CHEST™

Improved Criterion for Assessing Lung Function Reversibility

Helen Ward, MBChB; Brendan G. Cooper, PhD; and Martin R. Miller, MD

BACKGROUND: Consensus on how best to express bronchodilator reversibility (BDR) is lacking. We tested different BDR criteria against the null hypotheses that BDR should show no sex or size bias. To determine the best criterion for defining BDR, we hypothesized that clinically important BDR should be associated with better survival in respiratory patients compared with that of patients without BDR.

METHODS: We used the first BDR test of 4,231 patients who had known subsequent survival status (50.8% male sex; mean age, 60.9 years; mean survival, 5.2 years [range, 0.1-16.5 years]). BDR for FEV_1 was expressed as absolute change, % baseline change, and change as % predicted FEV_1.

RESULTS: Having BDR defined from absolute change was biased toward men (male to female ratio, 2.70) and toward those with larger baseline FEV_1. BDR defined by % change from baseline was biased toward those with lower baseline values. BDR defined by % predicted had no sex or size bias. Multivariate Cox regression found those with FEV_1 BDR $>8\%$ predicted (33% of the subjects) had an optimal survival advantage (hazard ratio, 0.56; 95% CI, 0.45-0.69) compared with those with FEV_1 BDR $\leq 8\%$ predicted. The survival of those with FEV_1 BDR $>8\%$ predicted was not significantly different from that of those with FEV_1 BDR $>14\%$ predicted but was significantly better than that of those with FEV_1 BDR <0.

CONCLUSIONS: We have shown that expressing FEV_1 BDR as % predicted avoids sex and size bias. FEV_1 BDR $>8\%$ predicted showed optimal survival advantage and may be the most appropriate criterion to define clinically important reversibility. CHEST 2015; 148(4):877-886

Manuscript received September 30, 2014; revision accepted March 26, 2015; originally published Online First April 16, 2015.

ABREVIATIONS: BD = bronchodilator; BDR = bronchodilator reversibility; HR = hazard ratio; ICS = inhaled corticosteroid; +ve = positive; SR = standardized residual

AFFILIATIONS: From the Department of Respiratory Medicine (Drs Ward and Cooper), Queen Elizabeth Hospital; and the Institute of Occupational and Environmental Medicine (Prof Miller), University of Birmingham, Birmingham, England.

FUNDING/SUPPORT: The authors have reported to *CHEST* that no funding was received for this study.

CORRESPONDENCE TO: Martin R. Miller, MD, Institute of Occupational and Environmental Medicine, University of Birmingham, Birmingham, B15 2TT, England; e-mail: m.r.miller@bham.ac.uk

DOI: 10.1378/chest.14-2413

Reversibility testing of lung function using short-acting bronchodilators (BDs) is commonly undertaken in lung function laboratories, usually by recording any change in the FEV_1. There are numerous guidelines as to how to determine whether the observed change in FEV_1 is clinically important[1-9] (these are shown in Table 1) and "the lack of a standardized procedure for assessing bronchodilator reversibility has led to significant confusion surrounding this concept."[10] The change may be examined as absolute values, as percentage change from baseline, as change as a percentage of the subject's predicted value, or as combinations of these. Using change as a percentage of start value potentially biases the results to be positive in those with the lowest start value. Adding an absolute threshold change, based on observations about short-term variability in FEV_1,[11] was thought to overcome some of this criticism.

To determine how lung function reversibility test results should best be expressed, we tested the methods for expressing lung function reversibility against the null hypotheses that a suitable method should find no systematic difference in reversibility between the sexes and should not be size biased. A clinically important degree of reversibility is likely to be associated with the clinical condition of asthma and should mean that lung function may be improved by relevant drug therapy and that subsequent outcome may be improved. Lung conditions that usually do not have an important degree of reversibility (eg, COPD, pulmonary fibrosis, and bronchiectasis, to name a few) should have a worse prognosis than asthma. Therefore, we also tested the null hypothesis that a clinically important degree of BD reversibility (BDR) has no survival advantage.

Materials and Methods

All the records of patient visits to our hospital lung function laboratory were extracted from the database (N = 41,411) on January 11, 2012. The database was set up in 1996, and the use of anonymized data from it for research purposes was approved by the hospital's Caldicott Guardian, who was the Medical Director overseeing all regulatory issues. From these data, results were obtained for the first reversibility test for white subjects where salbutamol 4×100 μg was administered via a spacer device. The data for patients of Asian and Caribbean extraction were not included because of the continued debate around adjustments for ethnicity regarding lung function. There were valid data before and 40 min after BD for FEV_1 and FVC for 4,227 separate subjects aged > 20 years with $FEV_1 \geq 0.2$ L and $FEV_1/FVC \geq 0.2$ (2,147 men [50.8%]) whose National Health Service number was available to determine their survival up to June 13, 2012. There were data for an additional 1,124 subjects referred by family doctors who did not have an National Health Service number and so could not be analyzed further. These 1,124 included significantly more women (55.5% vs 49.2%, $P < .001$ [χ^2 test]) than did the study group and they were slightly older (mean ± SD, 62.9 ± 14.5 years vs 60.9 ± 13.6 years; $P < .001$ [Kruskal-Wallis test]).

All tests were conducted according to Association of Respiratory Technicians and Physiologists guidelines[12] and the ratio FEV_1/FVC was taken from the best FEV_1 divided by the best FVC (which could be from separate blow efforts). Subjects referred with FEV_1/FVC greater than the lower limit of normal were not routinely tested for BDR. Using the Global Lung Initiative (GLI) 2012 equations, baseline values were converted to standardized residuals (SRs), which are the same as *Z* scores.[13] The SR values were termed FEV_1 SR, FVC SR, and FEV_1/FVC SR and were derived from (recorded value − predicted value)/RSD, where RSD is the residual SD for the scatter of values in the normal reference population.[3]

The difference between the largest pre-BD FEV_1 and the largest post-BD FEV_1 was called ΔFEV_1. This was standardized in two ways: (1) as a percentage of the pre-BD (ie, start) value and (2) as a percentage of the subject's predicted value.

Statistical analysis was undertaken using Stata/SE 11.0 (StataCorp LP). Because the distribution of the change in FEV_1 after BD was skewed, comparisons were made using rank sum tests. Survival analysis was undertaken using Cox proportional hazards regression, with survival measured from the date of the BD test.

Results

Table 2 shows the median and interquartile range for pre-BD results for the 4,227 subjects, separated by sex and by survival status. By June 13, 2012, 444 subjects (10.5%) had died, with a significantly higher mortality in the men than in the women (11.6% vs 9.3%, $P < .02$ [χ^2 test]). The mean survival was 5.2 years (range, 0.1-16.5 years; median, 4.7 years; interquartile range, 2.7-7.5 years). For both sexes, those who died had, on average, worse lung function than did those alive at the end of the study, but the time at risk was no different in the women who died and was slightly shorter in the men who died. When considering only those who had died, the men had a significantly greater pack-year smoking exposure than did the women ($P < .05$ [rank sum test]), but for age, time at risk, and SR values for lung function, there were no differences between the sexes. In those alive at the end of the study, the men had significantly higher pack-year smoking exposure, shorter time at risk, better FVC SR, and worse FEV_1/FVC SR than did the women ($P < .05$ [rank sum test]).

On the request forms, the putative diagnosis or reason for requesting the test had asthma mentioned in 25% and COPD in 30% of subjects. These possible diagnoses could not be verified independently. Most of the remaining indications were symptoms (28%), and 5% were preprocedures such as bronchoscopy or general surgery. Most subjects were already receiving inhaled