-finding methodologies for undiagnosed asthma are sparse. The Asthma Screening Questionnaire (ASQ) is a 6-item questionnaire that contains questions pertaining to cough, wheeze, shortness of breath, and chest tightness provoked by activities such as laughing, physical activity, or talking on the phone[18]. The ASQ score range is 0 to 20. The sensitivity and specificity of the ASQ at various threshold values has not been studied [20].

The objectives of this study were to develop a population-based case-finding strategy to identify randomly-selected community-dwelling individuals with undiagnosed airflow obstruction in order to assess the prevalence of undiagnosed asthma or COPD amongst adults with respiratory symptoms in Canada. Secondary objectives were to assess the ability of the available case-finding questionnaires to accurately identify subjects with undiagnosed disease, and to determine whether a new composite questionnaire could improve predictability. Our study explicitly excluded subjects with a previous history of asthma, COPD, or other lung diseases, since the intent was to discover new cases of previously undiagnosed asthma or COPD.

Methods:

Participant Recruitment

Adults ≥18 years of age were recruited in a two-step process from June 19 2017 to May 16 2019 using random digit dialing of land lines and cell phones located within a 90-minute radius of the 10 most populous metropolitan areas in Canada. This technique has been previously used to recruit a random sample from the Canadian urban and rural population[21]. A sample initial contact telephone script follows:

"Hello, my name is Amanda and I'm calling from The Ottawa Hospital. We are conducting a research study and you would help us greatly by answering one question only; this will take less than 15 seconds of your time. We are looking to call back individuals who have experienced breathing problems within the past 6 months. Symptoms of potential breathing problems include shortness of breath, wheezing, increased mucus or sputum, or prolonged cough. Is there anyone in your home who is 18 years of age or older and has experienced one or more of these symptoms in the past 6 months?"

Those households who responded affirmatively received a call back from the local study coordinator and the individual within the household who was identified as having respiratory symptoms was verbally consented and screened for study entry over the telephone. Subjects were excluded from participating if they had: 1) a previous physician diagnosis of asthma, COPD, cystic fibrosis, bronchiectasis, pulmonary fibrosis, or lung cancer; 2) history of use of an inhaled respiratory medication other than as-needed salbutamol or inhaled nasal medications; 3) contraindications to spirometry: including history of MI, stroke, aortic/cerebral aneurysm, eye surgery or detached retina within the previous 3 months; 4) unable/refusal to provide informed consent; 5) third trimester of pregnancy, or 6) age < 18 years. All subjects determined to be eligible completed the ASQ via telephone. Those subjects ≥ 60 years old, as well as subjects < 60 years old who scored < 6 points on the ASQ were also administered the COPD-DQ. Those participants who scored ≥ 20 points on the COPD-DQ or ≥ 6 points on the ASQ were invited to visit the study site for spirometry.

Study Procedure

Once they arrived at the study site, subjects signed written informed consent and then all subjects proceeded to undergo pre and post bronchodilator spirometry administered by certified study personnel to confirm the presence of an obstructive lung disease. Asthma was diagnosed in subjects whose FEV1 improved by ≥ 12% and ≥ 200 ml following bronchodilator administration with 400 ug of salbutamol[7]. COPD was diagnosed in subjects whose post-bronchodilator FEV1/FVC ratio remained below the lower 95% confidence limit of normal (the lower 95% confidence limit for a healthy normal, adjusted for sex, age, and height)[8]. Subjects who met spirometry criteria for both conditions were considered to have COPD with some bronchodilator reversibility and were classified as COPD for the purposes of the present analysis.

The COPD Assessment Test (CAT) was administered to all participants to measure respiratory symptoms. The CAT is a validated questionnaire composed of 8 questions about the severity of respiratory symptoms and generates a summed total score ranging from 0 (no impact on daily activities) to 40 (very high impact on daily activities)[22]. A CAT threshold total score of >10 points has been

associated with symptomatic COPD.[23] We used The St. George's Respiratory Questionnaire (SGRQ) to assess quality of life specific for respiratory diseases.

Statistical Analysis

Descriptive measures for subjects classified with or without obstructive lung disease (OLD) were summarized using arithmetic means and standard deviations, where applicable. Receiver operating characteristic (ROC) curves for the disease outcomes of asthma, COPD, and OLD were computed for each of the ASQ, COPD-DQ and CAT questionnaires using published scoring formulas for these instruments. Areas under their ROC curves (AUC) were also computed using non-parametric procedures. Sensitivity and specificity values for the questionnaires were calculated using recommended diagnostic threshold scores reported in the literature, as cited earlier.

Multivariate stepwise logistic regression was used to find the combination of questions in the ASQ and CAT questionnaires which best predicted an outcome of OLD. The COPD-DQ questionnaire was not used in this analysis because this questionnaire was only administered to a subset of 499 subjects. Four descriptive variables, age, body mass index (BMI), sex and pack-years smoked, were included in the pool of candidate variables. One question from the ASQ was removed from the candidate pool because it was judged to lack clinical face validity- subjects who reported coughing when laughing were found to be less likely to have obstructive lung disease, and this question was thus removed. The stepwise procedure used entry and removal thresholds for P-values of 0.01 and 0.05, respectively. No variable was forced into the regression model. The set of questions selected by the stepwise logistic regression is referred to here as the 8-item questionnaire. In the first stage, stepwise regression selected the predictor variables from the candidate pool. The second stage involved taking the best predictor variables from the first stage and re-estimating the regression function for the set of selected predictors. Predicted logistic scores and estimated probabilities of disease for the 8-item questionnaire were computed. A ROC curve was generated for the 8-item regression model as well as sensitivity and specificity values for a selection of cut-points for the predicted probability of OLD.

Each of the eight CAT questions is a 0-5 rating scale for severity of a COPD symptom. Although the rating scales are implicitly ordinal (progressing from none to very severe), we discovered that the scales have non-monotonic associations with the disease outcome variable. To account for this tendency, a logistic regression function was computed for each symptom by regressing disease outcomes on subjects' rating responses. The predicted logistic score for each subject was then taken as the subject's response measure for the question. The online Supplement (Section S.3) provides additional technical detail and illustrations of this re-scoring procedure.

The OLD outcome consists of two clinically distinct component outcomes: asthma or COPD. To investigate the ability of the 8-item questionnaire to predict these component outcomes, multinomial stepwise logistic regression was used to estimate separate predictive functions for the two component outcomes. The predicted probabilities from these separate functions, when added together, give an estimate of the probability of OLD. ROC curves and AUC values for these separate predictive functions are reported.

A cross-validation of the 8-item model was conducted in which random draws of 80% of the study sample were selected to estimate the model and then the estimated model was used to predict outcomes of the remaining 20% of subjects who were withheld. The predictive check was repeated with 1000 random draws. The average predictive performance as measured by sensitivity and specificity were then assessed. All logistic regression models were also tested for goodness of fit using a Hosmer-Lemeshow test.

Statistical analysis was done using STATA 15 statistical software (StataCorp, College Station, TX, USA). The study was approved by The Ottawa Health Science Network Research Ethics Board as well as the Ethics Boards of the other 9 participating study hospitals.

Results:

Figure 1 depicts the results of the case-finding strategy. 12,117 individuals were contacted by random digit dialing and screened for entry into the study, and 10,411 were excluded. The most common reasons for exclusion were pre-existing physician-diagnosed COPD or asthma (n=3,936). Of the 1706 remaining participants 446 subjects (26%) were excluded because they did not score ≥ 6 points on the Asthma Questionnaire or ≥ 20 Points on the COPD Questionnaire. Ultimately, 1260 individuals were eligible to undergo spirometry, however 350 did not attend for spirometry or were unable to complete adequate spirometry. The remaining 910 individuals (72% of those eligible) successfully completed pre and post bronchodilator spirometry and were evaluated for a diagnosis of asthma or COPD. Of 910 participants, all but two individuals fully completed the ASQ and CAT questionnaires. According to study protocol, only 499 completed the COPD-DQ questionnaire. Of 910 symptomatic participants, 184 subjects (20%) had spirometry consistent with a diagnosis of obstructive lung disease; comprised of 73 cases of asthma (8.0%) and 111 cases of COPD (12.2%). The COPD cases include 24 who met spirometry criteria for both COPD and asthma.

Table 1 presents mean summary statistics for the 910 study subjects divided into two groups: those with OLD (n=184) and those without spirometric evidence of OLD (n=726). The group with OLD had a history of heavier smoking; mean total pack years of 31.4 versus 19.8 (P<0.001). The mean FEV1 predicted was 75.1% in the OLD group compared to 95.8% in the group without OLD (P<0.001), and the mean pre-bronchodilator FEV1/FVC ratio was significantly different between groups (0.63 versus 0.77, P<0.001). Subjects with OLD had poorer health-related quality of life compared those without OLD; mean SGRQ total score was 40.5 ± 18.5 in the group with OLD vs 37.0 ± 17.3 in those without OLD (P <0.001).

The mean COPD-DQ score in the group with OLD was 23.4 versus 20.8 in the group without OLD (P<0.001). The mean CAT score in the group with OLD was 18.2 versus 16.7 in the group without OLD (P<0.001). The mean ASQ score did not differ between groups.

Table 2 shows the values of sensitivity, specificity, and area under the ROC curve for the ASQ, COPD-DQ and CAT questionnaires when the questionnaires were used to identify the disease for which they were each designed, asthma or COPD accordingly, as well as when the questionnaires were used to identify the composite outcome of undiagnosed OLD. The areas under the ROC curve for identification of undiagnosed OLD were 0.49, 0.64, and 0.56 for the ASQ, COPD-DQ, and CAT respectively. These performances measures for the three questionnaires rely only on their individual questions and do not incorporate any descriptive variables, such as age, BMI, and sex, or pack years smoked. When the four descriptive variables (age, BMI, sex and pack years smoked) were considered alone as predictors, they produced a combined area under the ROC curve of 0.68 (Table 3). The ASQ and four descriptive variables, when considered together, had an area under the ROC curve of 0.69 for identifying subjects with undiagnosed OLD. The combined COPD-DQ and descriptive variables had an area under the ROC curve of 0.68 while the combined CAT and descriptive variables had an area under the ROC curve of 0.69 (Table 3).

Two-stage stepwise logistic regression was applied to questions found in the ASQ, CAT, and the four descriptive variables. The stepwise procedure selected eight questions (Table 4). This 8-item questionnaire demonstrated a modest improvement in identifying undiagnosed obstructive lung disease compared to the individual questionnaires (Table 3). The first four items of the 8-item questionnaire include demographic and smoking variables: age, BMI, sex, and pack years smoked. The remaining four items include two questions from the ASQ and two from the CAT (Table 4). Responses to the items are weighted by their regression coefficients and summed. The weighted sum represents a logistic score that can be converted mathematically into a probability estimate for obstructive lung disease. A cutoff of 10%

for this probability gives sensitivity and specificity values of 89% and 30% respectively (Supplemental Table S.1). The area under the ROC curve was 0.74 for the 8-item questionnaire. Supplemental Table S.2 displays a case demonstration calculating the estimated probability of obstructive lung disease for a random subject taken from our study population.

Internal cross validation of the 8-item questionnaire revealed mean sensitivity and specificity pairs of 87% and 29% respectively over 1000 repeated samples. A Hosmer-Lemeshow test showed an acceptable goodness of fit of the logistic regression model for the 8-item questionnaire (P = 0.244).

The 8-item questionnaire was evaluated against asthma and COPD disease outcomes independently. The questionnaire demonstrated an area under the ROC curve of 0.64 for identifying persons with undiagnosed asthma with only 1-item of the 8-item questionnaire having statistical significance (Table 5A). The questionnaire performed best at identifying undiagnosed COPD, with an area under the ROC curve of 0.84, and all 8 items in the questionnaire were statistically significant (Table 5B). The predictive capacity of the 8-item questionnaire is summarized in Figure 2 with the area under the ROC curve increasing from asthma, to OLD, to COPD.

Discussion:

Our case-finding strategy used population-based random sampling to identify symptomatic community-dwelling adults living with undiagnosed obstructive lung disease. Ultimately 20% of enrolled individuals were confirmed as having undiagnosed airflow obstruction due to asthma or COPD. Thus, an estimated 1 out of 5 Canadians with respiratory symptoms but without a previous diagnosis of lung disease, were found to have undiagnosed obstructive lung disease using our population-based case-finding strategy. Persons who had undiagnosed obstructive lung disease had worse lung function, a heavier smoking history, more severe respiratory symptoms, and poorer disease-specific quality of life, compared to those without obstructive lung disease.

Other approaches to case-finding COPD, but not asthma, are currently being developed [24]. A recently published study used machine-learning methods to select and validate variables most important in identifying patients with clinically significant COPD. The investigators used random forest analysis to reduce 44 identified variables down to a 5-item CAPTURE questionnaire (COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk)[25]. Using the CAPTURE questionnaire together with measurement of peak expiratory flow exhibited the best sensitivity (89.7%) and specificity (78.1%) for distinguishing COPD cases from control subjects in a primary care setting.

Similar pre-screening questionnaires have been developed in Latin America countries. The PUMA prescreening COPD questionnaire includes 7 items including gender, age, pack-years smoking, dyspnea, sputum, cough, and whether previous spirometry was performed. In a validation study the questionnaire had a ROC area under the curve of 0.73 and showed moderate accuracy for identifying subjects with COPD within a primary care setting[26]. However, both the PUMA and CAPTURE questionnaires do not address case-finding for asthma.

We used pre-existing questionnaires to identify subjects with respiratory symptoms in order to maximize our chances of finding subjects with undiagnosed obstructive lung disease. The COPD-DQ and ASQ questionnaires were used to exclude subjects at low risk, and ultimately 26% of potentially eligible subjects were excluded because both their ASQ and COPD-DQ scores fell below published thresholds associated with disease. However, neither of these 2 questionnaires, nor the CAT questionnaire, demonstrated good performance for case-finding undiagnosed obstructive lung disease. Moreover, the set of four descriptive variables (age, sex, BMI, and pack years smoked) displayed a stand-alone ability to identify undiagnosed obstructive lung disease (AUC = 0.68), that was better than the existing questionnaires. Adding the individual questionnaires (the ASQ, COPD-DQ, and CAT) to these four descriptive variables had little impact on the predictive value, with respective areas under the ROC curve of 0.69, 0.68, and 0.69. The results suggest that most of the predictive strength for identifying obstructive lung disease comes from these key demographic factors and an individual's smoking history. The screening questionnaires provide little, if any, value.

We acknowledge that the CAT is not validated as a COPD diagnostic tool and is not validated for measuring symptoms in asthmatics, nor is the COPD-DQ validated for use in diagnosing asthma. However, we used these tools in this exploratory study to determine whether we could create a new composite questionnaire that would improve prediction of undiagnosed COPD or asthma. As such we selected questions from these assessment tools that we found were highly predictive of undiagnosed asthma or COPD, and we incorporated these questions into our 8-item questionnaire.

We used stepwise logistic regression techniques to find the combination of questions in the ASQ and CAT questionnaires which best predicted an outcome of OLD. The resulting 8-item questionnaire displayed a superior ability to predict obstructive lung disease than the individual existing questionnaires. The area under the ROC curve for the 8-item questionnaire for identification of OLD was still not impressive (ROC area = 0.74). Furthermore, only 1-item of the 8-item questionnaire had significant predictive ability to identify undiagnosed asthma. These results suggest that more research is needed to build effective questionnaires for case-finding undiagnosed obstructive lung disease, especially asthma, in

communities. The challenge will be to find simple questions which can be posed to respondents and answered easily without requirement for lung function testing.

There are some limitations to our study. Our study identified more cases of undiagnosed COPD compared to undiagnosed asthma. This likely occurred because 24 individuals who met spirometry criteria for both conditions were considered to have COPD with some bronchodilator reversibility and were classified as COPD for the purposes of the present analysis, rather than as asthma-COPD overlap. It is possible that some of these subjects could have asthma, rather than COPD, and this might only become evident after treatment with bronchodilators and inhaled steroids fully resolves their airflow obstruction. This could ultimately result in re-classification of some subjects who met spirometry criteria for both conditions.

It is also possible that our study missed finding asthma in some subjects[19]. We did not employ bronchial challenge testing or FeNO testing in those who tested negative for airflow obstruction or bronchodilator reversibility (726 of 910 participants), since we wanted our case-finding strategy to be potentially translatable to community practices, and within communities' bronchial challenge testing is relatively expensive and difficult to access. Therefore, some symptomatic individuals having asthma with airway hyper-responsiveness, but without reversible airway obstruction, may have been missed. We are currently undertaking a sub-study to try to determine the prevalence of undiagnosed asthma in those subjects in our case-finding study who did not exhibit airflow obstruction or bronchodilator reversibility. Subjects who tested negative for obstructive lung disease using spirometry are being enrolled and they are undergoing bronchial challenge tests, sputum induction for eosinophils, blood eosinophilia measurements, and FeNO measurements to attempt to discover evidence of undiagnosed asthma. This sub-study is currently ongoing. Finally, individuals with asthma or COPD who were not aware of symptoms (e.g., those who had significant smoking histories and potential airflow obstruction but without symptoms) were not detected using this case-finding strategy.

Another potential limitation of our study is that over 12,000 phone calls had to be made, and 910 subjects tested with spirometry, in order to ultimately find 184 subjects with undiagnosed obstructive lung disease. The cost of this strategy, while not a focus of the current study, was considerable. Random-digit dialing is probably not a practical or affordable way to find cases of undiagnosed asthma or COPD, although it was the right research approach to ensure selection of a representative population-based sample of subjects with respiratory symptoms for this study, in order to determine prevalence of undiagnosed airflow obstruction in the population.

The results of our study bring up some important questions outside the scope of the current report, such as "why are these patients undiagnosed?" and "are there clinical benefits associated with case-finding and early diagnosis of obstructive lung disease?". We are in the process of recruiting a control group of subjects with physician-diagnosed asthma or COPD to determine what patient factors and health system factors may predispose these subjects to be less likely to be diagnosed compared to subjects with physician-diagnosed disease. In addition, we are currently undertaking a clinical trial to explore whether case-finding and diagnosing previously undiagnosed obstructive lung disease has value. Patients who are found to have undiagnosed asthma or COPD are being randomized to an intensive early treatment strategy based on GINA and GOLD guidelines versus usual care. This clinical trial will allow us to determine whether detection of undiagnosed obstructive lung disease, with subsequent provision of early intensive treatment, provides clinical benefit.

In summary, our study found that undiagnosed asthma and COPD can be identified in randomly-selected adults from the community. Our population-based case-finding strategy found that approximately 20% of Canadian adults without a previous history of diagnosed lung disease, who report recent or current respiratory symptoms, have undiagnosed asthma or COPD. Existing symptom questionnaires designed for case-finding subjects were relatively unhelpful and lacked predictive ability to accurately find subjects with undiagnosed asthma or COPD in the community. More research is needed to develop better case-finding tools, especially case-finding tools for identification of undiagnosed asthma.

Figure Legends:

Figure 1: Results of the case-finding strategy.

Figure 2: ROC curves derived from the 8-item questionnaire for separate predictive equations for asthma, COPD and the combined composite outcome of OLD.

The respective AUC values for the curves are 0.74 (OLD), 0.64 (asthma) and 0.84 (COPD).

Table 1. Characteristics of participants with or without obstructive lung disease.

	Obstructive Lung Disease (n=184) Mean (SD)	No Obstructive Lung Disease (n=726) Mean (SD)	P-Value
Demographic Characteristics			
Age, year	61.3 (14.3)	57.1 (15.2)	<0.001
Male sex, n (%)	62.3%	48.0%	<0.001
Body mass index	29.2 (6.9)	30.6 (6.8)	<0.001
Lung Function			
Pre-bronchodilator spirometry			
FEV1, L	2.26 (0.77)	2.87 (0.81)	<0.001
FEV1, % predicted	75.1 (15.6)	95.8 (15.3)	<0.001
FEV1/FVC, ratio	0.63 (0.09)	0.77 (0.06)	<0.001
Post-bronchodilator spirometry			
FEV1, L	2.51 (0.83)	2.94 (0.83)	<0.001
FEV1, % predicted	82.8 (17.7)	98.0 (16.2)	<0.001
FEV1/FVC, ratio	0.66 (0.11)	0.79 (0.06)	<0.001
% change in FEV1 post-bronchodilator	11.6% (10.1%)	2.9% (4.1%)	<0.001
Health Related Quality of Life (SGRQ total score)	40.5 (18.5)	37.0 (17.3)	<0.001
Diagnostic Questionnaires			
Asthma Screening Questionnaire Score	8.5 (3.6)	8.7 (3.6)	0.130
COPD Diagnostic Questionnaire Score*	23.4 (5.2)	20.8 (5.3)	<0.001
Symptoms Assessment			
COPD Assessment Test (CAT) Total Score	18.2 (6.9)	16.7 (7.0)	<0.001
Smoking Status			
Lifetime Non-Smokers	24.5%	43.4%	<0.001
Previous Smokers	40.8%	37.2%	
Current Smokers	34.8%	19.4%	
Smoking History			
Total Pack Years in Previous & Current Smokers	31.4 (24.3)	19.8 (18.5)	<0.001

*499 subjects completed this questionnaire (116 with OLD and 383 without OLD).

Table 2. Sensitivity, specificity and areas under the ROC curve (AUC) values for the study questionnaire scores using designated thresholds. Values are shown for both the target disease of the questionnaire as well as for the composite outcome of OLD.

Questionnaire	Disease	Subjects N	Range[1]	Threshold (≥)	Sensitivity[2] (%)	Specificity[2] (%)	AUC[3]
ASQ	Asthma	908	0 - 20	6	93	16	0.53
	OLD	908	0 - 20	6	83	14	0.49
COPD-DQ	COPD	499	0 – 38	20	89	34	0.71
	OLD	499	0 – 38	20	84	34	0.64
CAT	COPD	909	0 – 40	10	92	15	0.58
	OLD	909	0 – 40	10	90	16	0.56

1. Range refers to the minimum and maximum possible scores for the questionnaire
2. Sensitivity and specificity values are derived from 2 x 2 contingency tables for the designated thresholds. The values are not adjusted for demographic variables (age, sex, BMI) or pack years smoked.
3. AUC values are computed from non-parametric estimates of ROCs for questionnaire scores.

Table 3. Areas under the ROC curve for detecting OLD measured for the descriptive variables alone, the individual ASQ, COPD-DQ and CAT questionnaires combined with the descriptive variables, and the composite 8-item questionnaire which includes the descriptive variables.

Questionnaire	Area under the ROC curve
Descriptive Variables (Age, Sex, BMI, Pack years smoked)	0.68
Descriptive Variables + ASQ	0.69
Descriptive Variables + COPD-DQ	0.68
Descriptive Variables + CAT	0.69
8-item questionnaire (includes descriptive variables)	0.74

Table 4. The 8-item questionnaire and associated logistic regression coefficients for detecting undiagnosed obstructive lung disease.

Question	Possible Responses	Coefficient	P-value
Age	# in years	0.020	0.002
Male Sex	Yes/No	0.395	0.031
BMI	# as Kg/m^2	-0.044	0.003
Pack years smoked	# as packs/day*years smoked	0.024	0.000
Do you experience worsening of wheezing following physical activity?	Yes/No	0.604	0.001
Do you experience shortness of breath following physical activity?	Yes/No	0.526	0.044
I have no phlegm in my chest at all – my chest is completely full of phlegm	0 to 5	0.802	0.001
I sleep soundly – I don't sleep soundly because of my chest condition	0 to 5	1.045	0.023

The original 0-5 responses for the phlegm and sleep questions taken from the CAT questionnaire are re-scored before multiplying by the corresponding regression coefficient in the risk calculation. See Section S.3 of the Supplement for technical explanations.

Table 5

A) The 8-item questionnaire for detecting undiagnosed asthma.

Question	Possible Responses	Coefficient	P-value
Age	# in years	0.004	0.611
Male Sex	Yes/No	0.284	0.263
BMI	# as Kg/m^2	-0.007	0.697
Pack years smoked	# as packs/day*years smoked	-0.002	0.751
Do you experience worsening of wheezing following physical activity?	Yes/No	0.335	0.202
Do you experience shortness of breath following physical activity?	Yes/No	0.419	0.236
I have no phlegm in my chest at all – my chest is completely full of phlegm	0 to 5	1.042	0.003
I sleep soundly – I don't sleep soundly because of my chest condition	0 to 5	1.014	0.122

The original 0-5 responses for the phlegm and sleep questions taken from the CAT questionnaire are re-scored before multiplying by the corresponding regression coefficient in the risk calculation. See Section S.3 of the Supplement for technical explanations.

B) The 8-item questionnaire for detecting undiagnosed COPD.

Question	Possible Responses	Coefficient	P-value
Age	# in years	0.037	0.000
Male Sex	Yes/No	0.522	0.032
BMI	# as Kg/m^2	-0.086	0.000
Pack years smoked	# as packs/day*years smoked	0.037	0.000
Do you experience worsening of wheezing following physical activity?	Yes/No	0.861	0.001
Do you experience shortness of breath following physical activity?	Yes/No	0.721	0.047
I have no phlegm in my chest at all – my chest is completely full of phlegm	0 to 5	0.833	0.002
I sleep soundly – I don't sleep soundly because of my chest condition	0 to 5	1.419	0.037

The original 0-5 responses for the phlegm and sleep questions taken from the CAT questionnaire are re-scored before multiplying by the corresponding regression coefficient in the risk calculation. See Section S.3 of the Supplement for technical explanations.

Figure 1

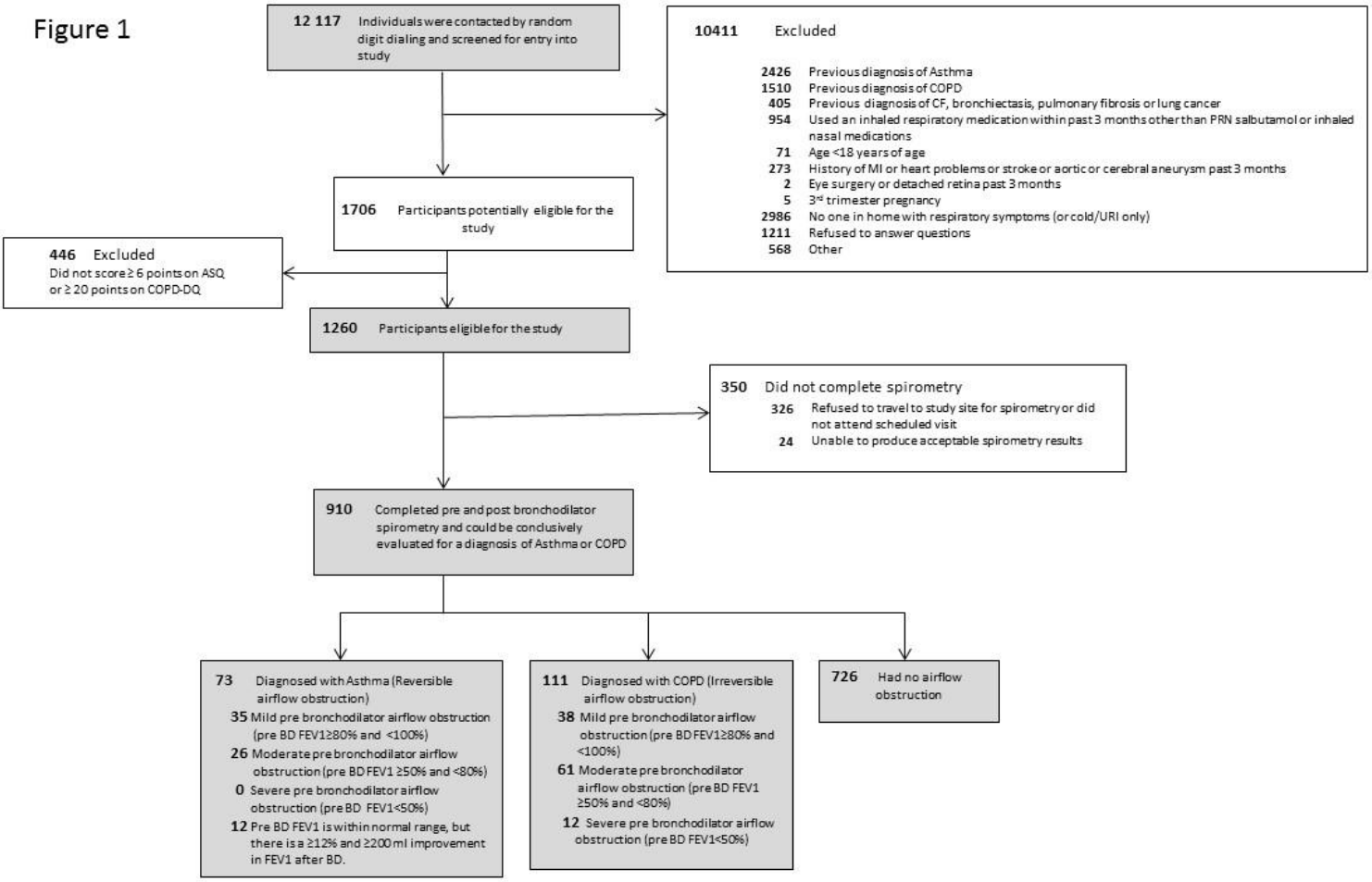
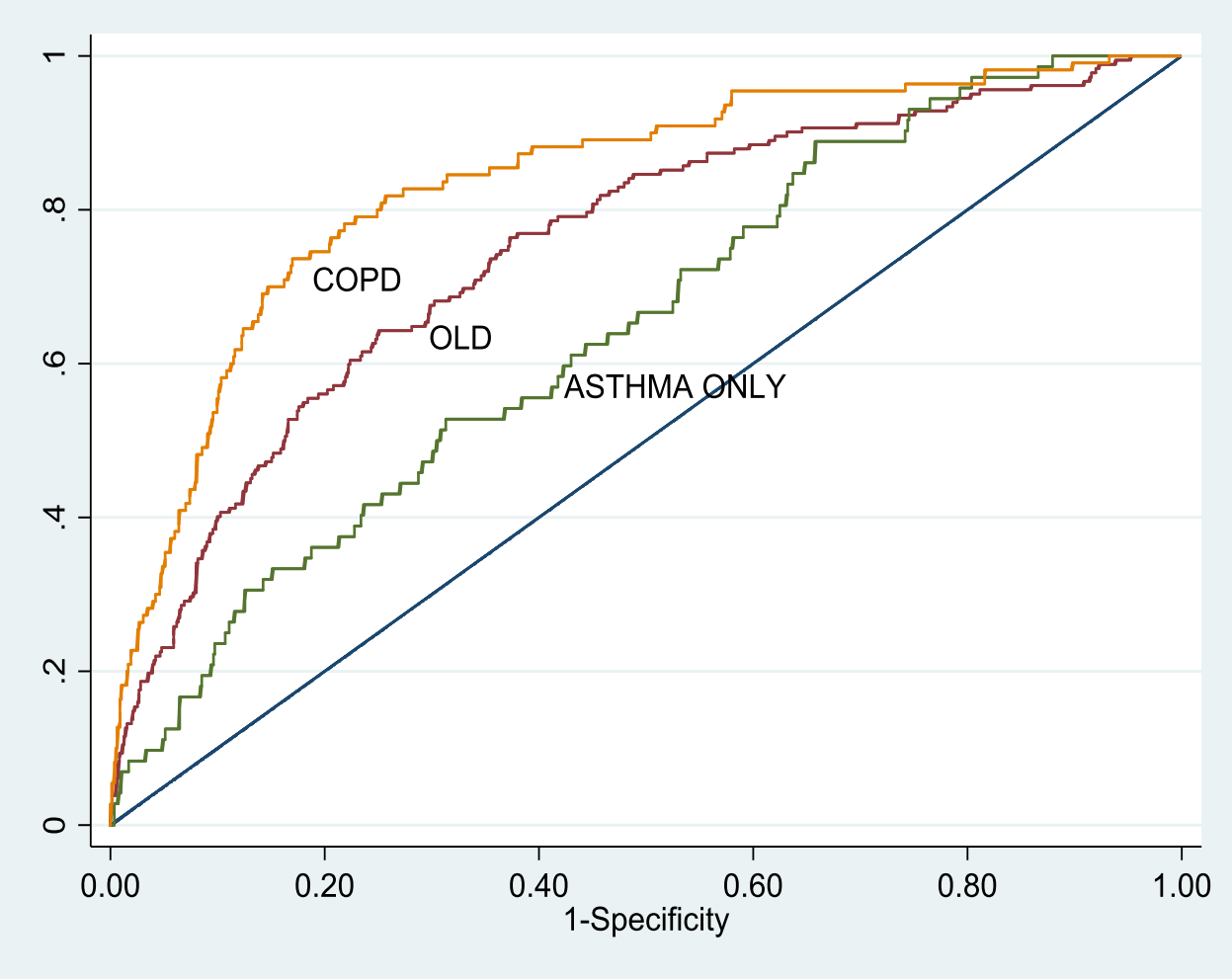

12 117 Individuals were contacted by random digit dialing and screened for entry into study
10411 Excluded
2426 Previous diagnosis of Asthma
1510 Previous diagnosis of COPD
405 Previous diagnosis of CF, bronchiectasis, pulmonary fibrosis or lung cancer
954 Used an inhaled respiratory medication within past 3 months other than PRN salbutamol or inhaled nasal medications
71 Age <18 years of age
273 History of MI or heart problems or stroke or aortic or cerebral aneurysm past 3 months
2 Eye surgery or detached retina past 3 months
5 3rd trimester pregnancy
2986 No one in home with respiratory symptoms (or cold/URI only)
1211 Refused to answer questions
568 Other
1706 Participants potentially eligible for the study
446 Excluded
Did not score ≥ 6 points on ASQ or ≥ 20 points on COPD-DQ
1260 Participants eligible for the study
350 Did not complete spirometry
326 Refused to travel to study site for spirometry or did not attend scheduled visit
24 Unable to produce acceptable spirometry results
910 Completed pre and post bronchodilator spirometry and could be conclusively evaluated for a diagnosis of Asthma or COPD
73 Diagnosed with Asthma (Reversible airflow obstruction)
35 Mild pre bronchodilator airflow obstruction (pre BD FEV1≥80% and <100%)
26 Moderate pre bronchodilator airflow obstruction (pre BD FEV1 ≥50% and <80%)
0 Severe pre bronchodilator airflow obstruction (pre BD FEV1<50%)
12 Pre BD FEV1 is within normal range, but there is a ≥12% and ≥200 ml improvement in FEV1 after BD.
111 Diagnosed with COPD (Irreversible airflow obstruction)
38 Mild pre bronchodilator airflow obstruction (pre BD FEV1≥80% and <100%)
61 Moderate pre bronchodilator airflow obstruction (pre BD FEV1 ≥50% and <80%)
12 Severe pre bronchodilator airflow obstruction (pre BD FEV1<50%)
726 Had no airflow obstruction

Figure 2:


0
.2
.4
.6
.8
1
0.00
0.20
0.40
0.60
0.80
1.00
1-Specificity
COPD
OLD
ASTHMA ONLY

References:

1. Ford, E. S., Mannino, D. M., Wheaton, A. G., Giles, W. H., Presley-Cantrell, L., & Croft, J. B. (2013). 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*, *143*(5), 1395-1406.
2. Gershon AS, Warner L, Cascagnette P, Victor JC, To T. (2011). Lifetime Risk of Developing Chronic Obstructive Pulmonary Disease: a longitudinal population study. *Lancet, 378*, 991-996.
3. Turner, R., DePietro, M., & Ding, B. (2018). Overlap of asthma and chronic obstructive pulmonary disease in patients in the United States: analysis of prevalence, features, and subtypes. *JMIR public health and surveillance*, *4*(3), e60.
4. Chronic Obstructive Pulmonary Disease (COPD). Center for Disease Control and Prevention. Feb 21st, 2018. Retrieved from: https://www.cdc.gov/copd/infographics/copd-costs.html
5. Asthma in the US. Center for Disease Control and Prevention. May 3rd, 2011. Retrieved from: https://www.cdc.gov/vitalsigns/asthma/index.html
6. Johnson, K. M., Bryan, S., Ghanbarian, S., Sin, D. D., & Sadatsafavi, M. (2018). Characterizing undiagnosed chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Respiratory research*, *19*(1), 26.
7. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention. 2019 [accessed 2019 Aug 1]. Available from: http://www.ginasthma.org/.)
8. Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2019. Global Strategy for the Diagnosis, Management and Prevention of COPD. [accessed 2019 Aug 1]. Available from: https://goldcopd.org/.
9. Gershon AS, Hwee J, Croxford R, Aaron SD, To T. (2014). Patient and physician factors associated with pulmonary function testing for COPD: A population study. *Chest, 145*(2), 272-281.
10. Gershon AS, Victor JC, Guan J, Aaron SD, To T. (2012). Pulmonary function testing in the diagnosis of asthma: A population study. *Chest, 141*(5), 1190-1196.
11. Gershon, A. S., Hwee, J., Chapman, K. R., Aaron, S. D., O'Donnell, D. E., Stanbrook, M. B., Bourbeau, J., Tan, W., Su, J., Victor, J.C., & To, T. (2016). Factors associated with undiagnosed and overdiagnosed COPD. *European Respiratory Journal*, *48*(2), 561-564.
12. Aaron SD, Boulet LP, Reddel H, Gershon A. Under-diagnosis and over-diagnosis of asthma. The American Journal of Respiratory and Critical Care Medicine 198 (8); 1012–1020, 2018.
13. Diab N, Gershon A, Sin D, Tan W, Bourbeau J, Boulet LP, and Aaron SD. Under-diagnosis and over-diagnosis of chronic obstructive pulmonary disease. The American Journal of Respiratory and Critical Care Medicine 198(9):1130-1139, 2018.
14. Guirguis-Blake JM, Senger CA, Webber EM, Mularski RA, Whitlock EP. (2016). Screening for chronic obstructive pulmonary disease: Evidence report and systematic review for the us preventive services task force. *JAMA, 315*(13), 1378-1393.
15. US Preventive Services. (2016). Screening for chronic obstructive pulmonary disease: Us preventive services task force recommendation statement. *JAMA, 315*(13), 1372-1377.
16. Backer V, Harmsen L, Lund T, et al. A 3-year longitudinal study of asthma quality of life in undiagnosed and diagnosed asthma patients. The International Journal of Tuberculosis and Lung Disease 2007;11:463-9.

17. Adams RJ, Wilson DH, Appleton S, et al. Underdiagnosed asthma in South Australia. Thorax 2003;58:846-50.
18. Stanley AJ, Hasan I, Crockett AJ, van Schayck OCP, Zwar NA. (2014). COPD Diagnostic Questionnaire (CDQ) for selecting at-risk patients for spirometry: a cross-sectional study in Australian general practice. *NPJ Prim Care Respir Med, 24*,14024.
19. Kotz D, Nelemans P, van Schayck CP, Wesseling GJ. (2008). External validation of a COPD diagnostic questionnaire. *Eur Respir J, 31*(2), 298.
20. Shin B, Cole SL, Park SJ, Ledford D, Lockey R. (2010). A new Symptom-Based Questionnaire for Predicting the Presence of Asthma. *Invest Allergol Clin Immunol, 20*(1), 27-34.
21. Aaron SD, Vandemheen K, FitzGerald JM, et al (2017). Reevaluation of Diagnosis in Adults with Physician-Diagnosed Asthma. JAMA 317(3):269-279.
22. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. (2009). Development and first validation of the COPD Assessment Test. *Eur Respir J, 34*(3), 648.
23. Tsiligianni IG, Alma HJ, de Jong C, Jelusic D, Wittmann M, Schuler M, Schultz K, Kollen BJ, van der Molen T, Kocks JW (2016). Investigating sensitivity, specificity, and area under the curve of the clinical COPD questionnaire, COPD assessment test, and modified Medical Research Council scale according to GOLD using St George's respiratory questionnaire cutoff 25 (and 20) as reference. Int J Chron Obstruct Pulmon Dis. 2016;11:1045–1052.
24. Labaki W and Han ML (2018). Improving Detection of Early Chronic Obstructive Pulmonary Disease. Ann Am Thorac Soc; 15 (Supplement 4): S243–S248.
25. Martinez FJ, Mannino D, Leidy NK, Malley KG, Bacci ED, Barr RG, et al (2017). A new approach for identifying patients with undiagnosed chronic obstructive pulmonary disease. Am J Respir Crit Care Med; 195:748–756.
26. Varela M, Montes de Oca M, Wehrmeister F, Rodriguez C, Ramirez L, Menezes A (2019). External validation of the PUMA COPD diagnostic questionnaire in a general practice sample and the PLATINO study population. International Journal of COPD:14 1901–1911.

Acknowledgements:

The authors gratefully acknowledge the invaluable assistance of the following individuals from the study sites:

Ottawa Hospital Research Institute, Ottawa, Ontario: Taylor Poulin, Joanne Cassidy RRT, Vicky Panteleakos RRT, Amanda Bergeron MSc, Susan Deveau RRT, Jennifer Biggs RN; Vancouver General Hospital, Vancouver, British Columbia: Shelley Abercromby BSc, Jana Caine, David Savage, Natasha Verzosa, Ravneet Mahal; Queen Elizabeth II Health Sciences Centre, Halifax, Nova Scotia: Scott Fulton RRT; Hôpital du Sacré Coeur de Montréal, Montréal, Québec: Simone Chaboillez MT, Meliza Benabdallah; St. Joseph's Hamilton, Hamilton, Ontario: Liz Johnson; St. Boniface Hospital, Winnipeg, Manitoba: Cheryl Noble RN; Institut Universitaire de Cardiologie et de Pneumologie de Québec-Université Laval, Québec, Québec: Johane Lepage BSc, Joanne Milot RN; University of Calgary, Calgary, Alberta: Lisette Machado MD; University of Alberta, Edmonton, Alberta: Miranda Bowen RRT, Fay Hartt, Angie Hillaby RRT; St. Micheal's Hospital, Toronto, Ontario: Stephanie Segovia, Carolyn Spiegel-Feld; Queen's University Kingston General Hospital, Kingston, Ontario: Ann Taite BSc; University of Saskatchewan Royal University Hospital, Saskatoon, Saskatchewan: Nancy Zacher, Janet Baran RN; London Health Sciences Centre - Victoria Hospital, London, Ontario: Katie Maguire, Heba Almadhoun; St. Clare's Mercy Hospital, St John's, Newfoundland: Sarah Anthony BNRN, Heather Dalton; Royal Victoria Regional Health Centre, Barrie, Ontario: Ashley Brown RRT; Michael Garron Hospital, Toronto, Ontario: Liz Lee

Data manager: Ottawa Hospital Research Institute, Ottawa, Ontario: Dong Vo

Random digit dialing: ASDE Survey Sampler, Inc (Gatineau, Quebec, Canada)

We would also like to thank the study participants who gave their time and came in for the study visits.

September 15, 2019

Supplement for:

Population-Based Case-Finding to Identify Subjects with Undiagnosed Asthma or COPD

S.1 Table of Sensitivity and Specificity Values for the 8-item Questionnaire

The 8-item questionnaire and associated logistic regression coefficients for assessing probability of OLD is given in Table 4 of the main article. The ROC curve for this questionnaire appears in Figure 2 of the main article. The AUC of this curve is 0.74. Table S.1 here gives the sensitivity and specificity values of this questionnaire for obstructive lung disease at a selection of cutoff values for the probability of OLD.

Probability of OLD (%)	Sensitivity (%)	Specificity (%)
5	99	7
10	89	30
15	79	54
20	67	69
25	58	79

Table S.1: Sensitivity and specificity values of the 8-item questionnaire for obstructive lung disease at a selection of cutoff values for the probability of OLD.

S.2 Illustration of Risk Scoring of OLD for a Demonstration Subject

Question	Response	Regression Coefficient	Score Contribution
What is your age?	36	0.0201	0.722
Are you male?	Yes (value=1)	0.395	0.395
What is your BMI?	33.52 kg/m^2	-0.0441	-1.478
How many pack years have you smoked?	15	0.0236	0.353
Do you experience worsening of wheezing following physical activity?	No (value=0)	0.604	0.000
Do you experience shortness of breath following physical activity?	Yes (value=1)	0.526	0.526
I have no phlegm in my chest at all (0) – my chest is completely full of phlegm (5)	"3" (value= -1.419)	0.802	-1.138
I sleep soundly (0) – I don't sleep soundly because of my chest condition (5)	"2" (value=-1.229)	1.045	-1.284
Constant			-0.054
Total Logistic Score			***-1.958***

Table S.2: Risk scoring for a demonstration subject using the 8-item questionnaire for obstructive lung disease.

Table S.2 illustrates the risk scoring calculations for a demonstration subject using the 8-item questionnaire for obstructive lung disease. By summing the score contributions of the subject's

responses to the eight questions and then adding the regression constant term, one obtains a total logistic score of -1.958. The risk is calculated from the total score using the following logistic formula:

exp(-1.958)/ [1+exp(-1.958)]=0.124.

This score corresponds to a probability of 0.124 or a 12% risk that the subject has OLD.

We note that the last two questions are taken from the CAT questionnaire and have respective rating responses of "3" and "2" for this demonstration subject. These two ratings are assigned scores of -1.419 and -1.229, respectively, for the risk score calculation. Refer to Section S.3 below for an explanation of how values are assigned to the CAT responses for risk assessment with the 8-item questionnaire. Table S.4 is a reference table for these assigned values.

S.3 Scoring of Questions in the CAT Questionnaire

The eight symptom questions in the COPD Assessment Test (CAT) questionnaire are each structured as a six-point ordinal scale. The scale is a progression of symptom severity from none (0) to extreme (5). In spite of the ordinal construction, however, the relationships of the ratings to the outcome probabilities for OLD, asthma, or COPD are not monotonic. Subjects in our study who have the disease tend to respond at intermediate levels of severity, avoiding the most extreme response. The CAT question about 'phlegm in my chest' provides a convincing example of this pattern. Phlegm symptoms are strongly indicative of OLD. The relative frequency of OLD among subjects tends to rise through ratings 0 to 4 on the phlegm question of CAT. Then, the relative frequency drops significantly at a rating of 5. This point is illustrated in Table S.3.

CAT phlegm rating	All Subjects	Subjects with OLD	Percentage with OLD
0	86	14	16.3
1	177	24	13.6
2	181	37	20.4
3	260	51	19.6
4	136	46	33.8
5	53	10	18.9
Total	***893***	***182***	***20.4***

Table S.3: Percentage of subjects with OLD among subjects giving different ratings to the CAT phlegm question. The table only includes subjects with complete records who were used to estimate the 8-item risk model (n=893).

The non-monotonic relationship between outcome risk for OLD and the rating response for each CAT question led us to score the CAT question using a fitted logistic function. In particular, logistic regression was used to relate disease outcome to the question response, treating the latter as a categorical variable. The fitted value of the logistic regression function for each subject is the subject's risk score for that question. Mathematically this score is the estimated log-odds for the disease outcome based on the subject's individual response to the CAT question. As an illustration, Table S.4 gives the predicted scores from these fitted logistic regression functions for the CAT questions for phlegm and sleep, both of which

feature in the 8-item risk model for OLD presented in Table 4 of the main article. These scores define predictor variables *zphlegm* and *zsleep* in the 8-item model. For example, a subject reporting ratings of "3" for phlegm and "2" for sleep would have scores of *zphlegm* = -1.419 and *zsleep* = -1.229 for these two symptoms.

	Phlegm		**Sleep**	
CAT rating	**All Subjects**	**Logistic Score**	**All Subjects**	**Logistic Score**
0	86	-1.651	171	-1.798
1	177	-1.872	176	-1.379
2	181	-1.386	185	-1.229
3	260	-1.419	197	-1.325
4	136	-0.650	112	-1.147
5	53	-1.459	52	-1.316
Total	***893***		***893***	

Table S.4: Logistic scores for phlegm and sleep computed from logistic regression functions that relate the OLD outcome to CAT question ratings for the 893 subjects whose records entered the 8-item risk model.