4

Standardization of Spirometry with Special Emphasis in Field Testing

REED M. GARDNER

University of Utah and LDS Hospital
Salt Lake City, Utah

HENRY W. GLINDMEYER III

Tulane University School of Medicine
New Orleans, Louisiana

JOHN L. HANKINSON

National Institute of Occupational Safety and Health
Morgantown, West Virginia

I. Introduction

The hazardous effect of some airborne pollutants on the respiratory system have resulted in increased interest in the measurements of lung function. The hazardous effects of dust from coal mines [1], hard rock mines [2], and the cotton industry [3] have been recognized and have resulted in an improved working environment in these industries. In addition, chemical vapors and gases, as well as particulates from other industries [4,5], have also been shown to adversely affect the lung. Determining dose-response relationships of these occupational stimuli requires an accurate medical and employment history as well as environmental characterization and quantitative pulmonary function testing. To be most useful, these pulmonary function tests must be accurate, precise (repeatable), sensitive to the environmental stimuli, and readily applicable at the occupational site. This chapter will focus on the standardization of pulmonary function tests and define variables which affect their accuracy and utility in the study of occupational medicine.

The usefulness of pulmonary function testing to quantify the hazardous effects of occupational pollutants have been limited by several factors.

Table 1 Pulmonary Function Test

	Factors influencing test			
Parameter	Pulmonary	Nonpulmonary	Advantages	Disadvantages
Spirometry				
FVC	Elastic recoil of lung Caliber of larger and smaller airways	Poor initial effort Insufficient recording time Poor sustained effort Leaks Calibration errors Calculation errors	Minimal equipment Repeatability Well-documented	Effort dependent, especially near TLC and RV
FEV_1	Same as FVC	Same as FVC Time zero determinations	Same as FVC	Effort dependent, especially near TLC
FEF25–75	Same as FEV_1 but measures the effect of smaller airways	Same as FVC	Same as FEV_1 but not as repeatable Sensitive to smaller airways than the FEV_1 [9–11]	Same as FVC since FVC determines location of 25% and 75% points
$\dot{V}$%FVC [29]	Same as FEF25–75 and sensitive to smaller airways when flow is measured at smaller volumes	Same as FVC	Same as FEF25–75 but not as repeatable	Same as FEF25–75 Necessary to display flow-volume loop

MVV	Same as FVC Air trapping	Poor sustained effort Frequency of breathing not standardized Leaks Calculation errors Calibration errors	Minimal equipment	Effort dependent Hard on patient
He-O_2 Flow-volume maneuver [32,33]				
$\dot{V}_{iso}$%FVC	Same as FEF25–75 but sensitive to smaller airways Affected by distribution of ventilation	Same as FVC Dependent on number of breaths of He-O_2	Minimal equipment	Effort dependent Not very repeatable
$V_{iso}\dot{V}$ [34]	Same as $\dot{V}$%FVC and sensitive to smaller airways	Same as $\dot{V}_{iso}$%FVC	Same as $\dot{V}_{iso}$%FVC	Same as $\dot{V}_{iso}$%FVC
FRC	Elastic recoil of total respiratory system	Patient anxiety affects level of FRC Leaks Calibration errors Calculation errors	Fairly repeatable	Different methods sometimes yield different results (i.e., washout, dilution) Tidal breathing often affected by subject anxiety

Table 1 (Continued)

Parameter	Factors influencing test: Pulmonary	Factors influencing test: Nonpulmonary	Advantages	Disadvantages
				Test time consuming (test can last 15 min before equilibration or washout occurs
RV	Air trapping	Poor effort Leaks Calibration errors Calculation errors	N_2 washout technique is quick and repeatable (patient can breathe at abnormally high frequency and volume)	Effort dependent in reaching RV when test is initiated
Closing volume [34,35]				
CC, CV, ΔN_2, slope of phase III	Distribution of ventilation	Same as FVC Sensitive to flow rate Sensitive to volume history	Looks at small airways	Hard to measure Effort dependent Often difficult to interpret
Diffusion capacity [36]				
$DL_{CO_{sb}}$	$\dot{V}/\dot{Q}$ ratio of lung Total lung capacity Affected by a change of cross-sectional area or thickness of air-blood interface	Same as FVC	Easier than steady-state method but not comparable	Not specific Expensive and sophisticated equipment that requires sophisticated and frequent calibration

	Can be affected by decrease of available hemoglobin			
Plethysmography [37,38]				
TGV	Not influenced by air trapping Same as FRC	Same as FRC and RV; however, box heightens anxiety, and this sophisticated system needs relatively sophisticated calibration methods	Test takes only a few seconds; however, calibration and set-up can be time consuming	Calibration difficult High anxiety of patient yields abnormal patient FRC level Expensive equipment
Raw [39–41]	Caliber of larger airways	Same as TGV Dependent on panting frequency Dependent on lung volume	Same as TGV	Not very repeatable Same as TGV Very difficult maneuver for patient Forced oscillation more repeatable [41]

1. Choice of test: There are a wide variety of tests available (see Table 1). Deciding which test to use and interpreting the results of the test presents a major methodological problem. Most pulmonary function tests are relatively nonspecific and they are dependent on many uncontrolled variables which affect not only population studies but even results on the same subject on repeated tests. The scientific community is still looking for an ideal test which will provide early detection and measurement of degradation of pulmonary function. At the present time, however, the forced expiratory spirogram appears to be the best single test available.

2. Instrument standardization: The literature is replete with information which is of marginal value because of inadequate instrumentation. Contradictory results obtained from similar studies are often caused by systematic differences between instruments.

3. Instrument calibration: Quality control to ensure instrument stability and accuracy is necessary in any research study but it becomes especially important in longitudinal studies which may continue for several years. Because of varying and often unfavorable conditions that prevail during testing outside the usual "laboratory" environment, careful and frequent calibration tests must be performed in the occupational setting.

4. Standardization of test procedure: Few lung function tests have been adequately standardized [6,7]. There is still much to be done in the standardization of testing procedures so that results can be interpreted and verified from similar testing done by different investigators in different locations. Utilization of a different testing procedure in itself may cause differences in results which could incorrectly be attributed to occupational exposure. Because of the transient situation and the demanding conditions usually encountered in field testing, technicians performing the tests must have special training.

5. Uniform measurement and computation methods: Even if tests are performed with adequate instruments and accepted procedures, the tests can be quickly invalidated by inadequate measurement and computational methods. In addition, the computerized options available with many instruments often produce erroneous results.

6. Data interpretation: Interpretation of pulmonary function results requires consideration of many factors such as smoking history, sex, height, and race, as well as occupational factors. Care must be used when integrating results from several pulmonary function tests, because when a greater number of tests is used, the probability that all test results will be normal decreases rapidly. This observation is especially important since some subjects never learn to perform the pulmonary tests adequately, and unfortunately most of the tests of lung function are dependent on subject cooperation and effort.

7. Special problems in field testing: Environmental factors such as temperature, electrical power source, and available space are usually suboptimal and

must be compensated for. Efficient test scheduling in occupational studies requires reliable instrumentation and competent technicians to minimize wasted time for the test subjects. In order to maintain cooperation of the study group and their management, it is important to develop efficient, effective testing procedures and perform them at a predetermined time interval.

Because the factors listed above are limiting research progress, and as a consequence are limiting our understanding of occupational lung disease, recommendations and standards are beginning to be made in each of these problem areas. Since spirometry is the most widely used, and one of the most practical and specific tests of lung function, the majority of this chapter will be devoted to spirometric testing.

II. Available Tests and What They Measure

One of the most difficult aspects of conducting occupational studies is deciding which test to employ. To evaluate the available tests, one must consider (1) what they measure (pulmonary versus nonpulmonary factors); and (2) advantages and disadvantages of the test, which includes factors such as equipment costs, testing time, and repeatability of results. Table 1 lists these factors for spirometry, the helium-oxygen flow-volume maneuver, lung volumes, closing volumes, diffusing capacity, and body plethysmography. Chapter 5 provides a review of the various pulmonary function tests [8]. In general, the more sensitive tests are less repeatable and require more sophisticated instrumentation. Spirometry, the least expensive in terms of equipment, requires minimal testing time, and some of the parameters measured from the forced vital capacity (FVC) maneuver are sensitive to small airway disease [9,10]. These parameters include the volume of air exhaled during the first second of the FVC maneuver (FEV_1), the FEV_1/FVC ratio, the average flow rate measured during the middle half of the FVC (FEF_{25-75}), and the instantaneous flow rates measured at specific lung volumes ($\dot{V}_{\%FVC}$). Of these parameters, the FEV_1 is the least sensitive to small airway abnormalities, but is the most repeatable [11].

The unique spirographic patterns produced by normal subjects and subjects with different types of pulmonary pathology make spirometry useful in discriminating between obstructive and restrictive lung disease. Table 2 describes the parameters measured from a forced vital capacity for subjects with obstructive and restrictive disease. Five of the six parameters discriminate between these defects; only the FEV_1 is substantially reduced in both. If the parameters of flow are based on total lung capacity instead of the forced vital capacity, they too would be substantially reduced in both diseases.

Table 2 Forced Spirographic Patterns

Parameter	Obstructive	Restrictive
FVC	Mild reduction with prolonged expiration	Reduced
FEV_t	Reduced	Reduced
FEV_1/FVC	Reduced	Normal or increased
FEF_{25-75}	Reduced	Near normal
FEF_{75-85}	Reduced	Near normal
$\dot{V}$%FVC	Reduced	Near normal

III. Standardization of Instrumentation

The value of spirometry as a test for following a worker's lung function over time is dependent upon standardization of spirometers, test procedures, and measurement techniques. Since the introduction of the FVC maneuver in 1947, instruments which were originally designed to record slow volume changes were being used to record dynamic events. Over the years these instruments have become more sophisticated and a wide variety of designs have been introduced. Presently, there are more than 100 different types of spirometers on the market. These instruments can be divided into two basic categories depending on whether they measure volume directly or detect flow indirectly. The designs for volume-measuring instruments include water-sealed, dry rolling-sealed; and bellows-type spirometers. The water-sealed spirometer consists of a bell seated in an annulus of water. As the subject blows into the spirometer, the bell rises in order to accommodate the volume blown into it. The bell is counterweighted or made of plastic in order to reduce the effect of gravity. The dry rolling-sealed spirometer consists of a canister within a canister sealed by a Teflon membrane. As the subject blows into the spirometer, the Teflon membrane rolls and the spirometer canister separates in order to accommodate the added volume. This design reduces the effect of gravity since the displacement of the piston is in the horizontal plane. The bellows-type spirometer also moves in a horizontal plane, thereby reducing the effect of gravity. For all three designs of volumetric spirometers, the simple expansion of the system as air is introduced into the spirometer is related to volume, and can be displayed as such on a recorder.

The designs for flow-detecting devices include the pneumotachygraph, hot wire, and turbine. The pneumotachygraph is a pressure sensor. Differences in pressure occur with changes in flow and can be related to changes in volume. With the hot wire, a wire is heated to several hundred degrees Fahrenheit. As air is blown across the wire, it is cooled, and this change in temperature changes the resistance of the wire. The resistance is then monitored electronically and

related to flow. The final design for flow-detecting devices is the turbine, which simply spins as air passes across it. The faster the rotation of the turbine, the greater the flow of air.

In addition to sensing devices, a spirometer should have a display of volume versus time or flow versus volume during the entire forced expiration. All six spirometer designs discussed previously can be linked to a recorder so that a graphical output can be obtained. Recorders commonly used with spirometers include kymographs, chart plates, XYT recorders, and oscilloscopes. These recorders can sometimes be connected to computers. Usually, volume or flow thresholds are used to initiate the computer memory. These thresholds vary from 10 to 200 ml for volume, or from 1 to 300 ml/sec for flow. Obviously, some of these may allow too much of the signal to escape, and therefore should be tested to assure that the entire forced vital capacity is captured and recorded by the computer. Finally, the computer can produce a digital display, printout, or some type of permanent storage. It is often assumed that the cost of these computers is related to the accuracy of their output; however, due to poor programming, one often finds little correlation between the two. Obviously, there are many different designs of spirometers attached to different types of recorders and computers. Some of these instruments have been evaluated and reported on in the literature. Though some reports are favorable, others have shown variations in spirometer volume measurements in excess of 20% [12]. Because of the wide variety of instruments and the variability of their accuracy, the American Thoracic Society initiated the Snowbird Workshop on the Standardization of Spirometry in January of 1977 [13]. This workshop produced a document on standardization of spirometry which has been updated and revised several times and is now an American Thoracic Society statement [14].

The following lists primary measurements made with a spirometer and the instrument requirements, recommended by the ATS, necessary for these measurements.

1. The vital capacity (VC) is the maximum volume of air exhaled from the point of maximum inspiration. Instruments which measure vital capacity should be able to accumulate volume for at least 30 sec. To make this measurement, spirometers should have volumes of at least 7 liters measured at body temperature and pressure saturated (BTPS) and should be capable of measuring this volume over a flow range from 0 to 12 liters/sec. The accuracy required for this measurement is at least ±3% of reading or ±50 ml, whichever is greater.

The rationale for this requirement is that vital capacity is a time-independent measurement. Thirty seconds is generally the maximum time in which a subject can extend exhalation. Studies of populations show that for vital capacity and forced vital capacity, a 7 liter volume will allow measurement of more than 95% of the population [14,15]. Accuracy of ±3% or ±50 ml, whichever is greater, was determined by intertest variability and the day-to-day variability for

the same subject [15,16]. Instrument errors should ideally be smaller than the subject variability, which is about 3%, and minimum resolution of 50 ml is a reasonable lower resolution limit for adults when very small volume test results are obtained.

2. The forced vital capacity (FVC) is the vital capacity obtained with a maximal forced expiratory effort. As with the vital capacity, the volume recommended for instrumentation is at least 7 liters with capacity for measuring flows between 0 and 12 liters/sec. Also, the instrument must be capable of accumulating volume for at least 10 sec [14]. A 10 sec interval is required so that most obstructed subjects have adequate time to complete their expiration.

3. The timed forced expiratory volume (FEV_t) is defined as the volume of air exhaled in a specified time during the performance of an FVC maneuver. For example, an FEV_1 is the volume of air exhaled during the first second of the FVC. This measurement and its ratio, obtained when it is divided by the forced vital capacity (FEV_1/FVC), are good indicators of obstructive lung disease. Requirements and their rationale for this measurement are similar to those for forced vital capacity measurement. Volume accuracy should be within ±3% of reading or ±50 ml, whichever is greater [14]. An important requirement of this measurement is that the start of the test, time zero, is to be determined by the back-extrapolation method. The start of test, by back extrapolation, is obtained by the extrapolation of peak flow to maximum inspiration.

This test is dependent on resistance to airflow, e.g., blowing through a drinking straw into a spirometer will yield a different FEV_1 than blowing through a 1 in. diameter tube. The resistance to airflow should be less than 1.5 cmH_2O per liter/sec at flow rates of 12 liters/sec [14].

4. The FEF_{25-75} measures the forced expiratory flow during the middle half of the FVC. Requirements for instrument accuracy for this test are ±5% of reading or ±100 ml/sec, whichever is greater. This measurement has a greater subject standard deviation than the FVC or FEV_1 because two, rather than one, volume-time measurements are made [14].

5. Flow or instantaneous forced expiratory flow, $\dot{V}$, can be measured either electronically or mechanically. When flow-volume loops or other measurements of flow are made, the flow measurement should be accurate to within ±5% of reading or ±0.2 liters/sec, whichever is greater [14]. Flow range should be 0 to 12 liters/sec.

6. A permanent graphical record of the forced vital capacity is one of the most important requirements which evolved from increasing recent interest in standardization of the instruments used for spirometry. The Snowbird Workshop recommended that instruments used to record the FVC and FEV_1 should provide at least a tracing of volume and time or volume and flow during the entire forced expiration. For the volume-time tracing, the recorder must be capable of displaying the entire FVC maneuver, at constant speed, from maxi-

mum inspiration for at least 10 sec after the start of the maneuver. If the paper record is made it must have at least the following characteristics: (1) paper speed of 2 cm/sec, with higher speeds preferable; (2) volume sensitivity of at least 10 mm of chart per liter of volume; and (3) flow sensitivity of at least 4 mm of chart per liter/sec of flow (all specifications are measured in BTPS). The participants at the Snowbird Workshop felt that the spirogram represented the best method of ensuring that this "effort-dependent" test was properly performed. Most forced vital capacity spirograms are displayed as a volume-time tracing. In order to determine the quality of the start of the FVC test and achieve reliable results by back extrapolation to determine time zero, the recorder should be "up to speed" before the forced expirogram is begun. The 10 sec record requirement is based on data which show that most obstructed subjects can complete the test in 10 sec [14,15]. The requirements of chart speed and volume sensitivity are based on earlier recommendations and the need to have accurate visual resolution on the record.

Although not specifically recommended, a thermometer installed in the spirometer is extremely important for volume-measuring devices [17] since the air blown into the relatively cool spirometer is condensed. This can yield a 1% volume decrease for every 2°C difference between body and spirometer temperature. Two conditions common in occupational testing have emphasized this point:

1. Uncontrolled temperature conditions in the occupational setting, such as cold temperatures in a mine or hot temperatures in a mill, can greatly affect results.
2. Rapid testing of many subjects can increase the spirometer temperature. Therefore, the spirometer can easily warm up several degrees during a testing session.

Even spirometers which comply with these standards can produce significantly different results. For instance, the accuracy for the FEV_1 requires a maximum measurement error of ±3% or ±50 ml, whichever is greater. Since the annual decline for the FEV_1 is approximately 1% per year, acceptable spirometers can produce results with a mean difference six times this predicted annual decline. Even instruments from the same manufacturer can produce systematically different data. Therefore, it is not only important to obtain an acceptable instrument, but to employ the same type of spirometer, and, if possible, the same instrument when repeated testings are required. This is especially true if the study protocol demands testing the subject before and after a work shift.

IV. Calibration Techniques

To obtain accurate results, spirometry system "calibration checks" should be conducted at least at the beginning of every day and "complete calibration" should be performed at least every week. The most important calibrating device currently available is a calibrating syringe with a volume of at least 3 liters with ±1% accuracy, traceable to the National Bureau of Standards. It may be necessary to deviate slightly from the calibration procedures listed below, depending on the type of spirometer. Since some flow-detecting devices are affected by humidity, the manufacturer should provide the appropriate "multiplying" factor to compensate for the effects of using "cool," "dry" air in the syringe rather than the warm, moist expired gas. It is also important with volume-measuring devices that the air inside the calibrating syringe be at the same temperature and relative humidity as the air inside the spirometer. Otherwise, some heat transfer may occur after the syringe is emptied into the spirometer and the gas inside the spirometer may expand or contract.

Calibration Check Procedure

1. The first step in evaluating a spirometer already owned, or where purchase is being considered, is comparison of manufacturer's specifications with the ATS recommendations. It is important to observe that a 10 liter spirometer which has an accuracy specification of ±3% of full scale, that is, ±300 ml, will not meet the ATS recommendation of ±3% of reading or ±50 ml, whichever is greater.

2. Check for any leaks in the tubing or spirometer; this is particularly important for volume-measuring devices.

3. Simulate a normal and obstructed patient by injecting the air from a 3 liter calibrating syringe into the spirometer in approximately 2 sec (normal) and 6 sec (obstructed). The spirometer "corrected" volume output should read between 2.91 and 3.09 for a 3 liter syringe. Also observe if there are adequate recorder volume and time sensitivities based on the spirometer recommendations.

4. A forced expiration should be performed with relatively low flow rates at the end of the maneuver to determine if the spirometer prematurely terminates its volume measurement or if it continues to show an increase in volume as you approach residual volume. Premature termination at low flows is a particular problem with currently available flow-measuring devices.

5. Check the "start of test" determination for any unusual sensitivity. These artefacts can occur when the subject is shaking the mouthpiece and tubing while straining to completely inhale at the start of the FVC maneuver. When this occurs, the FEV_1 may be zero or unusually low due to the false start.

6. The recorder's timing accuracy should be checked with a stopwatch simply by observing the time displacement over an appropriate time period.

7. The automatically determined FEV_1 should be compared with several hand-determined FEV_1 values using volume-time tracings and the back-extrapolation method. This comparison is necessary to test that the instrument is using a start-of-test determination method equivalent to the back-extrapolation method.

8. Perform any other calibration check procedures which may be recommended by the manufacturer. These procedures should be simple enough for a technician to follow and complete enough to ensure that the spirometer is functioning within the recommendations.

9. Finally, it is good practice to have a technician perform a few FVC maneuvers at the beginning of each session to serve as quality control values. These data can also provide information concerning the variability of the repeat tests performed in your laboratory.

V. Standardization of Test Procedures

The Snowbird Workshop also dealt with the need to standardize spirometric testing procedures. Standardized procedures and measurement techniques were suggested in order to produce spirometric data compatible between instruments and laboratories and from one time period to another.

Standardized methods of spirometric testing are as follows: The subjects are to be instructed in the FVC maneuver and the appropriate technique demonstrated. A minimum of three acceptable FVC maneuvers should be performed. Acceptability is determined by the technician's observation that the subject understood the instruction and performed the test properly. This includes observation of a smooth continuous exhalation with a good start and apparent maximal effort and without (1) coughing; (2) valsalva maneuver (closed glottis); (3) early termination of expiration (in a normal subject this will be before completion of the breath, and in an obstructed subject this should be assumed to take place if the expiratory time is less than 6 sec); (4) a leak; (5) an obstructed mouthpiece (obstruction due to tongue being placed in front of the mouthpiece, dentures falling in front of the mouthpiece); (6) an excessive variability among the three acceptable curves, e.g., the FVC of the two best of three acceptable curves should not vary by more than 5% or 100 ml, whichever is greater; and (7) an unsatisfactory start of expiration characterized by excessive hesitation or false starts. Unsatisfactory starts prevent accurate back extrapolation to determine time zero. To achieve accurate time zero the extrapolated volume on the volume-time tracing spirogram should be less than 10% or 100 ml, whichever is greater.

The FVC maneuver can be performed by either a closed or open circuit method. For the closed circuit method the subject inhales the maximal inspiration from the spirometer. For the open circuit method, the subject inspires

maximally, places the mouthpiece in his or her mouth, and then forcefully exhales. With the open circuit method, there is no display of inspiration and the subject can lose volume from full inspiration, prior to placement of the mouthpiece, without detection by the technician. Also, placement of the mouthpiece while holding at full inspiration is cumbersome and contributes to the systematic differences observed with these procedures [18]. Some instrumentation can also confound the procedural variability by employing the closed circuit method without displaying inspiratory volume or flow. Although the use of nose clips may not appreciably influence the forced vital capacity performed using the open circuit technique, some subjects breathe through the nose during the test when a closed circuit technique is used [14,18]. Also, adult subjects may be studied either sitting or standing [19], while for children under the age of 12 years the position should be indicated [14]. In any case, nose clips are recommended, and the same procedure (open or closed) and position should be used on repeat testing of the same subject.

VI. Importance of Technician Training

Perhaps the most difficult factor to control in the administration of pulmonary function testing is the technician's influence on the subject's performance. The FVC maneuver demands cooperation, and the subject must completely understand what is required. This responsibility rests totally with the technician, who must be aware of all pulmonary and nonpulmonary factors affecting the test. If a satisfactory series of tests cannot be obtained, the technician must report that the data are submaximal.

Adequate technician training and perhaps certification is an essential first step toward obtaining good quality pulmonary function data. Technicians must also be continuously evaluated to ensure that they continue to obtain the best possible performance from a subject. It is not unusual for some technicians to be incapable of mastering the art of coaching subjects properly on a continuing basis.

The pulmonary function technician should receive at least 16 hr of formal instruction followed by a period of monitoring either by direct observation or by review of time-volume or flow-volume tracings which they have collected. The formal instruction should consist of at least 6 hr of lectures and 10 hr of practical application. Due to the wide variety of instruments, it is important that the technician receive instruction on the type of spirometer which he or she will be using. The following topics should be included in the formal instruction.

Basic physiology of the forced vital capacity maneuver and the determinants of airflow limitation with emphasis on the reproducibility of results: Instruction in basic physiology is needed for the technician to understand why

the forced expiratory volume maneuver is reproducible and why in a few subjects it may not be reproducible.

Instrumentation requirements, including calibration procedures, sources of error, and their correction: The technician should know how to check the spirometer system for accuracy and proper operation. If this training is lacking, then a large amount of inaccurate data could be collected without being detected.

Performance of the testing, including subject coaching, recognition of improperly performed maneuvers, and corrective action: If the test is invalid, the technician should be capable of coaching the subject to give a more acceptable test result. Coaching of a subject must often be modified to accommodate the subject.

Data quality, with emphasis on reproducibility: The technician should understand all the criteria listed above for judging test acceptability.

Measurement of tracings and calculations of results: The technician should understand the BTPS correction factor and be capable of measuring the FVC, FEV_1, and FEV_{25-75} by hand from a volume-time tracing using the back-extrapolation technique. The technician should also be taught to obtain predicted values and express the results as a percentage of predicted.

Due to the Cotton Dust Standard requirements for technician certification, the National Institute of Occupational Safety and Health has applied these basic guidelines in order to approve qualified training courses. At this time, several approved coarses are available throughout the United States.

VII. Standardization of Measurement and Computation

Measurements of the spirogram should be made from a series of at least three acceptable forced expiratory curves [14]. The maximal FVC and the maximal FEV_1 recorded should be obtained after examining the data from *all* the acceptable curves even if the maximum FVC and the maximum FEV_1 do not come from the same curve. The beginning of time for the FEV_1 should be obtained by the method of back extrapolation. If the FEF_{25-75} and/or instantaneous maximal expiratory flows are to be obtained, they should be measured from the single acceptable test which yields the greatest sum of FEV_1 and FVC. This is defined as the "best" curve.

Best efforts cannot be determined by simple inspection of a spirogram. Measurement and computation are required to determine the largest values. There is little difference between the largest values and the mean values if data are properly collected. However, independently selecting the largest value for FVC and FEV_1 accounts for an occasional influence of learning and possible deterioration in performance due to fatigue or induced bronchospasm. There is

no need to discard the best FEV_1 value even if a maneuver is prematurely terminated.

The discriminating quality of spirometry is greatly influenced by the time duration of the forced expiratory spirogram for obstructive subjects. If a severely obstructed subject performs an FVC maneuver with flow continuing for 10 sec, the subject may produce a relatively normal FVC while abnormal values of FEV_1/FVC, FEF_{25-75} and $\dot{V}_{\%\ FVC}$ would be observed. If this same subject's spirogram were terminated after 6 sec, the FVC would be reduced, while the other parameters would be substantially increased. Though the discriminating quality of the 10 sec FVC is advantageous, it may be difficult for an obstructed subject to repeat the maneuver several times with equally sustained effort. Therefore, the FVC and those parameters influenced by it are extremely effort dependent as the subject approaches full expiration. One could conduct an FEV_6 or FEV_7 and base all other parameters on this volume. Though variability would be reduced, some of the discrimination between obstructive and restrictive disease is lost. This tradeoff is constantly encountered in epidemiological investigations where test repeatability and sensitivity must be optimized. However, if one decides to base these parameters on an FEV_6 or FEV_7, the maneuver should be labeled as such and not be considered a forced vital capacity.

One facet of measurement and computation which is often overlooked is the correction of volumes and flows to body temperature and pressure saturated with water vapor (BTPS). Approximately a 1% change in volume or flow (equivalent to the yearly predicted decline for some parameters) is introduced for every 2°change in temperature or 200 ml of mercury change in barometric pressure. Therefore, temperature and pressure should be monitored in order to make correct measurements. However, since the correction factors from ATPS to BTPS at 22°C are 1.0904, 1.0910, and 1.0915 for barometric pressures of 770, 760, and 750 mmHg, respectively, it is unnecessary to correct for small deviations from standard barometric pressure. Some instruments display the data on a graph which has two volume grids, one in ATPS (ambient temperatures and pressures saturated with water vapor) and the other in BTPS. It is important to note that the values in BTPS on these graphs are corrected only for one specific temperature and pressure, usually 25°C and 760 mmHg. Therefore, if the actual spirometer temperature or pressure is different, this BTPS output will be incorrect.

Manufacturers of mass flow meters (hot wires) often indicate that their instruments are not affected by temperature and therefore a correction to BTPS is not necessary. However, it is important to note that these instruments are significantly affected by differences in water vapor. For such instruments, calibration with a syringe of dry air will yield significantly different results than with a syringe containing air saturated with water vapor. Therefore, correction factors should be obtained from these manufacturers.

VIII. Data Interpretation

Once some measurement of ventilatory function has been obtained using adequate equipment and procedures, the next step is to interpret the results. Because measures of ventilatory function are dependent on age, height, sex, race, and many other factors, care must be taken to consider these factors in interpreting any measurements. There are two basic approaches to evaluating ventilatory tests. One approach is to use the subject as his or her own control and follow changes in ventilatory function with time. The second, more common method is to compare the measured values with "normal values" through the use of prediction equations. The sensitivity of any test of ventilatory function is dependent to a large extent on how well these various factors, such as age and height, can be removed from the inherent variability of the parameter both for a given subject and within a normal population, the latter being greater. For example, the within-subject coefficient of variation is approximately 3% for the FVC compared to approximately 14% within a population [20–22].

Using a subject as his or her own control and obtaining repeated measurements is considered by many as the preferred method; however, this is often impractical. The normal yearly decrement in ventilatory function is small (approximately 25 ml/year [23] for the FVC); however, the inherent variability of the parameter requires that an abnormality can only be detected if the observed changes are large or if the subject is followed over an extended period of time. In addition, while there have been many cross-sectional studies of ventilatory function of "normal" populations, which provide a reasonable estimate of the decrements of ventilatory function for age, there have been relatively few longitudinal studies to provide an estimate of the variability of this decrement within a given subject. Until more data are gathered by longitudinal studies, comparison of a subject's observed values with some "normal" or expected values will continue to be necessary.

Since age and height are strongly related to ventilatory function, most prediction equations include these parameters in their estimate of ventilatory function. Ideally, the prediction equations should be derived from a study of healthy, normal, nonsmoking individuals using procedures and equipment which conform to the requirements presented here. Many of the published predicted normal standards or equations have included nonsmokers, smokers, and former smokers. Since cigarette smoking has been shown to adversely affect pulmonary function, any population which includes smokers and former smokers could hardly be considered "normal." In addition, many studies have not considered or critically examined other important factors such as previous history of respiratory disease and exposure to occupational or environmental agents which may affect the respiratory system.

One study which attempted to use a healthy nonsmoking population was conducted by Morris and associates [23]. They tested 998 healthy nonsmokers who lived near Portland, Oregon, and were members of the Church of Jesus Christ of Latter-day Saints. In addition, Morris used a Stead-Wells spirometer which has been shown to meet the ATS instrumentation recommendations [14,17].

Although the Morris study provided some of the best predicted standards, there are still potential problems with adopting them. First, the methods of calculation were those of Kory and associates [24], and therefore their method of calculating the FEV_1 did not conform with the method of back extrapolation. Second, a single best effort was always selected for determination of the FVC, FEV_1, etc., which does not conform with the recommended method of using the largest FVC and FEV_1, regardless of the curve(s) on which they occur. Also, the FEF_{25-75} should be measured from the curve with the largest sum of FVC plus FEV_1. Finally, some question remains as to whether the population used by Morris adequately represents an average normal healthy population.

Morris performed an interesting comparison of various prediction equations and demonstrated large differences between them. For example, he reported as much as an 820 ml difference between their observed mean FVC and the predicted mean FVC of a previous study [23].

In a more recent study, Knudson and associates [25] reported prediction equations for men and women obtained from 746 healthy nonsmoking subjects who live in Tucson, Arizona. While they used equipment and procedures which appear to meet the recommendation, the authors reported the FVC and FEV_1 obtained from averaging the best two of five values. This method of averaging does not conform with the ATS recommended use of the largest FVC and FEV_1 in deriving their equations. However, their comparison of the largest of the first three FVC and FEV_1 values with the average of the best two of five did not show any statistically significant differences.

An additional consideration in the comparison of a subject's observed value with a predicted value is the ethnic background. For example, several studies [26–28] have shown that male blacks have a predicted FVC from 10 to 15% lower than their white counterparts of the same age and height. These reports have recommended multiplying the predicted value obtained from a white population by approximately 0.85. Obviously, a more desirable approach would be to develop prediction equations for every ethnic group. Only by using this approach will it be possible to fully compensate for ethnic differences in the relationship between ventilatory function, age, and height.

After selecting appropriate predicted values, one can determine whether the observed values are significantly different from normal. Measurements of ventilatory function with large standard errors (SE) and poor correlation with age and height must have a corresponding large departure from the predicted value in order to be significant. Morris and associates [23] found that males aged 30–39 years had a mean FVC of 5.38 liters with a standard deviation of

0.89 liters. Therefore, 16% of the normal individuals studied would have an FVC of 5.38 − 0.89 = 4.49 liters or less, and 2.5% of the normal individuals would have an FVC of 5.38 − 1.78 = 3.60 liters or less. If the observed values are expressed as a percentage of predicted, 2.5% of the normal individuals studied would have an observed value less than 67%, 69%, and 47% of predicted for FVC, FEV_1, and FEF_{25-75}, respectively. In a similar type of analysis, Knudson and associates [29] determined the percent of predicted value above which 95% of asymptomatic nonsmokers fell. They found that 95% of the males aged 16-35 years had an FEV_1 of 81.8% of predicted or greater and a $\dot{V}_{max_{50}}$ of 66.1% of predicted or greater. The lower percentage of prediction for the $\dot{V}_{max_{50}}$ is a result of its larger variability. Therefore, a patient must have observed values for the FEF_{25-75} and instantaneous flow rates considerably less than 80% of predicted to indicate significant abnormality.

It has been proposed that a good method of ensuring that the patient is giving the best possible effort is to compare the test results with some predicted value while the test is being conducted. If the subject's observed value is below the predicted, then the technician is encouraged to coach the subject to expend more effect. While on the surface this technique seems reasonable, there are potential problems with this approach. If coaching is dependent on the prediction equations used, the test results obtained will be biased toward these predicted values. Subjects with ventilatory function values above normal will not be encouraged to try harder, in contrast to the abnormal subject who will receive considerably more coaching. A better approach is to use criteria which depend on the reproducibility of the test results, and only after testing of the subject is completed should the observed values be compared with predicted values.

In many epidemiological studies, lung function is used to determine if a particular harmful agent has a detrimental effect on respiratory health. To detect these effects, mean function values in the exposed population are compared to expected values of a nonexposed group. Since it is often difficult and expensive to obtain values on a nonexposed group, the observed mean values are sometimes compared to published normal values. The observed values for each subject can be divided by the corresponding predicted value for every subject, and results are expressed as an average observed/predicted percentage. Considerable care must be taken in interpreting the results of these types of analyses. For reasons discussed previously, the results obtained by a study may not be directly comparable to a published predicted equation obtained using different equipment, procedures, and analysis. In addition, many normal studies have excluded smokers, exsmokers, and symptomatic subjects, and therefore the normal population may be considerably more healthy than an appropriate nonexposed or control group. Symptom status may be compared by using the Medical Research Council questionnaire commonly employed in epidemiological surveys [30].

Ideally, when any exposed group is studied, a matched nonexposed control group should be studied at the same time using the same equipment, procedures, and technicians. Taking care to match the two populations with respect

to age, height, sex, ethnic group, history, and smoking status should allow direct comparison of population mean values. Depending on how well the populations are matched, it is possible to attach significance to even small differences in population mean values and thereby increase the sensitivity of the study. Although prediction equations or covariance analysis can be used to compensate for imperfect matching of populations, these techniques require several assumptions which complicate interpretation of the study results.

IX. Special Problems of Field Studies

It is often advantageous to conduct pulmonary function tests at the industrial site. Time away from the job is reduced, and lung function can be assessed in conjunction with estimates of industrial exposure. In addition, on-site testing is required if an acute response is to be detected by pre- and postshift studies. Since rapid testing of large populations is often desirable, it becomes necessary to employ an automated data collection system. A completely automated system will reduce testing and data turnaround time by eliminating data transcription, computation, and reduction by the technician; however, it can pose serious problems in transport and setup in the industrial environment. One solution to this problem is the self-contained mobile laboratory. Testing equipment, calibration equipment, and any test gas cylinders required for the investigation remain together, avoiding loss or damage during shipment. At the site, the mobile laboratory provides adequate working space and a controlled environment. The success of a mobile laboratory requires careful installation of the test equipment. All equipment should be shock-mounted and shielded from spurious electrical inputs. An adequate electrical system must also facilitate the large capacity heating and cooling units essential for a controlled laboratory environment. Extreme temperatures are not only uncomfortable for personnel and subjects, but may adversely affect test equipment. Service access to the rear of equipment should be provided by the addition of doors or hatches to the vehicle. The location of equipment should facilitate testing and provide good balance for driving safety. Even the most sophisticated equipment can be suitably housed in a mobile unit. The laboratory currently employed in field studies nationwide by Tulane University School of Medicine is displayed in Figures 1 and 2 [31].* A Winnebago motor home houses a pulmonary laboratory, a computer interface for data transmission, a CRT (cathode ray tube) for display of data, four diskette drives for magnetic data storage, and a printer. The laboratory's capabilities include measuring both timed volumes and flow

*Design engineering by Henry Glindmeyer, D. Eng. Installation of equipment by Biomedical Associates, Inc.

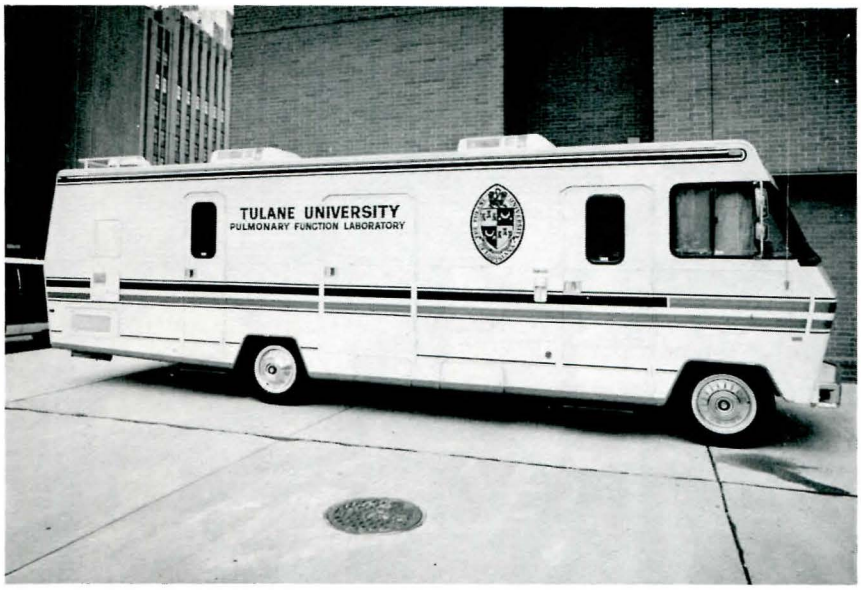

Figure 1 Exterior of mobile laboratory displays several access hatches and electronic leveling jacks. Test tanks are mounted on roll-out storage trays (not shown).

rates from the forced vital capacity maneuver, lung volumes and capacities, single-breath diffusing capacity, and parameters from the He-O_2 flow-volume maneuver. In addition, data reduced from pre- and postshift studies include across-shift isovolume-determined flow rates. All data are first displayed on the CRT for technician acceptance, and are then stored on IBM compatible diskettes and finally printed.

Because of the advantage of the automated system, many investigators are beginning to switch to more sophisticated equipment. The most important aspect of a transition is maintaining compatibility, or at least documenting the variability between past methods of data collection and those introduced by the new system. Subtle differences, e.g., computer initiation thresholds, data transmission resolutions, data reduction methods, can pose significant problems, especially if a new system is introduced during the course of a longitudinal study. Defining past operational procedures is a necessary first step in documenting the requirements of a new system.

Those who undertake to equip and operate a mobile laboratory will face a number of problems, some will be of a general nature; however, others will be unique to the equipment and vehicle chosen. Experience points to the need for good engineering support in the planning, establishment, and operation of a mobile laboratory.

Figure 2 View toward rear of laboratory displays Pulmolab, CRT, and printer. Compartment behind Pulmolab can be used for exercise testing equipment or plethysmography. Rear of unit contains generator, refrigerator, head, water storage, filing cabinet, and voltage regulator. Driving area is used for administering the questionnaires.

X. Conclusion

Spirometry has become a cornerstone of epidemiological studies of occupational lung disease. Progress is being made to resolve the six special problems outlined in the introduction. Many of these problems are being met by developing recommendations and standards for instrumentation, test procedures, measurements and calculation, and interpretation. Great care has been and should be taken when standards are developed to allow for reasonable flexibility and innovation. Therefore, the standards outlined in this chapter should be considered a starting point, not final and inflexible.

Development of recommendations for spirometry can serve as a model for development of other lung function testing techniques. However, as new techniques develop, opportunity for standardization should be taken sooner than it was with spirometry. The principal research laboratories using new methods should discuss uniformity of testing techniques, measurement, and interpretation, so that research data will be comparable, and implementation of the techniques in occupational health research can be more rapid.

References

1. Morgan, W. K. C., and N. L. Lapp, Respiratory disease in coal mines, *Am. Rev. Respir. Dis.*, **113**:531 (1976).
2. Ziskind, M., D. B. Ellithorpe, and H. Weill, State of the art—silicosis, *Am. Rev. Respir. Dis.*, **113**:643 (1976).
3. Jones, R. N., J. Carr, H. Glindmeyer, J. Diem, and H. Weill, Respiratory health and dust levels in cottonseed mills, *Thorax*, **32**:281–286 (1977).
4. Jones, R. N., and H. Weill, Occupational lung diseases, *Basic of RD*, Vol. 6, No. 3 (1978).
5. Becklake, M. R., Asbestos-related diseases of the lungs and other organs: Their epidemiology and implications for clinical practice, *Am. Rev. Respir. Dis.*, **114**:187 (1976).
6. Leith, D. C., and J. Mead, *Principles of Body Plethysmography*. National Heart and Lung Institute, Division of Lung Diseases, November, 1974.
7. Macklem, P. T., *Procedures for Standardized Measurement of Lung Mechanics*. National Heart and Lung Institute, Division of Lung Diseases, November, 1974.
8. See Chapter 5 (pages 87–98) in this monograph.
9. Fry, D. L., and R. E. Hyatt, Pulmonary mechanics, *Am. J. Med.*, **29**: 672–689 (1960).
10. Mead, J., J. M. Turner, P. T. Macklem, and J. B. Little, Significance of the relationship between lung recoil and maximum expiratory flow, *J. Appl. Physiol.*, **22**:95–108 (1967).
11. Weill, H., Pulmonary function testing in industry, *J. Occup. Med.*, **15**: 693–699 (1973).
12. Fitzgerald, M. X., A. A. Smith, and E. A. Gaensler, Evaluation of "electronic" spirometers, *N. Engl. J. Med.*, **289**:1283–1288 (1973).
13. Gardner, R. M. (Chairman), American Thoracic Society—Report of Snowbird workshop on standardization of spirometry, *ATS News*, **3**:20–24 (Summer 1977).
14. Gardner, R. M. (Chairman), ATS statement—Snowbird workshop on standardization of spirometry, *Am. Rev. Respir. Dis.*, **119**:831–838 (1978).
15. Hankinson, J. L., and M. R. Peterson, Data analysis for spirometry instrument standards (abstract), *Am. Rev. Respir. Dis.*, **115**(supplement):116 (1977).
16. Cochrane, G. M., F. Prieto, and T. J. H. Clark, Intra-subject variables of maximum expiratory flow volume curve, *Thorax*, **32**:171 (1977).
17. Gardner, R. M., J. L. Hankinson, and B. J. West, Evaluating commercially available spirometers, *Am. Rev. Respir. Dis.*, **121**:73–82 (1980).
18. Glindmeyer, H. W., S. T. Anderson, J. E. Diem, and H. Weill, A comparison of the Jones and Stead-Wells spirometers, *Chest*, **73**:596–602 (1978).
19. Pierson, D. J., H. P. Dick, and T. L. Petty, A comparison of spirometric values with subjects in standing and sitting positions, *Chest*, **70**:17 (1976).
20. McCarthy, D. S., D. B. Craig, and R. M. Cherniack, Intraindividual variability in maximal expiratory flow volume and closing volume in asymptomatic subjects, *Am. Rev. Respir. Dis.*, **112**:407–411 (1975).

21. Dawson, A., Reproducibility of spirometric measurements in normal subjects, *Am. Rev. Respir. Dis.*, **93**:264–268 (1966).
22. Black, L. F., K. Offord, and R. E. Hyatt, Variability in the maximal expiratory flow volume curve in asymptomatic smokers and in nonsmokers, *Am. Rev. Respir. Dis.*, **110**:282–292 (1974).
23. Morris, J. F., A. Kiski, and L. C. Johnson, Spirometric standards for healthy nonsmoking adults, *Am. Rev. Respir. Dis.*, **103**:57 (1971).
24. Kory, R. C., R. Callahan, H. G. Boren, and J. C. Syner, The Veterans Administration-Army cooperative study of pulmonary function. I. Clinical spirometry in normal men, *Am. J. Med.*, **30**:243 (1961).
25. Knudson, R. J., R. C. Slatin, M. D. Lebowitz, and B. Burrows, The maximal expiratory flow-volume curve: Normal standards, variability and effects of age, *Am. Rev. Respir. Dis.*, **113**:587 (1976).
26. Rossiter, C. E., and H. Weill, Ethnic differences in lung function: Evidence for proportional differences, *Int. J. Epidemiol.*, **3**:55 (1974).
27. Lapp, N. L., H. E. Amandus, R. Hall, and W. K. C. Morgan, Lung volumes and flow rates in black and white subjects, *Thorax*, **29**:185 (1974).
28. Damon, A., Negro-white differences in pulmonary function (vital capacity, timed vital capacity, and expiratory flow rates), *Hum. Biol.*, **38**:380 (1966).
29. Knudson, R. J., B. Burrows, and M. D. Lebowitz, The maximum flow volume curve: Its use in the detection of ventilatory abnormalities in a population study, *Am. Rev. Respir. Dis.*, **114**:871–879 (1976).
30. Ferris, B. G., Epidemiology Standardization Project II, recommended respiratory disease questionnaires for use with adults and children in epidemiologic research, *Am. Rev. Respir. Dis.*, **118**(supplement):7–53 (1978).
31. Glindmeyer, H. W., H. Weill, R. Jones, and M. Ziskind, Design and operation of a mobile pulmonary function laboratory, *J. Occup. Med.*, **16**:584–588 (1974).
32. Despar, P. J., M. Leroux, and P. T. Macklem, Site of airway obstruction in asthma determined by measuring maximal expiratory flow breathing air and helium-oxygen mixture, *J. Clin. Invest.*, **51**:3235–3243 (1972).
33. Schilden, D. P., A. Robert, and D. L. Fry, Effect of gas density and viscosity on the maximal expiratory flow-volume relationship, *J. Clin. Invest.*, **42**:1705–1713 (1963).
34. Cosio, M., H. Ghezzo, J. C. Hogg, R. Corbin, M. Loveland, J. Dosman, and P. T. Macklem, The relationship between structural changes in small airways and pulmonary function tests, *N. Engl. J. Med.*, **298**:1277–1281 (1977).
35. Fairman, R. D., J. Hankinson, H. Imbus, N. L. Lapp, and W. K. C. Morgan, Pilot study of closing volume in byssinosis, *Br. J. Ind. Med.*, **32**:235–238 (1975).
36. Cotes, J. E., *Lung Function*, 3d Ed. Blackwell Scientific Publications, London, 1975, pp. 238–259.

37. Bouhuys, A., *Breathing.* Grune & Stratton, New York, 1974.
38. Bouhuys, A., *Airways Dynamics,* Charles C. Thomas, Springfield, Illinois, 1970.
39. DuBois, A. B., A. W. Brody, P. H. Lewis, and B. F. Burgess, Jr., Oscillation mechanics of lungs and chest in man, *J. Appl. Physiol.,* **8**:587–594 (1956).
40. Fischer, A. B., A. B. DuBois, and R. W. Hyde, Evaluation of the forced oscillation technique for the determination of resistance to breathing, *J. Clin. Invest.,* **47**:2045–2057 (1968).
41. Frank, N. R., J. Mead, and J. L. Whittenberger, Comparative sensitivity of four methods for measuring changes in respiratory flow resistance in man, *J. Appl. Physiol.,* **31**:934–938 (1971).

OCCUPATIONAL LUNG DISEASES

RESEARCH APPROACHES AND METHODS

Edited by

Hans Weill

Tulane University School of Medicine
New Orleans, Louisiana

Margaret Turner-Warwick

Cardiothoracic Institute
Brompton Hospital
London, England

MARCEL DEKKER, INC. New York • Basel