# SELECTION OF THE BEST SPIROMETRIC VALUES FOR INTERPRETATION[1]


## Summary

Selection of spirometric test values for reporting and interpretation has recently received considerable attention. In 1977, the American Thoracic Society (ATS) Snowbird Workshop on Standardization of Spirometry recommended that the maximal values for FVC and $FEV_1$ be used for clinical interpretation, even if they came from different spirometric tracings. The Intermountain Thoracic Society (ITS) had recommended in 1975 that FVC and $FEV_1$ be reported from the single tracing, using the largest sum of FVC plus $FEV_1$ (*best test*). We evaluated the results of 1,853 spirometric test sessions in 1,101 subjects (923 hospital patients and 178 normal volunteers). The mean difference between the 2 test selection methods cited above was 5.8 ml for FVC and 8.4 ml for $FEV_1$. In 98.4% of the FVC comparisons and 95.7% of the $FEV_1$ comparisons, the differences were within the minimal instrument accuracy standard (± 50 ml or ± 3% of the reading) suggested by the ATS. Differences between maximal and *best test* FVC and $FEV_1$ were small. The selection of values for interpretation from the *best test* did not compromise accuracy, and was a simpler and more practical method for reporting clinical spirometric results.


An ATS workshop held at Snowbird, Utah, made recommendations concerning the standardization of spirometry (1). Included in this statement was a procedure for the selection of the best spirometric values for clinical interpretation. The committee suggested that the maximal forced vital capacity (FVC) and forced expiratory volume in one second ($FEV_1$) be used for interpretation, regardless of whether they both came from the same tracing. In a previous publication, the ITS suggested that the spirometric values for interpretation should be taken from a single tracing, which was referred to as the *best test*, and was the tracing with the maximal sum of FVC plus $FEV_1$ (2).

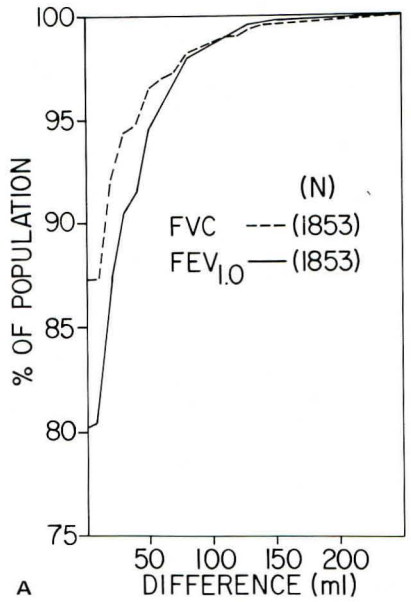
Goldman and Haley (3) found that *best test* selection in tests on 15 healthy nonsmokers allowed simple comparison of sequential tracings, and found no significant difference in daily variability when compared with other selection methods. The work of Nathan and co-workers (4) conducted on a randomly selected community population supported the ATS recommendation of doing only 3 trials (1). They evaluated methods using maximal values, but did not make comparisons with the *best test* criteria.

TABLE 1

FUNCTIONAL CLASSIFICATION OF SPIROMETRIC DATA USED FOR FVC AND $FEV_1$ COMPARISONS
(1,853 Test Sessions in 1,101 Adults)

| | | Bronchodilator | |
|---|---|---|---|
| | | Pre | Post |
| Hospitalized patients | | 923 | 752 |
| Obstruction | mild | 319 | 139 |
| | moderate | 209 | 166 |
| | severe | 147 | 137 |
| Restriction | | 82 | 106 |
| Normal | | 166 | 204 |
| Normal subjects | | 178 | — |

To determine the difference between the use of maximal and *best test* values under hospital pulmonary laboratory conditions, we analyzed data from a total of 1,853 spirometric test sessions on 1,101 adults (table 1); this included 923 hospital patients and 178 healthy nonsmoking volunteers in a normal population study. All of the hospital patients and healthy volunteers were tested




A


(*Received in original form May 15, 1980 and in revised form August 4, 1980*)


[1] Presented at the Western Society for Clinical Research, Carmel, California, February 9, 1980.

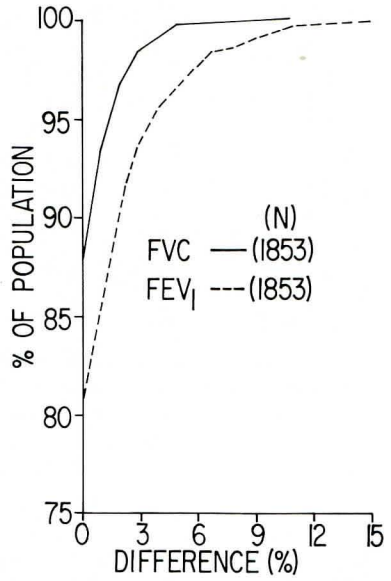
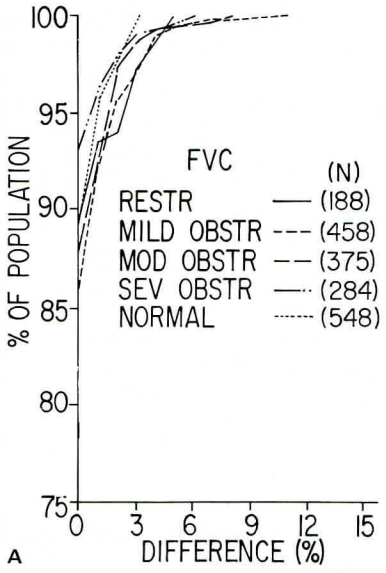
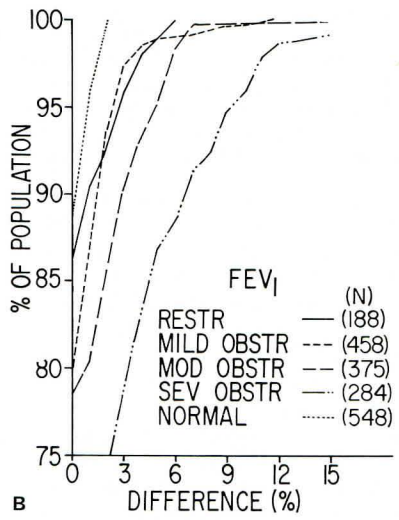



Fig. 1. A and B. Percentage of the population (*vertical axis*) with less than the difference (maximum—*best test*) expressed in ml (A) and as a % of maximum (B) (*horizontal axis*).

before the administration of a bronchodilator, and 752 of the hospital patients were also tested after. The patients were classified according to diagnostic categories (normal, obstructed, or restricted). Patients with airway obstruction were classified as mild, moderate, or severe, as designated by ITS diagnostic criteria (2). Patients with lung restriction (mild, moderate, or severe) were classed together because of the small number of them and the lack of significant differences among them.

We retrospectively analyzed patient data, which was collected in a hospital pulmonary laboratory. The data



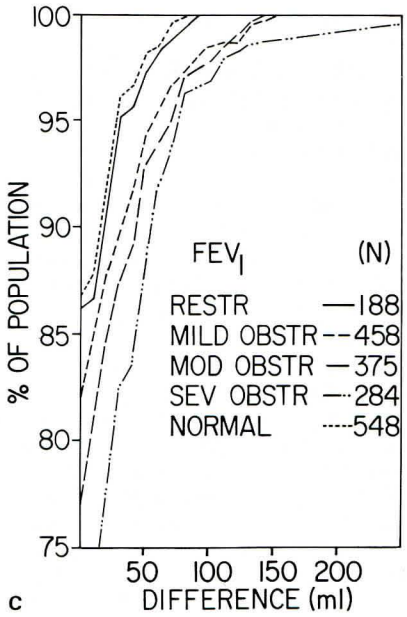




Fig. 2. A, B, and C. Percentage of the population (*vertical axis*) with less than the difference in % of the maximum (*horizontal axis*) for FVC (A) and for $FEV_1$ (B); in ml difference for $FEV_1$ (C) by patient group. Restr. = chest restriction; Obstr. = airflow obstruction (mild, mod = moderate, sev = severe); N = number of spirometric test sessions.

was collected over a 2-yr period by trained technicians who used computerized spirometric methods (5). They used the standard technique recommended by the ITS (2), and used a 13.5 L water seal spirometer (P1300; Warren E. Collins, Braintree, MA). Time zero was determined by the back extrapolation technique, and all methods and equipment conformed to ATS recommendations (1, 6). At least 3 reproducible spirograms were

obtained in each test session. The spirometric data were obtained from a retrospective analysis of computer records. For each set of data the differences between the maximal and the *best test* FVC and $FEV_1$ were expressed in milliliters (maximal value minus the *best test* value) or in percentages (difference in milliliters divided by the maximal value in milliliters multiplied by 100).

In 87% of the FVC comparisons and 80% of the $FEV_1$ comparisons, there were no differences between the maximal and *best test* results. In 68.5% of the comparisons, the maximal $FEV_1$ *and* FVC occurred in the same test; however, with the usual instrument/subject variability of ± 50 ml or ± 3% of the reading applied, we found that 94.3% of the comparisons resulted in no important differences in FVC or $FEV_1$. Ninety-seven per cent of the FVC comparisons and 95% of the $FEV_1$ comparisons had less than 50-ml difference (figure 1A). For both FVC and $FEV_1$, only 1% of the comparisons resulted in differences greater than 150 ml. Ninety-eight per cent of all FVC comparisons and 94% of all $FEV_1$ comparisons had less than a 3% difference, and only 1% of the comparisons had differences greater than 8% (figure 1B).

Very few differences in FVC were observed among the diagnostic categories (figure 2A). The percentage differences in $FEV_1$ were small, except for patients with severe airflow obstruction (figure 2B). In patients with severe obstruction, the absolute difference (in ml) was small, but it represented a larger fraction of the reduced $FEV_1$ value. Seventy-eight per cent of the $FEV_1$ comparisons in the severe obstruction category had less than 3% differences, whereas 89% had less than 50-ml differences (figure 2C).

For the 284 patients classified with severe obstruction, only 16 (5.6%) changed classification (by ITS criteria) if the maximal (Snowbird) criteria were used instead of *best test* criteria. Using the same methods, none of the 548 normal subjects changed classification. Because the classification criteria were somewhat arbitrary, and the variability of results in patients with severe obstruction was large, the fact that over 94% of the subjects were classified correctly, even the most ill one, showed that the simpler *best test* method is clinically applicable without important errors in classification.

The ATS Snowbird Workshop recommended a minimal spirometer accuracy of ± 50 ml or ± 3% of the reading, whichever was greater (1). The same accuracy requirement was also specified by the Division of Lung Diseases of the National Heart, Lung and Blood Institute Epidemiology Standardization Project (7). Because our study was on a hospital laboratory population, we make no recommendations for epidemiologic studies. Ninety-eight and four tenths per cent of the FVC differences and 95.7% of the $FEV_1$ differences were within these recommended limits. These limits were met more than 97% of the time in all patient categories for FVC, in all patient categories except moderate (93%) and severe (88%) airflow obstruction for $FEV_1$, and in 739 of 740 normal subjects (figure 3). The mean difference for all comparisons was 5.8 ml for FVC and 8.4 ml for $FEV_1$. The selection of FVC and

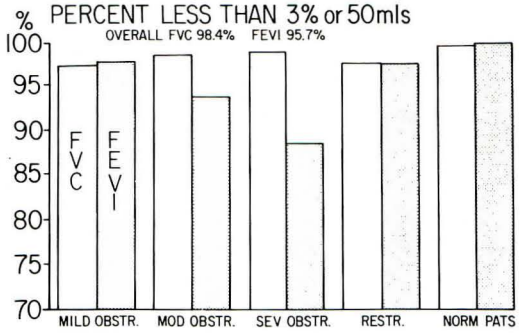



Fig. 3. A and B. Percentage of the population (*vertical axis*) with (maximum—*best test*) difference less than ± 50 ml *or* ± 3% of maximum by patient group. Same classification as in figure 2. Norm Pats = normal patients.

$FEV_1$ values for interpretation from the *best test* method presented no clinically significant deviation from the maximal values.

* * *

We concluded that both methods of selecting values would provide equally accurate and reproducible assessment of the patient; however, the *best test* method offered the advantage of evaluating a single curve that is internally consistent and allows easier comparison of sequential tracings. Because the ATS (1) recommended that other measurements, such as flow, $FEF_{25-75\%}$, and $FEF_{200-1200}$, be reported from the best tracing, reporting the FVC and $FEV_1$ from the same tracing is logical. Because most laboratories report the $FEF_{25-75\%}$, and thus need to determine the *best test*, it would be simpler and more practical to use the *best test* for all results. Reporting the spirometric results of the *best test* causes less confusion for the technician performing the test and for the physician interpreting the results. A single tracing representative of each test session would facilitate laboratory data reporting. The selection of FVC and $FEV_1$ values for interpretation from the *best test* does not significantly compromise accuracy and is a simple and practical method for reporting clinical spirometric results.

JOHN B. SORENSEN
ALAN H. MORRIS
ROBERT O. CRAPO
REED M. GARDNER[2]
*The Departments of Medical Biophysics and Computing and Medicine*
*LDS Hospital/University of Utah*
*Salt Lake City, Utah 84143.*

[2] Requests for reprints should be addressed to Reed M. Gardner, LDS Hospital, 325 Eighth Avenue, Salt Lake City, Utah 84143.

## References


1. Gardner RM, chairman. Snowbird workshop on standardization of spirometry. ATS statement. Am Rev Respir Dis 1979; 119:831–8.
2. Kanner RE, Morris AH. Clinical pulmonary function testing. Salt Lake City: Intermountain Thoracic Society, 1975.
3. Goldman AL, Haley JA. Influence of tracing selection technique on daily variation of $FEV_1$ and FVC (abstract). Am Rev Respir Dis 1979; 119: (No. 4, part 2) 218.
4. Nathan SP, Lebowitz MD, Knudson RJ. Spirometric testing: number of tests required and selection of data. Chest 1979; 76:384–8.
5. Dickman ML, Schmidt CD, Gardner RM, Marshall HW, Day WC, Warner HR. On-line computerized spirometry in 738 normal adults. Am Rev Respir Dis 1969; 106:780–90.
6. Smith AA, Gaensler EA. Timing of the forced expiratory volume in one second. Am Rev Respir Dis 1975; 112:882–5.
7. Ferris BG. Epidemiology standardization project. Recommended standardized procedures for pulmonary function testing. Am Rev Respir Dis 1978; 118: part 2.