Acceptability and Reproducibility Criteria of the American Thoracic Society as Observed in a Sample of the General Population[1,2]

JOHN L. HANKINSON and KI MOON BANG

SUMMARY An analysis of spirograms of 6,486 subjects from the general population, ages 8 to 90, was conducted to determine their ability to satisfy the American Thoracic Society's (ATS) acceptability and reproducibility criteria. The results indicate that both older and younger subjects had more difficulty satisfying the ATS acceptability and reproducibility criteria. The difficulty in satisfying the ATS reproducibility criterion, particularly in younger subjects, was in part associated with their smaller heights and lung volumes. A relatively uniform within-subject variability of FVC and FEV_1 in terms of the mean differences between the largest and second largest FVC and FEV_1, for all heights, was observed. In addition, unlike the ATS reproducibility criterion, when a constant 200-ml reproducibility criterion for FVC and FEV_1 was used, there was no longer a significant difference between the number of reproducibility criterion failures for the 14 different height groups used. These results suggest that the ATS reproducibility criterion, based on a percentage of the FVC and FEV_1, may inappropriately classify a higher percentage of subjects with smaller heights and lung volumes as having a nonreproducible test. In contrast, subjects with larger heights and lung volumes are much less likely to fail the ATS reproducibility requirement. These results emphasize the importance of following the ATS recommendation of using the reproducibility criterion only as a goal during data collection, not to classify a subject as having an invalid test.

AM REV RESPIR DIS 1991; 143:516-521

Introduction

In 1978 the National Institute for Occupational Safety and Health recommended to the Occupational Safety and Health Administration that certain criteria be established for the administration of pulmonary function examinations under provisions of the Cotton Dust Standard (1). These criteria included instrumentation and procedural and technician training requirements. The procedural recommendations included both acceptability and reproducibility criteria for the FVC and FEV_1. Specifically, the reproducibility criterion recommended that the largest FVC and FEV_1 and second largest FVC and FEV_1 be within 10% or 200 ml, whichever is greater. In 1979 the American Thoracic Society (ATS) published Spirometry Recommendations (2), which also included acceptability and reproducibility criteria. In 1987 the ATS published an update to their 1979 recommendation (3), which recommended a reproducibility criterion for the FVC and FEV_1 of 5% or 100 ml, whichever is greater. The ATS also recommended that at least three acceptable curves (no cough, excessively large extrapolated volume, and others) be obtained. One important change in the 1987 ATS update was the clarification that the reproducibility criterion was to be used only as a guide to whether more than three FVC maneuvers were needed. The reproducibility criterion was not to be used for excluding results from reports or for excluding subjects from a study. The rationale for this recommendation was based on several recent studies that have shown that the exclusion of subjects for poor reproducibility may inappropriately exclude subjects who may have a ventilatory abnormality (4-6).

The following analysis was conducted to investigate the ability of members of the general population to satisfactorily complete a spirometric examination, that is, their ability to satisfy the ATS reproducibility criterion as a goal during data collection.

Methods

Spirograms from 6,486 subjects were studied at 25 different locations as part of the National Health and Nutrition Examination Survey III (NHANES III). The NHANES III is the most recent in a series of studies designed to assess the health and nutrition status of adults and children in the United States through interviews and direct physical examinations. The sample design of the NHANES III is a stratified multistage probability sample of the U.S. population. Approximately 40,000 individuals, aged 2 months and older, were randomly selected to participate in the survey (7). The survey has been conducted by the National Center for Health Statistics since September 1988 and will continue approximately 6 yr at 88 locations across the United States. During the first 2 yr of the survey spirometry data have been collected for sample persons aged 8 yr and older in 25 sample areas.

Quality control of the spirometry data has been performed by the National Institute for Occupational Safety and Health (NIOSH), Morgantown, West Virginia, which serves as the quality control center. Each technician received at least 1 wk of formal training, satisfactorily completing a NIOSH-approved course on spirometry. Before the beginning of this study, four pilot studies were conducted (820 subjects; data not included in this analysis), during which the technicians practiced and received additional supervised instruction and monitoring.

Each subject attempted to perform at least five FVC maneuvers, with a goal of meeting the ATS acceptability and reproducibility criteria. For each FVC maneuver only the exhaled volume was accumulated in a dry-rolling seal spirometer. The displacement of the spirometer was measured using a digital shaft encoder and stored in digital memory using

(Received in original form May 18, 1990 and in revised form September 24, 1990)

[1] From the Division of Respiratory Disease Studies, National Institute for Occupational Safety and Health, Morgantown, West Virginia, and the Division of Health Examination Statistics, National Center for Health Statistics, Hyattsville, Maryland.

[2] Correspondence and requests for reprints should be addressed to John L. Hankinson, Ph.D., Division of Respiratory Disease Studies, National Institute for Occupational Safety and Health, 944 Chestnut Ridge Road, Morgantown, WV 26505.

a dedicated microprocessor. The spirometer system has a volume resolution of approximately 2.6 ml and a sampling interval of 10 ms and has been independently tested (8) and found acceptable with regard to ATS spirometry recommendations. The entire uncorrected volume-time curve (displacement from the digital shaft encoder), up to a maximum of 20 s, was saved on digital tape for each FVC maneuver performed by the subject.

Before each subject began his or her spirometric examination the technician explained and demonstrated the performance of the FVC maneuver. During the performance of the FVC maneuver real-time displays of the flow-volume and volume-time curves, as well as a 6-s exhalation indication, were provided for the technician to monitor the subject's performance. At the completion of each FVC maneuver the computer displayed, for all maneuvers, the flow-volume curves, the calculated parameters of FVC, FEV_1, peak flow, and expiratory time, and the percentage difference between each value of FVC, FEV_1, and peak flow and the corresponding largest value. The computer also determined whether the last curve was acceptable (no cough, excessive extrapolated volume, and late peak flow) and whether additional maneuvers were needed to meet the ATS acceptability and reproducibility criteria. Using these displays and results the technicians modified their coaching to obtain acceptable and reproducible results, usually obtaining a minimum of five and a maximum of eight maneuvers. The technicians were instructed to ensure that subjects produced the highest possible peak flows and that maximal exhalations continued for at least 6 s and until there was no appreciable change in volume (40 ml) for at least 2 s. For Spanish-speaking subjects a Spanish-speaking technician administered the test or an interpreter was provided. Unless there was a valid reason, nose clips were worn and the test was performed in the standing position. A note was made if these procedures could not be followed.

For all but the first few study sites the completed volume-time curves were uploaded at the end of a day's testing to a computer (VAX 11/780) located at quality control center. All spirograms were reviewed by a senior quality control technician, and appropriate follow-up instructions were provided to the technicians. In addition, the technicians were periodically monitored by a senior quality control technician who traveled to the field to observe and provide additional instructions. At the completion of each study at a particular location (25 total), a quality control report evaluating each technician's performance was generated and used to determine whether additional training or monitoring was warranted.

In accordance with ATS (2) guidelines all unacceptable maneuvers (with coughs or extrapolated volumes greater than 5% of the FVC) were excluded before the reproducibility calculations were performed. The technician could also classify curves as unacceptable, and these field classifications were reviewed by a senior technician at the quality control center. The largest FVC and FEV_1 were selected from acceptable maneuvers, regardless of whether they came from the same maneuver. All volumes were reported at BTPS. Subjects with fewer than two acceptable maneuvers (301 subjects) were excluded from further consideration in our reproducibility analysis, and unless otherwise noted all means and percentages were obtained excluding these 301 subjects (N = 6,185). All acceptable curves (no cough or large extrapolated volume) were used in the reproducibility analysis, even if they represented a submaximal effort (low peak flow). Mean differences between the largest FVC and second largest FVC (defined as ΔFVC) and mean differences between the largest FEV_1 and second largest FEV_1 (ΔFEV_1) were calculated for each height group using values of height within ± 5 cm of those shown in figures 3 and 4.

The ATS committee's rationale for using a percentage of FVC and FEV_1 for the reproducibility criterion is presumed to be based on the assumption that the within-subject variability is proportional to lung size. An analysis to determine whether the variability of FVC and FEV_1 (ΔFVC and ΔFEV_1) increases with increasing FVC and FEV_1 should resolve this issue. However, FVC and FEV_1 variabilities have been shown to increase with obstructive lung disease (9). To investigate the effect of lung size on reproducibility, therefore, comparisons were made versus height as a surrogate for lung size under the assumption that respiratory diseases should be independent of height.

The subject's ability to understand and perform the FVC maneuver can also influence the test reproducibility. This may occur in younger and older subjects, who as a group had a higher percentage of unacceptable curves and a higher percentage of abnormal FVC and FEV_1/FVC%. Therefore, a separate analysis was conducted on subjects between the ages of 18 and 55 to reduce these effects.

All statistical analyses were performed on an IBM PC 386 computer using the Statistical Analysis System from the SAS Institute (10). All comparisons of group means were conducted using *t* tests, and tests of independence were conducted using a chi-square test. To investigate other factors that may influence test reproducibility the SAS GLM procedure was used with continuous variables of age, height, percentage of predicted FVC and FEV_1/FVC%, and indicator variables of sex, race, and Hispanic origin. For classification of an abnormal pulmonary function test result each subject's observed FVC and FEV_1/FVC% was compared to his or her predicted values to determine if either were below the lower 95th percentile or lower limit of normal (LLN) of Knudson's (11) reference equations. Predicted FVC for black subjects were adjusted by multiplying the Knudson's value for white subjects by 0.85 (12). The relatively few subjects who were neither white nor black (48), or for whom no ethnic group code was assigned (78), were grouped into the nonwhite category. Subjects of Hispanic origin were grouped into the white group for the white versus nonwhite comparisons.

Results

Of the 6,486 subjects studied 301 (4.6%) subjects had fewer than two acceptable curves and were excluded from the reproducibility analysis, leaving 6,185 subjects. The mean height, number of subjects with two or more acceptable curves, and number of subjects with an FVC or FEV_1/FVC% less than the LLN for eight different age groups are listed in table 1. The subjects older than 55 had a greater tendency to fall below Knudson's (10) 95th percentile or the lower limit of normal for FVC and FEV_1/FVC%, and a slightly lower height.

The results by age group for several measures of the quality of the FVC maneuver are shown in figures 1 (males) and

TABLE 1

MEAN HEIGHT AND NUMBER OF SUBJECTS WITH TWO OR MORE ACCEPTABLE CURVES AND WITH FVC OR FEV_1/FVC% BELOW LOWER LIMIT OF NORMAL FOR TOTAL POPULATION EXAMINED BY AGE GROUPS

Age Groups	Total N	Height		≥ Two Acceptable Curves		FVC < LLN		FEV_1/FVC% < LLN	
		Mean	SEM	N	%	N	%	N	%
8–15	1,087	149.4	0.42	1,052	96.8	14	2.6	10	1.8
16–25	1,155	167.4	0.28	1,128	97.7	15	2.6	25	4.3
26–35	993	167.2	0.30	970	97.7	8	1.5	30	5.6
36–45	868	167.0	0.34	843	97.1	14	2.8	54	10.8
46–55	613	167.6	0.39	590	96.3	7	2.2	39	12.5
56–65	722	165.8	0.36	676	93.6	19	5.1	75	20.0
66–75	591	164.8	0.39	539	91.2	24	8.3	58	20.0
> 75	457	162.8	0.47	387	84.7	38	16.7	57	25.0
Total	6,486	163.6	0.15	6,185	95.4	139	4.1	348	10.3

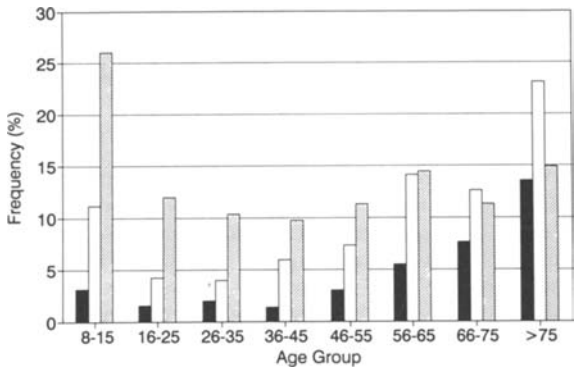

Fig. 1. Percentage of male subjects with fewer than three acceptable curves, nonreproducible test, and more than seven FVC maneuvers by age group (N = 3,119). Solid bars = < 3 acceptable; open bars = nonreproducible; hatched bars = > 7 curves.

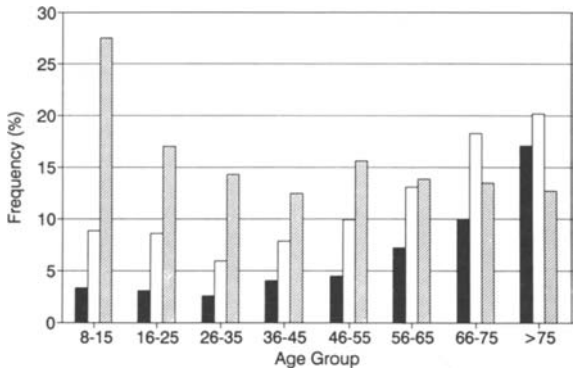

Fig. 2. Percentage of female subjects with fewer than three acceptable curves, nonreproducible test, and more than seven FVC maneuvers by age group (N = 3,367). Solid bars = < 3 acceptable; open bars = nonreproducible; hatched bars = > 7 curves.

2 (females). Both male and female older subjects had more difficulty meeting the ATS recommended reproducibility (largest and second largest FVC and FEV_1 within 5%) and acceptability criteria (at least three acceptable maneuvers). However, it appears that the technician appropriately responded to the lack of a reproducible or acceptable test result by obtaining more maneuvers from these subjects, particularly in the younger subjects. Note the higher percentage of the younger subjects with more than seven FVC maneuvers. During the data collection the technicians reported more difficulty explaining the test procedure to the older subjects and the necessity of obtaining additional maneuvers. For subjects older than 65 yr 55 (45.1%) of those with fewer than three acceptable trials simply declined to perform more than two maneuvers.

TABLE 2

MEAN DIFFERENCES BETWEEN LARGEST FVC AND SECOND LARGEST FVC AND LARGEST FEV_1 AND SECOND LARGEST FEV_1, NUMBER WITH FEWER THAN THREE ACCEPTABLE CURVES, AND NUMBER FAILING ATS REPRODUCIBILITY CRITERION*

Age	Total Number	Study Number	ΔFVC (*ml*) Mean	ΔFVC (*ml*) SEM	ΔFEV_1 (*ml*) Mean	ΔFEV_1 (*ml*) SEM	< Three Acceptable Curves N	< Three Acceptable Curves %	Non-reproducible N	Non-reproducible %
8–90										
Males	3,119	3,057	74.8*	1.78	64.5*	1.62	122*	3.9	287†	9.2
Females	3,367	3,274	65.6	1.65	56.3	1.58	179	5.3	348	10.3
18–55										
Males	1,558	1,542	69.9*	2.23	63.8*	2.33	13†	0.8	81*	5.2
Females	1,801	1,766	61.7	2.05	55.0	2.12	27	1.5	140	7.9
8–90										
White	4,593	4,483	76.9*	2.43	65.0*	2.20	221†	4.8	378*	8.2
Nonwhite	1,893	1,848	67.2	1.38	58.3	1.31	80	4.2	257	13.6
18–55										
White	2,297	2,263	59.9*	1.53	54.8*	1.71	59†	2.6	102*	4.4
Nonwhite	1,062	1,045	77.6	3.43	68.4	3.29	32	3.0	119	11.2
8–90										
Hispanic	2,147	2,074	68.4†	2.05	64.4*	2.28	136*	6.3	193†	9.0
Non-Hispanic	4,339	4,257	70.8	1.50	58.2	2.29	165	3.8	442	10.2
18–55										
Hispanic	1,170	1,141	64.3†	2.58	59.8†	2.95	50*	4.3	71†	6.1
Non-Hispanic	2,189	2,167	66.1	1.87	58.7	1.82	41	1.9	150	6.8

* Males and females, whites and nonwhites, and subjects of Hispanic and non-Hispanic origin. Statistically significant different between categories, $p < 0.01$.

† Not statistically significant different between categories, $p > 0.05$.

The mean ΔFVC and ΔFEV_1 and the standard error of the mean (SEM) for males and females, subjects of Hispanic origin and non-Hispanic origin, and white and nonwhite subjects are listed in table 2. Overall a slightly higher percentage of female subjects failed to meet the ATS recommended acceptability criterion than males, and males showed a significantly larger mean difference (ΔFVC and ΔFEV_1).

The mean ΔFVC and ΔFEV_1 and SEM for 14 (males) and 13 (females) in different height groups are shown in figures 3 (males) and 4 (females). From these figures it appears that the mean ΔFVC and ΔFEV_1 are approximately uniform for all heights. In addition, a uniform or slightly decreasing mean ΔFVC and ΔFEV_1 were observed with respect to FVC and FEV_1 (not shown). These findings were unexpected since the ATS reproducibility criterion for volumes greater than 2 L is 5% of the FVC or FEV_1, indicating an expectation that these differences (measure of reproducibility) would increase with increasing lung size or its surrogate height. The dashed lines in figures 3 and 4 represent the ATS reproducibility limits for FEV_1 based on 5% of the mean FEV_1 for the respective height groups. The actual mean ΔFVC and ΔFEV_1 are lower than these limits and do not increase with increasing height. Therefore taller subjects are less likely to exceed these ATS reproducibility limits than shorter subjects.

The percentage of subjects who would fail to meet the ATS reproducibility criterion (applied to both FVC and FEV_1) as a function of their height group is shown in figure 5. Note that as height increases the number of subjects failing to meet the ATS reproducibility criterion decreases. This decreasing trend is observed even when the analysis is limited to subjects between the ages of 18 and 55 (dashed lines in figure 5). This decreasing trend is not explained by a lack of trials, as a higher percentage of both male and female shorter subjects had more than seven maneuvers. For example, 15%

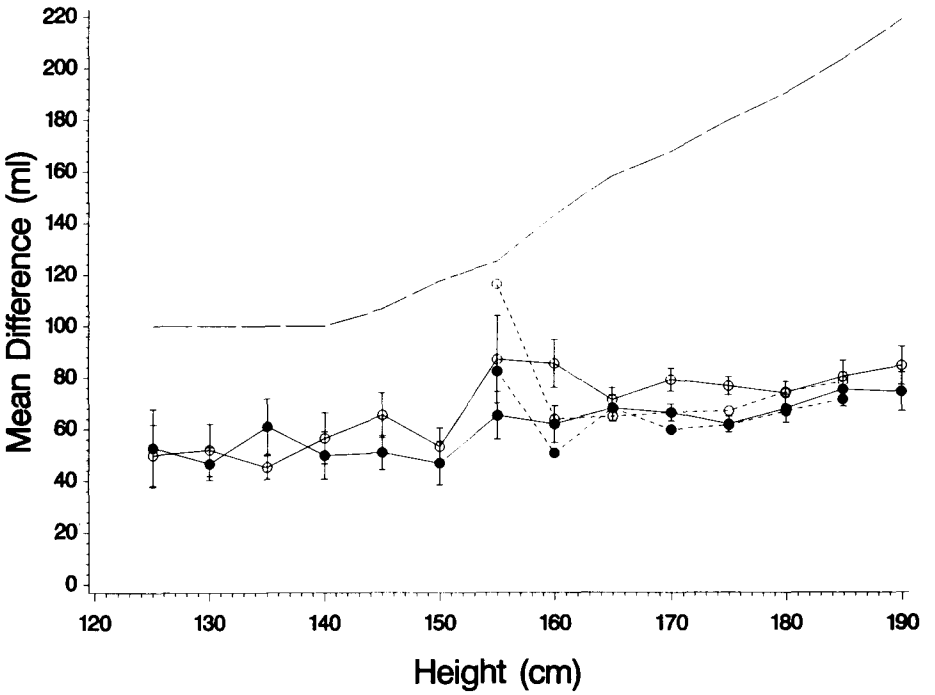

Fig. 3. For males the mean ΔFVC (ml; *open circles*) and mean ΔFEV_1 (*filled circles*) by height, ages 8 to 90 yr. Dashed lines indicate the results for the adult age group (ages 18 to 55). Long dashed line indicates the ATS reproducibility criterion limit of 5% of the FEV_1 or 100 ml, whichever is greater. The ATS reproducibility limit for FVC is not shown. High-low bars indicate ± 1 SEM.

(12 of 80) of the adult male subjects with a height of less than 162.5 cm had more than seven maneuvers, but only 6.4% (13 of 203) of male adult subjects with a height of greater than 182.5 cm had more than seven maneuvers.

When a constant-volume reproducibility criterion of 200 ml (based on the 95th confidence interval calculations of the ΔFVC and ΔFEV_1) is applied to FVC and FEV_1, there appears to be less of a relationship between the percentage of subjects with a nonreproducible test and height (figure 6). Shorter subjects fail to meet the 200-ml reproducibility criterion at about the same rate or less than the taller subjects.

A separate analysis was conducted using a linear model of the ΔFVC and ΔFEV_1 as a function of the continuous variables of age, height, percentage of predicted FVC, percentage of predicted $FEV_1/FVC\%$, and the indicator variables of sex, race, and Hispanic origin. Of all these parameters only the relationship between height and ΔFVC and ΔFEV_1 was not statistically significant. This linear model analysis was repeated leaving the percentage of predicted values (FVC and $FEV_1/FVC\%$) out of the equation and the same results were obtained.

When the linear model analysis was conducted using only the ΔFVC and ΔFEV_1 for the adults between the ages of 18 to 55, the relationship between height and ΔFVC and ΔFEV_1 was not statistically significant. In addition, the relationship between age and ΔFEV_1 was no longer statistically significant but was for the ΔFVC ($p < 0.05$). These linear model results appear to be consistent with the results shown in table 2 and figures 3 and 4.

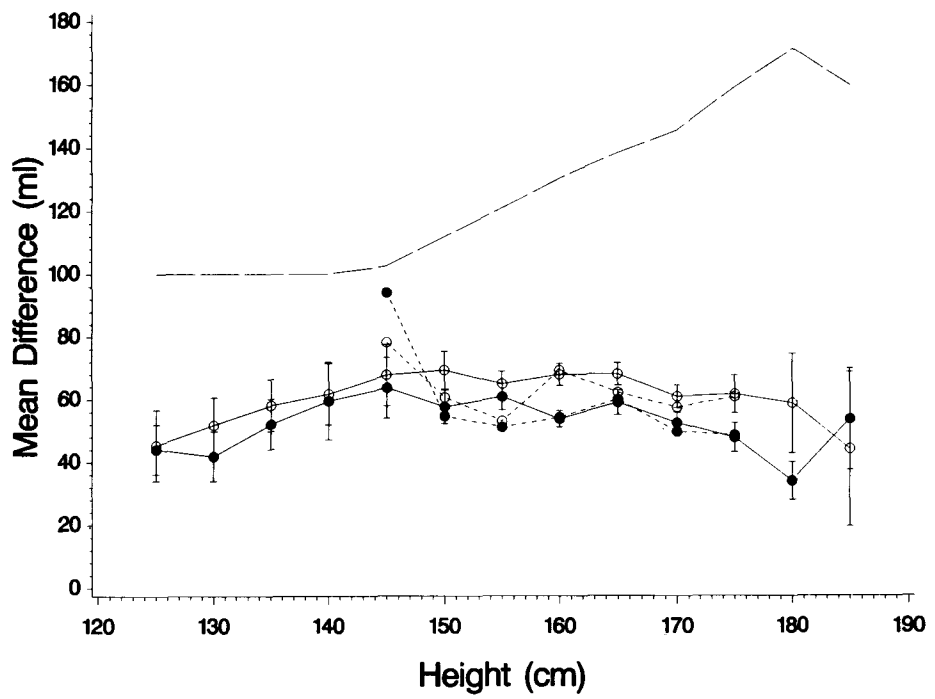

Fig. 4. For females the mean ΔFVC (ml, *open circles*) and mean ΔFEV_1 (*filled circles*) by height, ages 8 to 90 yr. Dashed lines indicate the results for the adult age group (ages 18 to 55). Long dashed line indicates the ATS reproducibility criterion limit of 5% of the FEV_1 or 100 ml, whichever is greater. The ATS reproducibility limit for FVC is not shown. High-low bars indicate ± 1 SEM.

Discussion

Although our analyses were based on preliminary data of 6,486 subjects, these results indicate that younger and especially older subjects have more difficulty meeting the ATS acceptability and reproducibility criteria. At least a portion of the failure to meet the ATS reproducibility criterion may be due to the smaller heights and lung volumes observed in these subjects, particularly the younger subjects. Factors other than ability to understand the test procedure may be related to test reproducibility. Specifically, there appears to be a difference in test variability between males and females and whites and nonwhites in the absence of a difference in the percentage of subjects with fewer than three acceptable curves (table 2). In contrast to these groups, subjects of Hispanic origin had a higher percentage of fewer than three acceptable curves, but test variability (ΔFVC and ΔFEV_1) was not significantly different.

A particularly important finding of this study is the relatively uniform within-

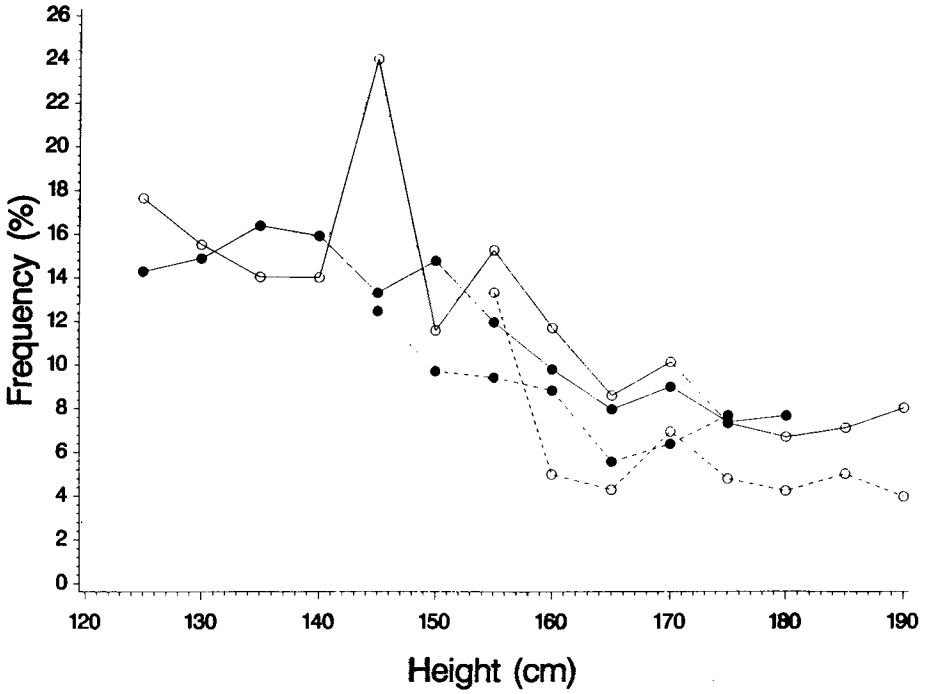

Fig. 5. Percentage of subjects failing to meet the ATS reproducibility (applied to FVC and FEV_1) criterion by height, ages 8 to 90 yr. Open circles indicate males, closed circles indicate females, and dashed lines represent the results for the adult age group (ages 18 to 55).

subject variability of FVC and FEV_1 in terms of the mean ΔFVC and ΔFEV_1 versus height (figures 3 and 4) and lung volumes. These results suggest that the ATS reproducibility criterion based on a percentage of the FVC and FEV_1 and a constant 100 ml for volumes less than 2 L may inappropriately classify a higher percentage of subjects with smaller heights and consequently smaller lung volumes as having a nonreproducible test. In contrast, taller subjects with larger lung volumes are much less likely to fail the ATS reproducibility requirement.

For comparison we tested a 200-ml reproducibility criterion (largest FVC and second largest FVC within 200 ml and largest FEV_1 and second largest FEV_1 within 200 ml). When this 200-ml reproducibility criterion is used the percentage of subjects failing to meet the reproducibility criterion does not appear to be related to height (figure 6), whereas this relationship was observed with the current ATS recommended reproducibility criterion (figure 5).

That the ATS reproducibility criterion was used in the collection of these data does not explain these results, as application of the ATS reproducibility criterion would tend to diminish the effects seen in figure 5. Since taller subjects with larger lung volumes are less likely to receive additional coaching because of their failure to meet the ATS criterion, one would expect the mean differences to be greater in these subjects, not less as was observed. Likewise, the shorter subject with smaller volumes received more coaching and therefore should have a smaller difference. We observed approximately the same mean ΔFVC and ΔFEV_1 for all heights or volumes, despite any coaching effect.

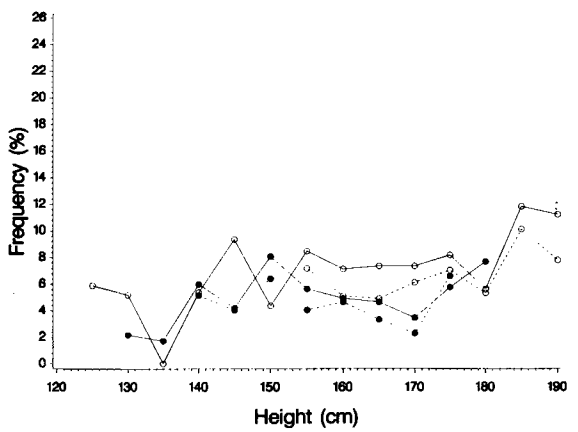

Fig. 6. Percentage of subjects failing to meet a 200-ml reproducibility criterion by height, ages 8 to 90 yr. Open circles males, closed circles females, and dashed lines represent the results for the adult age group (ages 18 to 55).

One possible explanation for the uniformity in the variability of FVC and FEV_1 is that subjects with smaller heights (primarily younger subjects) have more difficulty understanding and performing the test and therefore have a higher variability, even if they performed more maneuvers than the adults. To investigate this possibility an analysis of adults between the ages of 18 and 55 was conducted and produced results similar to those for the total population. Also, in the results using the linear model height was not significantly related to the ΔFVC or ΔFEV_1 for either the total population or the adult group.

Several studies have questioned the elimination of subjects for failure to meet a reproducibility criterion because such practice may introduce a population bias (4–6). Subjects with lung disease, particularly obstructive lung disease, have a greater test variability and may be preferentially eliminated. Our analysis using a linear model supported this observation, in that the FVC and FEV_1 variabilities were found to be related to the percentage of predicted FVC and the percentage of predicted FEV_1/FVC%. In addition, our results suggest that elimination of subjects for failure to meet a reproducibility criterion may also introduce a population bias with respect to subjects with smaller heights and lung volumes. Since subjects with moderate or severe obstructive lung disease tend to have lower FEV_1, it is possible that the preferential elimination of these subjects may be due in part to their lower FEV_1 combined with the ATS reproducibility criterion, rather than entirely to the larger within-subject variability of FVC and FEV_1 associated with disease. In our study a higher percentage of the subjects older than 55 fell below Knudson's lower limit of normal for FVC and FEV_1/FVC%, which could partially explain the relatively high rate of failure to meet the ATS reproducibility criterion in older subjects.

One natural consequence of a uniform ΔFVC and ΔFEV_1 expressed in milliliters is that the ΔFVC and ΔFEV_1, expressed as a percentage of FVC or FEV_1, would increase with decreasing FVC and FEV_1.

This means that subjects with smaller volumes (e.g., FVC and FEV_1) will have a much higher variability expressed as a percentage than subjects with larger volumes.

Although the ATS does not recommend excluding subjects for failure to meet their reproducibility criterion, it does recommend that reproducibility be considered in the interpretation of results. Therefore the impact of an inappropriate use of the reproducibility criterion may have consequences greater than simply having the subject perform additional unnecessary FVC maneuvers. A 200-ml reproducibility criterion may provide a commensurable level of difficulty for all subjects, regardless of age or lung volume. The practice of classifying a patient as having an invalid test based on failure to meet the ATS reproducibility criterion may nevertheless place shorter subjects with smaller lung volumes at a disadvantage.

References

1. Federal Register 1980; 45(42):13695–8.
2. Gardner RM (chairman). ATS statement on standardization of spirometry. Am Rev Respir Dis 1979; 119:831–8.
3. American Thoracic Society. Standardization of spirometry – 1987 update. Am Rev Respir Dis 1987; 136:1285–98.
4. Eisen EA, Robins JM, Greaves IA, Wegman DH. Selection effects of repeatability criteria applied to lung spirometry. Am J Epidemiol 1984; 120:734.
5. Eisen EA, Oliver LC, Christiani DC, Robins JM, Wegman DH. Effects of spirometry standards in two occupational cohorts. Am Rev Respir Dis 1985; 132:120.
6. Kellie SE, Attfield MD, Hankinson JL, Castellan RM. The ATS spirometry variability criteria: associations with morbidity and mortality in an occupational exposed cohort of coal miners. Am J Epidemiol 1987; 125:437–44.
7. National Center for Health Statistics. Third national health and nutrition examination survey (NHANES III); final OMB clearance package, 0920–0237. January 1990.
8. Nelson SB, Gardner RM, Crapo RO, Jensen RL. Performance evaluation of contemporary spirometers. Chest 1990; 97:288–97.
9. Pennock BE, Rogers RM, McCaffree DR. Changes in measured spirometric indices – what is significant? Chest 1981; 80:97.
10. SAS Institute Inc. SAS/STAT guide for personal computers, Version 6 ed. Cary, NC 1987.
11. Knudson RJ, Lebowitz MD, Holbert CT, *et al.* Changes in the normal maximal expiratory flow-volume curve with growth and aging. Am Rev Respir Dis 1983; 127:725–34.
12. Hankinson JL. Pulmonary function testing in the screening of workers: guidelines for instrumentation, performance, and interpretation. J Occup Med 1986; 28:1081–92.