Repeated Measures of FEV_1 Over Six to Twelve Months: What Change is Abnormal?

Mei-Lin Wang, MD, MPH
Edward L. Petsonk, MD

Monitoring change in FEV_1 (ΔFEV_1) is useful for assessing adverse respiratory effects in an individual, but high variability impedes reliable recognition of accelerated decline. The American Thoracic Society (ATS) recommends a $\geq$15% year-to-year FEV_1 decline for clinical significance. To evaluate the applicability of this criterion in health monitoring programs, we examined the mean, lower 5th percentile, and lower 5% cutoff value of ΔFEV_1 determined from 2 tests at 6- and 12-month intervals using data obtained with ATS-recommended equipment and procedures in 389 white male workers, each with 3 to 11 spirometry tests over 5 years. Results indicate that when healthy working males perform spirometry according to ATS standards, a yearly decline in FEV_1 greater than 8% or 330 mL should not be considered normal, whereas the 15% ATS criterion could be appropriate in clinical settings. (J Occup Environ Med. *2004;46:591–595*)

Spirometry has become one of the most useful methods for assessing lung health. Over the last 25 years, improved knowledge regarding equipment, methods, and interpretation has resulted in the development of professional standards, including a series of American Thoracic Society (ATS) statements, which have been widely disseminated and adopted in research, clinical testing, medical monitoring, and screening.[1–4] Monitoring of lung function change is recommended to detect air flow limitation in its early stage, when interventions are most likely to be successful.[5] However, the reliable detection of excessive longitudinal change in spirometry measurements is complicated by the presence of various sources of random and systematic variation.[6–9] These difficulties are particularly challenging when the follow-up duration is short. Results from longitudinal studies have indicated that even when professional standards and recommendations are closely followed, the observed individual rates of change still suffer from substantial variability, with short-term repeated measurements in particular.[10–12] This could preclude reliable determination of a subject's longitudinal change in FEV_1 without prolonged follow up.

The ATS 1991 Statement on interpretive strategies for spirometry recommended that a year-to-year change in FEV_1 of $\geq$15% be used to identify a clinically important change.[13] We sought to evaluate the applicability of this criterion in health monitoring programs and to further investigate the normal limits

From the Division of Respiratory Disease Studies, National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention, Morgantown, West Virginia.

Address correspondence to: Edward L. Petsonk, MD, FCCP, NIOSH, Mail Stop H-G900.2, 1095 Willowdale Road, Morgantown, WV 26505; E-mail address: elp2@cdc.gov.

DOI: 10.1097/01.jom.0000128159.09520.2a

of FEV_1 changes over 6- and 12-month intervals (ΔFEV_1) in a working population by analyzing an actual longitudinal spirometry dataset. The results from this study provide additional guidance for the interpretation of short-term changes of FEV_1 in occupational settings.

Methods

The study population and data collection methods have been previously reported[14]; they are briefly reviewed here. Study methods were approved by the Institutional Review Board, and all participants gave written informed consent.

Subjects and Health Measurements

Data were from our previous study on bronchial responsiveness and 5-year FEV_1 decline among a cohort of 411 underground coal miners and working nonminers. In the present study, we used data available from the 389 white male subjects. Standardized respiratory symptom questionnaires and methacholine testing were completed at the initial and final health survey over 5 years. Spirometry was performed at the worksite at 6-month intervals using an 8-L water-sealed survey spirometer with an attached microprocessor following ATS 1979 recommendations.[1] One hundred sixty subjects had completed 11 measurements, whereas the others had performed 3 to 10 tests. Initial values for FEV_1 and FVC were compared with published reference values.[15] Subjects were labeled responders if their FEV_1 dropped $\geq$15% from baseline after inhaling methacholine aerosol ($\leq$25 mg/mL). Individual 5-year FEV_1 slopes (mL/year) were calculated by linear regression using all the FEV_1 measurements available (from 3–11) for each study participant. In contrast to 5-year FEV_1 slope, we defined the short-term ΔFEV_1 as changes in FEV_1 between any 2 measurements over 6- and 12-month intervals.

Grouping by Health Conditions

In estimating the expected changes and variability in FEV_1 over 6 and 12 months, we grouped subjects with various health conditions as follows: 1) subjects who developed methacholine responsiveness over the 5-year study period (DHR); 2) subjects who developed any of 5 respiratory symptoms (cough, phlegm, dyspnea, persistent wheeze, attacks of wheeze with dyspnea) over the 5 years (DSX); and 3) subjects with accelerated long-term decline in lung function (rapid = 5-year FEV_1 slope in the lowest quartile, ie, equal to or more negative than -90 mL/year). For estimating the "normal" or expected change and variability in ΔFEV_1, we used 2 different reference groups: stable, comprising subjects who did not experience excessive long-term functional declines (ie, 5-year FEV_1 slope less negative than -90 mL/year); and healthy, comprising nonsmoking individuals who did not respond to methacholine at both initial and final survey, and had not developed any of the respiratory symptoms over the 5-year study period.

Estimate the Normal Limits of Short-Term ΔFEV_1

The change between any 2 FEV_1 measurements over both 6- and 12-month intervals was examined in 2 ways: the absolute change and the percentage change. The absolute change in FEV_1 was calculated by subtracting the first FEV_1 value (baseline) from the subsequent test value (at either a 6- or 12-month interval), expressed in milliliters.

For ΔFEV_1 over a 6-month interval, the calculation was:

$$\Delta FEV_{1(6\text{-month})} = FEV_{1(6\text{-month})} - FEV_{1(\text{baseline})};$$

and, for ΔFEV_1 over a 12-month interval, the calculation was:

$$\Delta FEV_{1(12\text{-month})} = FEV_{1(12\text{-month})} - FEV_{1(\text{baseline})}$$

The percentage change in FEV_1 was calculated as the absolute change divided by the baseline FEV_1 multiplied by 100 (ie, $\%\ \Delta FEV_{1\ (6\text{-month})} = (\Delta FEV_{1\ (6\text{-month})} / FEV_{1\ (\text{baseline})}) \times 100$ for a 6-month interval; and $\%\ \Delta FEV_{1\ (12\text{-month})} = (\Delta FEV_{1\ (12\text{-month})} / FEV_{1\ (\text{baseline})}) \times 100$ for a 12-month interval, respectively).

The lower limit of normal was defined in 2 ways, as the ΔFEV_1 value at the 5th percentile and as the lower 5% cutoff, assuming a 1-sided test (lower 5% limit = mean – 1.645 $\times$ SD). These 2 estimates should be very close to each other if the distribution of ΔFEV_1 is approximately normal.

We used Statistical Analysis Software (SAS® version 8.2, 1999–2001; SAS Institute Inc., Cary, NC) to test normality of the distribution for ΔFEV_1 and to compute the mean, standard deviation, and the values of ΔFEV_1 at the 95th, 75th, 50th, 25th, and 5th percentiles. We estimated the lower limit of normal for ΔFEV_1 based on the results from the previously described stable and healthy groups.

Validity by Other Datasets

Two other datasets were used for comparison. The first dataset was obtained from our previous study of asthma-like symptoms in wood product plant workers.[16] This dataset contains 2 years of spirometry measurements in 6-month intervals from 160 white male workers. The second dataset was collected during routine annual medical surveillance by the medical department of a steel company for 391 steel workers (5% female, 4% black) over a 3-year period from 1991 to 1994. For both datasets, spirometry testing was performed using a National Institutes for Occupational Safety and Health (NIOSH) dry rolling-seal spirometer and in conformance with then-current ATS recommendations. Comparison datasets included all workers without elimination of those reporting respiratory symptoms or conditions.

TABLE 1
Changes in FEV_1 by Health Condition

Health Condition	Subjects (no.)	6-Mo ΔFEV_1 (mL)		12-Mo ΔFEV_1 (mL)		5-Yr FEV_1 Slope (mL/yr)	
		Mean	SD	Mean	SD	Mean	SD
DHR	37	−41	173	−73	188	−74	38
DSX	65	−44	198	−73	203	−71	54
Rapid	97	−66	187	−116	198	−122	34
Stable	292	−24	169	−44	175	−36	44
Healthy	71	−28	167	−55	171	−56	45

SD, standard deviation.

TABLE 2
Estimates of "Normal Limits" for ΔFEV_1

	Subjects (no.)	Interval (months)	Observations (no. of ΔFEV_1s)	5th Percentile		Lower 5% Cutoff	
				(mL)	(%)	(mL)	(%)
All	389	6	2843	−320	−7.8	−321	−8.2
		12	2511	−350	−9.0	−363	−9.3
Stable	292	6	2157	−300	−7.1	−301	−7.8
		12	1911	**−330**	**−8.1**	**−331**	**−8.5**
Healthy	71	6	555	−280	−6.5	−302	−7.0
		12	507	**−320**	**−7.4**	**−335**	**−7.7**

Results

The study subjects were 389 white male underground miners (194) and nonmining blue collar workers (195). The median age at the initial survey was 37 years (range, 19–65 years), with approximately 28% current smokers and 28% exsmokers, averaging 11 pack-years at baseline. The mean FEV_1 was 4.15 L (99% predicted) at baseline, and at the final survey, the mean FEV_1 was 3.83 L (94% predicted). Methacholine responders at the initial and final survey comprised 27% and 30% of the total, respectively.

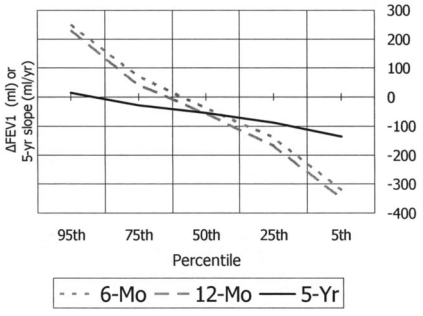

Fig. 1. Distribution of change in FEV_1: 6- and 12-month ΔFEV_1 versus 5-year slope.

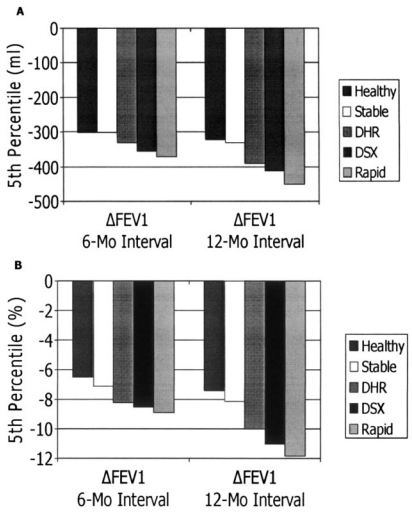

Fig. 2. The 5th percentiles of ΔFEV_1 in milliliters (**A**) and percentage (**B**) by health conditions. **HEALTHY**—Subjects who had never smoked, did not respond to methacholine at both initial and final survey, and had not developed any of the respiratory symptoms (cough, phlegm, dyspnea, persistent wheeze, attacks of wheeze with dyspnea) over 5 years, $n = 71$; **STABLE**—Subjects with 5-year FEV_1 slope less negative than -90 mL/year, $n = 292$; **DHR**—Subjects who developed methacholine responsiveness over 5-years, $n = 37$; **DSX**—Subjects who had new onset of any of the 5 respiratory symptoms (cough, phlegm, dyspnea, persistent wheeze, attacks of wheeze with dyspnea) over 5 years, $n = 65$; **RAPID**—Subjects with a 5-year FEV_1 slope equal to or more negative than -90 mL/year, $n = 97$.

The normality test showed that the distribution of ΔFEV_1 is approximately normal. Figure 1 illustrates that the ΔFEV_1 distributions for 12-month intervals are broadly similar to those for 6-month intervals, with a tendency toward higher overall mean declines. Although the medians are all similar, the greater range in ΔFEV_1 values for 6- and 12-month intervals, compared with the 5-year slope values, is obvious.

Figure 2 demonstrates the ΔFEV_1 5th percentile by health conditions. For all groups, the absolute values of the ΔFEV_1 5th percentile were slightly lower for 6-month than for 12-month intervals. The values for the 2 reference groups (healthy and stable) and for those who did not develop airway responsiveness or symptoms (data not shown) were all similar. In contrast, the values for those developing airway responsiveness (DHR) or symptoms (DSX) were generally closer to the rapid decliners (RAPID). Whether expressed as absolute change (Fig. 2A) or as percentage change (Fig. 2B), the relative magnitudes of the 5th percentile of ΔFEV_1 were similar for the various groups.

Table 1 shows the mean and the standard deviation of 6- and 12-month ΔFEV_1, and 5-year slopes, by health condition. Groups with adverse health conditions had greater mean declines in FEV_1 compared with the reference groups. Comparing the 12-month ΔFEV_1 with the 5-year slopes, the means were similar, whereas the standard deviations were more than 3 to 5 times greater. Table 2 summarizes the estimates of normal limits for short-term changes in FEV_1. Among the 2511 observations of ΔFEV_1 over a 12-month interval, only 0.6% was ≥15% and therefore met the ATS-recommended criterion as a clinically important change. In contrast, for the healthy and stable study subjects, "normal limits" of 12-month ΔFEV_1

TABLE 3
Distribution of 12-Month ΔFEV_1 for Current Study and Comparison Data Sets

	Subjects (no.)	Observations (no. of ΔFEV_1s)	ΔFEV_1 5th Percentile		ΔFEV_1 Lower 5% Cutoff	
			(mL)	(%)	(mL)	(%)
Wood product workers	160	506	−370	−8.6	−400	−9.5
Steel workers	391	468	−308	−9.2	−327	−9.2
Present study	389	2511	−350	−9.0	−363	−9.3

ranged from -320 to -335 mL and from -7.4 to -8.5%, respectively.

Comparisons with the 2 other datasets are shown in Table 3. The values of the 5th percentile and the lower 5% cutoff of 12-month changes in FEV_1 from these 2 datasets were quite close to those from the present study.

Discussion and Conclusion

Pulmonary function testing has evolved from a tool confined to the physiology research laboratory into one of the most useful health assessments for clinical diagnosis and management, as well as for the investigation of potential adverse effects of environmental and occupational exposures. The importance of objective measurements in the evaluation of individuals who are at risk for chronic obstructive pulmonary disease from smoking or exposure to other hazardous materials and environments has recently been emphasized.[5]

Periodic testing at 6- to 12-month intervals is often recommended for monitoring of lung health, but the interpretation of changes in spirometric indices over the short-term with only 2 measurements has been problematic.[17,18] To assure the appropriate selection of reference values for cross-sectional interpretation of spirometry, clinicians must not only take into account factors that are recognized to affect lung function such as age, height, gender, and ethnicity, but also the equipment, technician training, and the various technical considerations such as method of calculation, ambient and equipment temperatures, and instrument calibration. Selecting reference values for interpretation of longitudinal spirometry involves additional considerations in comparison to cross-sectional spirometry. The changes observed over time in an individual's spirometry results could be influenced by a number of technical and measurement factors, seasonal and diurnal variation, as well as short-term illnesses or chronic alterations in lung health. It is known that the variation between tests could be greater among individuals with lung disorders, particularly those associated with airway secretions and bronchospasm.[19–21]

Consistent approaches to techniques, procedures, and equipment for performing spirometry, as was attempted during the collection of the data analyzed in the current study, can improve the precision of results. Even with these precautions, unexplained systematic differences between health surveys (a survey effect) could be seen.[22] The longitudinal reference value selected should be based on data with a frequency of testing, an interval of observation, and a method of calculating change (eg, regression slopes vs. 2-point differences) similar to the procedures used for obtaining the value to be interpreted.[12] The current ATS statement on interpretation of spirometry recommends that a year-to-year change of 15% is required to be clinically meaningful. This recommendation appears to have been based on data collected, at least in part, before the establishment of the current professional recommendations for equipment, procedures, and training.[10] An alternative criterion recommended for determining a meaningful longitudinal FEV_1 decline is 85% of the baseline FEV_1 value minus the expected decline resulting from aging.[4] When spirometry results span multiple years, this approach could result in earlier identification of individuals with accelerated FEV_1 declines, compared with the ATS criterion, but it is less sensitive for year-to-year changes.

In clinical practice, a significant abnormality is often defined as a value that would be expected in less than 5% of healthy persons. However, previous studies have not reported the distribution of change in FEV_1 measured over 6- to 12-month intervals among generally healthy individuals. To address this issue, the current study was designed to offer guidance in the interpretation of individual longitudinal spirometry results by providing estimates of the variability in test results over the 6- and 12-month intervals commonly used when monitoring lung health. The short-term changes of FEV_1 (between any 2 measurements over 12 months) that we observed during a 5-year study in this working population suggest that the ATS recommendation requiring a $\geq$15% change from year-to-year could be too strict in some circumstances. In our study, only 0.6% of the 1-year intervals showed a $\geq$15% change, and our results suggest that a year-to-year decline of $>$8% or $>$330 mL is unlikely to occur by chance among healthy working males. (See bolded results for the stable and healthy groups in Table 2.)

To assure that these lower normal limits for year-to-year ΔFEV_1 were derived from a truly "normal" group, individuals who developed adverse health outcomes over the 5 years of observation (ie, respiratory symptoms, airway responsiveness, large long-term losses in FEV_1) were excluded from the reference groups. The validity of the findings is

strengthened by the similarity of the study data to spirometry results collected for both research and surveillance purposes in other working populations. However, even among workers in our study who met this >8% or >330 mL criterion, many did not show accelerated declines over the entire 5 years of follow up (data not shown), emphasizing that a finding of an increased year-to-year decline in an individual requires further assessment and confirmation.

We conclude that short-term repeated measurements of FEV_1 are associated with a variability that is large in comparison to the long-term FEV_1 slope (mL/year). The variability of ΔFEV_1 is similar for 6- and 12-month intervals. Individuals with health conditions tend to have both greater declines and greater variability in ΔFEV_1. The ≥15% threshold for recognizing clinically important FEV_1 changes recommended by the ATS could be appropriate in a clinical setting in light of the greater variation typically seen among patients with respiratory health conditions. However, when spirometry is performed to monitor lung health among working populations using procedures, equipment, and technician training that meet or exceed ATS recommendations, our study indicates that a year-to-year decline in FEV_1 greater than 8% or 330-mL decline should not be considered normal.

Acknowledgments

Supported by the National Institute for Occupational Safety and Health.

References

1. American Thoracic Society. Standardization of spirometry. *Am Rev Respir Dis*. 1979;119:4–11.
2. American Thoracic Society. Standardization of spirometry: 1987 update. ATS statement. *Am Rev Respir Dis*. 1987;136: 1285–1298.
3. American Thoracic Society. Standardization of spirometry: 1994 update. ATS statement. *Am J Respir Crit Care Med*. 1995;52:1107–1136.
4. Hankinson J, Wagner GR. Medical screening using periodic spirometry for detection of chronic lung disease. *Occup Med*. 1993;8:353–361.
5. NHLBI/WHO workshop summary. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2001;163:1256–1276.
6. Glindmeyer HW, Jones RN, Diem JE, et al. Useful and extraneous variability in longitudinal assessment of lung function. *Chest*. 1987;92:877–882.
7. Pennock BE, Rogers RM, McCaffree DR. Changes in measured spirometric indices: what is significant? *Chest*. 1981; 80:97–99.
8. Enright PL, Connett JE, Kanner RE, et al. Spirometry in the lung health study: II. Determinants of short-term intraindividual variability. *Am J Respir Crit Care Med*. 1995;151:406–411.
9. Buist AS, Vollmer WM. The use of lung function tests in identifying factors that affect lung growth and aging. *Stat Med*. 1988;7:11–18.
10. Burrows B, Lebowitz MD, Camilli AE, et al. Longitudinal changes in forced expiratory volume in one second in adults. *Am Rev Respir Dis*. 1986;133: 974–980.
11. Clement J, Van De Woestijne KP. Rapidly decreasing forced expiratory volume in one second or vital capacity and development of chronic airflow obstruction. *Am Rev Respir Dis*. 1982;125:553–558.
12. Wang ML, Gunel E, Petsonk EL. Design strategies for longitudinal spirometry studies: study duration and measurement frequency. *Am J Respir Crit Care Med*. 2000;162:2134–2138.
13. American Thoracic Society. Lung function testing: selection of references values and interpretative strategies. ATS statement. *Am Rev Respir Dis*. 1991;144: 1202–1218.
14. Hodgins P, Henneberger PK, Wang ML, et al. Bronchial responsiveness and five-year FEV_1 decline: a study in miners and nonminers. *Am J Respir Crit Care Med*. 1998;157:1390–1396.
15. Morris J, Koski A, Johnson L. Spirometric standards for healthy nonsmoking adults. *Am Rev Respir Dis*. 1971;103:57–67.
16. Petsonk EL, Wang ML, Lewis DM, et al. Asthma and asthma-like symptoms in wood product plant workers exposed to methylene diphenyl diisocyanate. *Chest*. 2000;118:1183–1193.
17. Pennock BE, Cottrell JJ, Rogers RM. Pulmonary function testing. What is 'normal'? *Arch Intern Med*. 1983;143:2123–2127.
18. Townsend MC. ACOEM position statement. Spirometry in the occupational setting. American College of Occupational and Environmental Medicine. *J Occup Environ Med*. 2000;42:228–245.
19. Camilli AE, Burrows B, Knudson RJ, et al. Longitudinal changes in forced expiratory volume in one second in adults: effect of smoking and smoking cessation. *Am Rev Respir Dis*. 1987;135:794–799.
20. Anthonisen NR, Wright EC, and the IPPB Trial Group. Bronchodilator response in chronic obstructive pulmonary disease. *Am Rev Respir Dis*. 1986;133: 814–819.
21. Nickerson BG, Lemen RJ, Gerdes CB, et al. Within-subject variability and per cent change for significance of spirometry in normal subjects and in patients with cystic fibrosis. *Am Rev Respir Dis*. 1980; 122:859–866.
22. Lebowitz MD. Age, period, and cohort effects: influences on differences between cross-sectional and longitudinal pulmonary function results. *Am J Respir Crit Care Med*. 1996;154:5273–5277.