Variation Among Spirometry Interpretation Algorithms

Katrina A D'Urzo, Florence Mok, and Anthony D D'Urzo

Several algorithms exist to facilitate spirometric interpretation in clinical practice, yet there is a lack of consensus on how spirometric criteria for asthma, COPD, and restrictive disorders should be incorporated into spirometry interpretation algorithms suitable for use in day-to-day primary care management. The purpose of this review was to identify and describe the variability that exists among spirometry interpretation algorithms and how this might be relevant to the interpretation of spirometric data of common conditions encountered in primary care. MEDLINE, Embase, and mainstream search engines were used to identify all English-language spirometry interpretation algorithm–related material between January 1990 and December 2018. Eight variations in spirometry interpretation algorithms were identified via specific a priori assumptions that each spirometry interpretation algorithm should contain content consistent with national and international guidelines related to spirometry interpretation. Of the 26 spirometry interpretation algorithms identified, 5 were deemed impractical for day-to-day use in primary care (19%), 23 lacked a logic string leading to the postbronchodilator FEV_1/FVC (88%), 4 relied on postbronchodilator change in FEV_1 to distinguish between asthma and COPD (15%), 24 lacked a prompt for bronchodilator challenge when FEV_1/FVC was considered to be at a normal level (92%), 12 did not indicate whether the data represented a prebronchodilator or postbronchodilator scenario (46%), 7 did not include a logic string that considers mixed obstructive/restrictive defect (27%), 23 did not contain a prompt to refer for methacholine challenge testing when spirometry appeared normal (88%), and 2 spirometry interpretation algorithms did not include a logic string leading to restrictive disorder (8%). Our review suggests that there is considerable variability among spirometry interpretation algorithms available as diagnostic aids and that there is a need for standardization of spirometry interpretation algorithms in primary care. *Key words: spirometry interpretation algorithms; variation; primary care; clinician tools; family medicine; respiratory illness.* [Respir Care 2020;65(10):1585–1590. © 2020 Daedalus Enterprises]

Introduction

Guidelines on the management of common conditions like asthma and COPD stress the importance of objective testing for diagnostic confirmation.[1,2] Guidelines suggest that an improvement in FEV_1 of at least 200 mL and 12% from baseline after bronchodilator administration (FEV_1 reversibility criteria) is consistent with a diagnosis of asthma, whereas a persistent reduction in $FEV_1/FVC < 0.70$ or the lower limit of normal after bronchodilator administration is in keeping with COPD.[3,4] Typically, a normal FEV_1/FVC associated with a reduction in the FVC below the predicted normal level is considered to be suggestive of restrictive ventilatory disorder.[2]

Clinical algorithms (ie, flow charts) are information resources specifically designed to represent a sequence of clinical decisions that guide patient care and facilitate clinical decision-making.[5] Although several spirometry interpretation algorithms exist, there is little information in the literature that describes the variability that may exist among these algorithms to guide clinicians as they interpret spirometry data. This information would be useful for evaluating how the content of available spirometry interpretation algorithms conform to current guidelines dealing with asthma, COPD, and restrictive defects and how any differences among spirometry interpretation algorithms may influence decision making.

Previous companion reports have recently described differences in 2 spirometry interpretation algorithms promoted for use in primary care.[6-8] Indeed, relevant differences in the interpretation of the same spirometric data were observed when the 2 algorithms were utilized by family physicians as stand-alone documents.[9] Although these findings do not highlight how differences in spirometry interpretation algorithms may influence the management of patients with asthma, COPD, and restrictive disorders, they reinforce concerns about how variations in spirometry interpretation algorithms may influence decision making and promote the potential for disease misclassification.[6,8]

The reports comparing the 2 spirometry interpretation algorithms noted above identified important limitations of a spirometry interpretation algorithm promoted for adoption in primary care that warrant consideration.[6,7] First, they note the lack of a logic string leading to a reduction in FEV_1/FVC (ie, the spirometric criteria for objective confirmation of COPD). Second, investigators describe that the lack of a logic string prompting bronchodilator challenge when FEV_1/FVC is normal may represent a lost opportunity for asthma confirmation because in some patients FEV_1 reversibility can be demonstrated under these circumstances.[7] The latter finding is particularly relevant because most patients with asthma encountered in primary care often present with normal lung function.[10,11] This review has 2 central goals: (1) to describe the variability that exists among available spirometry interpretation algorithms, and (2) to describe how variations in spirometry interpretation algorithms may guide the interpretation of spirometric data. Examining the variability between existing spirometry interpretation algorithms may provide a foundation for the standardization of spirometry interpretation algorithm use in primary care.

Are There Variations Among Spirometry Interpretation Algorithms Reported in the Medical Literature?

Given the nature of this review, the spirometry interpretation algorithms referred to herein were identified based on specific a priori assumptions such that individual spirometry interpretation algorithms should contain core spirometric diagnostic content that was consistent with national and international guideline recommendations related to the diagnosis of asthma, COPD, and restrictive lung disease.[12] Spirometric diagnostic criteria consistent with asthma, COPD, and restrictive lung disease include: (1) an improvement in postbronchodilator FEV_1 of at least 12% and 200 mL from baseline levels, regardless of whether FEV_1/FVC was above or below 0.70 or the lower limit of normal; (2) a persistent reduction in the postbronchodilator $FEV_1/FVC <$ 0.70 or lower limit of normal; and (3) a normal FEV_1/FVC associated with a reduction in the FVC below the predicted normal level.[2]

To meet the criteria of what constituted a spirometry interpretation algorithm,[5] documents needed to describe the interpretation process in a step-by-step fashion using identified logic strings and decision nodes in a flow chart format that contained spirometric diagnostic criteria outlined previously.[2] Documents that were composed primarily of

Ms D'Urzo is affiliated with the School of Medicine, Royal College of Surgeons in Ireland, Dublin, Ireland. Drs Mok and D'Urzo are affiliated with the Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada.

This work was supported by the Comprehensive Research Experience for Medical Students, Faculty of Medicine, Department of Family and Community Medicine, University of Toronto. The authors have disclosed no conflicts of interest.

Supplementary material related to this paper is available at http://www.rcjournal.com.

Correspondence: Anthony D D'Urzo MD MSc. E-mail: tonydurzo@sympatico.ca.

DOI: 10.4187/respcare.07294

Table 1. Variations in Spirometry Interpretation Algorithms Identified

Category	Description
A	Impractical for use in primary care for point of care management; included data that (1) could not be captured same day in the primary care setting, or (2) are not required for asthma and COPD diagnosis as outlined in published guidelines (eg, total lung capacity, diffusion capacity)
B	Lack of a logic string leading to postbronchodilator FEV_1/FVC
C	Use of postbronchodilator change in $FEV_1 > 12\%$ and > 200 mL to distinguish between asthma and COPD
D	Lack of prompt for bronchodilator challenge when $FEV_1/FVC > 0.70$ or the lower limit of normal
E	Unclear whether the data represent a prebronchodilator or postbronchodilator scenario
F	Lack of logic string leading to consideration of mixed obstructive/restrictive defect
G	Lack of prompt to refer for methacholine challenge testing when spirometry appears normal (ie, when FEV_1/FVC is > 0.70 or the lower limit of normal and both FEV_1 and FVC are normal)
H	Lack of logic string leading to postbronchodilator FEV_1/FVC that is normal with a reduced FVC ($< 80\%$)

written text without logic strings, decision nodes, or flow chart format were not included in this review. MEDLINE, Embase, and mainstream search engines (ie, PubMed and Google Scholar) were used to identify all English-language spirometry interpretation algorithm-related material from January 1990 to December 2018. Key words used were "spirometry," "algorithm" and "interpretation."

Intra-rater Assessment and Reliability

One author with expertise in spirometry interpretation assessed each spirometry interpretation algorithm twice for the variations described in Table 1. Each assessment was separated by 1 week. Each spirometry interpretation algorithm may include one or more of the variations identified (Table 2). Intrarater differences were then reviewed and resolved following the 2 independent assessments. As indicated in Table 3, the level of intrarater agreement between the 2 assessments varied from fair to perfect agreement. The lowest level of agreement was observed for Category F, and the highest level of agreement was observed among Categories C and H. The observed levels of agreement between intrarater assessments highlight spirometry interpretation algorithm features that warrant additional discussion. In particular, the strong-to-almost perfect agreement seen among categories A, B, C, G, and H may suggest that these variations had more clearly defined assessment criteria; however, it is also possible that these variations are more consistently represented within existing spirometry interpretation algorithms.

Analysis

All statistical analyses were conducted with SPSS 24 (IBM, Armonk, New York). Frequency counts of the number of variations identified for each spirometry interpretation algorithm were computed. Cohen's kappa coefficients were calculated to assess intrarater reliability coding of each spirometry interpretation algorithm.

Table 2. Variations Among Identified Spirometry Interpretation Algorithms Based on Category

Algorithm	A	B	C	D	E	F	G	H
National Asthma Council and COPDX[13,14]		X		X	X	X	X	
Pellegrino et al[2]	X	X		X	X		X	
Enright and Hyatt[15]	X	X		X	X		X	
Peters et al (A)[16]	X	X		X	X	X	X	
Peters et al (B)[16]		X		X	X	X	X	
Gildea and McCarthy[17]		X		X			X	
Levy et al[18]		X		X	X		X	
Petty (A)[19]		X		X			X	
Petty (B)[20]		X		X			X	
Barreiro[21]		X		X		X	X	
Lowry[22]		X	X	X			X	
Hughes[23]	X	X		X	X		X	
Koegelenberg et al[24]		X		X	X		X	
Johnson and Theurer[25]		X		X			X	
Collen et al[26]		X		X	X		X	
Raghunath et al[27]		X		X			X	X
Jenkins et al[28]			X	X		X	X	
Maestu and de Pedro[29]		X		X	X		X	
Schneider et al[30]	X	X	X	X		X		X
Lung Function Practical[31]		X		X			X	
O'Connor and Manning[32]		X	X	X			X	
NAC Spirometry Quick Guide[33]		X		X		X	X	
Ra et al (A)[34]		X		X	X		X	
Ra et al (B)[34]		X		X	X		X	
D'Urzo et al[8]								
The Lung Association[8(Revised)]								
Sum	**5**	23	4	24	12	7	23	2

The letter "X" indicates that this feature (A–H from Table 1) is present in the spirometry interpretation algorithms identified in the reference cited.
The references beyond 14 are available online in supplementary materials at http://www.rcjournal.com.

Twenty-six ($n = 26$) spirometry interpretation algorithms meeting criteria outlined above were identified. Among these spirometry interpretation algorithms, 8 relevant variations were identified and agreed upon by the authors (Table 1).

Table 3. Intrarater Reliability

Category	Cohen's kappa (95% CI)	*P*
A	0.87 (0.84–0.89)	.001
B	0.84 (0.82–0.85)	.001
C	1.00 (1.00–1.00)	.001
D	0.62 (0.60–0.64)	.001
E	0.62 (0.55–0.70)	.001
F	0.50 (0.43–0.58)	.001
G	0.84 (0.82–0.85)	.001
H	1.00 (1.00–1.00)	.001

Algorithm Variability

Among the 26 spirometry interpretation algorithms, 5 were deemed impractical for day-to-day use in primary care (19%; Category A), and 23 lacked a logic string leading to the postbronchodilator FEV_1/FVC (88%; Category B). Four spirometry interpretation algorithms relied on the postbronchodilator change in FEV_1 to distinguish between asthma and COPD (15%; Category C). Twenty-four spirometry interpretation algorithms lacked a prompt for bronchodilator challenge when FEV_1/FVC was considered to be at a normal level (92%; Category D). In 12 of the spirometry interpretation algorithms, it was unclear whether the data represented a prebronchodilator or postbronchodilator scenario (46%; Category E). Seven spirometry interpretation algorithms did not include a logic string that considers mixed obstructive/restrictive defect (27%; Category F). Twenty-three spirometry interpretation algorithms did not contain a prompt to refer for methacholine challenge testing when spirometry appears normal (88%; Category G). Two spirometry interpretation algorithms (8%; Category H) did not contain a logic string leading to a restrictive disorder.

This review describes considerable variability that exists among spirometry interpretation algorithms available for adoption in primary care. The variability observed among the identified spirometry interpretation algorithms is driven to a great extent by the inclusion or exclusion of various logic string items, some of which are not in keeping with guideline recommendations for asthma and COPD diagnosis.

How Variations in the Identified Spirometry Interpretation Algorithms may Influence Interpretation of Spirometric Data?

Five of the spirometry interpretation algorithms identified would be considered impractical for point of use (Category A). Indeed, several of the spirometry interpretation algorithms recommended unnecessary tests (eg, diffusion capacity for carbon monoxide) for interpretation purposes, which is not feasible within the primary care setting at point of care. Furthermore, despite the clinical practice of β_2-agonist challenge testing, 12 spirometry interpretation algorithms failed to indicate whether the data represented was prebronchodilation or postbronchodilation scenarios (Category E). This limits the clinician's ability to make any meaningful decisions about the data being interpreted. Together, these findings are particularly important as diagnostic aids are more likely to be adopted by time-constrained clinicians if they can provide relevant and clear information in a timely manner.

Variations That Hinder COPD Diagnosis and Fail to Consider the Spirometric Overlap Between Asthma and COPD

Variations outlined in Categories B and C have important clinical implications because these variations result in clinicians being unable to confirm the diagnosis of COPD objectively: 23 lacked a logic string leading to the postbronchodilator FEV_1/FVC (88%; Category B), and changes in FEV_1 after bronchodilation (Category C) are used to distinguish asthma from COPD. Given the spirometric overlap that exists between asthma and COPD, spirometry interpretation algorithms that fulfill Category B may lead physicians to consider a spirometric interpretation of asthma diagnosis over COPD. This possibility was confirmed in previous work and emphasizes the importance of utilizing spirometry interpretation algorithms that take into account the spirometric overlap between asthma and COPD.[6,8] For example, one of the logic strings included in a new spirometry interpretation algorithm leads the clinician to consider both asthma and COPD when the FEV_1/FVC is reduced < 0.70 or the lower limit of normal and when the FEV_1 improves by 12% and 200 mL.[7,8] This is relevant because there is evidence to suggest that many if not most patients with moderate COPD meet the FEV_1 reversibility criteria required for asthma diagnosis.[9] The latter is underscored by reports that different spirometry interpretation algorithms result in differences in the interpretation of the same spirometric data among primary care physicians, including a consideration of asthma when the spirometry data contains features of both asthma and COPD.[6] Appreciating how to navigate through this type of clinical spirometric overlap is facilitated by understanding how disease characteristics may factor into data interpretation and decision making. For example, the response to a bronchodilator in asthma is described by an inverted U-shaped curve, where patients with normal or near normal baseline function typically exhibit little or no significant response, whereas those with moderate baseline obstruction tend to demonstrate the greatest response. Finally, the blunted bronchodilator

response in patients with severe asthma is thought to be related to increased airway wall thickness and lumen occlusion rather than a defect in smooth muscle response to a bronchodilator.[13]

Absence of Prompt for Bronchodilator Challenge When FEV_1/FVC Is Normal

Twenty-four of the spirometry interpretation algorithms identified did not prompt a bronchodilator challenge when FEV_1/FVC was > 0.70 or the lower limit of normal (Category D), a feature that would eliminate the opportunity to identify asthmatic patients with well-preserved lung function at the time of testing.[6,8] Indeed, normal FEV_1/FVC is not synonymous with the absence of heightened airway tone because administration of β_2-agonist challenge could result in clinically relevant improvements in FEV_1 under such circumstances.[8] Because most asthma patients encountered in primary care present with normal lung function based on FEV_1, FVC, and FEV_1/FVC criteria,[10,11] this common spirometric feature represents an important limitation that should be addressed in new spirometry interpretation algorithms.

Absence of Prompt for Methacholine Challenge When Spirometry Appears Normal

Almost all of the spirometry interpretation algorithms failed to prompt referral for methacholine challenge testing when spirometry appears normal (Category G), suggesting that the spirometry interpretation algorithms might lead clinicians to underidentify asthmatics with airway hyperresponsiveness despite normal spirometry. The suggestion in most spirometry interpretation algorithms that spirometry is "normal" may lead to a false sense of security among clinicians who might consider this finding to mean the absence of disease activity.

Absence of a Logic String Suggesting a Mixed Obstructive/Restrictive Defect

Many of the spirometry interpretation algorithms were structured in ways that do not include a logic string to consider a mixed obstructive/restrictive defect (Category F). A mixed defect scenario is depicted by a reduction in both FEV_1/FVC and FVC; however, it is not clear whether FVC is reduced due to air trapping associated with air flow obstruction or reduced lung volume. An improvement in FVC after bronchodilator challenge is helpful in distinguishing between air trapping and restrictive disease. Although a mixed obstructive/restrictive defect is described in some spirometry interpretation algorithms, it is important to highlight that this may be more accurately described as an obstructive defect associated with decreased FVC.

Normal FEV_1/FVC and Reduced FVC < 80%

Although restrictive disease is much less commonly encountered in primary care, only 2 of the spirometry interpretation algorithms lacked a logic string leading to postbronchodilator FEV_1/FVC that is normal with reduced FVC ($< 80\%$; Category H). By contrast, 23 spirometry interpretation algorithms lacked a logic string leading to the postbronchodilator FEV_1/FVC (88%; Category B) needed to objectively confirm a diagnosis of the more commonly encountered condition, COPD. This finding has particular clinical relevance within the context that COPD is significantly underdiagnosed in the primary care setting.[14]

Summary

Important limitations of this review are worth noting. First, the criteria used to define the variations described in the defined categories have not been objectively validated. However, the a priori assumptions used to identify variations in spirometry interpretation algorithms are in keeping with guideline principles promoting spirometry testing and interpretation and with the spirometric overlap that is known to exist between asthma and COPD.[9] Second, the spirometry interpretation algorithms identified here largely relate to interpretation strategies using spirometric indices that apply primarily to adults. Few, if any, spirometry interpretation algorithms exist specifically for younger patients; therefore, our findings cannot be generalized to pediatric management. In addition, it is possible that our search was not exhaustive and that other relevant spirometry interpretation algorithms were not included for evaluation. However, because we included spirometry interpretation algorithms outlined in well recognized international guidelines and publications, we believe that this type of omission would not minimize the clinical implications of our findings. Finally, the last 2 algorithms in Table 2 were intentionally developed in accordance with guidelines to address the problematic variations noted among the other spirometry interpretation algorithms included here.

Our review suggests that there is considerable variability among spirometry interpretation algorithms available for adoption as diagnostic aids in primary care. Available evidence suggests that the variability among some spirometry interpretation algorithms may lead to different interpretations of the same data.[6] Further studies are required to evaluate whether differences in the interpretation of similar data translates into differences in patient care in the clinical setting, including disease misclassification. We believe that our review may represent an important starting point toward the standardization of spirometry interpretation algorithms in primary care.

REFERENCES

1. Yang IA, Brown JL, George J, Jenkins S, McDonald CF, McDonald VM, et al. COPD-X Australian and New Zealand guidelines for the diagnosis and management of chronic obstructive pulmonary disease: 2017 update. Med J Aust 2017;207(10):436-442.
2. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. Eur Respir J 2005;26(5):948-968.
3. Bourbeau J, Bhutani M, Hernandez P, Marciniuk DD, Aaron SD, Balter M, et al. CTS position statement: pharmacotherapy in patients with COPD—an update. Can J Respir Crit Care, Sleep Med 2017;1(4):222-241.
4. Lougheed MD, Leniere C, Ducharme FM, Licskai C, Dell SD, Rowe BH, et al. Canadian Thoracic Society 2012 Guideline Update: Diagnosis and management of asthma in preschoolers, children and adults: executive summary. Can Respir J 2012;19(6):e81-e88.
5. Margolis CZ. Uses of clinical algorithms. JAMA 1983;249(5):627-632.
6. He X-O, D'Urzo A, Jugovic P, Jhirad R, Sehgal P, Lilly P. Differences in spirometry interpretation algorithms: influence on decision making among primary-care physicians. NPJ Prim Care Respir Med 2015;25(1):15008.
7. D'Urzo AD, Tamari I, Bouchard J, Jhirad R, Jugovic P. Limitations of a spirometry interpretation algorithm. Can Fam Physician 2011;57 (10):1153-1156.
8. D'Urzo AD, Tamari I, Bouchard J, Jhirad R, Jugovic P. New spirometry interpretation algorithm: Primary Care Respiratory Alliance of Canada approach. Can Fam Physician 2011;57 (10):1148-1152.
9. Hanania NA, Celli BR, Donohue JF, Martin UJ. Bronchodilator reversibility in COPD. Chest 2011;140(4):1055-1063.
10. Lusuardi M, De Benedetto F, Paggiaro P, Sanguinetti CM, Brazzola G, Ferri P, Donner CF. Randomized controlled trial on office spirometry in asthma and COPD in standard general practice. Chest 2006;129 (4):844-852.
11. Aaron SD, Vandemheen KL, Boulet L-P, McIvor RA, Fitzgerald JM, Hernandez P, et al. Overdiagnosis of asthma in obese and nonobese adults. Cmaj 2008;179(11):1121-1131.
12. Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2013;187(4):347-365.
13. Tashkin DP. Is it asthma, COPD, or something in between, and does it matter? Respir Care 2012;57(8):1354-1356.
14. Bednarek M, Maciejewski J, Wozniak M, Kuca P, Zielinski J. Prevalence, severity and underdiagnosis of COPD in the primary care setting. Thorax 2008;63(5):402-407.