Original Article

Diagnostic Performance of a Machine Learning Algorithm (Asthma/Chronic Obstructive Pulmonary Disease [COPD] Differentiation Classification) Tool Versus Primary Care Physicians and Pulmonologists in Asthma, COPD, and Asthma/COPD Overlap

Janwillem W.H. Kocks, MD, PhD[a,b,c], Hui Cao, MD, PhD[d], Björn Holzhauer, PhD[e], Alan Kaplan, MD[f], J. Mark FitzGerald, MD[g,*], Konstantinos Kostikas, MD, PhD[h], David Price, FRCGP[c,i], Helen K. Reddel, MBBS, PhD[j], Ioanna Tsiligianni, MPH, PhD[k], Claus F. Vogelmeier, MD[l], Sebastien Bostel, PhD[e], and Paul Mastoridis, PharmD[d]

Groningen, the Netherlands; Singapore; Basel, Switzerland; Toronto, Ontario; and Vancouver, British Columbia, Canada; Ioannina, Greece; Aberdeen, United Kingdom; Sydney, New South Wales, Australia; Crete, Greece; Marburg, Germany; and East Hanover, NJ

What is already known about this topic? Misdiagnosis of asthma and chronic obstructive pulmonary disease (COPD) can have many negative health consequences. Machine learning has an increasing role in diagnostic medicine and potential use for health care professionals in the accurate diagnosis of chronic respiratory diseases.

What does this article add to our knowledge? The Asthma/COPD Differentiation Classification machine learning–based diagnostic tool demonstrated superior diagnostic accuracy compared with primary care physicians and pulmonologists in the diagnosis of asthma and COPD in patients aged 35 years and older.

How does this study impact current management guidelines? The Asthma/COPD Differentiation Classification tool has the potential to aid in the differential diagnosis of patients with asthma or COPD and provides a valuable additional resource to supplement the decision-making of practicing physicians.

BACKGROUND: The differential diagnosis of asthma and chronic obstructive pulmonary disease (COPD) poses a challenge in clinical practice and its misdiagnosis results in inappropriate treatment, increased exacerbations, and potentially death.

OBJECTIVE: To investigate the diagnostic accuracy of the Asthma/COPD Differentiation Classification (AC/DC) tool compared with primary care physicians and pulmonologists in asthma, COPD, and asthma-COPD overlap.

METHODS: The AC/DC machine learning-based diagnostic tool was developed using 12 parameters from electronic health records of more than 400,000 patients aged 35 years and older. An expert panel of three pulmonologists and four general practitioners from five countries evaluated 119 patient cases from a

[a]General Practitioners Research Institute, Groningen, the Netherlands
[b]University Medical Centre Groningen, GRIAC Research Institute, University of Groningen, Groningen, the Netherlands
[c]Observational and Pragmatic Research Institute, Singapore
[d]Novartis Pharmaceuticals Corporation, East Hanover, NJ
[e]Novartis Pharma AG, Basel, Switzerland
[f]Family Physician Airways Group of Canada, University of Toronto, Toronto, Ontario, Canada
[g]Department of Medicine, Division of Respiratory Medicine, University of British Columbia, Vancouver, British Columbia, Canada
[h]Respiratory Medicine Department, University of Ioannina School of Medicine, Ioannina, Greece
[i]Centre of Academic Primary Care, Division of Applied Health Sciences, University of Aberdeen, Aberdeen, United Kingdom
[j]Woolcock Institute of Medical Research and University of Sydney, Sydney, New South Wales, Australia
[k]Health Planning Unit, Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Crete, Greece
[l]Department of Medicine, Pulmonary and Critical Care Medicine, University Medical Centre Giessen and Marburg, Philipps-Universität Marburg, Member of the German Centre for Lung Research (DZL), Marburg, Germany
*The author was deceased at the final stage of manuscript development.

This study was funded by Novartis Pharmaceuticals Corporation, East Hanover, NJ.

Conflicts of interest: J.W.H. Kocks reports grants, personal fees, and nonfinancial support from AstraZeneca, Boehringer Ingelheim, and GSK; grants and personal fees from Chiesi Pharmaceuticals and Novartis; and grants from Mundi Pharma and Teva, outside the submitted work, all paid to his institution; and holds 72.5% of shares in the General Practitioners Research Institute. A. Kaplan is a medical advisor or is on the speaker's bureau for AstraZeneca, Behring, Boehringer Ingelheim, Cipla, Covis, GSK, Merck Frosst, Pfizer, Novartis, Novo Nordisk, Sanofi, Eisai, Teva, Trudel, and Valeo. J.M. FitzGerald has received honoraria as a member of speaker bureaus and advisory boards for AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline (GSK), Novartis, and Sanofi Regeneron, outside the area of work. K. Kostikas has received honoraria for presentations and consultancy fees from AstraZeneca, Boehringer Ingelheim, CSL Behring, Chiesi, ELPEN, GSK, Menarini, Novartis, and Sanofi (paid to the University of Ioannina). His department has received funding and grants from AstraZeneca, Boehringer

Abbreviations used
AC/DC- Asthma/COPD Differentiation Classification
ACO- Asthma/COPD overlap
ACQ- Asthma Control Questionnaire
CCQ- Clinical COPD Questionnaire
COPD- Chronic obstructive pulmonary disease
CrI- Posterior credible interval
PCP- Primary care physician

prospective observational study and provided a confirmed diagnosis (n = 116) of asthma (n = 53), COPD (n = 43), asthma-COPD overlap (n = 7), or other (n = 13). Cases were then reviewed by 180 primary care physicians and 180 pulmonologists from nine countries and by the AC/DC tool, and diagnostic accuracies were compared with reference to the expert panel diagnoses.

RESULTS: Average diagnostic accuracy of the AC/DC tool was superior to that of primary care physicians (median difference, 24%; 95% posterior credible interval: 17% to 29%; $P < .0001$) and was noninferior and superior (median difference, 12%; 95% posterior credible interval: 6% to 17%; $P < .0001$ for noninferiority and $P = .0006$ for superiority) to that of pulmonologists. Average diagnostic accuracies were 73%, 50%, and 61% by AC/DC tool, primary care physicians, and pulmonologists versus expert panel diagnosis, respectively.

CONCLUSION: The AC/DC tool demonstrated superior diagnostic accuracy compared with primary care physicians and pulmonologists in the diagnosis of asthma and COPD in patients aged 35 years and greater and has the potential to support physicians in the diagnosis of these conditions in clinical practice. **(J Allergy Clin Immunol Pract 2023;11:1463-74)**

Key words: *Asthma; COPD; Differential diagnosis; Machine learning; AC/DC tool; Asthma/COPD overlap; Primary care physician; Pulmonologist; Accuracy*

INTRODUCTION

Asthma and chronic obstructive pulmonary disease (COPD) are heterogeneous chronic respiratory diseases that have overlapping diagnostic criteria and sometimes similar clinical presentations. This poses a challenge in their differential diagnoses, especially in smokers, ex-smokers, and older adults.[1-4] Asthma/COPD overlap (ACO) is composed of patients with characteristics of both asthma (eg, variability of airway limitation, allergies) and COPD (eg, age at onset of 40 years or greater, chest x-ray with severe hyperinflation).[1,4,5] Chronic respiratory diseases are major causes of morbidity and mortality, and an incorrect diagnosis may lead to negative consequences in disease management.[6] For example, underdiagnosis of asthma leads to increased hospitalizations, emergency room visits, risk for death, and health care resource costs.[7-9] Misdiagnosis can result in adverse events owing to incorrect treatment (particularly when asthma is treated with long-acting bronchodilators alone) and increased treatment costs.[7-12] The overlapping diagnosis of asthma and COPD was reported to be 15% to 32%.[4] Hence, an accurate diagnosis is important for therapeutic decision-making.

Artificial intelligence, especially machine learning, has an increasing role in diagnostic medicine and might be useful for primary care physicians (PCPs) and other health care professionals in the accurate and differential diagnosis of chronic respiratory diseases.[13-16] The Asthma/COPD Differentiation Classification (AC/DC) tool employs a machine learning–based algorithm and was developed to aid PCPs and other physicians in the fast and

Ingelheim, Chiesi, Innovis, ELPEN, GSK, Menarini, Novartis, and NuvoAir (paid to the University of Ioannina). D. Price has board membership with Amgen, AstraZeneca, Boehringer Ingelheim, Chiesi, Circassia, Mylan, Mundipharma, Novartis, Regeneron Pharmaceuticals, Sanofi Genzyme, Teva Pharmaceuticals, and Thermo Fisher; consultancy agreements with Amgen, AstraZeneca, Boehringer Ingelheim, Chiesi, GSK, Mylan, Mundipharma, Novartis, Pfizer, Teva Pharmaceuticals, Theravance; grants and unrestricted funding for investigator-initiated studies (conducted through Observational and Pragmatic Research Institute Pte Ltd) from AstraZeneca, Boehringer Ingelheim, Chiesi, Circassia, Mylan, Merck, Zentiva (Sanofi Generics), Mundipharma, Novartis, Pfizer, Regeneron Pharmaceuticals, Respiratory Effectiveness Group, Sanofi Genzyme, Teva Pharmaceuticals, Theravance, and UK National Health Service; payment for lectures/speaking engagements from AstraZeneca, Boehringer Ingelheim, Chiesi, Cipla, GSK, Kyorin, Mylan, Mundipharma, Novartis, Regeneron Pharmaceuticals, Sanofi Genzyme, and Teva Pharmaceuticals; payment for the development of educational materials from Mundipharma and Novartis; payment for travel, accommodations, and meeting expenses from AstraZeneca, Boehringer Ingelheim, Mundipharma, Mylan, Novartis, and Thermo Fisher; funding for patient enrollment or completion of research from Novartis; stock or stock options from AKL Research and Development Ltd, which produces phytopharmaceuticals; owns 74% of the social enterprise Optimum Patient Care Ltd (Australia and UK) and 74% of Observational and Pragmatic Research Institute Pte Ltd (Singapore), 92.61% of Observational and Pragmatic Research Institute Pte Ltd (Singapore), and 5% shareholding in Timestamp, which develops adherence monitoring technology; is a peer reviewer for grant committees of the Efficacy and Mechanism Evaluation programme and Health Technology Assessment; receives personal fees from Airway Vista Secretariat, EPG Communication Holdings Ltd, FIECON Ltd, Fieldwork International, OM Pharma SA, PeerVoice, Phadia AB, Spirosure Inc, Strategic North Limited, Synapse Research Management Partners S.L, Talos Health Solutions, and WebMD Global LLC, outside the submitted work; and was an expert witness for GSK. H.K. Reddel has provided independent advice on advisory boards or steering committees for AstraZeneca, Chiesi, GSK, Novartis, and Sanofi/Genzyme; has provided independent medical education at symposia funded by AstraZeneca, Boehringer Ingelheim, Chiesi, Getz, GSK, Novartis, Sanofi/Genzyme and Teva; and has provided independent consulting for AstraZeneca and Novartis; and her institute has received unrestricted research grants from AstraZeneca, GSK, and Novartis. I. Tsiligianni has received fees for consultancies from Novartis, Boehringer Ingelheim, AstraZeneca, and GSK and grants for research activities from GSK, AstraZeneca, Boehringer Ingelheim, and ELPEN. C.F. Vogelmeier gave presentations at symposia and/or served on scientific advisory boards sponsored by AstraZeneca, Boehringer Ingelheim, CSL Behring, Chiesi, GSK, Grifols, Menarini, Novartis, Nuvaira, and Aerogen. H. Cao, B. Holzhauer, S. Bostel, and P. Mastoridis are employees of Novartis and hold stock in Novartis Pharma AG. Authors received no compensation related to development of the manuscript.

Received for publication July 4, 2022; revised January 4, 2023; accepted for publication January 5, 2023.

Available online January 28, 2023.

Corresponding author: Paul Mastoridis, PharmD, Novartis Pharmaceuticals Corporation, East Hanover, NJ 07936. E-mail: paul.mastoridis@novartis.com.

2213-2198

https://doi.org/10.1016/j.jaip.2023.01.017

accurate diagnosis of asthma, COPD, or ACO, in conjunction with spirometry, and to reduce delays in symptomatic patients receiving appropriate therapy.[17] This study investigated the diagnostic accuracy of the AC/DC tool compared with PCPs and pulmonologists in the differential diagnosis of asthma, COPD, ACO, and other respiratory diseases using patient cases from a prospective observational study in general practice.[17] Pulmonologists and PCPs were selected as the medical professions for evaluation because they were the professionals most likely to interact initially with patients with a respiratory disease, they manage patients with respiratory diseases, which is not be the case with other groups such as allergists for COPD, and they are potential primary users of the AC/DC tool.

METHODS

Study design

This was a noninterventional, multinational, multiple-rater, multiple-case study that used deidentified patient cases from a prospective observational study (FOCUS), which recorded data for patients presenting with respiratory symptoms to general practices in the Netherlands.[18] Cases were included when patients were aged 35 years or older at the time of data collection and if the critical data required for the AC/DC tool had been recorded. Further details about methods are provided in the Supplemental Text (in this article's Online Repository at www.jaci-inpractice.org).

The AC/DC tool was initially developed using the clinical characteristics of more than 400,000 patients aged 35 years and older with a diagnosis of asthma, COPD, or ACO by specialists (pulmonologists/allergists), as identified from the Optum (Eden Prairie, MN) deidentified electronic health records dataset between 2010 and 2017 (for index date definitions, see the Supplemental Text). In an internal validation, the model achieved sensitivities of 0.98, 0.98, and 0.78, precision of 0.97, 0.97, and 0.92, and F1 scores of 0.98, 0.98, and 0.84 in diagnosing asthma, COPD, and ACO, respectively (Supplemental Text).[17] From the data on more than 400,000 patients, 12 variables were identified as the most impactful, and hence were used by the AC/DC tool (Table I).

The performance (external validation) of the AC/DC tool (using the 12 most impactful variables) was then compared with that of pulmonologists and PCPs in the diagnosis of patients from the FOCUS study; the findings are reported here.

Written informed consent obtained from each patient during the observational study[18] permitted secondary use of their data for this study. An independent ethics committee or institutional review board reviewed the study protocol. The study was conducted in accordance with the International Conference on Harmonization Guidelines for Good Clinical Practice and the Declaration of Helsinki. The study was essentially composed of four steps (Figure 1).

Step 1 Expert panel diagnosis of each case (reference standard). A panel of seven experts (composed of three PCPs and four pulmonologists from five countries, who were involved in developing the AC/DC tool) reviewed the clinical data of 119 deidentified (eligible; n = 116) patient cases from the observational study.[18] Each expert determined a diagnosis of asthma, COPD, ACO, or disease other than asthma, COPD, and ACO for each patient and recorded the difficulty of diagnosis on a 6-point Likert scale from 0 to 5, in which 0 to 1 was described as easy to diagnose and 4 to 5 was hard to diagnose for each. The observational study database included variables such as patients' demographics and baseline clinical characteristics; current inhaled medication (yes or no); medical history questionnaire including Medical Research Council dyspnea scale, Asthma Control Questionnaire (ACQ-7) (0-6) and Clinical COPD Questionnaire (CCQ) (0-6), as well as spirometry results. Two symptom definitions were used. Symptom definition 1 included symptoms present during the previous 7 days if the ACQ Q4 score (shortness of breath) was greater than 0, ACQ Q5 (wheeze) greater than 0 or CCQ Q5 (cough) greater than 0. Symptom definition 2 included an ACQ Q4 score greater than 1, ACQ Q5 greater than 1, or CCQ Q5 greater than 1. Symptoms (yes or no) were fed into the algorithm and shown to the physicians.

TABLE I. Most impactful variables used by Asthma/Chronic Obstructive Pulmonary Disease Differentiation Classification tool

Variable
FEV_1
FEV_1/FVC
Smoking pack-year
Age at onset of respiratory disease
Body mass index
Dyspnea
Wheeze
Cough
Diagnosis of allergic rhinitis
Current smoker
Never smoked
Diagnosis of chronic rhinitis

For a diagnosis to be considered an expert panel diagnosis, five of seven experts had to provide the same diagnosis. The primary case set included patients with a diagnosis of asthma, COPD, or ACO. The exploratory case set included patients with a diagnosis of asthma, COPD, ACO, and disease other than asthma, COPD, and ACO.

Step 2 Diagnosis of clinical cases by PCPs and pulmonologists. Primary care physicians and pulmonologists were recruited from nine countries (the United States, Canada, the United Kingdom, France, Germany, Spain, Australia, China, and India). Participating PCPs and pulmonologists were included if they were licensed and practicing at the time of study with 3 years or more in practice and had ever provided a diagnosis to or treated one or more patients with a respiratory disease.

Each physician reviewed 30 expert panel diagnoses of combined primary and exploratory clinical cases (24 cases and six rereviews to assess intra-rater variability) and assigned a diagnosis of asthma, COPD, ACO, or other, together with the level of confidence in the diagnosis, from 1 (not at all confident) to 7 (very confident), using a cross-sectional, 60-minute Web-based electronic case review system.

Step 3 Diagnosis of clinical cases by AC/DC tool. For the AC/DC tool, a total of 100 algorithms were trained, each with recall (true positive diagnosis) and precision (percentage of true positive diagnoses) of 80% or greater for asthma, COPD, and ACO, and overall accuracy of 95% or greater to characterize fully the stability of the model training process and the model performance.

The AC/DC tool assessed each expert panel diagnosis case and either rejected the expert-assigned diagnosis or assigned a probability to diagnoses of asthma, COPD, or ACO. The algorithm rejected cases when clinical characteristics were beyond the range at which the algorithm was trained. The diagnosis assigned by the AC/DC tool was the one with the highest predicted probability. Figure E1

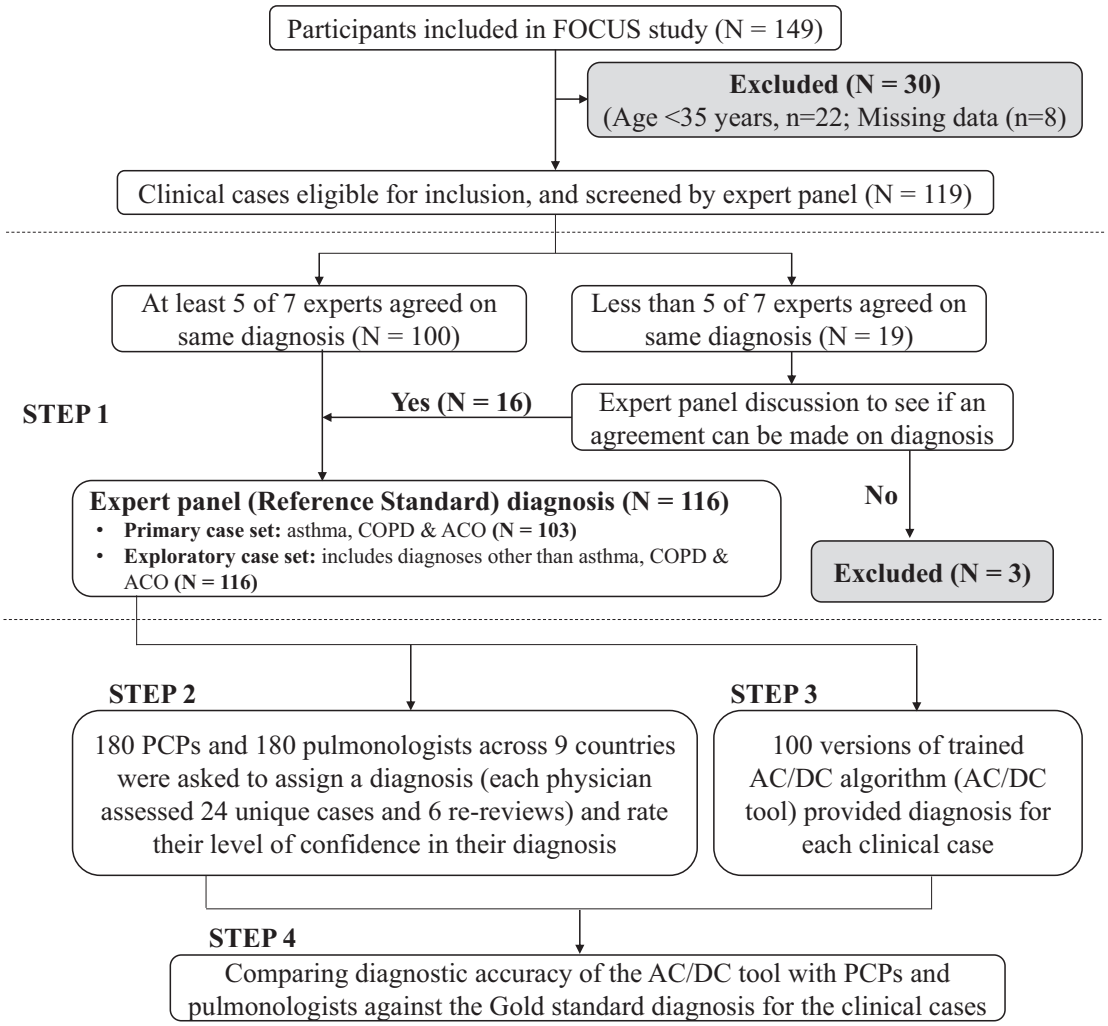

FIGURE 1. Study design and patients flow for Asthma/Chronic Obstructive Pulmonary Disease (COPD) Differentiation Classification (AC/DC) validation study. *ACO,* asthma/COPD overlap; *PCP,* primary care physician.

(in this article's Online Repository at www.jaci-inpractice.org) shows the confusion matrix of the panel diagnosis versus the diagnosis by algorithms and physicians.

Step 4 Outcome. The primary objective was to compare the average diagnostic accuracy of the AC/DC tool with those of PCPs and pulmonologists when evaluating clinical cases of asthma, COPD, or ACO in the primary case set. The diagnostic accuracy of the AC/DC tool, PCPs, and pulmonologists was defined as the correct diagnoses of clinical cases, expressed as a percentage, compared with confirmed diagnoses assigned by the expert panel (reference standard). Differences in average overall diagnostic accuracy were analyzed between the AC/DC tool and PCPs and between the AC/DC tool and pulmonologists with reference to the expert panel diagnoses.

Secondary objectives included (1) a comparison of diagnostic accuracy (sensitivity [recall], precision [positive predictive value], the F1 score [harmonic mean of sensitivity and precision], and the negative predictive value and specificity of the AC/DC tool in diagnosing asthma, COPD, and ACO cases compared with PCPs and pulmonologists in the primary case set; and (2) a determination of interrater and intra-rater agreement among PCPs and pulmonologists in the primary case set using Fleiss' kappa. The F1 score is the harmonic mean of precision and recall; therefore, this score considers both false positives and false negatives.

Key exploratory objectives included (1) an examination of the diagnostic accuracy of the AC/DC tool compared with PCPs and pulmonologists in diagnosing asthma, COPD, and ACO in the primary case set subgroups based on the expert panel scores for the difficulty of diagnosis (lower tertile [easy], middle tertile [moderately hard], and upper tertile [hard]); and (2) an examination of the diagnostic accuracy of PCPs and pulmonologists in diagnosing cases in the exploratory case set.

Statistical analysis

Expert panel diagnosis cases were divided into the primary case set (asthma, COPD, and ACO cases), and the exploratory case set (including cases of diseases other than asthma, COPD, and ACO). We analyzed the primary outcome using a Bayesian model that jointly modeled each patient's true disease status (ie, expert panel diagnosis) as a categorical random variable and the diagnoses given for each patient by each physician or algorithm (determined using

multinomial logistic regression). The multinomial logistic regression model included a separate intercept term for each combination of disease (asthma, COPD, and ACO) and group (PCPs, pulmonologists, and the AC/DC tool), as well as a random case and random rater effect. The primary analysis included the first diagnosis for each case by the physicians. Repeated diagnoses by the same physician of the same patient were excluded but were considered for an estimation of intra-rater reliability.

The key objective of a superiority trial is to demonstrate that a new treatment or device is better than an active control, a placebo, or a conventional method, whereas a noninferiority trial is designed to show that treatments are not unacceptably worse than (or noninferior to) the comparator.[19] A machine learning tool could be valuable to PCPs if it is superior to PCPs without necessarily needing to be superior to pulmonologists. Hence, the superiority of the AC/DC tool was tested against PCPs. However, after testing its superiority to PCPs, noninferiority to pulmonologists was tested followed by superiority to pulmonologists. The null hypothesis was tested against the alternative hypothesis for superiority of AC/DC tool versus PCPs and pulmonologists for the primary outcome, and a 10% noninferiority margin versus pulmonologists was used. A similar margin (10%) was previously used in the literature.[20] The main analysis used the first primary symptom definition. A sensitivity analysis was also performed using the second symptom definition for the primary outcome. Point estimates and their 95% credible intervals for the primary analysis were obtained as the medians, and the 2.5th and 97.5th percentiles of the posterior distribution for the average diagnostic accuracy of the AC/DC tool, PCPs, and pulmonologists, as well as for the differences in average accuracy among them. The calculation of differences and their 95% credible intervals allows a quantification of uncertainty regarding the diagnostic performance of each group (AC/DC, PCPs, and pulmonologists) and a judgment about whether between-group differences are likely due to chance.

Power calculations were based on simulations in which it was assumed that pulmonologists provide 60% correct diagnoses and the AC/DC tool 82% or greater, with the assumption that pulmonologists would perform better than PCPs. With at least 30, 30, and 20 patient cases with a panel diagnosis of asthma, COPD, and ACO, respectively, approximately 90% power was achieved for a comparison of algorithms compared with 50 pulmonologists. Because all pulmonologists could not review all clinical cases, the number of pulmonologists was increased to achieve a similar total number of reviewed cases. The number of PCPs chosen was the same as the number of pulmonologists. We conducted statistical analyses in R software (version 3.6.1, R Foundation for Statistical Computing, Vienna, Austria) using the RStan R package for the primary analysis.[21-23] Details of missing data imputation are presented in the Supplemental Text.

RESULTS

Baseline Demographics and Clinical Characteristics

This analysis included 116 patient (asthma, n = 53; COPD, n = 43; ACO, n = 7; and other, n = 13) assigned an expert panel diagnosis (ie, n = 103 [53 + 43 + 7] in the primary case set (the diagnosis of other was not included) and 116 in the exploratory case set; consensus was not achieved for three cases) (Figure 1). Table II lists baseline demographics and clinical characteristics of patients.

Of the 116 patients with an expert panel diagnosis used to evaluate the AC/DC tool, 95 (82%) had no missing data for the 12 variables, whereas of the remaining 21 patients (18%), 12 were missing information on dyspnea (10%), 13 were missing information on wheeze (11%), and eight were missing cough symptom information (7%; some of the 21 patients had missing information on more than one symptom). Other variables such as demographic information, spirometry results, smoking information, and comorbidity information were completely available for all 116 patients.

In total, 360 physicians (180 PCPs and 180 pulmonologists) from nine countries (20 PCPs and 20 pulmonologists from each country) were included with mean postresidency practice times comparable between PCPs (8.4-27.3 years) and pulmonologists (9.7-22.3 years).

Average diagnostic accuracy

Average diagnostic accuracy of the AC/DC tool was superior to that of PCPs (median difference, 24%; 95% posterior credible interval [CrI]: 17% to 29%; $P < .0001$) and was noninferior and superior to that of pulmonologists (median difference, 12%; 95% CrI: 6% to 17%; $P < .0001$ for noninferiority and $P = .0006$ for superiority) in the correct diagnosis of asthma, COPD, and ACO (based on expert panel diagnosis). The average diagnostic accuracy of pulmonologists was superior to that of PCPs (median difference, 12%, 95% CrI: 8% to 15%) (Figure 2). Sensitivity analyses showed results similar to those of the main analysis (AC/DC tool vs PCPs, median difference, 24%, 95% CrI: 17% to 30%; AC/DC tool vs pulmonologists, median difference, 11%, 95% CrI: 5% to 17%; pulmonologists vs PCPs, median difference, 12%, 95% CrI: 9% to 16%).

Secondary measures of diagnostic performance

For sensitivity (the percentage of true positive diagnoses [based on the expert panel] made from all diagnoses with each disease), the AC/DC tool correctly identified higher proportions of asthma and COPD patients, whereas PCPs and pulmonologists correctly provided a diagnosis for more ACO patients. The precision (percentage of true positive diagnoses from the total positive diagnoses made) results for the diagnosis of asthma was similar between the AC/DC tool and pulmonologists and only slightly lower for PCPs, whereas precision results for COPD and ACO were higher for PCPs and pulmonologists compared with the AC/DC tool. The F1 score (a measure of accuracy that combines sensitivity and precision) for the AC/DC tool was higher than that for PCPs and pulmonologists for the diagnosis of asthma, better than for PCPs and similar to pulmonologists for diagnosing COPD, and less than PCPs and pulmonologists for ACO. The negative predictive value (percentage of true negative diagnoses given from the negative diagnoses made) was higher for the AC/DC tool for asthma and COPD than those for PCPs and pulmonologists, whereas the values for PCPs and pulmonologists were slightly higher than the AC/DC tool for ACO. Specificity (percentage of true negative diagnoses made from all diagnoses for patients each of whom did not have a diagnosis) values were similar for the AC/DC tool, PCPs, and pulmonologists for the diagnosis of asthma, lower for the AC/DC tool versus PCPs and pulmonologists for the diagnosis of COPD, and higher for the AC/DC tool versus PCPs and pulmonologists for the diagnosis of ACO (Table III).

Diagnostic accuracy by case difficulty

The proportion of cases that PCPs and pulmonologists correctly diagnosed declined with increasing case difficulty, as

TABLE II. Baseline demographics and clinical characteristics of clinical cases of patients aged 35 years and older included in this analysis, by diagnosis assigned by expert panel (primary and exploratory case sets)

Characteristic	Asthma (n = 53)	Chronic obstructive pulmonary disease (n = 43)	Asthma/chronic obstructive pulmonary disease overlap (n = 7)	Others* (n = 13)	Total (n = 116)
Age, y	58.0 ± 11.88	66.4 ± 10.14	64.4 ± 3.82	56.2 ± 9.96	61.3 ± 11.43
Age at onset of lung problems, y	28.1 ± 22.44	55.0 ± 17.31	25.2 ± 26.71	45.1 ± 16.87	39.8 ± 23.78
Female, n (%)	32 (60.4)	13 (30.2)	4 (57.1)	7 (53.8)	56 (48.3)
Body mass index, kg/m^2	29.2 ± 5.32	27.3 ± 5.07	27.9 ± 6.46	25.0 ± 3.86	28.0 ± 5.27
Smoking status, n (%)					
Current smoker	2 (3.8)	20 (46.5)	3 (42.9)	2 (15.4)	27 (23.3)
Former smoker	19 (35.8)	23 (53.5)	4 (57.1)	7 (53.8)	53 (45.7)
Never smoked	32 (60.4)	0 (0.0)	0 (0.0)	4 (30.8)	36 (31.0)
Number of pack-years†	13.9 ± 16.09	35.3 ± 16.82	29.3 ± 11.22	20.9 ± 17.35	27.5 ± 18.51
Family history of respiratory diseases/problems, n (%)	33 (62.3)	17 (39.5)	3 (42.9)	6 (46.2)	59 (50.9)
Current respiratory medications, n (%)					
None	14 (26.4)	7 (16.3)	1 (14.3)	10 (76.9)	32 (27.6)
Reliever therapy	7 (13.2)	6 (14.0)	0	2 (15.4)	15 (12.9)
Maintenance therapy	32 (60.4)	30 (69.8)	6 (85.7)	1 (7.7)	69 (59.5)
Allergy, n (%)	24 (45.3)	5 (11.6)	—	1 (7.7)	30 (25.9)
Prebronchodilator FEV_1, L	2.7 ± 0.84	2.0 ± 0.62	1.5 ± 0.54	3.0 ± 0.77	2.4 ± 0.85
Reversibility (%)	7.7 ± 7.28	2.8 ± 7.79	21.8 ± 10.86	3.5 ± 4.76	6.6 ± 8.39

Data are presented as means ± SDs, unless otherwise specified. Additional data provided to the expert panel to make the diagnosis were the Asthma Control Questionnaire score, Clinical Chronic Obstructive Pulmonary Disease Questionnaire score, current symptoms, and disease history.
*Patient cases with a diagnosis of others were excluded from the primary case set.
†Analysis included only current and former smokers.

assessed by the expert panel (Figure 3). The AC/DC tool showed a notably higher percentage of accuracy for the hardest cases across all three categories compared with PCPs and pulmonologists (Figure 3) (the study was not powered to determine statistical significance in the diagnostic accuracy of the tool, PCPs, and pulmonologists by case difficulty).

Interrater and intra-rater agreement

Fleiss' kappa for inter-rater agreement for diagnostic consensus was higher among pulmonologists than PCPs across all diagnoses (0.29 [95% CrI: 0.25-0.33] and 0.19 [95% CrI: 0.16-0.22], respectively), as was intra-rater reliability (0.55 [95% CrI: 0.51-0.59] and 0.48 [95% CrI: 0.44-0.52], respectively). Interrater agreement was high for both definitions used in the AC/DC algorithm (Figure 4).

Performance of AC/DC tool and physicians in exploratory case set

The diagnostic accuracy for both PCPs and pulmonologists in the exploratory case set was the same as that in the primary set, whereas the diagnostic accuracy of the AC/DC tool was lower in the exploratory case set (Figure 5).

DISCUSSION

This multinational, noninterventional, observational study used deidentified, real-life clinical practice case data to determine the diagnostic accuracy of the AC/DC tool (developed by machine learning from data in an electronic medical record database) versus PCPs and pulmonologists in the diagnosis of asthma, COPD, and ACO in patients aged 35 years and older. The primary objective of this study was met; the average diagnostic accuracy of the AC/DC tool for these diagnoses was superior to that of PCPs and noninferior and superior to that of pulmonologists. Furthermore, the diagnostic accuracy of pulmonologists was superior to that of PCPs, as might be expected by virtue of their medical specialization.

In this study, the AC/DC tool displayed greater sensitivity for diagnosing asthma and COPD cases compared with PCPs or pulmonologists, and accuracy and precision values for the AC/DC tool were similar to those reported elsewhere for other machine learning models.[17,24] However, when diagnosing ACO, the diagnostic performance of the AC/DC tool was considerably lower than that of PCPs and pulmonologists. The small sample size (n = 7) might have contributed to this finding, but several other reasons might explain these results. First, machine learning algorithms can struggle when faced with class imbalance, and patients with ACO were the least common class of patients in the training data.[17] Second, pulmonologists also had the lowest sensitivity for this diagnosis, perhaps because of variations in definitions and perceptions of this disease among physicians and countries.[4] Indeed, some countries do not have a specific definition for ACO in their guidelines, and neither the Global Initiative for Asthma nor the Global Initiative for Chronic Obstructive Lung Disease considers ACO to be a specific diagnosis.[3,4] Third, the features used to develop the AC/DC tool may not be ideally suited for distinguishing ACO from COPD. One of the characteristics that the clinicians noted this study was the age at onset of respiratory symptoms. A younger age at onset of symptoms is one of the features that, in a patient with persistent airflow limitation, drives a clinical decision toward ACO. However, data for age at onset of respiratory symptoms were not recorded in the Optum database, and thus they were not included in the development of AC/DC. Moreover, although ACO is an interesting construct, double-blind randomized clinical studies are lacking on the treatment of ACO, and current safety recommendations are based on observational studies.[11,12,25]

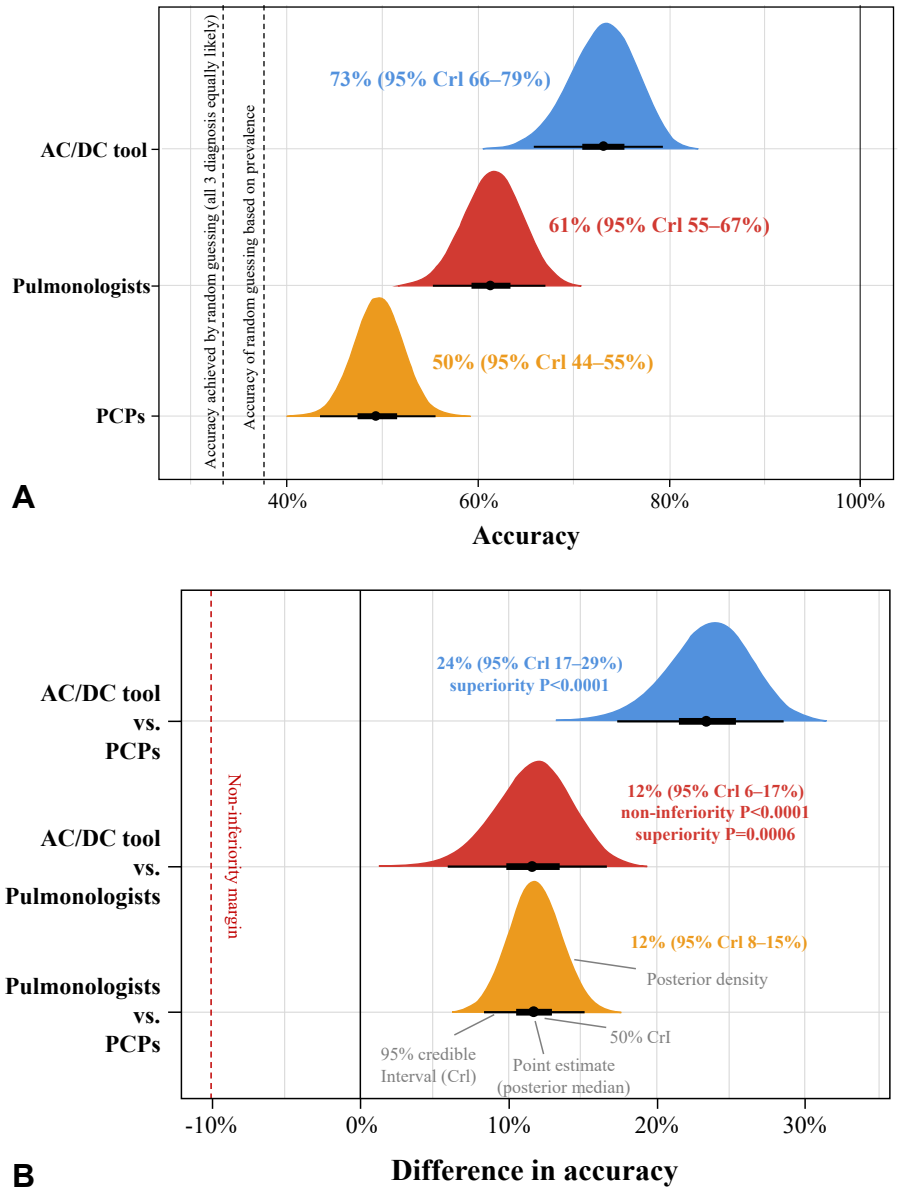

FIGURE 2. (**A**) Overall diagnostic accuracy and (**B**) difference in diagnostic accuracy of Asthma/Chronic Obstructive Pulmonary Disease (COPD) Differentiation Classification (AC/DC) tool, primary care physician (PCPs), and pulmonologists in diagnosis of clinical cases of asthma, COPD, and asthma/COPD overlap (primary case set). This analysis is based on primary symptom definition 1: a score greater than 0 on Asthma Control Questionnaire (ACQ) Q4 (dyspnea)/ACQ Q5 (wheeze)/clinical COPD questionnaire Q5 (cough). Differences in average overall diagnostic accuracy between the AC/DC tool and PCPs, and between the AC/DC tool and pulmonologists with reference to the expert panel diagnoses were analyzed. *CrI*, posterior credible interval.

The diagnostic consensus (interrater and intra-rater agreement) was higher among the pulmonologists than PCPs, but both were lower than the AC/DC tool. As expected, it was more likely that two pulmonologists would agree on a diagnosis than two PCPs because of their specialization. In contrast, the AC/DC tool was extremely consistent because all of the algorithms always produced a similar result from the same inputs. In addition, the difference in diagnostic accuracy between the AC/DC tool and both PCPs and pulmonologists increased with an increase in the difficulty of diagnosis, as assessed by the expert panel. These results suggest that the AC/DC tool has the potential to improve accuracy and specificity in the differential diagnosis of asthma and COPD, especially for cases that are more difficult to diagnose. These results are from estimates from a single point in time rather than longitudinal data. The use of longitudinal data in a primary care setting would allow PCPs to determine a response to treatment that could support a clinical diagnosis. This might partly explain the diagnostic accuracy of PCPs compared with the AC/DC tool and pulmonologists.

When the AC/DC tool misdiagnosed cases, it categorized patients with asthma as having a high probability of COPD and tended to assign a COPD diagnosis to ACO. However, this was

TABLE III. Sensitivity, precision, F1 score, negative predictive value, and specificity of AC/DC tool, PCPs, and pulmonologists in diagnosis of asthma, COPD, and ACO (primary case set)

Measures	Asthma* (n = 36)	COPD* (n = 36)	ACO* (n = 7)
Sensitivity (recall) (median [95% CrI])			
AC/DC	74% (65 to 81)	87% (80 to 91)	1% (0 to 6)
PCPs	39% (31 to 48)	62% (54 to 69)	41% (26 to 58)
Pulmonologists	52% (44 to 60)	73% (66 to 80)	47% (31 to 64)
AC/DC vs PCPs	35% (30 to 39)	25% (20 to 30)	−40% (−54 to −25)
AC/DC vs. pulmonologists	21% (17 to 26)	13% (10 to 18)	−46% (−60 to −30)
Pulmonologists vs PCPs	13% (9 to 18)	11% (7 to 16)	6% (−1 to 14)
Precision (positive predictive value) (median [95% CrI])			
AC/DC	88% (81 to 93)	66% (54 to 77)	5% (1 to 24)
PCPs	82% (73 to 89)	74% (63 to 82)	14% (6 to 27)
Pulmonologists	88% (80 to 93)	81% (71 to 88)	22% (10 to 39)
AC/DC vs PCPs	6% (1 to 12)	−7% (−12 to −3)	−8% (−17 to 4)
AC/DC vs pulmonologists	1% (−3 to 5)	−15% (−20 to −10)	−15% (−28 to −3)
Pulmonologists vs PCPs	5% (1 to 10)	7% (4 to 11)	7% (3 to 13)
F1 score (median [95% CrI])			
AC/DC	80% (74 to 85)	75% (66 to 82)	2% (1 to 10)
PCPs	53% (44 to 61)	67% (60 to 73)	21% (10 to 35)
Pulmonologists	65% (58 to 72)	77% (70 to 82)	30% (15 to 46)
AC/DC vs PCPs	27% (22 to 33)	8% (2 to 13)	−18% (−31 to −8)
AC/DC vs pulmonologists	15% (11 to 19)	−2% (−7 to 3)	−27% (−42 to −13)
Pulmonologists vs PCPs	13% (8 to 17)	10% (6 to 13)	8% (4 to 14)
Negative predictive value (median [95% CrI])			
AC/DC	81% (71 to 88)	85% (75 to 92)	91% (84 to 96)
PCPs	65% (54 to 74)	72% (61 to 81)	93% (87 to 97)
Pulmonologists	70% (60 to 79)	79% (70 to 87)	94% (89 to 98)
AC/DC vs PCPs	16% (12 to 20)	13% (8 to 18)	−2% (−5 to 0)
AC/DC vs pulmonologists	10% (7 to 14)	5% (2 to 10)	−3% (−6 to −1)
Pulmonologists vs PCPs	6% (4 to 8)	7% (4 to 11)	1% (0 to 3)
Specificity (median [95% CrI])			
AC/DC	92% (88 to 95)	63% (53 to 72)	98% (96 to 99)
PCPs	93% (91 to 95)	82% (77 to 85)	77% (74 to 80)
Pulmonologists	94% (92 to 96)	86% (81 to 89)	84% (81 to 87)
AC/DC vs PCPs	−1% (−4 to 2)	−19% (−27 to −12)	21% (18 to 24)
AC/DC vs pulmonologists	−2% (−5 to 1)	−23% (−30 to −16)	13% (11 to 16)
Pulmonologists vs PCPs	1% (−1 to 3)	4% (1 to 7)	7% (5 to 10)

AC/DC, Asthma/COPD Differentiation Classification; *ACO*, asthma/COPD overlap; *CrI*, posterior credible interval; *COPD*, chronic obstructive pulmonary disease; *PCP*, primary care physician.

Data are presented as posterior medians (95% CrI).

Combined posterior from 100 multiple imputations, each for patients accepted by one or more algorithm for that multiple imputation. This analysis is based on primary symptom definition 1: score greater than 0 on Asthma Control Questionnaire (ACQ) Q4 (dyspnea), ACQ Q5 (wheeze), or Clinical COPD Questionnaire Q5 (cough). Sensitivity (recall) is the percentage of true positive diagnoses from all diagnoses for patients who had the disease (expert panel diagnosis). Specificity refers to the percentage of true negative diagnoses given from all diagnoses for patients who did not have a disease. The F1 score is a measure of accuracy that combines recall and precision using the harmonic mean. Positive predictive value (precision) is the percentage of true positive diagnoses given from the positive diagnoses that were made. The negative predictive value is the percentage of true negative diagnoses given from the negative diagnoses that were made.

*n includes only asthma, COPD, and ACO patients who were not rejected by at least one algorithm for at least one multiple imputation.

not observed with PCPs or pulmonologists, who more often misclassified patients with asthma as having ACO, and patients with COPD as having ACO or asthma. The Global Initiative for Asthma[4] recommends that patients with asthma or ACO receive an inhaled corticosteroid–based therapy because it reduces the risk for hospitalization or death,[11-13,25] and many COPD patients may be safely treated with bronchodilators alone.[3] Thus, for safety reasons, modifying the training of the algorithm is essential to reflect the consequences of misdiagnosis.

Currently, some biomarkers are used as surrogates for the diagnosis of airway diseases owing to limitations and the availability of spirometry and to guide pharmacotherapy.[4] An analysis of exhaled nitric oxide and sputum or blood eosinophil count are sometimes used to determine corticosteroid responsiveness and adjust anti-inflammatory therapies in patients with asthma. Similarly, club cell secretory protein-16, surfactant protein D, and fibrinogen can predict the severity and risk for exacerbations in patients with COPD.[26] The blood eosinophil count was evaluated during the development of the AC/DC tool, but it had no impact on the outcomes, so it was not included in the tool. In addition, the literature does not report the use of biomarkers in machine learning models for the diagnosis of airway diseases.

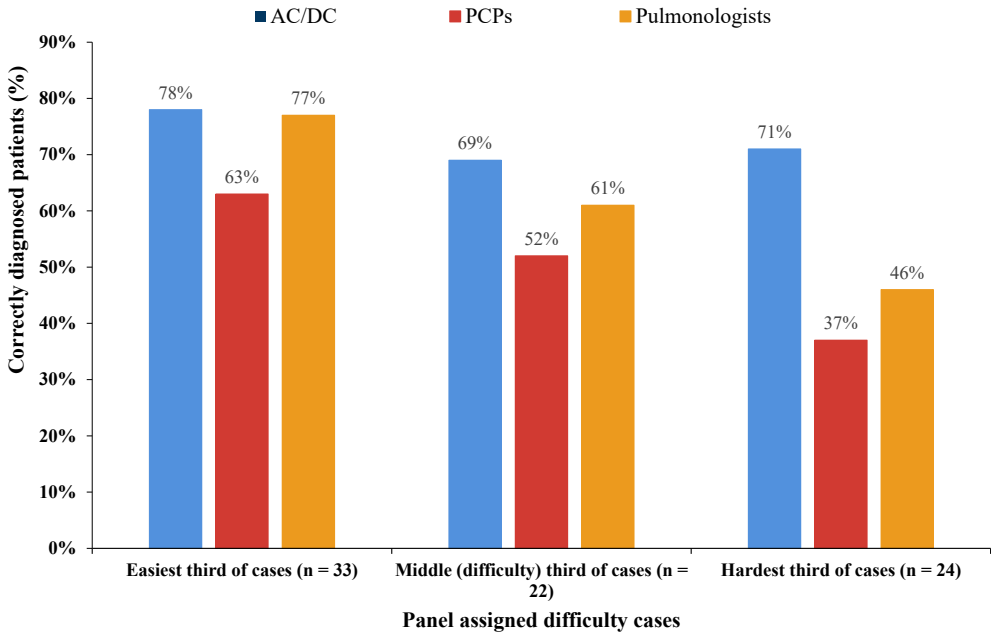

FIGURE 3. Diagnostic accuracy of Asthma/Chronic Obstructive Pulmonary Disease (COPD) Differentiation Classification (AC/DC) tool, primary care physicians (PCPs), and pulmonologists by case difficulty (assigned by expert panel) in primary case set. Based on tertiles of average of difficulty ratings of panel members: 1 = easy; 2 = moderately hard; 3 = hard to diagnose. The diagnosis was based on symptom definition 1, which includes an Asthma Control Questionnaire (ACQ) Q4 score greater than 0 (dyspnea), ACQ Q5 score greater than 0 (wheeze) and Clinical COPD Questionnaire Q5 score greater than 0 (cough); and symptom definition 2, which includes ACQ Q4 score greater than 1, ACQ Q5 greater than 1, and Clinical COPD Questionnaire Q5 score greater than 1.

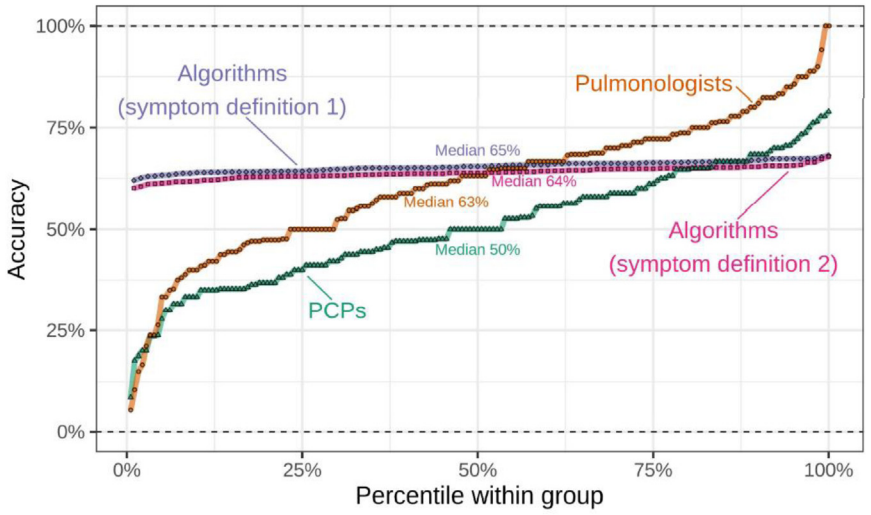

FIGURE 4. Variations in performance within Asthma/Chronic Obstructive Pulmonary Disease Differentiation Classification tool (algorithms), pulmonologists, and primary care physicians (PCPs) in primary case set.

The tool that is evaluated here was based on current specialist practice rather than diagnostic guidelines, which could theoretically lead to the reinforcement of common clinical errors in diagnosis. However, this external validation study has allowed an assessment of model performance based on the consensus diagnosis of a panel of experts who are highly familiar with existing guidelines and highly experienced in the field, which should avoid the routine application of diagnostic criteria and provide the most reliable diagnosis possible given the available information.

The AC/DC tool was designed to aid physicians in asthma and COPD diagnosis with higher accuracy after other diseases have been ruled out. Thus, this tool is not intended as a standalone model to rule out all diseases and accept only asthma, COPD, or ACO. This was particularly evident when the performance of the AC/DC tool was compared between the two

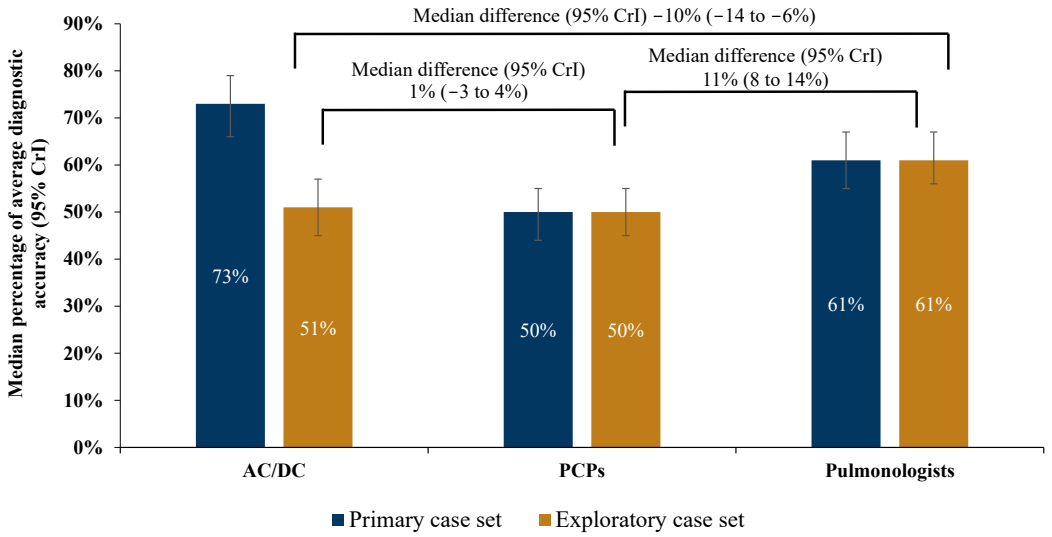

FIGURE 5. Comparison of performance of Asthma/Chronic Obstructive Pulmonary Disease Differentiation Classification (AC/DC) tool and physicians in primary and exploratory case sets. Data are presented as medians; error bars represent posterior credible interval (CrI) values. Combined posterior from 100 multiple imputations, each for patients accepted by one or more algorithm for that multiple imputation. This analysis is based on primary symptom definition 1: score greater than 0 on Asthma Control Questionnaire Q4 (dyspnea), Asthma Control Questionnaire Q5 (wheeze), and Clinical COPD Questionnaire Q5 (cough). *PCP*, primary care physician.

(primary and exploratory) case sets. The diagnostic accuracy of the AC/DC tool declined from the primary to exploratory case set because the tool does not have the option to distinguish a patient as other, and so it either rejects the patient or misdiagnoses the case as asthma, COPD, or ACO. An advantage of the AC/DC tool is that it provided a higher interrater agreement for the diagnosis of a specific disease across all diagnoses, whereas there was greater variability in decisions reported by both PCPs and pulmonologists.

This analysis had several limitations that need to be considered during clinical decision-making. (1) The study included only patients aged 35 years and older, and the AC/DC tool cannot be used in younger patients. (2) The AC/DC tool is not intended to provide a diagnose for patients on its own, but rather should be used in addition to spirometry to aid physicians in the differential diagnosis of asthma, COPD, and ACO, so it might have been worthwhile to include a group of physicians aided by the AC/DC tool in this study. (3) Although an electronic review of a case file may differ substantially from a face-to-face diagnosis in a physician's practice, the performance of the two physician groups aligned reasonably well with the published literature and the expectation that pulmonologists would outperform PCPs.[24,27] (4) The assumption that the clinicians had already ruled out all other potential causes of respiratory symptoms may not be clinically relevant, given (for example) the high cost of cardiac investigations for patients presenting with breathlessness. (5) The distinction between a history of allergic rhinitis and a history of chronic rhinitis, both of which were significant in developing the AC/DC tool, may not be clear in clinical practice. (6) Limited data on ACO were available in the database to train the AC/DC tool; however, the provision of data such as age of onset and reversibility to PCPs and pulmonologists did not improve their diagnostic accuracy versus the AC/DC tool. (7) Physicians included in this analysis were from the Ipsos (Paris, France) database rather than from random sampling for PCPs and pulmonologists. (8) Primary care physicians often also rely on social determinants of health, rather than spirometry.

The results of this study should be considered within the context of two assumptions: (1) the expert panel diagnoses of asthma, COPD, ACO, and other diseases are accurate for each patient, whereas the expert panel members were provided with only brief clinical details for the purpose of assigning a diagnosis; and (2) the AC/DC tool classified each case only to asthma, COPD, or ACO and did not have the other diagnostic option, unlike PCPs and pulmonologists for the primary analysis.

Further validation and assessment of the AC/DC tool is required given the lower performance for diagnosing ACO and the risk for hospitalization and death if patients with asthma or ACO are given a diagnosis of COPD and treated with bronchodilators alone.[11,12,25] The AC/DC tool accurately separates asthma from COPD, whereas the ACO diagnosis is not sensitive or specific, so it should prompt a reconsideration by clinicians to put patients in the asthma pathway to be safe.

The tool is currently under development and options for meeting regulatory requirements to make it available to physicians in the form of software as a medical device are currently being evaluated. Current discussions suggest that once the physician rules out other diagnoses and concludes that the patient has either asthma or COPD, the physician will ask the patient a series of five questions about symptoms that are entered into a smartphone, computer, or tablet, together with spirometry data (using a portable or other spirometer) from which the physician will then obtain the output of the AC/DC tool, which may be considered in the diagnosis. It is anticipated that this will take place at the physician's office and require no longer than 3 to 5 minutes to complete (Figure 6).

Subject to these validation and safety considerations, the tool has the potential to support a range of clinicians including nurse practitioners, PCPs, pulmonologists, and respiratory experts functioning across health care facilities such as mini-clinics,

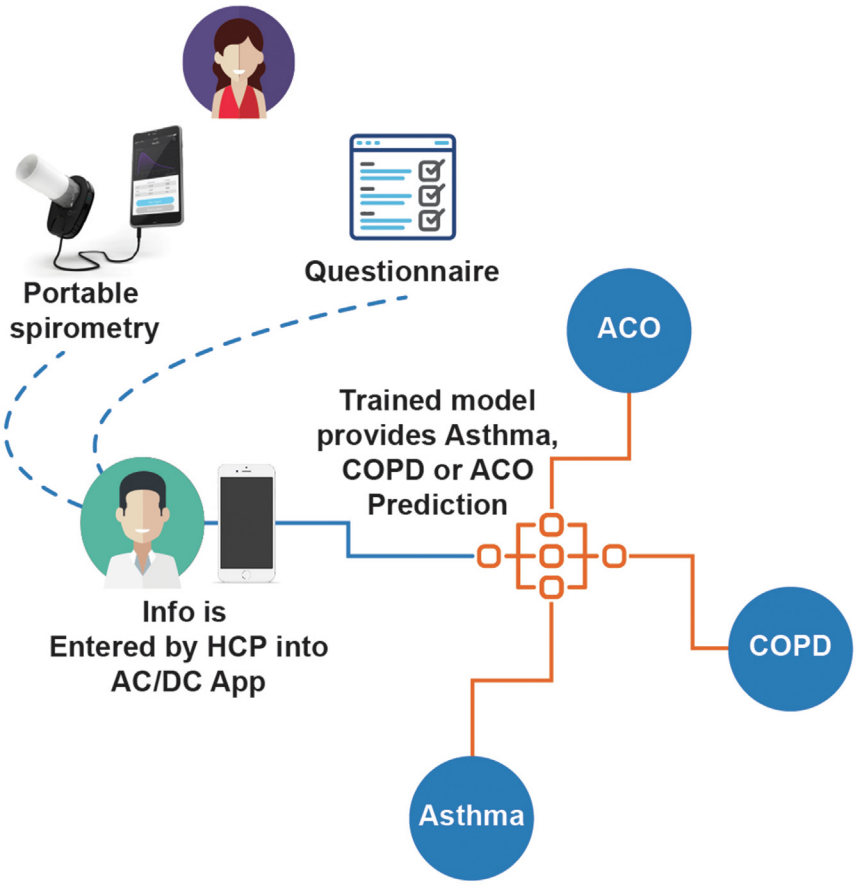

FIGURE 6. Potential clinical utility of Asthma/Chronic Obstructive Pulmonary Disease (COPD) Differentiation Classification (AC/DC) digital diagnostic tool. *ACO,* Asthma/COPD overlap; *HCP,* health care provider.

outpatient or satellite care centers, and large hospitals, in distinguishing between asthma and COPD in patients aged 35 years and older, in whom other causes of respiratory symptoms have been excluded. The noninvasive tool takes about 3 to 5 minutes to evaluate a patient, which benefits clinicians with busy schedules, and FEV_1 values generated through any spirometer can serve as input for the tool. It could be cost-effective and time-saving because fewer patient visits could be required to arrive at the diagnosis of asthma or COPD.

CONCLUSIONS

Overall, the AC/DC tool demonstrated superior diagnostic accuracy compared with PCPs and pulmonologists for correctly providing a diagnosis for patients with asthma and COPD, but not patients with ACO, as long as other' diagnoses can be ruled out before applying the AC/DC tool. The AC/DC tool has the potential to aid in the differential diagnosis for patients aged 35 years and older who have asthma and COPD and provide a valuable additional source of information to supplement final decision-making by practicing physicians.

DATA SHARING

Novartis is committed to sharing access to patient-level data and supporting documents from eligible studies with qualified external researchers. These requests are reviewed and approved by an independent review panel on the basis of scientific merit. All data provided are anonymized to respect the privacy of patients who have participated in the trial in line with applicable laws and regulations. The authors contributed to the preparation of the manuscript draft, along with critical review and approval of manuscript for submission to the journal. All authors contributed to intellectual content of the manuscript and approved for publication. Under the direction of authors Rabi Panigrahy, Preethi B and Ian Wright (professional medical writers; Novartis) assisted in the preparation of this article in accordance with the third edition of Good Publication Practice guidelines (http://www.ismpp.org/gpp3).

REFERENCES

1. Llanos JP, Ortega H, Germain G, Duh MS, Lafeuille MH, Tiggelaar S, et al. Health characteristics of patients with asthma, COPD and asthma-COPD overlap in the NHANES database. Int J Chron Obstruct Pulmon Dis 2018;13:2859-68.
2. Buist AS. Similarities and differences between asthma and chronic obstructive pulmonary disease: treatment and early outcomes. Eur Respir J Suppl 2003;39:30-5s.
3. Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for The Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Diseases. Accessed June 1, 2022. https://goldcopd.org/2022-gold-reports-2/
4. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention. Accessed June 1, 2022. https://ginasthma.org/wp-content/uploads/2022/07/GINA-Main-Report-2022-FINAL-22-07-01-WMS.pdf
5. Hosseini M, Almasi-Hashiani A, Sepidarkish M, Maroufizadeh S. Global prevalence of asthma-COPD overlap (ACO) in the general population: a systematic review and meta-analysis. Respir Res 2019;20:229.

6. Ho T, Cusack RP, Chaudhary N, Satia I, Kurmi OP. Under- and over-diagnosis of COPD: a global perspective. Breathe (Sheff) 2019;15:24-35.
7. Aaron SD, Boulet LP, Reddel HK, Gershon AS. Underdiagnosis and overdiagnosis of asthma. Am J Respir Crit Care Med 2018;198:1012-20.
8. Heffler E, Madeira LNG, Ferrando M, Puggioni F, Racca F, Malvezzi L, et al. Inhaled corticosteroids safety and adverse effects in patients with asthma. J Allergy Clin Immunol Pract 2018;6:776-81.
9. Heffler E, Crimi C, Mancuso S, Campisi R, Puggioni F, Brussino L, et al. Misdiagnosis of asthma and COPD and underuse of spirometry in primary care unselected patients. Respir Med 2018;142:48-52.
10. Broder MS, Raimundo K, Ngai KM, Chang E, Griffin NM, Heaney LG. Cost and health care utilization in patients with asthma and high oral corticosteroid use. Ann Allergy Asthma Immunol 2017;118:638-9.
11. Gershon AS, Campitelli MA, Croxford R, Stanbrook MB, To T, Upshur R, et al. Combination long-acting beta-agonists and inhaled corticosteroids compared with long-acting beta-agonists alone in older adults with chronic obstructive pulmonary disease. JAMA 2014;312:1114-21.
12. Kendzerska T, Aaron SD, To T, Licskai C, Stanbrook M, Vozoris NT, et al. Effectiveness and safety of inhaled corticosteroids in older individuals with chronic obstructive pulmonary disease and/or asthma. A population study. Ann Am Thorac Soc 2019;16:1252-62.
13. Kaplan A, Cao H, FitzGerald JM, Iannotti N, Yang E, Kocks JWH, et al. Artificial intelligence/machine learning in respiratory medicine and potential role in asthma and COPD diagnosis. J Allergy Clin Immunol Pract 2021;9:2255-61.
14. Topalovic M, Das N, Burgel PR, Daenen M, Derom E, Haenebalcke C, et al. Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. Eur Respir J 2019;53:1801660.
15. Badnjevic A, Gurbeta L, Custovic E. An expert diagnostic system to automatically identify asthma and chronic obstructive pulmonary disease in clinical settings. Sci Rep 2018;8:11645.
16. Abdel-Aal A, Lisspers K, Williams S, Adab P, Adams R, Agarwal D, et al. Prioritising primary care respiratory research needs: results from the 2020 International Primary Care Respiratory Group (IPCRG) global e-Delphi exercise. NPJ Prim Care Respir Med 2022;32:6.
17. Kaplan A, Cao H, Fitzgerald JM, Yang E, Iannotti N, Kocks JWH, et al. Asthma/COPD Differentiation Classification (AC/DC): machine learning to aid physicians in diagnosing asthma, COPD and asthma-COPD overlap (ACO). Am J Resp Crit Care Med 2020:A6285-A. American Thoracic Society (ATS) Conference 2020; May 15-20, 2020; Philadelphia, PA.
18. van de Hei SJ, Flokstra-de Blok BMJ, Baretta HJ, Doornewaard NE, van der Molen T, Patberg KW, et al. Quality of spirometry and related diagnosis in primary care with a focus on clinical use. NPJ Prim Care Respir Med 2020;30:22.
19. Head SJ, Kaul S, Bogers AJ, Kappetein AP. Non-inferiority study design: lessons to be learned from cardiovascular trials. Eur Heart J 2012;33:1318-24.
20. Meaden C, Joshi M, Hollis S, Higham A, Lynch D. A randomized controlled trial comparing the accuracy of general diagnostic upper gastrointestinal endoscopy performed by nurse or medical endoscopists. Endoscopy 2006;38:553-60.
21. R Core Team. R. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Accessed November 30, 2022. https://www.R-project.org/
22. Stan Development Team. RStan: the R interface to Stan. R package version 2.19.2. Accessed November 30, 2022. http://mc-stan.org/
23. Honaker J, King G, M. B. Amelia II: a program for missing data. J Stat Softw 2011;45:1-47.
24. Spathis D, Vlamos P. Diagnosing asthma and chronic obstructive pulmonary disease with machine learning. Health Inform J 2019;25:811-27.
25. Suissa S, Ernst P. Observational studies of inhaled corticosteroid effectiveness in COPD: lessons learned. Chest 2018;154:257-65.
26. Leung JM, Sin DD. Biomarkers in airway diseases. Can Respir J 2013;20:180-2.
27. Himes BE, Dai Y, Kohane IS, Weiss ST, Ramoni MF. Prediction of chronic obstructive pulmonary disease (COPD) in asthma patients using electronic medical records. J Am Med Inform Assoc 2009;16:371-9.

ONLINE REPOSITORY

SUPPLEMENTAL TEXT

Development and validation of Asthma/Chronic Obstructive Pulmonary Disease Differentiation Classification tool

The Asthma/chronic Obstructive Pulmonary Disease (COPD) Differentiation Classification (AC/DC) tool employs a machine learning–based algorithm using 12 variables (FEV_1, FEV_1/FVC, smoking pack-years, age, body mass index [BMI], dyspnea, wheeze, cough, diagnosis of allergic rhinitis, current smoker, never smoked, and diagnosis of chronic rhinitis) and was developed to aid primary care physicians and other physicians in the faster and more accurate diagnosis of asthma, COPD, or asthma/COPD overlap (ACO) and to reduce any delay in symptomatic patients receiving appropriate therapy.

The AC/DC tool was developed using the clinical characteristics of 400,000 or more patients aged 35 years or greater with asthma, COPD, and ACO (the data source was a US electronic health records database [Optum[E1]]). The tool has two components: an unsupervised model (it rejects a case not like trained cases), and supervised model that provides a diagnostic probability for these three diseases. In an internal validation, the supervised classification model provided considerable performance and achieved sensitivities of 0.98, 0.98, and 0.78, precision of 0.97, 0.97, and 0.92, and F1 scores of 0.98, 0.98, and 0.84, in diagnosing asthma, COPD, and ACO, respectively.[E2]

The index date was the date of the first diagnosis of asthma, COPD, or combined asthma and COPD (ACO) as given by the pulmonologist or allergist or reported in a hospitalization admission, if it occurred in the identification period (January 2010 to September 2017). Pre-index and post-index periods were 1 year from the index date. All patients needed to have at least 1 year before and after the index date to be included in the study.

Patients were split into three mutually exclusive cohorts: asthma, COPD, and ACO. All patients were required to be aged 35 years or older at the index date.

Asthma

Criteria defining the asthma cohort were:

1a. Had an inpatient or emergency room visit with a primary diagnosis of asthma in the identification period

1b. Had two outpatient visits with a diagnosis of asthma in the identification period

1c. Had either 1a or 1b

1d. Had no COPD diagnosis in the identification period

Chronic obstructive pulmonary disease

Criteria defining the COPD cohort were:

2a. Had an inpatient or emergency room visit with a primary diagnosis of COPD in the identification period

2b. Had two outpatient visits with a diagnosis of COPD in the identification period

2c. Had either 2a or 2c

2d. Had no asthma diagnosis in the identification period

Asthma/COPD overlap

Criteria defining the ACO cohort were:

3a. In both 1c and 2c

3b. Had a continuous diagnosis within 2 years after the index date

Use of diagnosis codes

Diagnosis codes for COPD were:

- International Classification of Diseases (ICD), Ninth Revision: 491.*, 492.*, 493.2*, and 496*
- ICD, 10th Revision: J41*, J42*, J43*, and J44*

Diagnosis codes for asthma were:

- ICD, Ninth Revision: 493*
- ICD, 10th Revision: J45* and J46*

Over 60 clinical features including spirometry and blood test results, comorbidities, and symptoms were extracted from patients' electronic health records (partly based on natural language processing of source notes, only the outcomes of which were provided by Optum):

- Patient characteristics
  - Cohort (asthma/COPD/ACO)
  - Age at diagnosis
  - Sex
  - Race
  - Ethnicity
  - BMI
  - Height
  - Weight
  - Smoking category (never smoked, current smoker, not currently smoking, other smoking status, and unknown status)
  - Pack-years (smoking)
- Spirometry
  - FEV_1; no distinction between pre- and postbronchodilator values, so given values could be either)
  - FVC; no distinction between pre- and postbronchodilator values, so given values could be either)
  - FEV_1 to FVC ratio (no distinction between pre- and postbronchodilator values, so given values could be either)
- Symptoms
  - Wheeze symptoms
  - Cough symptoms
  - Dyspnea symptoms
  - Tight chest symptoms
  - Sputum symptoms
  - Rhinitis symptoms
- Laboratory tests
  - IgE
  - Red blood cell count
  - Hemoglobin
  - Hematocrit
  - Mean corpuscular volume
  - Mean corpuscular hemoglobin
  - Mean corpuscular hemoglobin concentration
  - Platelet count
  - Mean platelet volume
  - Red blood cell distribution width
- Selected medications
  - ICS
  - ICS plus long-acting β-agonist combination medications
  - Leukotriene receptor antagonists
  - Methylxanthines
  - Other bronchodilators

- Monoclonal antibodies
- Phosphodiesterase-4 inhibitors
- Mast cell stabilizer
- Short-acting β-agonists
- Short-acting muscarinic antagonists
- Long-acting muscarinic antagonists
- Short-acting β-agonists plus short-acting muscarinic antagonists
- Long-acting muscarinic antagonist combinations
- Oral systemic corticosteroids

- Comorbidity score
  - Charlson comorbidity index
- Comorbidities (binary)
  - Charlson comorbidities
  - Chronic rhinitis
  - Allergic rhinitis
  - Hemoptysis
  - Gastroesophageal reflux disease

Diagnosis of reference standard cases by AC/DC tool

In total, 100 versions of the AC/DC algorithm (each with recall and a precision of 80% or greater for asthma, COPD, and ACO, and an overall accuracy of 95% or greater) were trained with different pseudorandom number seeds using data from an electronic health records database and 12 clinical features. Multiple versions of the AC/DC algorithm were trained to assess variations in its performance. The details of clinical cases were fed into all 100 versions of the AC/DC algorithm. Each version could either reject a case or diagnose the case as asthma, COPD, or ACO.

A total of 12 variables required by the AC/DC tool were extracted from each clinical case and prepared as the input for the algorithm. The AC/DC tool provided a diagnostic probability for asthma, COPD, and ACO. The disease with the highest prediction probability was considered to be the diagnosis by the AC/DC tool. Although in practice physicians will decide the final diagnosis considering the output of the AC/DC tool, for the purposes of this study the disease with the highest predicted probability was considered to be the diagnosis of the AC/DC tool. The confusion matrix of panel diagnosis versus diagnosis by algorithms and physicians is presented in Figure E1.

Participants

This analysis included deidentified clinical cases from the FOCUS study. These were real-life cases obtained from 15 general practitioners in the Netherlands, each of whom contributed 10 consecutive clinical cases of patients visiting the clinic with respiratory symptoms, who underwent systematic assessment in the FOCUS study and were maintained in a database.[E3] The database from the FOCUS study included more than 50 variables, including patients' demographics and baseline clinical characteristics, current inhaled medication, a medical history questionnaire including the Medical Research Council dyspnea scale, the Asthma Control Questionnaire (ACQ) (score of 0-6), and the Clinical COPD Questionnaire (CCQ) (score of 0-6), and spirometry results; these formed the expert panel cases. Cases were included in this evaluation study if (1) patients were aged 35 years or greater at the time of data collection; and (2) data were available for key variables (age, BMI, smoking status, spirometry, and diagnosis of allergic or chronic rhinitis) to run the AC/DC tool.

Secondary outcomes

Sensitivity (recall) is the percentage of true positive diagnoses from all diagnoses for patients who had the disease. Specificity refers to the percentage of true negative diagnoses given from all diagnoses for patients who did not have a disease. The F1 score is a measure of accuracy that combines recall and precision using their harmonic mean. Precision (positive predictive value) is the percentage of true positive diagnoses given of all positive diagnoses that were made. The negative predictive value is the percentage of true negative diagnoses given from all negative diagnoses that were made.

Statistical analysis

We analyzed the sensitivity, positive predictive value, F1 score, specificity, and negative predictive value using the primary analysis model. Interrater and intra-rater agreement were analyzed using Fleiss' kappa. The interrater analysis included data from the first diagnosis whereas the intra-rater analysis included data from a repeat diagnosis. All primary care physicians and pulmonologists were given repeated cases to analyze the reproducibility of the diagnosis and evaluate intra-rater variability. The number and percentage of cases from the primary case set that the AC/DC tool rejected for classification were summarized, and a prediction interval for a new algorithm used in new patients was given. For this and other analyses of binomial outcomes, a separate analysis with a conjugate β (0.5, 0.5) Jeffreys prior was conducted for each algorithm and the equally weighted mixture distribution of the resulting posteriors was used as the predictive distribution.

The diagnostic accuracy of the AC/DC tool and physicians was also assessed in subgroups of cases by difficulty assigned by the expert panel. Finally, the diagnostic performance of the AC/DC tool and physicians was compared using the primary combined with the exploratory case set. For this analysis, a diagnosis of asthma, COPD, or ACO by the AC/DC tool for a case from the other category was treated as a wrong diagnosis.

If variables necessary to run the AC/DC tool contained missing data, multiple imputation with 100 imputations was used under a missing at random assumption using the latent multivariate normal model.[E4] For each multiple imputation, only patients not rejected by at least one algorithm were included in the analysis. Demographic- and disease-related information including FEV_1, the FEV_1/FVC ratio, and individual answers to patient questionnaires were included in the imputation model.

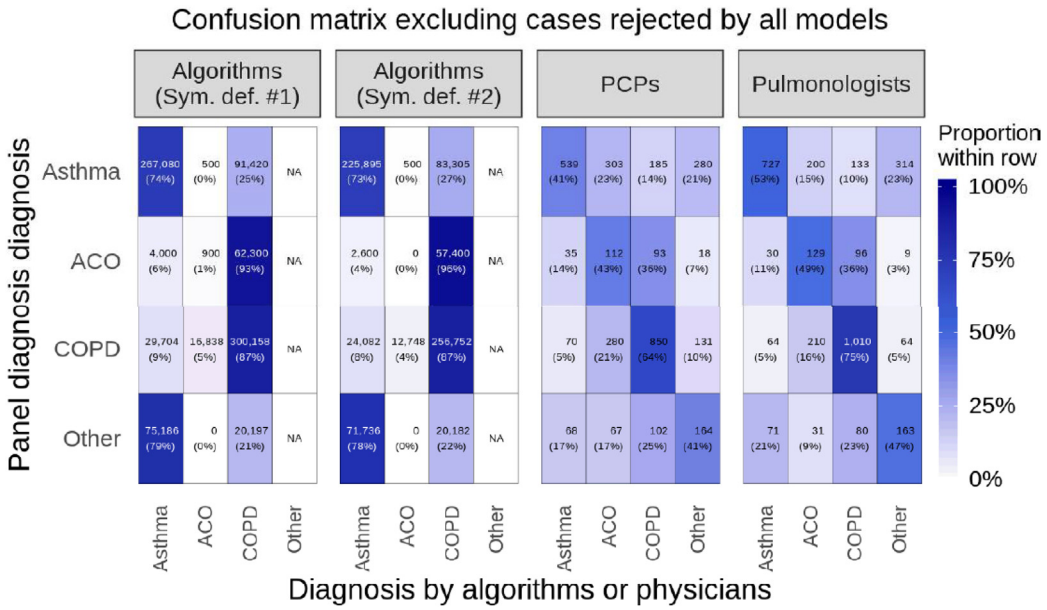

FIGURE E1. Confusion matrix of panel diagnosis versus diagnosis by algorithms and physicians. Sym. Def. 1: score greater than 0 on Asthma Control Questionnaire (ACQ) Q4 (dyspnea), ACQ Q5 (wheeze), or Clinical COPD Questionnaire Q5 (cough). Sym. Def. 2: greater than 1 on ACQ Q4, ACQ Q5, Clinical COPD Questionnaire Q5. Percentages are per row excluding rejected cases. Confusion matrices for physicians exclude repeat diagnoses of the same case. *ACO*, asthma COPD overlap; *COPD*, chronic obstructive pulmonary disease; *PCP*, primary care physician; *Sym. Def.*, symptom definition.

REFERENCES

E1. Optum. Electronic Health Records (HER) Data. Accessed November 30, 2022. https://www.optum.com/business/life-sciences/real-world-data/ehr-data.html

E2. Kaplan A, Cao H, Fitzgerald JM, Yang E, Iannotti N, Kocks JWH, et al. Asthma/COPD Differentiation Classification (AC/DC): machine learning to aid physicians in diagnosing asthma, COPD and asthma-COPD overlap (ACO). Am J Resp Crit Care Med 2020:A6285-A. American Thoracic Society (ATS) Conference 2020; May 15-20, 2020; Philadelphia, PA.

E3. van de Hei SJ. Flokstra-de Blok BMJ, Baretta HJ, Doornewaard NE, van der Molen T, Patberg KW, et al. Quality of spirometry and related diagnosis in primary care with a focus on clinical use. NPJ Prim Care Respir Med 2020;30:22.

E4. Honaker J, King G, Amelia MB II. A program for missing data. J Stat Softw 2011;45:1-47.