

# Factors associated with undiagnosed and overdiagnosed COPD

*To the Editor:*

Worldwide, studies have shown that about 60–86% of people with chronic obstructive pulmonary disease (COPD) have not been diagnosed, which represents a missed opportunity to decrease disease burden through optimal management, including smoking cessation support and prescription medications [1–3]. Overdiagnosed COPD is also common, with prevalence estimates ranging from 4% to 64% in the general population and primary care settings [4, 5]. Overdiagnosis can lead to unnecessary COPD treatments with their own risks and costs, poor health-related quality of life and missed detection and treatment of other diseases [6].

Misdiagnosed COPD indicates that an inaccurate label has been applied, either by not diagnosing COPD (false negative) or diagnosing COPD in someone who does not have it (false positive). Pulmonary function testing helps to avoid such errors [1]. Previous studies have examined factors that differentiate people with diagnosed COPD (true positives) from people without COPD (true negatives) and from people with undiagnosed COPD (false negatives) [7–9]. However, a clinician meeting a patient for the first time would benefit from being aware of factors that differentiate overdiagnosed from correctly diagnosed COPD (false from true positives) and undiagnosed from no COPD (false from true negatives). The objective of the current study was to determine these factors.

A population-based longitudinal cohort study using clinical and health administrative data was conducted. Ethics approval was obtained from the institutional review board at Sunnybrook Health Sciences Centre, Toronto, ON, Canada.

The Canadian Obstructive Lung Disease (COLD) study (a forerunner of the Canadian Cohort of Obstructive Lung Disease (CanCOLD) study) collected clinical information, including spirometry, from a population-based cohort of adults aged  $\geq 40$  years living across Canada between 2007 and 2011 [8]. Participants from Ontario, the largest province of Canada, were included in the current study. Their clinical information was individually linked to health information collected to administer Ontario's universal healthcare system and analysed at the Institute for Clinical Evaluative Sciences (Toronto, ON, Canada). Details of the databases used can be found elsewhere [10].

Participants were categorised into four groups based on them having true COPD, as determined by spirometry, and/or physician-diagnosed COPD, as determined by health administrative data: correctly diagnosed, undiagnosed, overdiagnosed or no COPD by either of the criteria. True COPD was determined, as in Burden of Lung Disease (BOLD) studies, as a post-bronchodilator fixed ratio of forced expiratory volume in 1 s to forced vital capacity of  $< 0.7$  [2]. In sensitivity analysis, lower limit of normal criteria, an alternate method used to diagnose COPD, were used in its place. Physician-diagnosed COPD was determined using a health administrative data case definition of one COPD ambulatory care visit or one COPD hospitalisation, which was previously shown to have 85% sensitivity and 78% specificity in individuals aged  $\geq 35$  years when compared to a clinical reference standard [11].

To determine factors associated with undiagnosed COPD among people without a previous COPD diagnosis, patient and physician factors were compared between the undiagnosed and no COPD groups. To determine factors associated with overdiagnosed COPD among people with a COPD diagnosis, factors were compared between the overdiagnosed and correctly diagnosed COPD groups. Multivariable logistic regression models with generalised estimating equations were used to adjust for clustering by recruitment location.

There were 1586 COLD participants from Ontario, of whom 96 (6.1%) had missing spirometry and 87 (5.5%) had data that could not be linked to health administrative records. Of the 1403 participants available for analysis, 192 (13.7%) had undiagnosed, 72 (5.1%) overdiagnosed and 52 (3.7%) correctly diagnosed COPD.

In multivariable analysis, older age, male sex, a  $> 20$ -pack-year smoking history, respiratory symptoms, a co-diagnosis of asthma and low overall comorbidity were associated with undiagnosed COPD among those without physician-diagnosed COPD. Younger age, not smoking and older physician age were associated with overdiagnosed COPD among those with physician-diagnosed COPD (table 1). Repeating the analysis using the lower limit of normal revealed the same associations except that male sex and low overall comorbidity were no longer significantly associated with undiagnosed COPD and age was no longer significantly associated with overdiagnosed COPD (data not shown).

TABLE 1 Patient and physician characteristics and odds of undiagnosed and overdiagnosed chronic obstructive pulmonary disease (COPD)

Characteristic	Undiagnosed COPD <sup>#</sup>	Overdiagnosed COPD <sup>¶</sup>
<b>Age years</b>	<b>1.08 (1.06–1.10)</b>	<b>0.93 (0.89–0.98)</b>
<b>Male</b>	<b>1.47 (1.02–2.08)</b>	0.79 (0.31–2.00)
<b>Socioeconomic status quintile</b>		
1 (lowest)	Reference group	Reference group
2	1.29 (0.67–2.48)	<b>0.23 (0.06–0.91)</b>
3	1.10 (0.56–2.15)	1.39 (0.32–6.05)
4	1.59 (0.86–2.96)	0.68 (0.16–2.89)
5 (highest)	1.23 (0.67–2.25)	1.15 (0.30–4.45)
<b>Body mass index kg·m<sup>-2</sup></b>	<b>0.97 (0.94–1.00)</b>	0.98 (0.92–1.05)
<b>Cigarette smoking pack-years</b>		
None	Reference group	Reference group
>0 to <20	1.33 (0.90–1.98)	<b>0.17 (0.04–0.63)</b>
≥20	<b>3.77 (2.38–5.97)</b>	<b>0.13 (0.03–0.50)</b>
<b>Cough, sputum, wheeze or breathlessness</b>	<b>1.69 (1.17–2.44)</b>	0.57 (0.17–1.94)
<b>Cardiac disease</b>	0.88 (0.59–1.31)	0.98 (0.38–2.54)
<b>Asthma</b>	<b>2.91 (1.78–4.78)</b>	0.45 (0.16–1.29)
<b>Overall level of comorbidity*</b>		
Low	Reference group	Reference group
Intermediate or high	<b>0.56 (0.34–0.91)</b>	1.15 (0.11–11.95)
<b>Male primary care physician</b>	0.81 (0.56–1.17)	1.06 (0.37–3.08)
<b>Primary care physician age group years</b>		
≤50	Reference group	Reference group
>50	0.92 (0.65–1.31)	<b>2.79 (1.09–7.15)</b>
<b>Specialist care in previous 5 years</b>	1.08 (0.75–1.55)	0.89 (0.34–2.34)

Data are presented as OR (95% CI). <sup>#</sup>: non-COPD (those without COPD by pulmonary function testing or physician diagnosis) is the reference group; <sup>¶</sup>: correctly diagnosed COPD is the reference group; \* : determined using the Johns Hopkins Adjusted Clinical Group Case-Mix System (Johns Hopkins HealthCare Solutions, Baltimore, MD, USA). Bold indicates statistical significance with an alpha error of 0.05.

Our results extend previous studies showing that older age and smoking are predictors of diagnosed COPD, to undiagnosed COPD as well [1]. They are not consistent with one previous population-based study that found that people with overdiagnosed COPD were more likely to be overweight and have allergic rhinitis [12]. While we did not specifically examine allergic rhinitis, we did not find a higher risk of asthma (another allergic condition) among this group. Our results are also not consistent with a study that found increased body mass index (BMI) to be associated with higher risk of overdiagnosis in veterans [13]. These inconsistencies could be due to differences in populations, the exclusion of nonsmokers and people with asthma in the other studies, how BMI was considered (as a continuous *versus* categorical variable) and variability in other factors adjusted for.

Our study had limitations. First, our health administrative data case definition of COPD might have been subject to misclassification. However, it was previously validated, is reflective of real-world practice and is an improvement over the self-reported COPD used to identify misdiagnosis in the past [11, 14]. Secondly, our study was underpowered to detect smaller associations between some factors and overdiagnosed or undiagnosed COPD. Thirdly, we were not able to determine the relationship between previous pulmonary function testing and misdiagnosis because, although the health administrative data revealed if testing occurred, it did not contain its results, quality and whether it was correctly interpreted.

A final point is that diagnosing COPD is complicated and understanding misdiagnosis involves an understanding of many more issues than are presented. For example, a single diagnostic spirometry test is recommended by COPD guidelines, is the method used to identify COPD in landmark COPD trials, and is the way COPD is often diagnosed in the “real world”; however, it may be insufficient to detect and categorise all chronic airflow obstruction [2]. As a result, it is possible that some people with asthma, who would have had their spirometry normalise over time, were incorrectly classified as having overdiagnosed COPD. This would have caused us to overestimate the rate of overdiagnosed COPD. It is also possible, however, that these people developed fixed obstruction as a result of their asthma, which falls within the spirometric definition of COPD [1]. Finally, it is possible that people without spirometric changes included in our overdiagnosed group had clinical and/or radiological signs related to smoking that accounted for COPD-like respiratory exacerbations [15].

In summary, un- and overdiagnosed COPD are common and there are factors that can be used to distinguish affected people in the general and COPD populations, respectively. Our results can be used by physicians to target people at higher risk for both un- and overdiagnosed COPD for pulmonary function testing, in order to reduce COPD misdiagnosis and improve quality of care and health outcomes.



@ERSpublications

Several factors are associated with undiagnosed and overdiagnosed COPD <http://ow.ly/4mW6lu>

Andrea S. Gershon<sup>1,2,3,4</sup>, Jeremiah Hwee<sup>2</sup>, Kenneth R. Chapman<sup>5</sup>, Shawn D. Aaron<sup>6</sup>, Denis E. O'Donnell<sup>7</sup>, Matthew B. Stanbrook<sup>2,3,5</sup>, Jean Bourbeau<sup>8</sup>, Wan Tan<sup>9</sup>, Jiandong Su<sup>2</sup>, J. Charles Victor<sup>2,3</sup> and Teresa To<sup>2,3,4</sup>

<sup>1</sup>Dept of Medicine, Sunnybrook Health Sciences Centre, Toronto, ON, Canada. <sup>2</sup>Institute for Clinical Evaluative Sciences, Toronto, ON, Canada. <sup>3</sup>Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada. <sup>4</sup>The Hospital for Sick Children, Toronto, ON, Canada. <sup>5</sup>Asthma and Airway Centre, Toronto Western Hospital, Toronto, ON, Canada. <sup>6</sup>Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada. <sup>7</sup>Respiratory Investigation Unit, Kingston General Hospital, Kingston, ON, Canada. <sup>8</sup>McGill University, Respiratory Division, Royal Victoria Hospital, Montreal, QC, Canada. <sup>9</sup>University of British Columbia, The UBC James Hogg Research Centre, Institute for Heart and Lung Health, St Paul's Hospital, Vancouver, BC, Canada.

Correspondence: Andrea S. Gershon, Institute for Clinical Evaluative Sciences, G1 06, 2075 Bayview Avenue, Toronto, ON, M4N 3M5, Canada. E-mail: andrea.gershon@ices.on.ca

Received: Nov 18 2015 | Accepted after revision: April 10 2016 | First published online: June 23 2016

Support statement: Funding for this project was made available through the Government of Ontario, Canada. This study was also supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). Funding partners were the Canadian Institutes of Health Research (CIHR) (CIHR/Rx&D Collaborative Research Program Operating Grants 93326), the Respiratory Health Network of the Fonds de la recherche du Québec, and industry partners AstraZeneca Canada Ltd, Boehringer Ingelheim Canada Ltd, GlaxoSmithKline (GSK) Canada Ltd, Novartis and Pfizer Canada Ltd. The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the MOHLTC is intended or should be inferred. A.S. Gershon was supported by a New Investigator Award funded by team grant OTG-88591 from the CIHR Institute of Nutrition, Metabolism and Diabetes while working on this study and is currently supported by a Physicians' Services Incorporated Foundation Fellowship in Translational Medicine and has received a CIHR New Investigator Award. T. To is supported by The Dales Award in Medical Research from the University of Toronto. K.R. Chapman is supported by the GSK-CIHR Research Chair in Respiratory Health Care Delivery at the University Health Network (Toronto, ON, Canada). Funding information for this article has been deposited with FundRef.

Conflict of interest: Disclosures can be found alongside this article at [erj.ersjournals.com](http://erj.ersjournals.com)

Acknowledgements: Parts of this material are based on data and information compiled and provided by the Canadian Institute for Health Information (CIHI). However, the analyses, conclusions, opinions and statements expressed herein are those of the authors, and not necessarily those of CIHI. Author contributions were as follows. A.S. Gershon, J. Hwee, J.C. Victor and T. To conceived and designed the study. A.S. Gershon, J. Hwee and J.C. Victor acquired the health administrative data and D.E. O'Donnell, K.R. Chapman, S.D. Aaron, J. Bourbeau and W. Tan acquired the clinical data. All authors designed the study. J. Hwee and J. Su carried out the statistical analysis. All authors interpreted the data. A.S. Gershon and J. Hwee drafted the manuscript. All authors critically revised the manuscript for important intellectual content and approved the submitted version. A.S. Gershon obtained funding. A.S. Gershon, J. Hwee and J.C. Victor were involved in administrative support. A.S. Gershon, J. Hwee and J. Su had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of the analysis. A.S. Gershon is guarantor.

## References

- 1 Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. Updated 2013.
- 2 Buist AS, McBurnie MA, Vollmer WM, *et al*. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet* 2007; 370: 741–750.
- 3 Soriano JB, Zielinski J, Price D. Screening for and early detection of chronic obstructive pulmonary disease. *Lancet* 2009; 374: 721–732.
- 4 Tálamo C, de Oca MM, Halbert R, *et al*. Diagnostic labeling of COPD in five Latin American cities. *Chest* 2007; 131: 60–67.
- 5 Walters JA, Walters EH, Nelson M, *et al*. Factors associated with misdiagnosis of COPD in primary care. *Prim Care Respir J* 2011; 20: 396–402.
- 6 Miravittles M, Andreu I, Romero Y, *et al*. Difficulties in differential diagnosis of COPD and asthma in primary care. *Br J Gen Pract* 2012; 62: e68–e75.
- 7 Hvidsten SC, Storesund L, Wentzel-Larsen T, *et al*. Prevalence and predictors of undiagnosed chronic obstructive pulmonary disease in a Norwegian adult general population. *Clin Respir J* 2010; 4: 13–21.
- 8 Labonté LE, Tan WC, Li PZ, *et al*. Undiagnosed COPD contributes to the burden of health care utilization: data from the CanCOLD study. *Am J Respir Crit Care Med* 2016; in press [DOI: 10.1164/rccm.201509-1795OC].
- 9 Hill K, Goldstein RS, Guyatt GH, *et al*. Prevalence and underdiagnosis of chronic obstructive pulmonary disease among patients at risk in primary care. *CMAJ* 2010; 182: 673–678.
- 10 Gershon AS, Campitelli MA, Stanbrook MB. Long-acting  $\beta$ -agonists with vs without inhaled corticosteroids for COPD – reply. *JAMA* 2015; 313: 305–306.

- 11 Gershon AS, Wang C, Guan J, *et al.* Identifying individuals with physician diagnosed COPD in health administrative databases. *COPD* 2009; 6: 388–394.
- 12 Zwar NA, Marks GB, Hermiz O, *et al.* Predictors of accuracy of diagnosis of chronic obstructive pulmonary disease in general practice. *Med J Aust* 2011; 195: 168–171.
- 13 Collins BF, Feemster LC, Rinne ST, *et al.* Factors predictive of airflow obstruction among veterans with presumed empirical diagnosis and treatment of COPD. *Chest* 2015; 147: 369–376.
- 14 Huzel L, Roos LL, Anthonisen NR, *et al.* Diagnosing asthma: the fit between survey and administrative database. *Can Respir J* 2002; 9: 407–412.
- 15 Tan WC, Bourbeau J, Hernandez P, *et al.* Exacerbation-like respiratory symptoms in individuals without chronic obstructive pulmonary disease: results from a population-based study. *Thorax* 2014; 69: 709–717.

Eur Respir J 2016; 48: 561–564 | DOI: 10.1183/13993003.00458-2016 | Copyright ©ERS 2016

# Handling missing items in the Exacerbations of Chronic Pulmonary Disease Tool



*To the Editor:*

The Exacerbations of Chronic Pulmonary Disease Tool (EXACT) is a 14-item, self-administered daily symptom diary designed to identify and characterise exacerbations of chronic obstructive pulmonary disease (COPD). It provides a reliable, valid and standardised measure of exacerbation symptoms, and is sensitive to changes during recovery [1]. Scores are expressed on a 100-point scale, with higher values indicating worse symptoms or health state. In addition, the EXACT-derived E-RS (EXACT Respiratory Symptoms) provides valid daily COPD symptom scores [2, 3]. Electronic administration is recommended and has several advantages, notably in preventing item omission [4]. However, the expense of electronic solutions may prove prohibitive, particularly in noncommercial studies, when a pen-and-paper version may be used instead. In this context, it is important to have a method to deal with missing items. This is yet to be established.

In psychometrically validated instruments with high internal consistency, such as the EXACT (Cronbach- $\alpha$   $\geq 0.9$  [1, 5]), missing items may be imputed from the remaining answered items [6]. Used appropriately, this is preferable to list-wise deletion of incomplete records, which reduces power and risks introducing bias if data are missed not at random, and to substituting values from neighbouring records (“last observation carried forward” or “next value carried backward”), as this assumes symptoms are in steady state, which is unlikely during an exacerbation [6]. However, imputing items increases random error and, if items are missed systematically, may introduce bias. These factors limit the number and combination of items that can be imputed without excessively compromising reliability and accuracy. We sought to define the parameters under which this may be done by simulating item imputation on complete EXACT records from a recent study.

The study was an investigator-led, multicentre, randomised, double-blind, placebo-controlled trial ([www.clinicaltrials.gov](http://www.clinicaltrials.gov) identifier number NCT01247870 and [www.isrctn.com](http://www.isrctn.com) identifier number ISRCTN66148745). Its methods and results are detailed elsewhere [7]. Briefly, the trial tested metformin in 52 patients admitted to hospital for COPD exacerbations, primarily to establish its antihyperglycaemic effect. Secondary end-points included symptomatic recovery, as determined by the EXACT. Eligible patients were aged  $\geq 35$  years, had established COPD and had been admitted for an exacerbation with an expected inpatient stay  $\geq 48$  h. Participants completed the EXACT on paper every evening for 1 month, including in hospital. Guidance was provided by investigators in person during the inpatient phase and telephone support was available following discharge.

The effect of imputing items was simulated on EXACT diary records from the first 17 participants, representing all participants enrolled by the time of this analysis. This dataset comprised 361 EXACT diary records, of which 302 (84%) were complete. In the first simulation, one randomly selected item was deleted from each complete record and an imputed score substituted in its place. Imputed scores were calculated as the mean raw score from the remaining items, rounded to the nearest integer and capped at the maximum available for the item being imputed. The total imputed and actual raw scores were transformed to a 100-point linear scale for analysis and interpretation [4]. The degree to which systematic error (bias) was introduced by imputation was quantified by the mean difference between imputed and