QUALITY OF LIFE MEASURED BY THE ST GEORGE'S RESPIRATORY QUESTIONNAIRE AND SPIROMETRY

1Mark Weatherall, 2Suzanne Marsh, 2Philippa Shirtcliffe, 2Mathew Williams,
2Justin Travers, 2Richard Beasley

1University of Otago Wellington, Wellington, New Zealand
2Medical Research Institute of New Zealand, Wellington, New Zealand

Short title: Quality of life and spirometry

Word count: 3,074

Correspondence:
Professor Richard Beasley
Medical Research Institute of New Zealand
PO Box 10055, Wellington 6143
NEW ZEALAND
Tel: +64-4-4729199
Fax: +64-4-4729224
Email: Richard.Beasley@mrinz.ac.nz
ABSTRACT [203 WORDS]

We aimed to determine if the criteria for the diagnosis of COPD and its classification by severity as recommended by the Global Initiative for Obstructive Lung Disease are supported by measurements of respiratory health-related quality of life.

A community-based sample of adults aged 25 to 75 years had pre- and post-bronchodilator spirometry and completed the St George’s Respiratory Questionnaire (SGRQ). Loess scatter plot smoothers of SGRQ versus post-bronchodilator FEV₁/FVC ratio and post-bronchodilator FEV₁ as percent predicted together with receiver operating characteristic (ROC) curve analysis were used to determine the relationship between spirometric variables and clinically important differences in the SGRQ score.

The scatter plot smoother and ROC curve analyses supported the value of 0.7 for post-bronchodilator FEV₁/FVC, which was about four units higher than the nadir of the SGRQ. To represent a distance of eight units on SGRQ, the cut-points for post-bronchodilator FEV₁ that delimit COPD severity stages were 80%, 60% and 40% predicted.

To diagnose COPD the use of a post-bronchodilator FEV₁/FVC of 0.7 is supported by health-related quality of life measurements. There may be advantages in using FEV₁ cut-points of 80%, 60% and 40% predicted for the classification of mild, moderate and severe COPD, similar to the approach recommended for asthma.

Key words: COPD, quality of life, spirometry
INTRODUCTION

The Global Initiative for Obstructive Lung Disease (GOLD) defines Chronic Obstructive Pulmonary Disease (COPD) as a post-bronchodilator forced expiratory volume in one second to forced vital capacity (FEV$_1$/FVC) ratio of less than 0.7, not explained by another lung disease [1]. However, use of this fixed ratio has been criticised for not reflecting physiological changes of this variable with aging, potentially leading to over diagnosis of COPD particularly in the elderly [2-5].

The utility of spirometry based definitions of COPD is supported if other measurements of the impact of lung disease show a definite change in relation to spirometry. This constitutes a form of external validity for whatever measurement of lung function is chosen, be it FEV$_1$, FEV$_1$ as a percent of its predicted value, or absolute definitions based on the FEV$_1$/FVC ratio. The St George’s Respiratory Questionnaire (SGRQ), which is a measurement of respiratory quality of life, has been validated in COPD [6-8]. However, it was developed prior to the GOLD guidelines with their rigorous spirometric definition of COPD. Furthermore these guidelines include severity cut off criteria based on FEV$_1$ as a percent of predicted value, which are acknowledged to be based on simplicity rather than clinical validation [1]. Large longitudinal epidemiological studies would be required to validate the choice of such severity cut-off criteria, however corroboration of these with meaningful differences in health-related quality of life could potentially serve as a useful surrogate for this. For the purposes of randomised controlled trials a change in the total score for the SGRQ that represents a clinically important difference is four units [6].

This aims of this paper are to examine if the value for FEV$_1$/FVC ratio of 0.7 can be supported by the measurement of respiratory quality of life in the SGRQ. Secondly we
explore whether the boundaries for severity of COPD based on the GOLD criteria, using cut points of FEV₁ as a percent predicted, are supported by clinically important quality of life changes reflected in the SGRQ.
METHODS

Subjects
Participants in the Wellington Respiratory Survey (WRS), n=3,500, were randomly selected from the electoral register, equally distributed by gender across the five decade age groups from 25 to 75 years. Subjects were sent a simple postal questionnaire seeking demographic, respiratory and smoking history data. All subjects who completed and returned questionnaires were invited to undertake more detailed questionnaires, pulmonary function tests and CT scanning as described in detail previously [9]. Subjects who completed all investigative modules formed the study group for this analysis. The reference range population comprised 212 subjects who were never smokers with no diagnosis of respiratory disease, no recent respiratory symptoms and no use of inhaled medication. The survey was approved by the Wellington Ethics Committee and written informed consent was obtained from each subject.

Pulmonary function testing
Pulmonary function tests in the WRS have been described in detail elsewhere [9,10]. In brief spirometry measurements were carried out according to ATS guidelines [11] with a minimum of three acceptable manoeuvres carried out and the best FEV₁ and FVC selected for analysis. Spirometry was repeated 45min after the administration of 400µg of salbutamol (Ventolin, GlaxoSmithKline NZ Ltd, Auckland, New Zealand) via a spacer (Space Chamber, Medical Developments International Ltd, Springvale, Australia). FEV₁ and FVC values were expressed as a percentage of predicted normal values [10].

Quality of life questionnaire
The St Georges Respiratory Questionnaire (SGRQ) is a disease specific quality of life assessment tool validated in both COPD and asthma [12-14]. The questionnaire consists of 76 items divided into three parts measuring symptoms, activity limitation and social and emotional impact of disease. Each item is accorded a weight determined by the degree of distress accorded to each symptom or state described. Overall scores range from 0 (no effect on quality of life) to a maximum score of 100 (maximum perceived distress), thus a higher score means a worse quality of life and the questionnaire is suitable for administration in healthy persons [8]. The English for New Zealand version was self completed by the subjects during the 45 minute bronchodilation time. Results were entered into an Excel spreadsheet, supplied by the questionnaire producers, and pre-programmed with formulae to calculate a total score and scores for each of the individual components of the SGRQ (symptoms, activity, impacts). A change in the total score for the SGRQ that represents a clinically important difference is four units [6].

**Statistical methods**

Quantile regression [15] was used to find a formula to predict the upper 90 and 95% quantiles for the SGRQ based on age. The dataset used for this was the 178/212 subjects from a reference range dataset which had values for the total score on the SGRQ (data for the SGRQ was missing in 32 subjects from the reference population sourced from outside of the WRS dataset [10] and from 2 of the WRS cohort).

Loess lines were fitted to the overall WRS dataset for all 713 subjects who had values for the total score on the SGRQ, plotting the SGRQ versus the post-bronchodilator FEV₁/FVC ratio and post-bronchodilator FEV₁ as a percent predicted. Loess plots are a form of robust scatter plot smoothing suitable for noisy data [16], providing information on the form of a relationship between variables where it may not be easily described by conventional
regression. The loess technique uses a weighted average of a set of data points to fit curves rather than the individual points as in conventional regression. The technique is relatively resistant to outlying values and does not require a prespecified global linear function to be chosen. The degree of smoothing is determined by the width of the window determining the weighted average. Optimal smoothing is obtained by balancing the residual sum of squares for each point against the degree of smoothness.

It was evident that both curves had a point of inflection, where the value of the SGRQ increased from a minimum level, and we examined the value of the SGRQ four and eight units above this point of inflection. Receiver operating characteristic (ROC) curves for two cut-points for post-bronchodilator FEV$_1$/FVC ratio, 0.7 and 0.65, and the total score on the SGRQ, were fitted to determine discriminatory values, sensitivity and specificity, for the SGRQ. SAS version 9.1 was used for the loess analysis.

The relationship of the post-bronchodilator FEV$_1$ as a percent of predicted, based on equations from the WRS [10], and the loess lines were used to examine the cut-points for FEV$_1$ as a percent predicted in relation to the GOLD severity cut-points of 80%, 50% and 30% predicted.
RESULTS

Subjects

There were 2319 responders to the postal questionnaire of whom 795 agreed to attend our research facility for further testing and 713 were able to complete all WRS investigative modules and form the study group. Subjects in the study group were broadly similar to the 1606 postal questionnaire responders not in the study group. Subjects in the study group were more likely to be of male gender (54% vs. 44%, p=<0.001), smoke tobacco (13% vs. 10%, p=0.007), have a history of chronic cough (15% vs. 11%, p=0.006) or breathing trouble (28% vs. 24%, p=0.04) and had a greater mean age (53.9 vs. 51.6 years, p<0.001) than postal questionnaire responders not in the study group.

Quantile regression prediction equations for SGRQ

The quantile regression equations for the 95 and 90% quantiles for the reference range subjects are (1.53 + Age×0.24) and (5.76 + Age×0.096) respectively. Using these equations, the 95% or 90% quantile is found by inserting the age in years into the equation, for example the 90% quantile for a 50 year old is 5.76 plus 50 times 0.096, equals 10.56. The predicted 95% and 90% quantiles ranged from 7.9 to 19.0 and 8.3 to 12.8 respectively across the age range of 27 to 74 years. Sex was not an important predictor for these two quantiles. We observed different coefficients for different quantiles as is allowed by the method of quantile regression.

SGRQ versus post-bronchodilator FEV₁/FVC ratio

Figure 1 shows a scatter plot of the SGRQ total score versus post-bronchodilator FEV₁/FVC ratio with the superimposed loess smoothed line and 90% confidence interval. The point of inflection, the minimum level for the SGRQ, in the loess plot was at a post-
bronchodilator FEV₁/FVC ratio of about 0.8. The SGRQ predicted at this value is 8.1 units. Four units higher than this value (SGRQ score 12.1 units), corresponds to a post-bronchodilator FEV₁/FVC ratio of 0.71. Eight units higher (SGRQ score 16.1 units) corresponds to a value of the post-bronchodilator FEV₁/FVC ratio of 0.65.

**Receiver operating characteristic (ROC) curve analysis**

119/713 (16.7%) subjects had a post-bronchodilator FEV₁/FVC ratio of less than 0.7. The area under the ROC curve for the SGRQ for this value of post-bronchodilator FEV₁/FVC ratio was 0.692 suggesting moderate discrimination of the SGRQ for this cut-point. 72/713 (10.1%) subjects had a post-bronchodilator FEV₁/FVC ratio of less than 0.65. The area under the ROC curve for the SGRQ for this value of post-bronchodilator FEV₁/FVC ratio was 0.812 suggesting moderate to good discrimination.
Table 1 shows the sensitivity and specificity values and the corresponding SGRQ total scores for the two post-bronchodilator FEV₁/FVC ratio cut-points of 0.7 and 0.65. Sensitivity refers to the number of subjects correctly allocated to below the cut-off value by the corresponding value of SGRQ total score divided by the total number of subjects who in fact had a ratio below the cut-off. Specificity refers to the number of subjects correctly allocated to above the cut-off divided by the total number who in fact had a ratio above the cut-off. For example a SGRQ total score of 12.0 has sensitivity for a post-bronchodilator FEV₁/FVC ratio of 0.65 of 66.7% and specificity of 70%.

Figure 2 shows the full ROC curves for the SGRQ and two values of the post-bronchodilator FEV₁/FVC ratio.

**SGRQ versus post-bronchodilator FEV₁ as percent predicted**

Figure 3 shows a scatter plot of the SGRQ total score versus post-bronchodilator FEV₁ as percent predicted with superimposed loess smoothed line and 90% confidence interval. The point of inflection in the loess plot is a post-bronchodilator FEV₁ of about 100% predicted. The SGRQ predicted at this value is 7.9 units. Four units higher than this value (SGRQ score 11.9 units) corresponds to a value of the post-bronchodilator FEV₁ as percent predicted of 88%. Eight units higher (SGRQ score 15.9 units) corresponds to a value of the post-bronchodilator FEV₁ as percent predicted of 80%.

From the plot it can be seen that the relationship between the SGRQ and post-bronchodilator FEV₁ as percent predicted is approximately linear from a FEV₁ value of 90% with a 10% change in FEV₁ associated with a four unit change in the SGRQ. The differences in SGRQ score associated with the GOLD boundaries for severity are shown in Table 2. These data show that there are major differences in the magnitude of change in
SGRQ score between the GOLD severity boundaries of 80%, 50% and 30% predicted, which define mild, moderate and severe COPD. In contrast, the magnitude of change in SGRQ score between the severity boundaries of 80%, 60% and 40% predicted (as recommended in asthma guidelines to define mild, moderate and severe disease [17]) were similar, corresponding to about twice the clinically important difference for SGRQ (eight units).
DISCUSSION

We have shown that the suggested boundary of normality for a post-bronchodilator FEV$_1$/FVC ratio of 0.7 for COPD is supported by analysis of our data set. This FEV$_1$/FVC value corresponded to a value of the SGRQ score close to the 90% quantile in the reference range sample representing a clinically important difference in the SGRQ from the nadir of the relationship. Our findings also suggest that the current GOLD cut-off values of 80%, 50% and 30% to define mild, moderate and severe COPD do not result in a similar magnitude of changes in SGRQ. If the boundaries of severity are to correspond to approximately equal distances on the scale of the SGRQ, boundaries of 80%, 60%, and 40% of FEV$_1$ as percent predicted would each correspond to about twice the clinically important difference for the SGRQ.

Use of a fixed ratio of FEV$_1$/FVC of 0.7 to diagnose COPD has been widely criticised for failing to take account of physiological changes in the value of the ratio with aging, which can result in over-diagnosis of COPD [3]. Our own group has shown a considerable difference in the prevalence of COPD using a fixed ratio compared to age determined predicted values for FEV$_1$/FVC [9]. In its defence, use of a fixed ratio results in guideline simplicity [1,18], has been supported by some epidemiological data [19], and values below this have been shown to be predictive of poorer outcomes despite apparent good health [20]. Since in practical terms use of a fixed ratio is likely to continue it is useful to show that this ratio corresponds to clinically significant changes in health status on a validated respiratory questionnaire with reasonable specificity.

We derived normal reference range values for the SGRQ based on its distribution in healthy non-smokers, identifying that the 90% quantile for the SGRQ across the age range...
in our reference range group was 12.8 units. This is not dissimilar to values reported in a previous study although different methods of reporting makes direct comparison difficult [8]. Our results also reflect the previously acknowledged effect of age, irrespective of respiratory disease, on health related quality of life (HRQL) [21]. Using smoothed plots of SGRQ versus post-bronchodilator FEV\textsubscript{1}/FVC ratio the nadir was 8.1 for the SGRQ and 0.8 for the post-bronchodilator FEV\textsubscript{1}/FVC ratio. An increase in the SGRQ score by four units, an amount that has been proposed to represent a clinically important difference [6,8], resulted in a value of 12.1 which corresponded to a post-bronchodilator FEV\textsubscript{1}/FVC ratio of 0.7. An increase in the SGRQ score by a further four units corresponded to an FEV\textsubscript{1}/FVC ratio of 0.65. Construction of ROC curves identified that a value of SGRQ of between 12 and 13, corresponding to the 90% quantile in our reference range sample, had moderate to good discriminative power for a cut-off value for post-bronchodilator FEV\textsubscript{1}/FVC ratio of 0.7. Not surprisingly use of a lower fixed ratio of 0.65 improves the specificity associated with the same SGRQ score, at the cost of reduced sensitivity. Thus, a post-bronchodilator FEV\textsubscript{1}/FVC ratio of 0.7 was represented in our dataset by a clinically important difference in the SGRQ compared to the nadir of the relationship and a value of the SGRQ close to the 90% quantile in a reference range sample.

We also used our data to examine the validity of the severity criteria cut off points for FEV\textsubscript{1} espoused by GOLD. Whilst it is known that lower FEV\textsubscript{1} values are associated with poorer health status the strength of this correlation is not always that strong [21]. Such poor correlations indeed strengthen the case for holistic patient assessment using tools such as quality of life questionnaires and the more recently developed BODE index [22]. Nonetheless, an FEV\textsubscript{1} cut point is simple to apply and has been used in guidelines to recommend preferred therapy and in some settings to control access to specific treatments [23]. Whilst the arbitrary nature of cut-points is acknowledged, it is reassuring that the
GOLD severity criteria broadly correlate with the SGRQ, a tool which takes into account the multifaceted impact of COPD. However, our findings suggest that the current cut-off values of 80%, 50% and 30% to define mild, moderate and severe COPD do not result in a similar magnitude of changes in SGRQ as the relationship is essentially linear below 80%. If the boundaries of severity are to correspond to approximately equal distances on the scale of the SGRQ, boundaries of 80, 60, and 40% of FEV₁ as percent predicted would each correspond to about twice the clinically important difference for the SGRQ. The use of the 80%, 60% and 40% cut-points would also have the advantage of consistency with the grading of severity for asthma [17]. In support of this approach, many older patients with asthma will meet the spirometric-defined criteria for COPD [24].

There are methodological issues to consider in this study. There were minor differences in the characteristics of the study group compared to postal questionnaire responders who did not complete all investigative modules, raising the possibility of selection bias. It is unlikely however that these differences would systematically affect the relationship between the SGRQ and spirometry. There is a high prevalence of asthma in New Zealand and some of the subjects identified as having COPD from post-bronchodilator spirometry may actually have asthma with incomplete bronchodilator reversibility rather than COPD. In a separate analysis of the WRS data, approximately 25% of subjects identified as having COPD had features of asthma and no apparent chronic bronchitis or emphysema [24]. As we cannot determine what proportion of these individuals have COPD with an asthma-like phenotype and what proportion have ‘pure’ asthma, we have not excluded individuals with an asthma-like phenotype of airway disease from the analysis. There were few subjects in the study group with severe COPD and thus our analysis primarily applies to those with mild or moderate COPD. The confidence intervals for the relationship between FEV₁ and SGRQ are consequently wider at a low FEV₁ and could support a
greater, lesser or even non-linear relationship between FEV$_1$ and SGRQ in advanced disease. The SGRQ measurements were 'noisy' and although the loess scatter plot smoother is useful for examining the underlying pattern of the data set it may give a spurious impression of precision. Reducing the degree of smoothing in the algorithm fitting the loess scatter plot smoother may better reflect the noisiness of the data but would reduce the ability to see the overall pattern we have demonstrated.

A considerable strength of our study was that it was a community based random sample and so may not be subject to biases created by using diseased populations such as those attending hospital clinics. The SGRQ is an accepted method of measuring respiratory health quality of life and the definition of a clinically important difference with this instrument means that the analysis reflects associations of clinical importance.
CONCLUSION

We conclude that in the diagnosis of COPD the use of a post-bronchodilator FEV\(_1\)/FVC ratio of 0.7 is supported by health-related quality of life measurements. There may be advantages in using post-bronchodilator FEV\(_1\) cut-points of 80%, 60% and 40% predicted for the classification of mild, moderate and severe COPD, similar to the approach recommended for asthma.
REFERENCES


Table 1: Comparison of the sensitivity and specificity of the SGRQ to identify a low post-bronchodilator FEV₁/FVC ratio using two different post-bronchodilator FEV₁/FVC ratio cut points.

<table>
<thead>
<tr>
<th>Post-bronchodilator FEV₁/FVC ratio = 0.7</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>SGRQ score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>41/119 (34.5%)</td>
<td>534/594 (89.9%)</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td>58/119 (48.7%)</td>
<td>475/594 (80.0%)</td>
<td>13.6</td>
</tr>
<tr>
<td></td>
<td>71/119 (59.7%)</td>
<td>416/594 (70.0%)</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>83/119 (69.7%)</td>
<td>325/594 (54.7%)</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>95/119 (79.8%)</td>
<td>216/594 (36.4%)</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>108/119 (90.8%)</td>
<td>139/594 (23.4%)</td>
<td>2.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Post-bronchodilator FEV₁/FVC ratio = 0.65</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>SGRQ score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32/72 (44.4%)</td>
<td>577/641 (90.0%)</td>
<td>23.1</td>
</tr>
<tr>
<td></td>
<td>43/72 (59.7%)</td>
<td>513/641 (80.0%)</td>
<td>14.0</td>
</tr>
<tr>
<td></td>
<td>48/72 (66.7%)</td>
<td>449/641 (70.0%)</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td>51/72 (70.8%)</td>
<td>441/641 (68.8%)</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td>58/72 (80.6%)</td>
<td>345/641 (53.8%)</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>65/72 (90.3%)</td>
<td>224/641 (34.9%)</td>
<td>3.9</td>
</tr>
</tbody>
</table>
Table 2: Difference in SGRQ between GOLD-defined COPD severity boundary points.

<table>
<thead>
<tr>
<th>Boundary points in GOLD severity: Range of post-bronchodilator FEV₁ as percent predicted</th>
<th>Number of units of SGRQ difference for range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild: 80% to 100%</td>
<td>8</td>
</tr>
<tr>
<td>Moderate: 50% to 80%</td>
<td>15</td>
</tr>
<tr>
<td>Severe: 30% to 50%</td>
<td>10</td>
</tr>
</tbody>
</table>
Figure 1: SGRQ Total score versus post-bronchodilator $FEV_1/FVC$ ratio with superimposed loess scatter plot smoother and 90% confidence interval. The loess smoother is a form of regression line using a weighted average of a set of data points at each part of the curve and is robust to outlying values.
Figure 2: Receiver operating characteristic plot showing the ability of the SGRQ to identify a low post-bronchodilator FEV1/FVC ratio for post-bronchodilator FEV1/FVC ratios of 0.65 (gray) and 0.7 (black).
Figure 3: SGRQ versus post-bronchodilator FEV₁ as percent predicted with superimposed loess scatter plot smoother and 90% confidence interval. The loess smoother is a form of regression line using a weighted average of a set of data points at each part of the curve and is robust to outlying values.