The Accuracy of Patient History, Wheezing, and Laryngeal Measurements in Diagnosing Obstructive Airway Disease

Sharon E. Straus, MD
Finlay A. McAlister, MD
David L. Sackett, MD
Jonathan J. Deeks, MSc
for the CARE-COAD1 Group

Context The accuracy of the clinical examination in detecting obstructive airway disease (OAD) is largely unknown because of a paucity of methodologically rigorous studies.

Objective To determine the accuracy of patient history, wheezing, laryngeal height, and laryngeal descent in the diagnosis of OAD.

Design Comparison study conducted from November 3, 1998, to December 4, 1998, evaluating 4 clinical examination elements for diagnosis of OAD vs the gold standard of forced expiratory volume in 1 second (FEV_1) and FEV_1–forced vital capacity (FVC) ratio less than the fifth percentile (adjusted for patient height, age, and sex).

Setting Twenty-five sites, including primary care and referral practices, in 14 countries.

Participants A total of 309 consecutive patients were recruited (mean age, 56 years; 43% female), 76 (25%) with known chronic OAD, 114 (37%) with suspected chronic OAD, and 119 (39%) with neither known nor suspected OAD.

Main Outcome Measures Sensitivity, specificity, and likelihood ratios (LRs) for each of the 4 elements of the clinical examination compared with the gold standard.

Results Mean FEV_1 and FVC values were 2.1 L/s and 2.9 L; 52% had an FEV_1 and FEV_1-FVC ratio less than the fifth percentile. The LR for wheezing was 2.7 (95% confidence interval [CI], 1.7-4.2) and was not statistically significant in the multivariate model. The LR for laryngeal descent ranged from 0.9 (95% CI, 0.5-1.4) to 1.2 (95% CI, 0.4-3.4), depending on the cut point chosen, and did not enter the multivariate model. Only 4 of the history or physical examination elements we tested were significantly associated with the diagnosis of OAD on multivariate analysis: smoking for more than 40 pack-years (LR, 8.3), self-reported history of chronic OAD (LR, 7.3), maximum laryngeal height of at least 4 cm (LR, 2.8), and age at least 45 years (LR, 1.3). Patients having all 4 findings had an LR of 220 (ruling in OAD); those with none had an LR of 0.13 (ruling out OAD). The area under the receiver operating characteristic curve for the model incorporating these 4 factors was 0.86.

Conclusions Further research is needed to validate our model, but in the meantime, our data suggest that less emphasis should be placed on the presence of individual symptoms or signs (such as wheezing or laryngeal descent) in the diagnosis of OAD.

JAMA. 2000;283:1853-1857 www.jama.com

Despite the central importance of the initial clinical examination in the care of patients, its elements have rarely been subjected to rigorous evaluation. The evaluation of obstructive airway disease (OAD) is a typical example: a systematic review of the literature[1] identified 29 articles evaluating a total of 32 clinical signs for the detection of OAD (median of 1 sign, 2 clinicians, and 93 patients per study). However, only 1 of these studies[2] fulfilled standard criteria[3] for classification as a methodologically rigorous study (an independent, blind comparison with a reference standard among an appropriate spectrum of consecutive patients). In that study, 2 physicians examined 164 consecutive patients in a preoperative evaluation clinic. They reported likelihood ratios (LRs) for several elements of the clinical examination, but none of the maneuvers was sufficiently sensitive to allow their absence to rule out OAD or sufficiently specific for their presence to rule in OAD.

The reported accuracies of commonly cited signs for OAD vary greatly between studies. For example, the de-

See also Patient Page.

Author Affiliations: The Centre for Evidence-Based Medicine, Nuffield Department of Medicine, Oxford, England (Drs Straus, McAlister, and Sackett); The Division of General Internal Medicine, Mt Sinai Hospital, University Health Network, Toronto, Ontario (Dr Straus); The Division of General Internal Medicine, University of Alberta, Edmonton (Dr McAlister); and The Imperial Cancer Research Fund/National Health Service Centre for Statistics in Medicine, Institute of Health Sciences, Oxford, England (Mr Deeks).
Members of the CARE-COAD1 Group are listed at the end of this article.
Corresponding Author and Reprints: Finlay A. McAlister, MD, Division of General Internal Medicine, 2E3.24 Walter Mackenzie Centre, University of Alberta Hospital, 8440 112 St, Edmonton, Alberta, Canada T6G 2R7 (e-mail: Finlay.McAlister@ualberta.ca).

©2000 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/25/2017

tection of wheezing on auscultation has been evaluated in 7 studies: sensitivity ranged from 9% to 100%, specificity from 37% to 100%, and positive LRs varied from 0.9 to infinity.[1] Even the accuracy of the overall clinical impression (formed after obtaining complete patient history and conducting physical examination) for the detection of OAD is unclear, with sensitivity (50%-64%), specificity (64%-93%), and positive (1.4-7.3) and negative (0.4-0.8) LRs varying sharply between studies.[1] This situation led to calls in THE JOURNAL[4,5] for larger, better studies of the clinical examination.

In an effort to obtain reliable information on the accuracy of the history and physical examination in diagnosing OAD, a multinational study involving investigators at various levels (primary, secondary, and tertiary care) was designed. In this study, the accuracy of several elements of the clinical examination in predicting OAD were investigated: patient self-reported history of chronic OAD, smoking history (yes/no, number of pack-years), wheezing on auscultation, laryngeal height (maximum and minimum), and laryngeal descent. A secondary objective was to assess whether it was possible to do large, fast, multicenter studies of the clinical examination using the Internet for clinician recruitment and data collection.

METHODS

Investigators were recruited from various centers around the world via the Internet using the study group Web site (http://www.carestudy.com) and the evidence-based health care e-mail discussion group. All investigators joined the study in groups of 2 or more (at least 1 clinician and 1 spirometrist) and took responsibility for obtaining local ethics approval for the study. Investigator enrollment and data entry were done via a secure Internet-based data entry system, and data collation and analysis were done at the Centre for Evidence-Based Medicine at the University of Oxford in England.

Twenty investigator groups (46 investigators) enrolled consecutive patients (from November 3 to December 4, 1998) within 3 broad categories: patients who were known to have chronic OAD, patients who were suspected of having OAD, and patients who were neither known nor suspected of having OAD. Investigators were asked to enroll a minimum of 4 consecutive patients from each category. Known chronic OAD was defined as prior pulmonary function test results demonstrating forced expiratory volume in 1 second (FEV_1) less than the fifth percentile, FEV_1–forced vital capacity (FVC) ratio less than the fifth percentile, or FEV_1-FVC ratio less than 0.7; or patient self-report of a prior diagnosis of chronic OAD, emphysema, or chronic bronchitis; or patient taking inhaled bronchodilators and/or inhaled steroids for long periods. A case was defined as "suspected OAD" if the patient did not fulfill any criteria for known chronic OAD but was referred for suspected chronic OAD, or if the participating clinician thought that OAD was a diagnostic possibility before the structured examination. Patients with known or suspected OAD were eligible for enrollment during exacerbations of their disease if no bronchodilator treatment was given between the clinical examination and spirometry. Excluded were patients with purely reversible airway obstruction (ie, asthma); patients with a terminal illness whose goals of therapy were confined to comfort and dignity; patients younger than 18 years; patients with respiratory distress so severe that bronchodilators could not be withheld safely until after spirometry; patients who were medically unstable from other causes (eg, acute myocardial infarction, drug overdose); and patients who were unable to cooperate for the clinical examination or spirometry (eg, impaired cognition, level of consciousness, or language).

All patients underwent clinical examination and independent, blinded spirometry. The items chosen for this study were based on a review of the literature and consensus among the investigators. The items assessed included self-reported history of chronic OAD, smoking history, laryngeal height (the distance between the top of the thyroid cartilage and the suprasternal notch), laryngeal descent, and wheezing. Maximum laryngeal height was measured at the end of expiration, minimum laryngeal height at the end of inspiration.[6,7] The difference between the maximum and minimum laryngeal heights is the laryngeal descent. A videotape of the laryngeal examination was provided on the study Web site for investigator training. Investigators listened for wheezes during respiration over 4 standardized areas (bilateral upper and lower back).[8] Each patient also underwent spirometry within 30 minutes of the clinical examination (without intercurrent bronchodilator use) to assess FEV_1 and FVC values. A standard protocol for spirometry was used and the better result of 2 attempts was recorded. The spirometrists and clinicians were blind to the results of the others' investigations.

Sensitivity, specificity, and LRs for each element of the clinical examination were calculated using spirometry as the gold standard (OAD was defined as an FEV_1 and FEV_1-FVC ratio less than the fifth percentile).[9] Percentile flow rates, adjusted for age, sex, and height, were calculated using the regression equation of Crapo et al.[10] Continuous measurements (age, pack-years of smoking, and measurements of laryngeal position and descent) were categorized, either according to cut points previously published or values derived from noting obvious inflection points on receiver operating characteristic (ROC) curves. Cut points were chosen such that the slopes of the ROC curve based on the selected cut points mirrored those in the full ROC curve. The relationships between each diagnostic element and OAD were tested using χ^2 tests and the Fisher exact test for dichotomous features, the χ^2 test for trend for categorical variables, and the *t* test for continuous variables.

Multivariate analyses were carried out using the method of Spiegelhalter and Knill-Jones (which adjusts for confounding from related diagnostic elements),[11,12] and a reduced multivariate model was produced by grouping categories with similar LRs within each

©2000 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/25/2017

element and only selecting diagnostic elements with adjusted LRs greater than 2 or less than 0.5. All analyses were done using statistical software.[13]

RESULTS

A total of 332 patients were recruited by 25 investigator groups from 14 countries. Twenty-three patients were excluded from further analysis because they had a primary diagnosis of asthma; no other protocol violations were identified. Thus, the final sample size was 309. After the closing of the study, a subset of investigators (chosen because of outlying results) were asked to submit their original data collection sheets to be checked against the database—11 of 680 data points were incorrectly entered (error rate, 1.6%).

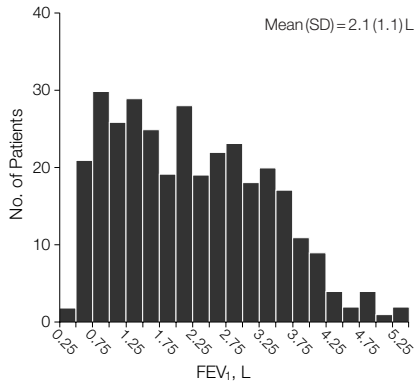
Patient demographics are outlined in TABLE 1 and the distribution of FEV_1 values are illustrated in the FIGURE. On objective testing, more than half (162 [52%]) of the patients had an FEV_1 and FEV_1-FVC ratio less than the fifth percentile. The accuracy of the various elements of the clinical examination assessed is outlined in TABLE 2. Whereas we used FEV_1 and FEV_1-FVC ratio less than the fifth percentile as the reference standard, given the controversy regarding the spirometric definitions of OAD,[5] the LRs for each element of the clinical examination also were calculated using other reference standards (FEV_1-FVC ratio <0.7, FEV_1-FVC ratio <0.8, FEV_1 value < the fifth percentile, or FEV_1-FVC ratio <the fifth percentile alone). The accuracies reported in Table 2 did not change appreciably: for example, the positive LRs for wheezing were 1.9, 3.1, 2.6, and 2.1, respectively, using the alternate reference standards. The mean minimum and maximum laryngeal

Table 1. Patient Demographics (N = 309)*

Variable	Value
Age, mean (SD), y	56.2 (17.6)
Female sex, No. (%)	133 (43)
White, No. (%)	275 (89)
Smoking status, No. (%) [n = 292]†	
Never smoked	119 (39)
Smoked <20 pack-years	80 (26)
Smoked 20-40 pack-years	49 (16)
Smoked >40 pack-years	44 (14)
Chronic OAD status, No. (%)	
Known	76 (25)
Suspected	114 (37)
Neither known nor suspected	119 (39)
Point of contact, No. (%)	
Primary	156 (50)
Secondary/tertiary	153 (50)
Clinical findings, mean (SD)	
Maximum laryngeal height, cm	5.8 (1.9)
Minimum laryngeal height, cm	4.1 (1.8)
Laryngeal descent, cm	1.7 (0.9)
FEV_1, L/s	2.1 (1.1)
FVC, L	2.9 (1.2)

*OAD indicates obstructive airway disease; FEV_1, forced expiratory volume in 1 second; and FVC, forced vital capacity.
†Data on number of pack-years smoked unavailable for 17 smokers.

Figure. Distribution of FEV_1 Values (n = 332)



FEV_1 indicates forced expiratory volume in 1 second.

Table 2. Accuracy of Elements of the Clinical Examination in Diagnosing OAD (Univariate Analysis)*

	All Patients (N = 309)			Patients Without Known COPD (n = 233)		
Diagnostic Element	No.	Crude Likelihood Ratio (95% CI)	*P* Value	No.	Crude Likelihood Ratio (95% CI)	*P* Value
Chronic OAD history						
Known	76	12.9 (5.4-31.0)	<.001	. . .		.02
Not known but suspected	114	0.9 (0.6-1.1)		114	1.5 (1.1-1.9)	
Not known or suspected	119	0.4 (0.3-0.5)		119	0.7 (0.5-0.9)	
Smoking						
Smoked >1 y	190	1.2 (1.0-1.5)	.01	128	1.1 (0.8-1.4)	.81
Never smoked, or smoked <1 y	119	0.7 (0.5-0.9)		105	1.0 (0.8-1.2)	
Amount smoked, pack-years (n = 292)†						
>40	44	19.1 (4.7-77.3)	<.001	17	11.7 (2.7-50.0)	<.001
20-40	49	0.9 (0.6-1.6)		30	0.8 (0.4-1.6)	
<20	80	0.5 (04-0.8)		64	0.5 (0.3-0.9)	
Age, y						
≥65	118	2.1 (1.5-2.9)	<.001	74	1.9 (1.3-2.8)	<.001
45-64	105	1.5 (1.1-2.0)		75	1.5 (1.1-2.2)	
<45	86	0.2 (0.1-0.3)		84	0.3 (0.2-0.5)	
Sex						
Male	176	1.3 (1.1-1.6)	.01	124	1.2 (1.0-1.5)	.41
Female	133	0.7 (0.6-0.9)		109	0.8 (0.6-1.1)	
Wheezing						
Present	79	2.7 (1.7-4.2)	<.001	44	2.1 (1.2-3.5)	.002
Absent	230	0.7 (0.6-0.8)		189	0.8 (0.7-1.0)	
Maximum laryngeal height, cm						
≤4	74	3.6 (2.1-6.0)	<.001	45	4.2 (2.3-7.9)	<.001
>4	235	0.7 (0.6-0.8)		188	0.7 (0.5-0.9)	
Minimum laryngeal height, cm						
≤4	168	1.5 (1.2-1.8)	<.001	116	1.3 (1.1-1.6)	.008
>4	141	0.6 (0.5-0.8)		117	0.7 (0.5-0.9)	
Laryngeal descent, cm						
>3	57	0.9 (0.5-1.4)	.50	22	0.9 (0.5-1.5)	.21
2.01-3	137	1.0 (0.8-1.3)		108	1.0 (0.8-1.3)	
1.01-2	99	1.1 (0.8-1.5)		91	1.1 (0.8-1.5)	
≤1	16	1.2 (0.4-3.4)		12	0.9 (0.3-3.2)	

*OAD indicates obstructive airway disease; COPD, chronic obstructive pulmonary disease; and CI, confidence interval. The reference standards were forced expiratory volume in 1 second (FEV_1) and FEV_1–forced vital capacity ratio less than the fifth percentile.
†Data on number of pack-years smoked missing for 17 smokers.

©2000 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/25/2017

Table 3. Multivariate Likelihood Ratios*

Diagnostic Element	Adjusted Likelihood Ratios	
	Factor Present	Factor Absent
All Patients (N = 309)		
Self-reported history of chronic OAD	7.3	0.5
Smoked >40 pack-years	8.3	0.8
Age ≥45 y	1.3	0.4
Maximum laryngeal height ≤4 cm	2.8	0.8
All 4 Factors	220.5	0.13
Patients Without Known Chronic OAD (n = 233)		
Smoked >40 pack-years	11.6	0.9
Age ≥45 y	1.4	0.5
Maximum laryngeal height ≤4 cm	3.6	0.7
All 3 Factors	58.5	0.32

*Area under the receiver operating characteristic curve for full model = 0.86 and for model in those patients without known chronic obstructive airway disease (OAD) = 0.77. The reported likelihood ratios (LRs) have been adjusted to take into account the effect of other factors in the model. Thus, these LRs can be multiplied to determine patient-specific LR in a patient with more than 1 factor. The other factors listed in Table 2 were not significant in either multivariate model. Confidence interval estimation for adjusted likelihood ratios requires evaluation of variance-covariance matrix for each estimate and is not presented.

heights were significantly smaller in patients with FEV_1 values and FEV_1-FVC ratio less than the fifth percentile than those with higher FEV_1 values (3.8 vs 4.6 cm, $P<.001$, and 5.4 vs 6.3 cm, $P<.001$). Laryngeal descent was not significantly associated with OAD diagnosis, even when the analysis was restricted to subgroups of patients with more severe obstruction (data not shown). There was no heterogeneity in accuracy across countries, investigator groups, examiner experience), or point of contact (primary or secondary/tertiary care). Moreover, the accuracy of the tested elements was similar even after exclusion from analysis of the 76 patients with known chronic OAD (Table 2).

The ROC curve for smoking demonstrated that the most appropriate cut point was at 40 pack-years (data available on request).

The reduced multivariate model included 4 items (TABLE 3). In patients with all 4 items (self-reported history of chronic OAD, smoked >40 pack-years, older than 45 years, and maximum laryngeal height ≤4 cm), the LR for the diagnosis of OAD is 220 (essentially ruling in the diagnosis). In patients without any of these 4 characteristics, the LR is 0.13 (essentially ruling out the diagnosis). A multivariate model derived from the 233 patients without known chronic OAD included the same 3 items (Table 3). In particular, wheezing and laryngeal descent did not enter the model even after exclusion of known chronic OAD patients.

COMMENT

We evaluated the accuracy of several elements of the clinical examination in diagnosing OAD. In terms of history, the most useful points to rule in a diagnosis of OAD are self-reported history of chronic OAD and smoking in excess of 40 pack-years. Age younger than 45 years virtually ruled out the diagnosis of OAD (given that patients with asthma were excluded). On physical examination, auscultated wheezing and maximal laryngeal height of 4 cm or less increased the likelihood that OAD was present but did not do so sufficiently to resolve the diagnostic process. For example, in a patient with a prior likelihood of 10%—the prevalence of chronic OAD among smokers[14]—the presence of wheezing raises the probability of OAD to only 23% (similarly, a maximum laryngeal height of ≤4 cm only increases the probability to 28%). Laryngeal descent was not helpful in either ruling in or ruling out the diagnosis of OAD. Although it may seem tautologous to include "history of chronic OAD" in a prediction rule for OAD, it must be acknowledged that clinicians usually collect history prior to physical examination or further diagnostic testing, and thus it is important to evaluate the accuracy of this element of the clinical assessment. Furthermore, in testing the accuracy of a symptom or sign, individuals representing a full spectrum of disease should be included. Finally, the accuracy of the tested elements did not change even after exclusion of patients with known chronic OAD, and inclusion of this factor more closely reflects actual practice.

Using multivariate analysis, we developed a 4-variable model for diagnosing OAD. The LRs for each of these variables can be multiplied (as they are adjusted to account for their nonindependence) to generate an LR for an individual patient.[11,12] For example, in a 65-year-old patient with self-reported chronic OAD, a 45-pack-year smoking history, and a maximum laryngeal height of 3 cm, the LR is 220. Thus, even if the pretest probability was only 10%, the constellation of symptoms and signs increases his/her posttest probability to 96%. Although this may obviate the need for spirometry for diagnostic purposes, it does play a useful role in identifying the severity of disease and the effects of therapy.

Our study adds substantially to the literature on the rigorous evaluation of the clinical examination for OAD (it triples the numbers of patients in such studies and increases the numbers of clinicians 10-fold). Furthermore, our findings are generally consistent with the literature. For example, Badgett and colleagues[14] found that the only useful items on history were self-reported history of chronic OAD (positive LR, 3.1) and smoking more than 70 pack-years (positive LR, 8.0). Others[5] have reported, as we did, that a history of never smoking significantly decreases the likelihood of OAD, but the negative LR is insufficient to allow the diagnosis to be definitively ruled out. Although our study contradicts previous studies that suggested increased tracheal descent was a useful sign in identifying OAD,[7,15] these were unblinded studies of nonconsecutive patients and thus subject to potential selection and measurement bias. In our study, the presence of wheezing was not as useful in diagnosing OAD as other investigators have reported it to be. In

©2000 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/25/2017

their study of 164 patients, Holleman and colleagues[2] reported a positive LR of 12 for wheezing; however, the 95% confidence interval ranged from 1.7 to 98. Our results are consistent with this estimate and, given our larger sample size, serve to refine the previously published estimates. The ability of our model to predict OAD is similar to previously published models (one[2] incorporated years of smoking exposure, patient-reported wheezing, and auscultated wheezing, and the other[14] incorporated pack-year smoking history, self-reported history of chronic OAD, and decreased breath sounds); however, our model is somewhat better at ruling out OAD than were those models. For example, in a 40-year-old nonsmoking patient without a prior diagnosis of chronic OAD and with a maximum laryngeal height of 7 cm, the LR of 0.13 virtually rules out the diagnosis of OAD.

This study established the feasibility of using the Internet to recruit investigators and conduct studies of the clinical examination. This approach allowed the rapid accrual of patients to this study (at a rate more than 20 times faster than that of the only other methodologically rigorous study[2] of the clinical examination in OAD) and has resulted in the development of a practice-based research network[16,17] of clinicians interested in performing other studies of the clinical examination.

However, there were some limitations to our study. First, we did not assess interrater reliability. This was a deliberate exclusion, as the primary focus of this initial study was to evaluate the accuracy of elements of the clinical examination and prove the feasibility of the study design. To achieve these objectives, the study was designed such that data collection would be brief. We decided a priori to defer assessment of interobserver variation for a future study in which we will only evaluate those signs that have been shown to be accurate. Second, to participate, investigators had to have access to the Internet. Critics might be concerned that this may affect the applicability of our results to patients or clinicians in other settings. However, our patients and results are similar to those in other studies,[5,14,15] suggesting that our findings are generalizable. Third, the applicability of our model for clinical practice has yet to be determined, although the preliminary observations from our data set suggest it holds significant promise. If it had been used to assess the patients in our study, a diagnosis (ie, probability of OAD >90% or <10%) could have been made and spirometry for diagnosis avoided in 48% of them. Furthermore, we derived LRs for the tested elements of the clinical examination (rather than sensitivity or specificity) to permit the ready extrapolation of our results to other settings with different prevalence of disease. Finally, our model must be validated in an independent sample of patients, and such a study is being planned.[18]

In summary, our results suggest that less emphasis should be placed on the presence of wheezing or exaggerated laryngeal descent in making a diagnosis of OAD. We found that a combination of 4 symptoms/signs (self-reported history of chronic OAD, pack-year smoking history, age, and maximum laryngeal height) can be used to predict airway obstruction. In those settings in which spirometry is readily available, it should be used because it takes only slightly longer to do than the clinical examination, definitively establishes the diagnosis of airway obstruction, and provides prognostic information. However, in those settings in which spirometry is unavailable, our model provides useful diagnostic support for the clinician. Future studies are under way to evaluate other signs and symptoms that have been described for the diagnosis of OAD and to test our model in an independent sample of patients.

The CARE-COAD1 (Clinical Assessment of the Reliability of the Examination–Chronic Obstructive Airways Disease) **Investigators** are M. Bermudez-Gomez, R. Dennis, Pontificia Universidad Javeriana, Bogotá, Colombia; S. Straus, F. McAlister, D. Sackett, John Radcliffe Hospital, Oxford, England; C. Baicus, M. Georgescu, N. Gh. Lupu Hospital, Bucharest, Romania; P. Sestini, R. Refini, R. Amerini, M. Vichi, Institute of Respiratory Diseases, University of Siena, Siena, Italy; D. Londono, A. Ruiz, Hospital San Ignacio, Universidad Javeriana, Bogotá; A. Ramos, A. de Pablo, Clinica Puerta de Hierro, Universidad Autonoma de Madrid, Madrid, Spain; M. Urtasun, M. Molinari, Hospital Regional Ushuaia, Ushuaia, Argentina; I. Scott, C. Mitchell, T. Maddocks, K. Loff, Princess Alexandra Hospital, Brisbane, Australia; A. Jelani, King Fahd National Guard Hospital, Riyadh, Saudi Arabia; O. Gajic, O. Yazdi, N. Anandarao, M. Oloomi, New York Methodist Hospital, Brooklyn, NY; E. M. Mutzig, M. Jelley, D. Bridges, University of Oklahoma College of Medicine, Tulsa; X. Cea, A. Baeza, Universidad de la Fontera, Temuco, Chile; D. Ross, Chorley, England; S. Salah, P. Badrinath, D. Obinerche, O. Lloyd, Al Ain Hospital and Faculty of Medicine and Health Sciences, Al Ain, United Arab Emirates; E. Etchells, M. Schreiber, The Toronto Hospital, Toronto, Ontario; K. Sharma, A. B. Fraser, Ottawa Hospital, Ottawa, Ontario; D. Newberry, I. England, Ashford Hospital, London, England; M. Kljakovic, S. Carr, General Practice Department, Wellington School of Medicine, Wellington, New Zealand; D. Teysseyre, St Sulpice de Royan, France; and J. Ibarra, Vitoria, Spain.

Funding/Support: Drs Straus and Sackett and Mr Deeks were supported by the National Health Service Research and Development Programme, United Kingdom. Dr McAlister is a population health investigator of the Alberta Heritage Foundation for Medical Research and was supported by the Medical Research Council of Canada.

Acknowledgment: We thank Les Irwig, MD, for his helpful comments on the data analysis.

REFERENCES

1. McAlister FA, Straus SE, Sackett DL, for the CARE-COAD1 Group. Why we need large, simple studies of the clinical examination. *Lancet*. 1999;354:1721-1724.
2. Holleman DR Jr, Simel DL, Goldberg JS. Diagnosis of obstructive airways disease from the clinical examination. *J Gen Intern Med*. 1993;8:63-68.
3. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine*. London, England: Churchill Livingstone; 1997.
4. Simel DL, Rennie D. The clinical examination. *JAMA*. 1997;277:572-574.
5. Holleman DR Jr, Simel DL. Does the clinical examination predict airflow limitation? *JAMA*. 1995;273:313-319.
6. Campbell EJ. Physical signs of diffuse airways obstruction and lung distention. *Thorax*. 1969;24:1-3.
7. Stubbing DG, Mathur PN, Roberts RS, Campbell EJ. Some physical signs in patients with chronic airflow obstruction. *Am Rev Respir Dis*. 1982;125:549-552.
8. Sapira JD. *The Art and Science of Bedside Diagnosis*. Munich, Germany: Urban & Schwarzenberg; 1990.
9. American Thoracic Society. Lung function testing. *Am Rev Respir Dis*. 1991;144:1202-1218.
10. Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am Rev Respir Dis*. 1981;123:659-664.
11. Feinstein AR. Clinical biostatistics, XXXIX: the haze of Bayes, the aerial palaces of decision analysis, and the computerized Ouija board. *Clin Pharmacol Ther*. 1977;21:482-496.
12. Spiegelhalter DJ, Knill-Jones RP. Statistical and knowledge-based approaches to clinical decision-support systems, with an application in gastroenterology. *J R Stat Soc A*. 1984;147:35-77.
13. StatCorp. *STATA Statistical Software: Release 5.0*. College Station, Tex: Stata Corp; 1997.
14. Badgett RG, Tanaka DJ, Hunt DK, et al. Can moderate chronic obstructive pulmonary disease be diagnosed by historical and physical findings alone? *Am J Med*. 1993;94:188-196.
15. Godfrey S, Edwards RH, Campbell EJ, et al. Repeatability of physical signs in airways obstruction. *Thorax*. 1969;24:4-9.
16. Green LA, Hames CG Sr, Nutting PA. Potential of practice-based networks: experiences from ASPN. *J Fam Pract*. 1994;38:400-406.
17. Nutting PA. Practice-based research networks. *J Fam Pract*. 1996;42:199-203.
18. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. *JAMA*. 1997;277:488-494.

©2000 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/25/2017

Supplemental Oxygen and Mountaineer Death Rates on Everest and K2

To the Editor: The use of supplemental oxygen by Himalayan mountaineers has been debated for more than 8 decades.[1] Although sometimes viewed as unsporting, supplemental-oxygen use may improve survival rates by increasing performance and lowering hypoxic stress.[1-3] Analyses of death rates of mountaineers descending from high summits may reveal an impact of supplemental oxygen on survival because descending mountaineers are often near exhaustion and vulnerable to accident, storm, or illness during their descent.

Methods. We analyzed interview data[4-6] and more recent data (Elizabeth Hawley, oral communication, May 4, 2000) on all mountaineers reaching the summit of the 2 highest peaks (Everest and K2) from 1978 (year both summits first reached without supplemental oxygen) through 1999. For "summit-team" analyses on Everest, we excluded recent data (1993-1999) to reduce the impact of guided expeditions, which may include inexperienced climbers. We used exact logistic regression (conditional maximum likelihood) with survival as the dependent variable and supplemental oxygen (used and not used) as a factor, stratified by mountain (Everest and K2). In a preliminary analysis, the year of summiting (covariate) was unrelated to individual death rates on Everest (either directly or via an interaction with supplemental oxygen, $P>.27$) and hence excluded from final analyses.

Results. Individual mountaineers not using supplemental oxygen had significantly higher death rates during descent than did those using supplemental oxygen (TABLE, $P<.001$). This pattern is especially evident on K2, where approximately 1 in 5 climbers not using supplemental oxygen died during descent (Table).

To control for nonindependence of climbers in a team, we used a "summit team" as a complementary unit of analysis and determined (for each team reaching the summit on a given day and route) whether supplemental oxygen was used and whether any descending mountaineer died. Number of summiters was a covariate because the probability of a death(s) may increase with the number of climbers exposed to risk. Even by this conservative analysis, teams not using supplemental oxygen had relatively high death incidences ($P=.03$).

Comment. Reaching the summit of Everest, and especially of K2, is dangerous. Overall, 1 in 29 climbers died during descent on Everest, and 1 in 7 died on K2 (Table). Reaching those summits without supplemental oxygen is associated with an even higher risk: 1 climber in 12 died on Everest, and approximately 1 in 5 died on K2 (Table). The survival impact of supplemental oxygen may be greater than suggested because mountaineers not using supplemental oxygen are probably relatively more experienced and therefore might be expected to have lower death rates. The association may be causal because supplemental oxygen decreases exposure time and reduces physical deterioration.[1,3] Nevertheless, alternative explanations (eg, mountaineers using supplemental oxygen are more risk averse) cannot be excluded. Moreover, a full risk assessment of supplemental oxygen use awaits incorporation of data on death rates during ascent, risk to porters ferrying oxygen canisters, actual causes of death, and weather conditions. In any case, Himalayan mountaineering is a dangerous activity[2] that balances adventure against risk. Mountaineers considering whether to use supplemental oxygen should consider the risk of death during descent.

Table. Use of Supplemental Oxygen and Death Rates of Individual Mountaineers and for Summit Teams* During Descent From the Summits of Everest and K2 Between 1978 and 1999

Mountain	Individual Mountaineers†		Summit Teams‡	
	Ascents, No.	Deaths, No. (%)	Teams Summiting, No.	Teams With Death, No. (%)
Everest				
Used supplementary oxygen	1077	32 (3.0)	93	8 (8.6)
No supplementary oxygen	96	8 (8.3)	28	4 (14.3)
K2				
Used supplementary oxygen	47	0 (0)	12	0 (0)
No supplementary oxygen	117	22 (18.8)	36	12 (33.3)

*Summit-team analyses for Everest restricted to 1978-1992, see "Methods" section.
†For comparison by exact logistic regression, stratified by mountain, $P<.001$.
‡For comparison by exact logistic regression, stratified by mountain and with number of summiters per team as a covariate (included because the probability of a death should increase with the number of climbers exposed to risk, all else being equal), $P = .03$.

Raymond B. Huey, PhD
University of Washington
Seattle

Xavier Eguskitza
Worcester, England

Funding/Support: This study was funded by a fellowship to Dr Huey from the J. S. Guggenheim Fellowship.

Acknowledgment: We thank T. Hornbein, MD, University of Washington, Seattle, and C. Houston, MD, University of Vermont, Burlington, for advice, and E. Hawley, Consul for New Zealand, Kathmandu, Nepal, for sharing data.

1. West JB. *High Life: A History of High-Altitude Physiology and Medicine*.Oxford, England: Oxford University Press; 1998:493.
2. Pollard A, Clarke C. Deaths during mountaineering at extreme altitude. *Lancet*. 1988;1:1227.
3. Peacock AJ, Jones PL. Gas exchange at extreme altitude: results from the British 40th Anniversary Everest Expedition. *Eur Respir J*. 1997;10:1439-1444.
4. Eguskitza X. Everest—the first forty years. In: Gillman P, ed. *Everest: The Best Writing and Pictures from Seventy Years of Human Endeavour*. Boston, Mass: Little Brown & Co; 1993:191-201.
5. Hawley E, Eguskitza X. Everest—die ersten 45 Jahre. In: Gillman P, ed. *Everest 8846m*. Munich, Germany: J Berg bei Bruckmann; 1998:191-202.
6. Eguskitza X. Appendices I-III. In: Curran J, ed. *K2: The Story of the Savage Mountain*. London, England: Hodder & Stoughton; 1996.

CORRECTION

Incorrect Wording: In the Original Contribution entitled "The Accuracy of Patient History, Wheezing, and Laryngeal Measurements in Diagnosing Obstructive Airway Disease" published in the April 12, 2000, issue of THE JOURNAL (2000;283:1853-1857), there was incorrect wording in the abstract. On page 1853, in the "Results" section, the sentence that read "maximum laryngeal height of at least 4 cm" should have read "maximum laryngeal height of 4 cm or less."

©2000 American Medical Association. All rights reserved.

Downloaded From: http://jama.jamanetwork.com/ on 05/25/2017