Pulmonary Function Test Quality in the Elderly: A Comparison With Younger Adults

Jeffrey M Haynes RRT RPFT

BACKGROUND: Elderly patients may be at greater risk for misdiagnosis and inappropriate treatment as a consequence of pulmonary function test underutilization and tests being conducted with low quality expectations. This study sought to determine if elderly patients are able to achieve both spirometry and diffusion capacity (DLCO) quality scores comparable to a younger adult population. METHODS: This was a retrospective review of pulmonary function data over a 22 month period. A list of every subject age ≥ 80 years (elderly group) and ages 40–50 years (control group) tested during the time period was compiled. The quality of spirometry and DLCO testing were examined. RESULTS: Overall, 92.6% (139/150) of the elderly group and 91.5% (163/178) of the control group spirometry tests satisfied all American Thoracic Society/European Respiratory Society acceptability and reproducibility criteria (P = .84), and 84.9% (96/113) of the elderly group and 88.5% (108/122) of the control group DLCO tests satisfied all the acceptability and reproducibility criteria (P = .45). CONCLUSIONS: Elderly patients referred to a hospital-based pulmonary function test lab can be expected to achieve spirometry and DLCO quality scores comparable to younger adult patients. Key words: pulmonary function tests; spirometry; diffusing capacity; test quality. [Respir Care 2014;59(1):16–21. © 2014 Daedalus Enterprises]

Introduction

Elderly patients with cognitive impairment and apraxia have difficulty performing spirometry correctly.1,2 However, several studies have shown that most elderly patients can produce quality spirometry data.3-7 There are limited and conflicting data on the effect of age on the ability to perform diffusion capacity of the lung for carbon monoxide (DLCO) correctly.8,9 Elderly patients are at risk for misdiagnosis and inappropriate treatment of respiratory disease,10,11 which may be compounded by pulmonary function test (PFT) underutilization and the inappropriate acceptance of suboptimal test quality, due to low performance expectations. This study sought to determine if elderly patients referred to a hospital-based PFT laboratory are able to achieve both spirometry and DLCO quality scores comparable to a younger adult population.

Methods

This study was a retrospective review of PFT data over a 22 month period (June 2011 through March 2013) in a hospital-based PFT laboratory (St Joseph Hospital, Nashua, New Hampshire). The St Joseph Hospital institutional review board approved this study. Using the “research query tools” function of the PFT system (ComPAS, Morgan Scientific, Haverhill, Massachusetts) a list of every subject age ≥ 80 years (elderly group) tested during the time period was compiled. Every subject age 40–50 years old tested during the same time frame served as the control group. Demographic data were collected and the quality of spirometry and DLCO testing were examined. The mean and median of the percent-of-predicted values and the percentage of test results below the lower limits of normal were recorded for DLCO, FVC, FEV1, and FEV1/FVC.
Our laboratory uses Global Lung Function Initiative (age range 3–95 years)\textsuperscript{12} and Cotes\textsuperscript{13} predicted equations for spirometry and \textit{D}_{\text{LCO}}\textsuperscript{16} respectively. American Thoracic Society/European Respiratory Society (ATS/ERS) quality standards were used to judge individual test quality.\textsuperscript{14,15} Our PFT system provides an on-screen display of effort-by-effort test acceptability and reproducibility, and detailed grading of test quality based on the ATS/ERS quality standards\textsuperscript{14-15} (Figure). In addition to computerized quality scoring, each pulmonary function test was examined in detail by the chief technologist to confirm the accuracy of computer grading. FVC and FEV\textsubscript{1} were graded on an A-B-C-D-F quality scale, as described in the Global Lungs Initiative Phase 2 Pilot Study\textsuperscript{16} (Table 1).

During this time period the laboratory was staffed by 3 technologists. Two technologists had > 20 years of testing experience, and the third technologist had 5 years of testing experience. The chief technologist of the laboratory is a registered pulmonary function technologist. The laboratory’s quality assurance program includes ongoing technologist performance surveillance and monthly and quarterly technologist performance feedback.\textsuperscript{17,18} Examination of the quality control records from the study period revealed no issues in terms of pneumotachograph or gas analyzer malfunction. Moreover, no out-of-control conditions were identified from the review of biologic control testing records.

Statistical Analysis

Commercially available software was used for statistical computations (Prism 4, StatMate 2.0, GraphPad Software, La Jolla, California). Differences in categorical data were examined with the Fisher exact test. Differences in continuous data were examined with the Student \( t \) test for unpaired means, or the Mann-Whitney test for unpaired medians. Data are reported as mean \( \pm \) SD or median and IQR. Post hoc power analysis of test success for this 22 month sample was performed. A 2-tailed \( P \) value of < .05 was considered significant.

Results

During the study period, 150 elderly subjects attempted 150 spirometry and 113 \textit{D}_{\text{LCO}}\textsuperscript{16} tests. During the same time frame, 178 control subjects attempted 178 spirometry and 122 \textit{D}_{\text{LCO}}\textsuperscript{16} tests. Demographic data are listed in Table 2. The elderly group had a lower FVC and FEV\textsubscript{1}, expressed as a percent of predicted, but not as a proportion of values below the lower limit of normal. The percentage of FEV\textsubscript{1}/FVC values below the lower limit of normal was not different between groups. \textit{D}_{\text{LCO}}\textsuperscript{16} was lower in the elderly group, both as a percent of predicted and as a proportion of values below the lower limit of normal.

\textbf{QUICK LOOK}

\textbf{Current knowledge}

Elderly patients may be at greater risk for misdiagnosis and inappropriate treatment as a consequence of underutilization of, and low quality expectations about, pulmonary function testing in the elderly. Comorbidities, including cognitive impairment and apraxia, may influence the quality of spirometry and diffusion-capacity testing.

\textbf{What this paper contributes to our knowledge}

This retrospective review suggests that the quality of spirometry and diffusion-capacity results in elderly patients is similar to that in young adults. Advanced age alone does not portend poor-quality spirometry or diffusion-capacity results.

\textbf{Spirometry Quality}

In the elderly group 92.6\% (139/150) of the spirometry tests satisfied all the ATS/ERS acceptability and reproducibility criteria (grade A).\textsuperscript{14} Of the 11 tests that failed to meet all ATS/ERS performance standards, 3 had both FVC and FEV\textsubscript{1} grades C or better. When those tests were included, 94.6\% of tests produced clinically meaningful FVC and FEV\textsubscript{1} data.\textsuperscript{16} Five tests had both FVC and FEV\textsubscript{1} grades worse than C, and 3 had an FEV\textsubscript{1} grade C or better coupled with an FVC grade worse than C. The most common reason for test failure were inability to exhale for \( \geq 6 \) seconds, and an extrapolated volume > 5\% of the FVC or > 150 mL.

In the control group 91.5\% (163/178) of spirometry tests satisfied all the ATS/ERS acceptability and reproducibility criteria (grade A).\textsuperscript{14} Of the 15 tests that failed to meet all ATS/ERS performance standards, 1 test had both FVC and FEV\textsubscript{1} grades C or better. When that test was included, 92.1\% of tests produced clinically meaningful FVC and FEV\textsubscript{1} data.\textsuperscript{16} Three tests had both FVC and FEV\textsubscript{1} grades worse than C, 10 tests had FEV\textsubscript{1} grades C or better coupled with FVC grades worse than C, and 1 test had an FVC grade C or better coupled with an FEV\textsubscript{1} grade worse than C. The most common reasons for test failure were inability to exhale for \( \geq 6 \) seconds, extrapolated volume > 5\% of the FVC or > 150 mL, and failure to reach the peak expiratory flow within 1.2 seconds. There was no difference in spirometry test quality between the groups (\( P = .84 \), Table 3). Post hoc analysis revealed that this comparison had 80\% power to detect an 8\% difference in success, with a significance level of .05 (2-tailed \( P \)).

\textbf{\textit{D}_{\text{LCO}}\textsuperscript{16} Quality}

In the elderly group 84.9\% (96/113) of tests satisfied all the ATS/ERS acceptability and reproducibility criteria.\textsuperscript{15} Of
the 17 $D_{LCO}$ tests that failed to satisfy all ATS/ERS criteria, 1 test was performed correctly and was reproducible, but only failed to satisfy the acceptability criteria because severe airflow obstruction precluded capture of the alveolar sample...
within 4 seconds. When that test was included, 85.8% of the DLCO tests were performed correctly and provided reproducible values.

In the control group 88.5% (108/122) of the tests satisfied all the ATS/ERS acceptability and reproducibility criteria.15 Of the 14 DLCO tests that failed to satisfy all ATS/ERS criteria, 2 tests were performed correctly and were reproducible, but only failed to satisfy acceptability criteria because severe air-flow obstruction precluded capture of the alveolar sample within 4 seconds. When those tests were included, 90.2% of the DLCO tests were performed correctly and provided reproducible values. There was no difference in DLCO test quality between the groups (P = .45, see Table 3). Post hoc analysis revealed that this comparison had 80% power to detect a 12% difference in success, with a significance level of .05 (2-tailed P).

**Discussion**

Undiagnosed respiratory disease and the underutilization of PFT in elderly patients are common, and the impact of undiagnosed disease is substantial.11,19 Indeed, undiagnosed COPD in older individuals has a significant effect on health-related quality of life.20 Dow et al10 reported that 84% of older patients with untreated asthma were found to have moderate to severe disease after spirometry testing was performed. In addition, elderly patients may more frequently be subjected to poorly conducted PFTs if the technologist has a preconceived notion that elderly patients are less likely to be able to perform PFTs correctly.

Diminished cognitive function and apraxia (inability to perform a motor activity from thought) in elderly patients with neurologic impairment can certainly make meaningful PFT data very difficult or impossible to obtain. Allen and Baxter1 studied the strength of cognitive tests to predict the ability of elderly subjects to perform spirometry correctly. Lower scores on the Mini Mental State Examination (overall cognition) and difficulty drawing intersecting pentagons were predictive of spirometry test failure. Interestingly, tests of executive function were found to be less predictive of outcome. Carvalhaes-Neto et al2 assessed the ability of 65 elderly institutionalized subjects with cognitive dysfunction to perform spirometry testing. Only 41% of the subjects were able to perform spirometry correctly. While it is clear that elderly patients with marked cognitive impairment and apraxia are less likely to perform spirometry (forced or slow21) correctly, most elderly patients without severe cognitive impairment are able to produce quality spirometry data.

Sherman and colleagues3 assessed the ability of 65 elderly subjects with only mild cognitive impairment to perform spi-
rometry. Of the group, 87.6% of subjects were able to produce acceptable spirometry data. The Salute Respiratoria Dell’Anziano et al. (SARA [Respiratory Health in the Elderly]) study collected spirometry data from patients with ages ranging from 65 to 100 years, using mostly inexperienced technicians. The SARA investigators reported that 78% of all the subjects produced 3 acceptable spirometry efforts. Pezzoli et al. studied 715 elderly patients with respiratory symptoms and found that 81.8% were able to perform spirometry correctly. These and other studies clearly show that cognitive function, and not age itself, is predictive of successful or unsuccessful spirometry in elderly patients. Indeed, there are no data to suggest that younger patients with severe cognitive impairment would perform spirometry any better than elderly patients with similar cognitive deficits.

There are far less data regarding the effect of age on the ability to perform DLCO correctly. In a general population study conducted in the 1980s, Welle et al. found that only 67% of adult subjects could perform DLCO testing correctly, and that younger age was an independent predictor of test failure. The Welle study differed from current practice because the minimally acceptable ratio of inspired vital capacity to FVC was 0.9, which is higher than the current standard 0.85. It is also notable that one technologist with unspecified experience performed 96% of the tests. Neas and Schwartz examined DLCO data from the first National Health and Nutrition Examination Survey (1971–1975). Missing DLCO data (it is assumed that the patient could not perform the test properly) occurred in 40% of subjects aged 65–74 years and in 22% of subjects aged 25–54 years. In contrast to the current study, both of those studies used data collected with older instrumentation that did not have quality control software, and analyzed data approximately 10 and 20 years, respectively, after it was originally collected.

In the present study 92.6% and 84.9% of the elderly subjects were able to perform both spirometry and DLCO tests to the satisfaction of ATS/ERS acceptability and reproducibility standards, respectively. Despite a higher prevalence of age-adjusted abnormal lung function (DLCO), there was no difference in the percentage of high quality tests between the elderly group (median age 83 years) and the much younger control group (median age 46 years). This study differs from many studies of spirometry quality in the elderly because this study used data from patients 80 years and older, whereas several previously published studies used data from patients aged 65 years and older.

A contributing factor to the high quality scores documented in this study is our laboratory’s use of a quality assurance program, which includes ongoing technologist performance surveillance and monthly and quarterly technologist performance feedback. The goal of our laboratory is that 90% of all tests meet ATS/ERS acceptability and reproducibility standards. Data supporting technologist performance feedback originated from the Lung Health Study. Enright and the Lung Health Study Research Group showed that technologist performance was vacillating until a program of technologist monitoring and feedback was instituted. The SARA study utilized a technologist monitoring and feedback program, and, as cited earlier, was able to obtain quality spirometry data from 78% of elderly patients, even with inexperienced technologists. More recently, Borg and colleagues documented that adherence to spirometry quality standards increased from 61% to 92% in a hospital-based PFT lab after a technologist performance feedback program was initiated. While technologist monitoring and feedback is an ATS/ERS quality assurance recommendation, as few as 30% of laboratories may have such a program in place.

PFTs should be conducted with the goal of collecting data of the highest quality; however, spirometry data with less than grade A quality should not necessarily be classified as invalid and go unreported. Hankinson et al. examined the impact of FVC and FEV1 quality grades (see Table 1) on predicted reference equations. Only FVC and FEV1 data with grades of D and F negatively impacted the formulation of predicted values and lower limits of normal. Average Z scores for FVC and FEV1 were similar from data with A-C quality grades. In clinical practice, spirometry tests with grades B and C, while not perfect, in most cases should be reported. In the current study the percentage of spirometry tests in the elderly group characterized as valid and meaningful increased from 92.6% to 94.6% with the inclusion of tests with B and C grades.

The present study has limitations. The subjects were from a population with little diversity in terms of culture and race. However, there is no reason to believe that the same level of PFT quality should not be achievable in any community. In a study of 6,193 DLCO tests, Punjabi and colleagues found no association between race and test reproducibility. Differences in regional PFT referral practices could significantly impact test quality. A laboratory that serves physicians who regularly refer patients with substantial cognitive impairment and apraxia may experience less success when testing an elderly population. In addition, this study reports data from a community hospital laboratory with 3 experienced staff members; results might differ in larger laboratories with more diversity in technologist experience and skill level. Similar studies in different communities and laboratory settings are warranted.

It has been suggested that the Cotes equation for DLCO reference equations may not be ideal for older patients, since, like most DLCO reference equations, there was limited sampling of elderly subjects. Garcia-Rio et al. published predicted equations for DLCO in patients aged 65–85 years, and a comparison with the Cotes equation suggested that the Cotes equation underestimates DLCO in elderly men and women. However, we have found the Cotes equations to function well in our patient population. The DLCO equation that best agreed with the Garcia-Rio equation was that of Crapo and Morris.
which we believe overestimates $D_LCO$ in both our young and elderly patients. It is noteworthy that the Garcia-Rio and Crapo equations were generated using data collected at altitudes of 655 meters and 1,400 meters above sea level, respectively, whereas Nashua, New Hampshire, has an altitude of 50 meters above sea level. $D_LCO$ is well known to be positively correlated with altitude.  

Conclusions

The overwhelming majority of elderly patients referred to a hospital-based PFT laboratory performed spirometry and $D_LCO$ testing to the satisfaction of ATS/ERS acceptability and reproducibility standards. Elderly patients are capable of producing quality spirometry and $D_LCO$ data comparable to younger adults. Advanced age alone should not discourage patient referrals for PFTs or lower expectations for a high quality test.

References