A Spirometry-Based Algorithm To Direct Lung Function Testing in the Pulmonary Function Laboratory*

Christine A. Glady, MD; Shawn D. Aaron, MD; Mary Lunau, MSc; Jennifer Clinch, MA; and Robert E. Dales, MD

Objective: To design a spirometry-based algorithm to predict pulmonary restrictive impairment and reduce the number of patients undergoing unnecessary lung volume testing.

Design: Two prospective studies of 259 consecutive patients and 265 consecutive patients used to derive and validate the algorithm, respectively.

Setting: A pulmonary function laboratory of a tertiary care hospital.

Patients: Consecutive adults referred to the laboratory for lung volume measurements and spirometry.

Measurements: The sensitivity of the algorithm for predicting pulmonary restriction and the cost savings associated with its use.

Results: Total lung capacity correlated strongly with FVC ($r = 0.66$) and showed an inverse correlation with the FEV$_1$/FVC ratio ($r = -0.41$). According to the algorithm, only patients with an FVC $< 85\%$ of predicted and an FEV$_1$/FVC ratio $> 55\%$ required lung volume measurements following spirometry. The algorithm had a high sensitivity for predicting restriction and a high negative predictive value (NPV) for excluding restriction (sensitivity, 96%; NPV, 98%). The diagnostic properties of the algorithm were reproducible in the validation study. Application of the algorithm would eliminate the need for lung volume testing in 48 to 49% of patients referred to the pulmonary function test (PFT) laboratory, reducing costs by 33%.

Conclusions: A spirometry-based algorithm accurately excludes pulmonary restriction and reduces unnecessary lung volume testing in the PFT laboratory almost in half.

(CHEST 2003; 123:1939–1946)

Key words: diagnosis; lung volumes; pulmonary function test; restriction; spirometry

Abbreviations: ATS = American Thoracic Society; NPV = negative predictive value; PFT = pulmonary function test; PPV = positive predictive value; SVC = slow vital capacity; TLC = total lung capacity; $V_a$ = alveolar volume

Pulmonary function tests (PFTs) are performed in order to diagnose and classify disease processes that impair lung function. Impairments in lung function can be broadly classified as those resulting in airflow obstruction, volume restriction, or a combination of obstructive and restrictive defects.

Typically, airflow obstruction can be diagnosed using spirometry alone by demonstrating a lower-than-predicted FEV$_1$/FVC ratio. Since affordable hand-held spirometers are now widely available, diagnosis of obstructive lung disease can easily be made in an outpatient setting. However, spirometry is less accurate at predicting pulmonary restriction. A low FVC seen on spirometry may be a clue to a restrictive impairment; however, this is not a specific finding since it can also be seen in patients with severe obstruction with air trapping.

If spirometry suggests a restrictive disorder, patients with this pattern are usually referred for additional PFTs to confirm the diagnosis. Diagnosis of a restrictive impairment depends on detecting a reduced total lung capacity (TLC) by lung volume measurement. Lung volumes can be measured by plethysmography (“body box”), by helium-dilution methods, or by the nitrogen-washout method. Measurement of lung volumes is almost exclusively done in licensed pulmonary function laboratories. With very few exceptions, measurements of lung volumes are ordered to diagnose restrictive pulmonary im-
painment, or to follow-up patients with suspected restrictive pulmonary impairment.8

The ability of spirometry to predict a restrictive disorder has been determined.3 The typical restrictive pattern (a reduced FVC below the 95th, and a normal or above-normal FEV1/FVC ratio) had a sensitivity of 68%. Of those with a restrictive spirometric pattern, < 60% had true restriction by lung volume measurements (positive predictive value [PPV], 58%).3 This lack of a sensitive and affordable means of screening for restriction has consequences on clinical practice and the health-care system. It creates a dependency on formal measurement of lung volumes to establish a diagnosis, resulting in additional health-care costs. Costs become even more considerable when we consider that plethysmographic lung volume testing is ordinarily only available at specialized referral centers, which necessitates patient travel and additional expense associated with the testing.9

The objective of our study was to design a simple algorithm to predict which patients require measurements of lung volumes following spirometry to diagnose restrictive lung disease. To maximize the potential information from spirometry, we examined individual variables as well as combinations of variables. To be practical for most busy clinicians, and to effectively direct patient management, this algorithm had to have clear guidelines for interpretation and be applicable to almost all patients presenting for spirometry. An impetus toward development of this algorithm was to significantly reduce the number of unnecessary lung volume measurements that were being ordered in our hospital to screen for restrictive disease.

**Materials and Methods**

Data for the spirometry-derived algorithm were prospectively collected from 259 white adults consecutively referred for PFTs to the Ottawa Hospital, General campus, from November 2000 to February 2001. Validation of the algorithm was obtained from an independent cohort of 265 patients, fitting the identical inclusion/exclusion criteria and methods. The algorithm generated in the derivation phase of the study was applied to the validation phase data, and the performance properties of the algorithm were reassessed.

**Results**

A total of 259 spirometric tests and lung volume measurements were performed during the derivation phase of the study. The mean (± SD) age of the patients was 52.8 ± 16.2 years, and 53.3% were male. Fifty-five patients (21.2%) had a restrictive defect by lung volume testing. Two hundred sixty-five PFTs and lung volume measurements were done during the validation phase of the study. The mean (± SD) age of the patients was 54.2 ± 15.6 years, and 55.5% were male. Sixty-three patients (23.8%) had a restrictive defect by lung volume testing.

Spirometric variables measured in the derivation data set were correlated with TLC. TLC correlated strongly with SVC and FVC (r = 0.71, p < 0.0001, and r = 0.66, p < 0.0001, respectively) and showed a moderate inverse correlation with the FEV1/FVC and FEV1/SVC ratios (r = −0.41, p < 0.0001, and r = −0.39, p < 0.001, respectively).

FVC and SVC were then comparatively analyzed for their ability to predict a restrictive impairment. Figure 1 shows the receiver operating curve comparing the predictive ability of FVC and SVC. Calculated areas under the curve for FVC (0.90) and SVC (0.89) illustrate that no difference exists between their ability to predict a restrictive impairment.
Because there was no advantage to pursuing SVC over FVC in determining variables that predicted restriction, we chose to continue our analysis with FVC. FVC is more familiar to clinicians, and current spirometric devices all measure it and include it as an output variable.

Multiple cut points of FVC were analyzed in a series of $2 \times 2$ tables to determine their sensitivity for predicting restriction. We narrowed the range of possible cut points by choosing those with a high sensitivity for restriction (target sensitivity between 94% and 100%). From these, we chose the cut point that required the fewest number of patients needing additional lung volume measurements to confirm the diagnosis. FVC < 85% of predicted performed the best in predicting restriction, with a sensitivity of 98%.

Aside from FVC, FEV$_1$/SVC and FEV$_1$/FVC ratios were the two spirometric variables that correlated most strongly, and equally, with the TLC. The FEV$_1$/FVC ratio is more familiar to clinicians, and it was chosen for further analysis. An FEV$_1$/FVC ratio $\geq$ 55% had the highest sensitivity for predicting restriction and required the fewest number of patients to continue to lung volume testing to confirm the diagnosis. An FEV$_1$/FVC ratio $\geq$ 55% therefore served as step 2 of our algorithm. Step 1 and step 2 were combined into the spirometric algorithm as shown in Figure 2. The combined sensitivity of this algorithm was 96% (specificity, 61%; PPV, 40%; negative predictive value [NPV], 98%). Use of the algorithm would have resulted in only 133 of 259 patients (51%) proceeding to lung volume testing for confirmation of restriction (Table 1).

Spirometric and lung volume measurements were measured prospectively over a subsequent 4-month period on 265 patients for the validation phase of our algorithm. We re-ran the algorithm on the validation data set. The overall sensitivity for predicting restriction in the validation set, using our algorithm, was 94% (specificity, 61%; PPV, 43%; NPV, 97%) (Table 1). Use of the algorithm required that 137 of 265 patients (52%) proceed to lung volume measurements to confirm the diagnosis (Fig 2).

To test whether our algorithm would perform if a different set of prediction equations were used, we re-ran the validation data using spirometry prediction equations from Morris et al$^{14}$ and lung volume prediction equations from Goldman and Becklake.$^{15}$ Of the 265 patients, 70 patients (26%) had a restrictive defect using Goldman and Becklake$^{15}$ lung volume criteria. An analysis of the accuracy of the algorithm with the alternative prediction equations yielded results consistent with those obtained when...
using the equations of Crapo et al\textsuperscript{12,13} (sensitivity, 94%; specificity, 67%; PPV, 51%; NPV, 97%) [Table 2].

The number of patients with a specific physiologic diagnosis based on spirometry and lung volume testing, using the predicted values of Crapo et al\textsuperscript{12,13} and ATS criteria,\textsuperscript{6} is shown in Table 3. Forty-four of 259 patients (17%) and 56 of 265 patients (21%) in the derivation and validation studies, respectively, had a pure restrictive defect. A combined restrictive

---

### Table 1—Ability of Algorithm To Predict Restrictive Pulmonary Impairment in Derivation and Validation Studies\textsuperscript{*}

<table>
<thead>
<tr>
<th>Variables</th>
<th>Restricted</th>
<th>Not Restricted</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derivation: accuracy of the algorithm to predict restriction†</td>
<td>Proceed to lung volumes</td>
<td>53</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Lung volumes unnecessary</td>
<td>2</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>55</td>
<td>204</td>
</tr>
<tr>
<td>Validation: accuracy of the algorithm to predict restriction‡</td>
<td>Proceed to lung volumes</td>
<td>59</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Lung volumes unnecessary</td>
<td>4</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>63</td>
<td>202</td>
</tr>
</tbody>
</table>

\textsuperscript{*}Data are presented as No. of patients.

\textsuperscript{†}Sensitivity, 53 of 55 patients (96%); PPV, 53 of 133 patients (40%); specificity, 124 of 204 patients (61%); NPV, 124 of 126 patients (98%); proceed to lung volume measurements, 133 of 259 patients (51%).

\textsuperscript{‡}Sensitivity, 59 of 63 patients (94%); PPV, 59 of 137 patients (43%); specificity, 124 of 202 patients (61%); NPV, 124 of 128 patients (97%); proceed to lung volume measurements, 137 of 265 patients (52%).

### Table 2—Validation Study: Accuracy of the Algorithm Using Alternative Prediction Equations\textsuperscript{*}

<table>
<thead>
<tr>
<th>Variables</th>
<th>Restricted</th>
<th>Not Restricted</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proceed to lung volumes</td>
<td>66</td>
<td>64</td>
<td>130</td>
</tr>
<tr>
<td>Lung volumes unnecessary</td>
<td>4</td>
<td>131</td>
<td>135</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>195</td>
<td>265</td>
</tr>
</tbody>
</table>

\textsuperscript{*}Data are presented as No. of patients. Sensitivity, 66 of 70 patients (94%); PPV, 66 of 130 patients (51%); specificity, 131 of 195 patients (67%); NPV, 131 of 135 patients (97%); proceed to lung volume measurements, 130 of 265 patients (49%). Alternate prediction equations are from Morris et al\textsuperscript{14} and Goldman and Becklake.\textsuperscript{15}
Data are presented as No. (%). ATS definitions are as follows: normal, pure restriction, pure obstruction, combined restrictive-obstructive, and obstructive defect were detected in 11 of 259 patients (4%) and 7 of 265 patients (3%), respectively, in the derivation and validation studies. A total of 21% and 24% of patients, respectively, in the derivation and validation studies were restricted.

The referring diagnosis of patients in the validation study is shown in Table 4. Seventeen of 265 patients (6%) were referred with a known diagnosis of interstitial lung disease, and 16 of 265 patients (6%) had suspected interstitial lung disease. An additional seven patients had suspected pleural or chest wall restrictive disease. In total, 40 of 265 patients (15%) were referred to the laboratory with suspected restrictive lung disease.

We tested the ability of the algorithm to detect restriction for the subgroup of patients who were referred to the laboratory with suspected parenchymal, pleural, or chest wall restriction. Of the 40 patients whose referral diagnosis indicated suspected restrictive impairment, 24 patients ultimately had a restrictive pulmonary impairment confirmed by body-box lung volume testing. All 24 patients were picked up by the algorithm and recommended for further lung volume testing. In this subgroup of patients with suspected restriction, the algorithm had a sensitivity of 100%, NPV of 100%, specificity of 75%, and PPV of 86%.

Patients with moderate-to-severe obstructive defects were also analyzed as a subgroup. A plot of TLC percentage of predicted vs FEV₁/FVC ratio for the subgroup of patients with an FVC < 85% of predicted and an FEV₁/FVC ratio < 55% in the validation study is shown in Figure 3. These are a group of patients with at least moderate airflow obstruction, for whom the algorithm does not recommend lung volume testing. Figure 3 shows four patients with a TLC between 85% and 100% of predicted, who may have a potential restrictive process that may be obscured by a higher-than-normal baseline TLC due to overinflation. Potential restricted cases such as these are not typically detected by current ATS criteria.6

A cost analysis comparing utilization of our algorithm to current practice patterns is shown in Table 5. In Ontario, adding lung volume measurements to spirometry increases the costs of PFTs by 66%.16 We determined, at our institution, that measurements of lung volumes following spirometry and lung volume measurements ordered by physicians in 2000. Assuming that all the tests were ordered to diagnose restriction, use of our algorithm would have prevented 48 to 49% of these patients from undergoing lung volume testing. Adopting our algorithm would have thus resulted in 588 fewer measurements of lung volumes following spirometry at our institution. This would have resulted in an overall cost reduction of 33%.

**Discussion**

The objective of this study was to design a simple algorithm that would help clinicians and pulmonary function laboratories to decide which patients should
undergo measurement of lung volumes following spirometry. Since lung volume measurements are almost always ordered to diagnose restrictive pulmonary disease, the algorithm was designed to predict the presence or absence of pulmonary restriction using spirometry. The goal was to design an algorithm that was highly sensitive for predicting a restrictive defect, and whose use would significantly reduce the number of unnecessary lung volume measurements currently being ordered by clinicians.

We have designed an algorithm that is sensitive for predicting pulmonary restriction and which reduces the number of unnecessary lung volume measurements by almost half. According to this algorithm, only patients with an FVC < 85% of predicted who also have an FEV1/FVC ratio ≥ 55% require further lung volume measurements to confirm the diagnosis of restrictive lung disease.

Physiologically, this algorithm makes intuitive sense. Those patients with an FVC > 85% predicted are unlikely to be restricted and do not require lung volume measurements. Similarly, patients with a low vital capacity, but who also have obstructive disease on spirometry (ie, FVC < 85% of predicted and an FEV1/FVC < 55%), have a reduced vital capacity because of air trapping, with impingement on their vital capacity by a high residual volume. These patients are also very unlikely to have true restrictive pulmonary disease and do not require measurements of lung volumes.

The sensitivity of this algorithm is 94 to 96%, and use of the algorithm avoids unnecessary lung volume tests in 48 to 49% of patients when compared with current clinical practice. At our institution, this results in a 33% reduction in the cost of PFTs. When validated in an independent and prospective study, and when tested using alternative prediction equations, these results were consistent and reproducible.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Spirometry Cost</th>
<th>Spirometry No. per Year</th>
<th>Lung Volume Cost</th>
<th>Lung Volume No. per Year</th>
<th>Total Yearly Cost, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current practice</td>
<td>16.75</td>
<td>1,200</td>
<td>33.20</td>
<td>1,200</td>
<td>59,940</td>
</tr>
<tr>
<td>Using the algorithm</td>
<td>16.75</td>
<td>1,200</td>
<td>33.20</td>
<td>1,200 ( \times 0.51 = 612 )</td>
<td>40,418</td>
</tr>
</tbody>
</table>

*Total potential savings are $19,522/859,940 per year (33% of total costs). Costs are calculated in Canadian dollars.
We foresee the algorithm being incorporated into clinical practice at the level of the practicing clinician. Clinicians would perform office or community spirometric testing and use the algorithm to decide which patients require subsequent referral to the PFT laboratory for lung volume testing.

Similarly, the algorithm can also be applied at the level of the PFT laboratory. Some referrals for combined spirometry and lung volume testing come into the PFT laboratory because the referring physician has simply “checked off all of the boxes” on the requisition form. For those patients for whom the algorithm does not recommend lung volume testing after spirometry, lung volumes can still be performed as requested, but the algorithm can be included and sent to the referring physician to help direct the necessity of future lung volume requests.

Conversely, for the occasional patient in whom the referring physician had only ordered spirometry, and not lung volume measurements, the laboratory could decide to perform measurements of lung volumes if the algorithm suggested that the patient has a relatively high probability of being restricted (ie, patients whose spirometry shows an FVC < 85% of predicted and an FEV1/FVC ratio ≥ 55%). Thus, adoption of this algorithm over the current clinical practice pattern could result in the following: (1) fewer patients unnecessarily undergoing lung volume measurements to rule out restriction, and (2) fewer missed cases of restriction, since pulmonary function laboratories could automatically measure lung volumes in patients whose spirometry showed an FVC < 85% and an FEV1/FVC ratio ≥ 55%.

Previous studies have investigated the relationship between single-breath alveolar volume (VA) and TLC, and have determined that VA accurately predicts TLC in patients with mild or no airflow obstruction. This suggests that an algorithm predicting restriction could be alternatively composed of VA and the FEV1/FVC ratio. However, measurements of VA must be made in the PFT laboratory. A substantial amount of decision making regarding which patients need referral for lung volume testing takes place in the clinician’s office, following spirometry; therefore, an algorithm that incorporated VA would not eliminate unnecessary referrals, whereas our algorithm is designed to be applied following spirometry, thus avoiding PFT laboratory referrals.

Race has consistently shown to be an important predictor of lung function. Compared to white subjects, values for other races usually show smaller lung volumes. The reason for the difference in race is unclear, but it is thought to be related to differences in body build or environmental and socioeconomic difference. In both the derivation and validation phases of our study, we utilized prediction equations from Crapo et al,12,13 which were derived from white patients. To be consistent with the prediction equations, we restricted our study population to white patients. Our algorithm can therefore only be reliably used on white patients, and cannot accurately predict results for other races.

There are a small number of patients for whom the algorithm may not be relevant. A small number of patients undergo lung volume measurements for reasons other than the diagnosis of restrictive lung disease. For example, lung volumes are occasionally measured to determine the residual volume and degree of hyperinflation in patients with severe emphysema prior to volume reduction surgery. Our algorithm would obviously not be relevant for use in this clinical situation. Similarly, the algorithm would not be applicable to those patients with known interstitial lung disease who undergo serial testing of lung volumes for purposes of monitoring disease progression.

Both a pure restrictive defect and a combined obstructive and restrictive disorder are described by a TLC below the 95% confidence interval, and it is this definition we chose for identifying restrictive lung diseases in our algorithm; however, there are potential limitations to using the ATS definition of restriction in cases of combined restriction and obstructive defects. Patients with severe obstruction with overinflation often have an enlarged residual volume and TLC. If there is a concurrent restrictive defect, TLC may then fall within the normal range. If one defines restriction strictly by the ATS criteria, then these cases may be missed.

In order to examine how this may affect our algorithm, we looked at those patients with obstruction for whom the algorithm does not recommend lung volume testing, defined as patients with an FVC < 85% of predicted and an FEV1/FVC ratio < 55%. A plot of the TLC percentage of predicted and FEV1/FVC ratio in this subgroup of patients shows four patients with a TLC between 85% and 100% of predicted (Fig 3). These are patients for whom the diagnosis of restriction cannot be made by current ATS criteria for interpreting lung volumes, but whose diagnosis must rely on radiologic criteria or clinical suspicion.

The algorithm was very accurate at detecting restriction in those in whom the disorder was either suspected clinically or previously diagnosed. A subgroup analysis of the 40 patients in the validation study referred with suspected restrictive lung disease showed that the algorithm performed with 100% sensitivity and a 100% NPV.

Although our algorithm has a high sensitivity, it missed a total of 6 patients (1.1%) with restriction from the combined data set of 524 patients. A
A detailed look at these patients showed that they all had a TLC that was only marginally low, suggesting that most patients with clinically important restriction will not be missed when applying our algorithm. A cost analysis of our algorithm compared to the current clinical practice pattern shows that adoption of our algorithm would cut the costs of combined spirometric and lung volume testing by 33%. At our institution, this amounted to a cost savings of almost $20,000 (Canadian) per year. These calculated cost savings might be expected to be even greater in centers in the United States, where the charge for performing and interpreting lung volume measurements can be up to $130 (US dollars) per test. Combining these potential cost savings with the improved sensitivity in diagnosing restriction makes adoption of our diagnostic algorithm both medically reasonable and economically wise.

REFERENCES
3 Aaron SD, Dales RE, Cardinal P. How accurate is spirometry at predicting restrictive pulmonary impairment? Chest 1999; 115:869–873
4 Chhabra SK. Forced vital capacity, slow vital capacity, or inspiratory vital capacity: which is the best measure of vital capacity? J Asthma 1998; 35:361–365
7 Killburn KH, Miller A, Warshaw RH. Measuring lung volumes in advanced asbestosis: comparability of plethysmographic and radiographic versus helium rebreathing and single breath methods. Respir Med 1993; 87:115–120