COMPARISON OF SPIROMETRY RESULTS BETWEEN TWO INTERPRETATION ALGORITHMS IN THE CONTEXT OF MEDICAL SCREENING

Thomas Czeczok[1], Marek Mikulski[2], Spencer Lourens[3], and Laurence Fuortes[2]

[1]The University of Iowa Carver College of Medicine M2 Student, [2]Department of Occupational and Environmental Health, College of Public Health, The University of Iowa, Iowa City, IA

[3]Department of Biostatistics, College of Public Health, The University of Iowa, Iowa City, IA

Abstract

Spirometry is the most common Pulmonary Function Test (PFT), and different reference standards exist for interpreting results. Percent predicted (%Pred) algorithms are most commonly used and are based on set percent values in diagnosing respiratory disease; however, this protocol may be responsible for misdiagnosing individuals, especially at the extremes of age. The lower limit of normal algorithm uses the bottom 5% of the statistical distribution in determining abnormalities and is thought to be a more valid method in screening for respiratory disease.

The two interpretation methods were compared using spirometry results obtained through medical screenings of former U.S. Department of Energy workers from two sites in Iowa

Agreement between the two methods was excellent (kappa=0.70), however declined with age, suggesting that algorithms differ more with increase in age. %Pred protocol was associated with increased diagnosis of COPD in smokers. LLN was more likely to predict restrictive disease in smokers. Since smoking is strongly associated with obstructive disease, %Pred seems to be more valid in predicting obstructive disease, but further studies are needed to determine if this is true.

Spirometry is a crude test used in medical screening programs, and despite using a theoretically more valid protocol, the test may be flawed due to improper reference curves, especially at the extreme of age.

Background

Spirometry is the most commonly used Pulmonary Function Test. Interpretation of spirometry results and spirometry reference standards have been subject to several updates over the years. The current recommendation for workplace screenings calls for use of the NHANES III based reference standard and 5th percentile of predicted values, a lower limit of normal (LLN), as a determinant of abnormality (Hankinson et al., 1999; Townsend, 2010). This recommendation is based on studies showing reduced misclassification by the LLN method in diagnosing obstructive airways impairment compared to fixed cut-off values for percent predicted indicators (Swanney et al., 2008). Previously the Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) standard was used. This standard relied on the fixed cut-off percent predicted value of FEV1/FVC ratio to diagnose abnormal spirometry. It has been demonstrated that this method may be responsible for misclassifying individuals as having obstructive airways physiology when in fact they do not, especially at the extremes of age (Aggarwal et al. 2006)

The purpose of this study was to examine concordance between the fixed percent predicted (%Pred) based spirometry interpretation protocol and LLN based protocol in determining abnormalities and attempt to validate the diagnoses.

Methods

The study cohort consisted of former nuclear weapons workers who participated in a U.S. Department of Energy (DoE)-funded medical screening program, which included spirometry testing.

Spirometry was performed by NIOSH certified personnel in compliance with ATS guidelines, and test equipment was calibrated daily. A reasonable effort was made to obtain three or more acceptable results. The most recent result for each individual was used, as screenings are repeated every three years.

Results were interpreted according to NHANES III reference standards[1], percent predicted (%Pred)[2,3] and lower limit of normal (LLN)[4] protocols for forced vital capacity (FVC), forced expiratory volume in the 1st second (FEV1) and FEV1/FVC ratio. Individuals were classified as normal, normal variant/borderline obstructive, obstructive, restrictive, mixed, or inconclusive.

Participant's age, weight, and height were self reported at time of testing. Body Mass Index (BMI) was calculated based on Keys et al.[5] Smoking history was obtained and individuals were categorized into never smoker, ex-smoker, and current smoker.

Statistical analyses were performed using Windows SAS 9.2. Concordance between spirometry protocols was evaluated using Fleiss and Cohen's weighted kappa statistics. Multinomial logistic regression models were built to assess validity of each protocol in diagnosing abnormal spirometry with known predictors including BMI, smoking and age.

Table 1: Interpretation Protocols for LLN and %Pred

Result	FEV1/FVC % ≥LLN	FVC ≥LLN	FEV1 ≥LLN	FEV1/FVC% ≥70%	FVC ≥80%Pred	FEV1 ≥80%Pred	FVC%Pred ≥FEV1%Pred
Normal	Yes	Yes	-	Yes	Yes	Yes	Yes/No
Normal variant or borderline obstructive	No	-	Yes	-	-	-	-
Obstructive	No	Yes	No	No	Yes/No	No	Yes
Restrictive	Yes	No	-	Yes	No	Yes	No
Mixed	No	No	No	Yes	No	No	No

Results

Table 2: Characteristics of Screened Population

Parameter	Male n=1,735 (75%)	Female n=584 (25%)	Total n=2,319
Age, mean (SD), range	67 (12); 21-98	61 (13); 22-94	65 (13); 21-98
Age, n (%)			
≤40	48 (2.8)	42 (7.2)	90 (3.9)
41-50	134 (7.7)	85 (14.6)	219 (9.4)
51-60	290 (16.7)	160 (27.4)	450 (19.4)
61-70	527 (30.4)	141 (24.1)	668 (28.9)
71-80	537 (30.9)	111 (19.0)	648 (27.9)
>80	199 (11.5)	45 (7.7)	244 (10.5)
HT (cm) mean (SD), range	177 (7); 150-198	164 (6); 147-190	174(9); 147-198
Race, n (%)			
White	1,659 (95.6)	551 (94.3)	2,210 (95.3)
African-American	34 (1.9)	17 (2.9)	51 (2.2)
Asian	30 (1.7)	13 (2.2)	43 (1.8)
Hispanic	12 (0.8)	3 (0.6)	15 (0.7)
Smoking, n (%)			
Never smoker	779 (44.9)	421 (72.1)	1,200 (51.7)
Ex-smoker	823 (47.4)	130 (22.3)	953 (41.1)
Smoker	127 (7.3)	33 (5.6)	160 (6.9)
Missing	6 (0.4)	-	6 (0.3)
BMI mean (SD), range	28(5); 17-58	28(7); 17-63	28(6);17-63
BMI, n (%)			
<25	447 (25.8)	208 (35.6)	655 (28.2)
25-29	736 (42.4)	190 (32.5)	926 (39.9)
≥30	552 (31.8)	186 (31.8)	738 (31.9)

Table 3: Agreement between Spirometry Protocols by Age Strata

LLN	%Pred	Normal			Obstructive			Restrictive			Mixed			Total
	Age	<60	60-80	>80	<60	60-80	>80	<60	60-80	>80	<60	60-80	>80	
Normal N Row % Col %	<60	504 98.6 97.1			6 1.2 21.4			1 0.2 3.3			-			511 100.0 -
	60-80		716 91.6 99.7			40 5.1 17.2			21 2.7 23.1			5 0.6 3.1		782 100.0 -
	>80			97 77.6 100.0			13 10.4 19.1			13 10.4 72.2			2 1.6 6.7	125 100.0 -
Obstructve N Row % Col %	<60	-			7 100.0 25.0			-			-			7 100.0 -
	60-80		-			36 100.0 15.5			-			-		36 100.0 -
	>80			-			17 100.0 25.0			-			-	17 100.0 -
Restrictive N Row % Col %	<60	15 14.6 2.9			5 4.9 17.9			29 28.2 96.7			54 52.4 100.0			103 100.0 -
	60-80		2 0.8 0.3			56 19.7 24.0			70 24.6 76.9			156 54.9 96.9		284 100.0 -
	>80			-			15 31.3 22.1			5 10.4 27.8			28 58.3 93.3	48 100.0 -
Mixed N Row % Col %	<60	-			10 100.0 35.7			-			-			10 100.0 -
	60-80		-			101 100.0 43.3			-			-		101 100.0 -
	>80			-			23 100.0 33.8			-			-	23 100.0 -
Total		519 - 100.0	718 - 100.0	97 - 100.0	28 - 100.0	233 - 100.0	68 - 100.0	30 - 100.0	91 - 100.0	18 - 100.0	54 - 100.0	161 - 100.0	30 - 100.0	2,047

Table 4: Multinomial Logistic Regression Models

Spirometry Result	Predictor	LLN Protocol				%Pred Protocol			
		Estimate	OR	P-value	Overall P-value	Estimate	OR	P-value	Overall P-value
Normal	Age	-	1.0	-	<0.0001	-	1.0	-	<0.0001
Normal/Borderline Obstructive		0.002	1.002	0.82		-	-	-	
Obstructive		0.04	1.04	**0.002**		0.07	1.08	**<0.0001**	
Restrictive		0.02	1.02	**<0.0001**		0.04	1.04	**<0.0001**	
Mixed		0.05	1.05	**<0.0001**		0.05	1.05	**<0.0001**	
Normal	BMI	-	1.0	-	<0.0001	-	1.0	-	<0.0001
Normal/Borderline Obstructive		-0.06	0..94	**0.02**		-	-	-	
Obstructive		-0.04	0.96	0.21		-0.005	1.00	0.73	
Restrictive		0.08	1.08	**<0.0001**		0.05	1.05	**0.0009**	
Mixed		-0.04	0.96	**0.05**		.10	1.10	**<0.0001**	
Normal	Smoking	-	1.0	-	<0.0001	-	1.0	-	<0.0001
Normal/Borderline Obstructive		0.40	1.50	**0.03**		-	-	-	
Obstructive		1.04	2.83	**<0.0001**		1.26	3.53	**<0.0001**	
Restrictive		0.37	1.45	**<0.0001**		-0.05	0.95	0.77	
Mixed		1.51	4.51	**<0.0001**		0.46	1.58	**0.0001**	

Analysis

- Analysis of concordance revealed excellent overall agreement between two protocols (weighted kappa statistic = 0.70)
- Concordance decreased with increasing age
  - <60 y.o. = 0.82 (almost perfect agreement)
  - 61-80 y.o. = 0.69 (substantial agreement)
  - >80 y.o. = 0.45 (moderate agreement)
- %Pred protocol more likely to diagnose obstructive disease in relation to normal spirometry than LLN in ever smokers
  - OR=3.53 for %Pred vs. OR=2.83 for LLN
- Ever smokers more likely to be diagnosed with mixed or restrictive disease in relation to normal spirometry using LLN protocol than %Pred protocol.
  - Restrictive: OR=0.95 %Pred vs. OR=1.45 LLN
  - Mixed: OR=1.58 %Pred vs. OR=4.39 LLN
- Increase in age corresponds with increase in likelihood of respiratory disease in both protocols (p<0.0001)
- Overweight individuals more likely to be diagnosed with restrictive airways disease in relation to normal spirometry using either protocol (p<0.0001)
  - OR=1.05 %Pred and OR=1.09 LLN

Discussion

- Neither of the protocols works well at the extreme of age (>80 years old).
- LLN was supposed to improve on the validity of interpretation protocol by using a more valid statistical method.
- Our data suggests that misclassification may similarly exist in both protocols.
- %Pred protocol appears to be more valid than LLN in categorizing obstruction.
  - Smoking has a strong association with obstructive disease.
  - Smokers were more likely to have restrictive disease using LLN, suggesting that LLN protocol is misclassifying elderly individuals with smoking history.
  - %Pred protocol associates smoking with obstructive disease
  - Further studies are needed.
- Spirometry is a crude test, and despite an improved statistical method, the test is flawed due to a flawed linear model in which both methods are based. This linear model fits individuals from age 8 to 80, but at the extremes of age, it is unlikely that a simple linear model is appropriate.
- Lung volumes and DLCO are a better test to characterize respiratory disease.
  - Restriction cannot be diagnosed without lung volumes.

References

1) Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. Am J Respir Crit Care Med. 1999;159:179-187.
2) Morris JF, Koski A, Johnson JC. Spirometric standards for healthy nonsmoking adults. Am Rev Respir Dis. 1971;103:57-67
3) Sobol BJ, Sobol PG. Per cent of predicted as the limit of normal in pulmonary function testing: a statistically valid approach. Thorax.1979;34:1-3
4) Townsend MC. ACOEM position statement. Spirometry in the occupational health setting-2010 update. East Grove Village, IL. American College of Occupational and Environmental Medicine. 2010
5) Keys A, Fidanza, F, Karvonen MJ, Kimura N, Taylor. HL. Indices of relative weight and obesity. *Chron Dis 1972*;25, 329-343.