Data Evaluation and Methods Research

Series 2
Number 86

Computer-Assisted Spirometry Data Analysis for the National Health and Nutrition Examination Survey, 1971-80

The equipment, procedures, and data reduction methods employed in the National Health and Nutrition Examination Survey for the collection and analysis of spirometric data are described. Data variability and testing methodology are discussed, as well as the influence of milieu and technician training. The computer programs that drive the data reduction and calibration are detailed, as are the algorithms used in the calculation of various spirometric parameters. The algorithms chosen for the determination of certain critical parameters are documented and validated.

DHHS Publication (PHS) 81-1360

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Office of Health Research, Statistics, and Technology
National Center for Health Statistics
Hyattsville, Md. October 1980

Library of Congress Cataloging in Publication Data

United States. National Center for Health Statistics.
Computer assisted spirometry data analysis for the Health and Nutrition Examination Survey, 1971-1980.

(Vital and health statistics : Series 2, Data evaluation and methods research ; no. 86) (DHHS publication ; no. (PHS) 81-1360)
Written by David P. Discher et al.
Includes bibliographical references.
1. Spirometry—Data processing. 2. Health and Nutrition Examination Survey. I. Discher, David P. II. Title. III. Series: United States. National Center for Health Statistics. Vital and health statistics : Series 2, Data evaluation and methods research ; no. 86. IV. Series: United States. Dept. of Health and Human Services. DHHS publication ; no. (PHS) 81-1360.
RA409.U45 no. 86 [RC734.S65] 312'.0723s 80-607930
ISBN 0-8406-0195-6

NATIONAL CENTER FOR HEALTH STATISTICS

DOROTHY P. RICE, *Director*

ROBERT A. ISRAEL, *Deputy Director*
JACOB J. FELDMAN, Ph.D., *Associate Director for Analysis*
GAIL F. FISHER, *Associate Director for Cooperative Health Statistics Systems*
ROBERT A. ISRAEL, *Acting Associate Director for Systems*
ALVAN O. ZARATE, Ph.D., *Acting Associate Director for International Statistics*
ROBERT C. HUBER, *Associate Director for Management*
MONROE G. SIRKEN, Ph.D., *Associate Director for Mathematical Statistics*
PETER L. HURLEY, *Associate Director for Operations*
JAMES M. ROBEY, Ph.D., *Associate Director for Program Development*
GEORGE A. SCHNACK, Ph.D., *Associate Director for Research*
ALICE HAYWOOD, *Information Officer*

DIVISION OF HEALTH EXAMINATION STATISTICS

ROBERT L. MURPHY, *Director*
JEAN ROBERTS, *Chief, Medical Statistics Branch*
KURT MAURER, *Acting Chief, Survey Planning and Development Branch*

COOPERATION OF THE U.S. BUREAU OF THE CENSUS

Under the legislation establishing the National Health Survey, the Public Health Service is authorized to use, insofar as possible, the services or facilities of other Federal, State, or private agencies.

In accordance with specifications established by the National Center for Health Statistics, the Bureau of the Census, under contractual agreement, participated in planning the survey and collecting the data.

Vital and Health Statistics-Series 2-No. 86

DHHS Publication (PHS) 81-1360
Library of Congress Catalog Card Number 80-607930

CONTENTS

LIST OF FIGURES

LIST OF TEXT TABLES

SYMBOLS

Data not available	- - -
Category not applicable	. . .
Quantity zero	-
Quantity more than 0 but less than 0.05	0.0
Figure does not meet standards of reliability or precision	*

COMPUTER-ASSISTED SPIROMETRY DATA ANALYSIS FOR THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY, 1971-80

David P. Discher, M.D.[a]; Alan Palmer, Ph.D.[b]; Gregory Hibdon[c]; Terence A. Drizd, M.S.P.H.[d]

INTRODUCTION

The wide acceptance of the Forced Expiratory Spirogram pulmonary function test in respiratory epidemiologic studies is evidenced by recent efforts of the Division of Lung Disease of the National Heart, Lung and Blood Institute and the American Thoracic Society to bring greater precision to this important test.[1,2] Spirometry provides both medical practitioners and epidemiologists with a simple yet objective method of following the course of chronic obstructive lung disease from its early inception to its more advanced states, thereby permitting the application of intervention measures and the monitoring of results. Furthermore, epidemiological studies can indicate early changes in function that can be related to various aspects of environmental pollution thus permitting development of control strategies to mitigate further degradation of function.

Unfortunately, spirometric testing is hampered by a lack of sound and sensitive data obtained from rigorous testing procedures on general population groups. These data are necessary for derivation of performance standards.

[a]Chairman, Department of Industrial and Environmental Medicine, San Jose Medical Clinic.

[b]Senior Epidemiologist, Center for Community Health Studies, Stanford Research Institute International.

[c]Computer Application Analyst, Stanford Research Institute International.

[d]Statistician, Division of Health Examination Statistics, National Center for Health Statistics.

The National Health and Nutrition Examination Survey of the National Center for Health Statistics is the largest ongoing examination survey in the world. Thus the National Health and Nutrition Examination Survey offers an opportunity to collect lung function data on various population groups representative of all socioeconomic groups, races, ages, sexes, and geographic areas. Because additional data are also collected on examinees that may be significant variables for spirometric function, this survey will lead to research on other variables.

Aware of the limitations of existing spirometry data, the staff of the Division of Health Examination Statistics, which conducts the National Health and Nutrition Examination Survey, have undertaken an extensive review of the existing spirometry data collection procedures and computer processing program criteria to ensure that data sensitivity is maximized.

This report details each of the steps taken to ensure the collection of optimal data. An identification of the multiple source of variability known to reduce the sensitivity of the data, a description of the subsequent operating procedures to minimize each of these sources of variance, a review of spirogram measurement criteria as currently used in the National Health and Nutrition Examination Survey program, and a comparative analysis of various alternative algorithms for increasing the accuracy of the measurements are presented. The development of alternative spirogram measurement techniques

was undertaken to further validate those techniques suggested in the recent National Heart, Lung and Blood Institute report[1] and, most important, to provide testable, documented logic for the National Health and Nutrition Examination Survey (NHANES) criteria used in quality control calibration, and measurement procedures.

These documented measurement criteria should provide a foundation for the analysis of current and future NHANES-collected data from which new regression equations will be developed for prediction on normative values.

BACKGROUND

Spirometry testing has been an integral part of the National Health Examination Survey (NHES) since 1963. During NHES Cycle II (1963-65), spirometry data were obtained by a Collins water-sealed spirometer, using the standard operating test procedures recommended in the spirometer instruction manual. Generally, technicians had little training in the theory and physiological meaning of spirometry. Measurements were made manually at great expense in time and money, and the limitations of this level of data collection became obvious.[3] During NHES Cycle III (1966-70), a spirometry testing module that used computerized data collection techniques was developed. Rigid standard operating procedures (SOP's) were developed, and concurrent technician and data surveillance programs were run to control for procedure and test data variability. Data were analyzed using the spirometry computer program[4,5] developed by the Public Health Service (PHS).

Further refinements were made in the spirometry data collection module in 1970 before the beginning of the National Health and Nutrition Examination Survey (NHANES I). The data acquisition hardware system that was used in NHANES I to collect spirograms is described in this report. Digital tape equipment was installed to replace the analog data systems used in NHES III, and general refinements of the SOP's were made to reflect the current methodology (e.g., the use of a standard set of five trials to ensure maximal values). While NHANES I data were being collected, the latest version of the PHS computer spirometry program was reevaluated and extensive program changes were made in calibration and quality control procedures and the logic used to define and compute the various spirometric measurements. Recently, new criteria have been developed and adopted by the National Heart, Lung and Blood Institute (NHLBI) to standardize the criteria used to compute and analyze spirometric data in epidemiologic studies.[1] The NHLBI criteria are comparable with those used by the NHANES programs except in "zero-time" and the "end-of-test" computations. These methods are compared and their strengths and weaknesses are documented.

The initial discussion in this report relates to nonsampling data errors that are caused by the host of variables that the NHANES planning group delineated as obstacles to collecting optimal data. This discussion is followed by a description of the instrumentation and the quality control programs that were developed to control for these errors and of the test procedures used during NHANES I. Finally an analysis of the test procedures is presented and various alternative methods for obtaining spirometric measurements are compared.

Spirometry Data Variability

In establishing testing uniformity, the variables that must be considered include selection and training of technicians, testing techniques, testing environment, spirometry equipment selection, data measurement and computation, and quality control.[6,7] Each of these areas is a potential cause of nonsampling error that diminishes or obscures any differences being sought in epidemiological studies as well as the validity of spirometry as a clinical-diagnostic tool.

Examinee sources of variance.—Submaximal expiratory effort during the performance of the Forced Expiratory Spirogram (FES) is attributable to a variety of factors. A common cause of poor test data is failure of the subject to comprehend the test instructions; in children this problem is often referred to as testing imma-

turity. This condition is a behavioral-social phenomenon exemplified by a lack of school readiness; the commands "Sit down and be quiet," "Raise your hand when you want to speak," "Pick up your pencil and copy the picture and the words in your book" all require understanding, willingess, and enough self-control for the pupil to perform properly and effectively.[8] Older children and adults of various ethnic and socioeconomic groups can also present problems of language and comprehension, and these frequently combine to frustrate meaningful data collection. Examinees with such problems are often performing the spirometry maneuver for the first time and this situation, coupled with anxiety regarding any medical procedure or its implications, often results in an unacceptable test despite the best efforts of the technician.

Technician sources of variance.—Spirometry testing requires maximum subject participation and an astute technician. Current practice dictates that vigorous verbal encouragement be given to the subject to stimulate maximal effort. An experienced technician is a combination of bully and cheerleader as he or she strives to elicit this maximal response from the subject. The technician must first explain the test, demonstrate the procedure, cheer on or goad the subject into putting forth his or her best effort, and evaluate the degree of cooperation obtained.

The methods used to administer the test not only vary from one technician to another, but also vary from trial to trial with the same subject. Not all technicians have equal abilities to perform all tasks well. Some work well only under supervision; if supervision is varied, technician performance also can vary.[9]

Any individual who is well motivated, interested, and reasonably intelligent and who has the equivalent of a high school education can be trained in spirometry.[7] Only 2 weeks of intensive training are required to learn how to administer the spirometry test, handle and calibrate the instruments, and perform the calculations. However, learning to obtain the best possible performances from examinees of all types and ages takes much longer—at least 6 months and perhaps a year. Such experience develops the many approaches necessary to instruct the examinee in a series of unfamiliar maneuvers, such as, taking in the deepest breath possible, inserting the mouthpiece and keeping the lips tightly around it, and exhaling into the spirometer as quickly, forcibly, and completely as possible.

The most important quality of a pulmonary function technician is the motivation to perform the very best test on every examinee. Initial enthusiasm after a while may turn into lack of interest. The intellectual ability of the technician becomes particularily important in discerning performance deficiencies of examinees and correcting these errors in maneuver.

The qualifications of personnel being hired to do spirometry are difficult to judge. This process may be accomplished though a personal interview with the prospective employee in which previous and related work experiences are reviewed and discussed. Each new technician should be evaluated to determine the level of training that will be required; and, if further training is needed, it should be done under the guidance of an experienced physician or pulmonary physiologist in a laboratory where ample testing is being performed with the highest standards of accuracy and quality control.

Equipment sources of variance.—Spirometers—much data are available on pulmonary function sensors that point to a basic set of desirable characteristics. The spirometers should be accurate and precise, have linear volume and flow rate response, be electronically (in electronic models) and pneumatically calibratable, have a frequency response of the signal being recorded (FES, 15 Hz), and have low inertia without oscillatory fluctuations.[2]

Portability and compactness, although desirable, should not be considered at the expense of any of the preceding characteristics.

Automation.—Hand measurements of spirometric data have been shown to be less precise than automatic systems. Studies have shown that when two trained pulmonary technicians analyzed a number of spirograms, interobserver differences were statistically significant.[10] Epidemiologic studies often require the combined efforts of two or more observers for the study of a large population; thus should one observer be more precise than the other, the quality of the better effort is diluted when the results are

pooled. The ability of measurements to discriminate between a normal and abnormal population is vitiated under such circumstances.

The expenditures of time and people for routine computations is no longer justifiable. The use of automated techniques conserves time, improves accuracy and precision, increases work capacity, and reduces cost. Through these means, the professional and technical staff become free to pursue more challenging activities.

Testing Methodology

The need for calibration.—Although spirometry equipment is extremely accurate, even the best equipment requires both careful attention and routine maintenance. For the electronic signals generated by moving the piston in the spirometer to be related to known volumes and known flows of air, the technician must perform periodic calibration checks. To detect minor signal fluctuations between pneumatic and electronic calibrations, the technician must perform a calibration as required by the SOP's.[11] Precise adjustments of the equipment are made that alter the volume-to-voltage relationship which are based on observations that the technician makes by using the pneumatic calibrations. The technician becomes aware of the need for electronic service to the equipment when the electronic calibrations show wide fluctuations of the standard electronic signal. Thus the first consideration in obtaining valid data on forced expiratory maneuvers by electronic spirometry is an understanding of the electronic principles inherent in calibrations and maintenance.

The need for technician-examinee rapport.—The second concept that the technician must understand is the requirement that the forced expiratory maneuver be correctly performed by the subject under the close observation and guidance of the technician. The technician can enhance this communication by developing an initial rapport, performing a good demonstration of the maneuver, and clearly stating the standard test instructions. The technician's skill is manifested by the subject's comprehension of the initial standard instructions, motivation to provide a maximal effort on a minimum of two of the five expiratory trials, and correct notation of procedural errors and redirection of test instructions accordingly. A number of barriers to a successful test can be identified as follows:

- Testing immaturity—the subject cannot follow directions.
- Inability to communicate—the subject cannot speak the language or dialect of the technician or any available interpreter.
- Pain or disability—the subject cannot take in a deep breath and/or rapidly exhale down to full expiration.
- Voluntary refusal—the subject will not participate because of fear or other reasons.

Instructions to subjects, therefore, are standardized for the initial trials and follow standard variations for subsequent trials depending on observations of the technicians—observations made by watching the subject perform the maneuver and by monitoring oscilloscope displays of the flow and volume signals of all completed trials for that subject. The skilled technician quickly perceives difficulties from these two sources and redirects the subject to perform a correct maneuver. A number of examinees tested in NHANES I did exhibit pain or discomfort while performing the test or indicated the presence of an upper respiratory infection. With the assistance of the resident physician, such subjects were disqualified from the examination. Regarding those who refused to take the test, their reasons were fully documented and will be examined for nonresponse bias.

In summary, the test requires both a technician-spirometer interaction to achieve accurate and reliable signals and a technician-subject interaction to achieve subject comprehension and motivation. These two interaction areas define technician skill. A review of technician performance in the field, however, revealed occasional drift in performance; therefore, retraining procedures were routinely implemented to reduce this source of error.

Spirometry data quality control.—The attending technician is responsible for spirometric

data quality: Direct observations can be made during the performance of the test and observed errors can be corrected during the procedure. Clearly, the technician has the cardinal role in data quality control because he or she provides clear and concise test instructions, coaches the examinee to perform a maximal expiratory maneuver, and provides an initial judgment of the acceptability of the data obtained.

The technician can carry out this role by proper use of the monitoring equipment and careful observation of the subject. A memory oscilloscope with an X-Y axis is regarded as a reasonably precise tool for monitoring patient's spirometric effort. Flow is registered on the Y (vertical) axis, and volume is measured on the X (horizontal) axis. Each respiratory effort results in a flow-volume curve, which is displayed on the oscilloscope and compared with subsequent curves (figure 1). The technician can thus monitor discreet changes in patient effort and cooperation by observing the shape of the curve and the height of the peak flow deflection. This monitoring information must be integrated with subject performance observations. Appendix I is a glossary of terms relating to this technician function and includes a diagram of the three phases in a normal spirogram trial (appendix figure I).

The following paragraphs describe the current criteria used by the NHANES technicians to judge data quality.[11]

Procedural error detection.—As a matter of conscientious workmanship, a technician examines each trial within a test set during its recording to identify the presence of any particular procedural error. Errors are a signal to the technician that the examinee is experiencing some problem with the test instructions either because the instructions were unclear or comprehension was inadequate. When a procedural error is identified, such as the absence of a terminal decay curve (as seen on both the flow and volume signal), the subject is reinstructed, with emphasis on that part of the instruction where the problem occurred, and a clear demonstration of the test procedure is given.

The common procedural errors that alert the technician to the possibility of an invalid trial are described below.[11] The best trials are those with the largest forced vital capacity (FVC) accompanied by the highest flow rates. Procedures for identifying the best trial are described within the section entitled "Reliability Error Detection."

Short baseline.—A short baseline can result when the technician starts the recording equipment too late, thereby not permitting establishment of a sufficient baseline, or when the subject initiates expiration before instructed to do so, thus obviating the baseline. A short baseline cannot be observed on the flow volume display but it is evident on the strip chart, as shown in figure 2.

No end-of-test plateau.—During the test procedure, subjects who have large vital capacities coupled with low terminal flow rates continue to increase their expired volumes beyond the preset recording time (9.19 seconds after the technician initiates the NHANES I recording system). This phenomenon typically occurs in subjects with chronic obstructive lung disease (COLD), although it can occur in subjects with no known disease. The strip chart, not the visual display, shows this phenomenon because the former is a 9.19-second record whereas the latter is a flow volume display that is independent of time (figure 3). The recording equipment described here does not have a manual override to permit recording volumes beyond 9.19 seconds; thus, the presence of a terminal flow is referred to as premature termination by the recorder.

Premature termination artifact.—The premature termination artifact is manifested by the absence of a typical phase III morphology of the spirometric curve, that is, a slow decay curve until residual volume is reached. Unlike a premature termination by the 9.19-second recorder, this phenomenon occurs within 9.19 seconds and is due to premature termination of the effort by the subject; thus the phenomenon is found on both the visual display and the strip chart (figure 4).

Inhalation artifact.—Inhalation artifacts are identified either by the flow-volume loop morphology depicted in figure 5 or by review of the flow signal on the recording paper and observing that the flow rate decreases below the baseline, which is followed by an increase of flow greater than 1 liter per second (1 l per

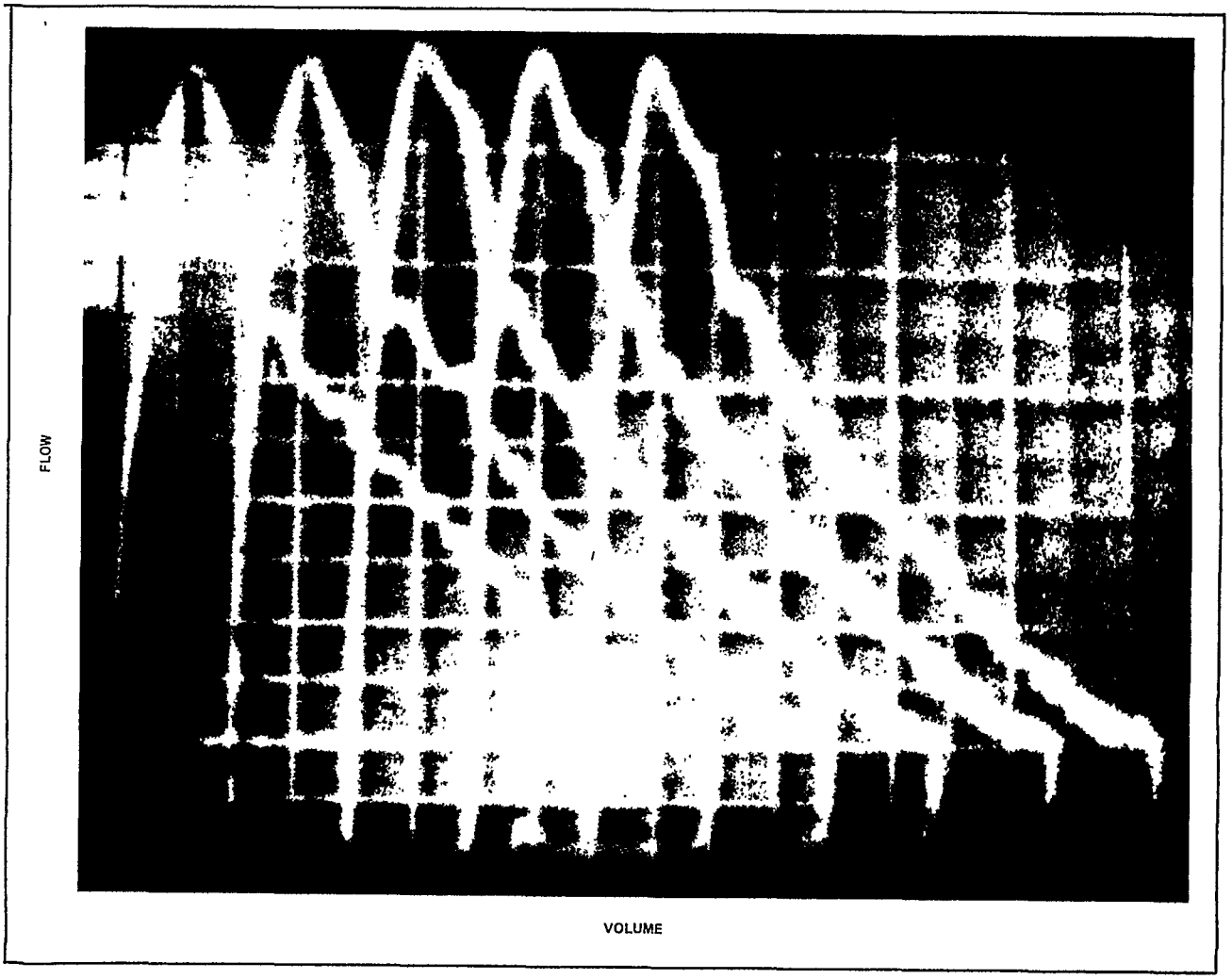


Figure 1. Typical subject flow-volume curve

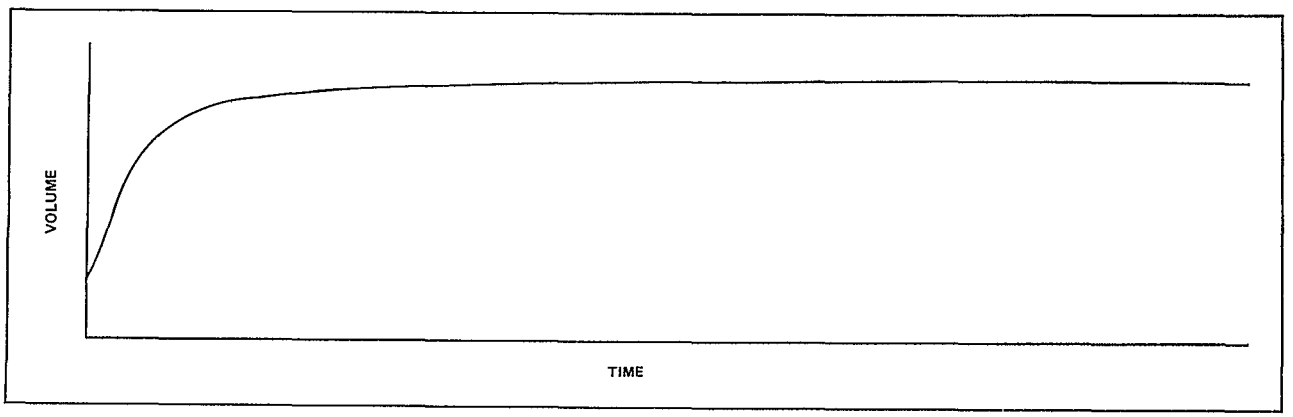


Figure 2. Sample spirogram demonstrating the short baseline procedural error

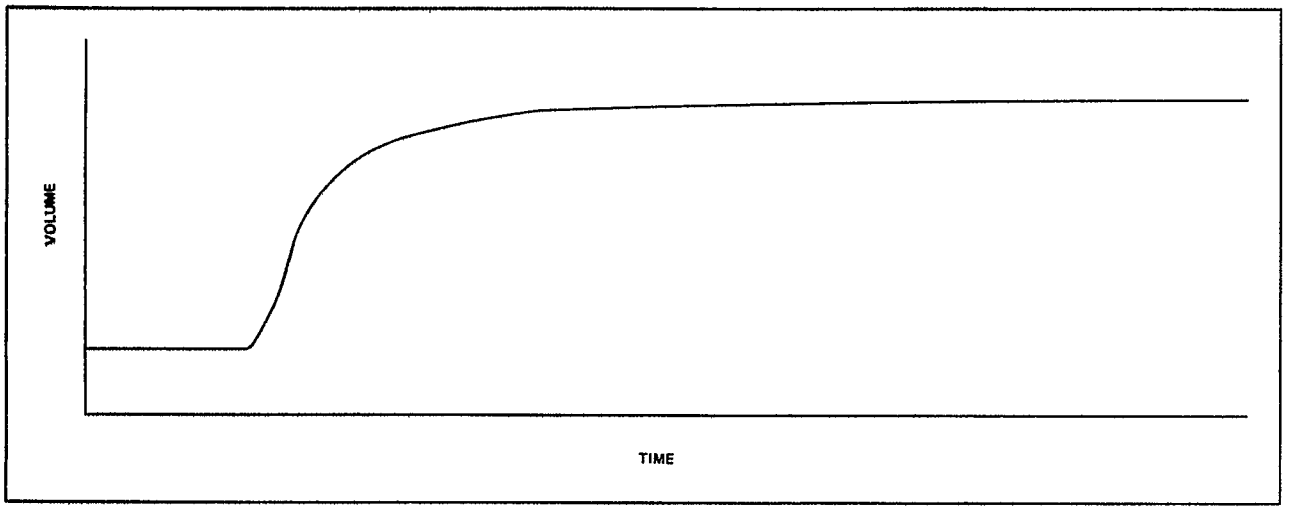


Figure 3. Sample spirogram demonstrating the no end-of-test plateau procedural error

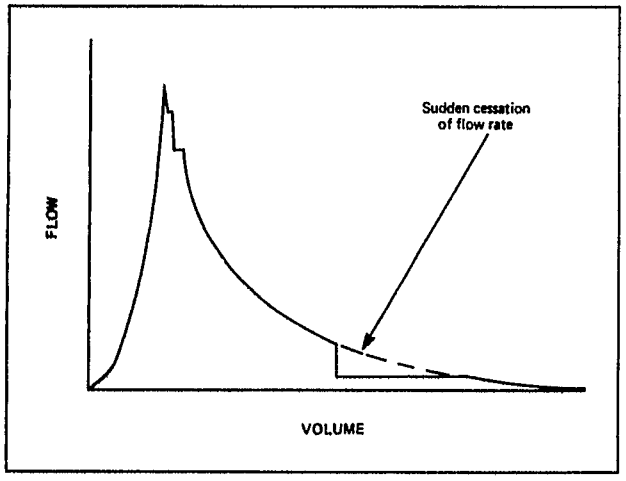


Figure 4. Sample oscilloscope tracing demonstrating the premature termination artifact procedural error

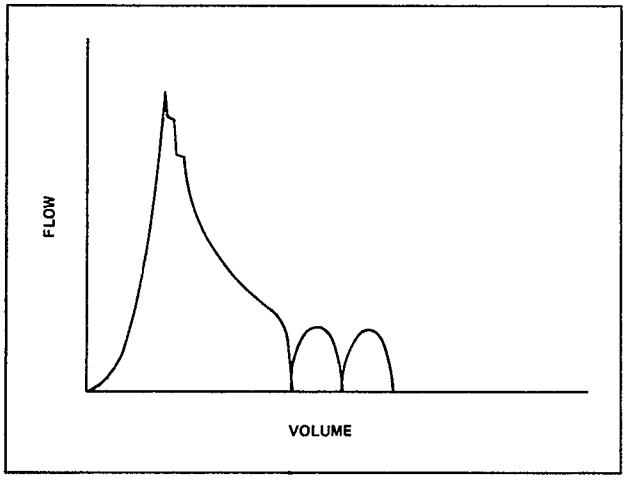


Figure 5. Sample oscilloscope tracing demonstrating the inhalation artifact procedural error

second). (These trials are automatically discarded as totally invalid and are not considered in a set of five.)

Venturi artifact.—The Venturi artifact is evident when FVC volumes and/or flow rate values are greater than clinically expected (figure 6). This phenomenon is caused by trumpeting into the mouthpiece with pursed lips, which causes room air to be drawn into the spirometer along with the expired air (figure 7). This situation occurs because of a vacuum effect from the high velocity of air movement from the pursed lips. The typical morphology of a Venturi trial is a rapid rise of flow rate to a high level that is sustained until residual volume is attained, followed by a rapid decrease to the zero line. Such an uncharacteristic trial is readily identified by a trained technician by review of the flow-volume display.

Because some members of the population (such as highly trained athletes) have extra large lung volumes and flow rates, large values can be obtained without any artifact; however, caution is required before accepting these readings. Again, if reliability criteria are met, which includes a careful review of the flow and volume histories, the test is valid.

Low peak flow artifact.—Peak flow rates of 50 percent of predicted value are sought as a measure of inital expiratory thrust. Lower values

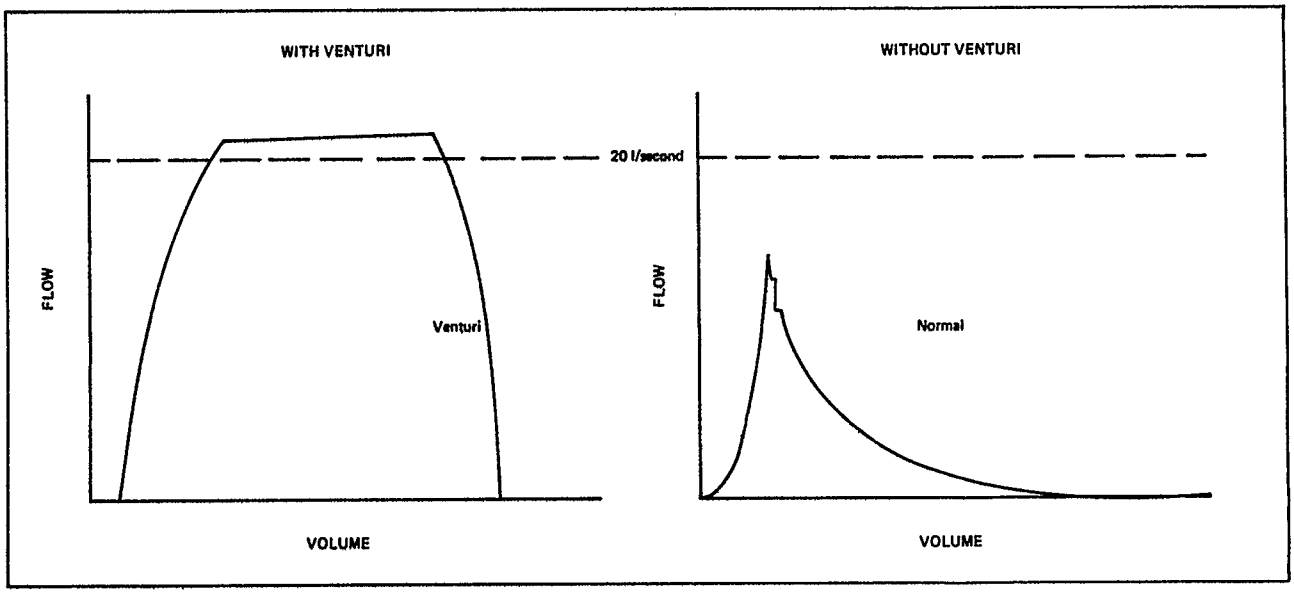


Figure 6. Sample oscilloscope tracings, one normal and one demonstrating the Venturi artifact procedural error.

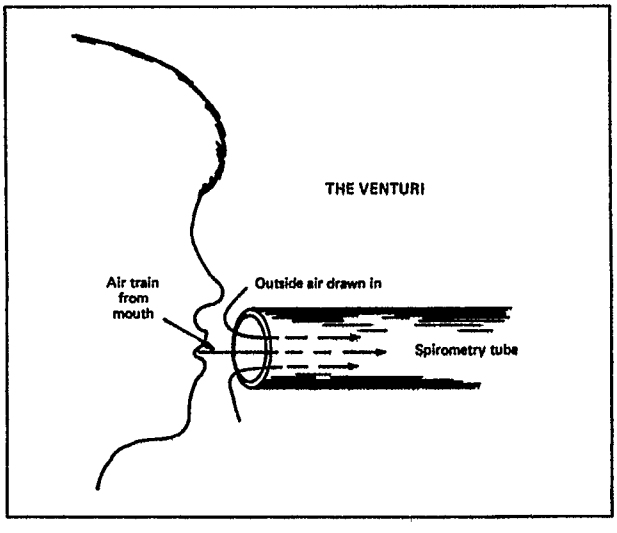


Figure 7. A depiction of the mechanics of a spirometric Venturi artifact

would indicate the possibility of malingering, not achieving total lung capacity before beginning to blow, trouble understanding the test instruction, or severe obstructive lung disease (figure 8). This possible error check is discarded if applied reliability criteria are met. No computer check is used for detecting this artifact: Detection is left to the technician who must observe both subject effort and peak flow estimates on the monitoring equipment.

Hesitation artifact.—The hesitation artifact should not occur during the three phases of the spirogram. If it does occur, the test may be considered acceptable only if the reliability criteria have been fulfilled and flow and volume histories are similar. This artifact (figure 9) is generally identified by the technician, and computer identification is limited to detection of a relatively large hesitation only at phases II and III.

Table A describes the output codes and criteria used by the computer program to flag the described procedural violations.

Reliability error detection.—Acceptable spirograms result in reproducible curves.[11,12] The technician makes an initial determination of reliability by using the monitoring equipment to superimpose one flow-volume curve over the other or, alternatively, to compare them side by side. At the conclusion of the fifth trial, the technician also examines the paper record for the two best trials. These trials are deemed reproducible if the estimates of the FVC and forced expiratory volume at 1 second ($FEV_{1.0}$) are within 5 percent, assuming that these volumes exceed 3 l or 10 percent for FVC and $FEV_{1.0}$ volumes of less than 3 l. If reproducibil-

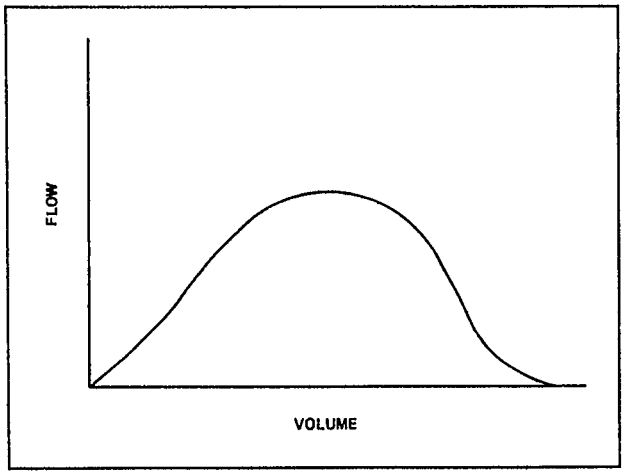


Figure 8. Sample oscilloscope tracing demonstrating the low peak flow artifact procedural error

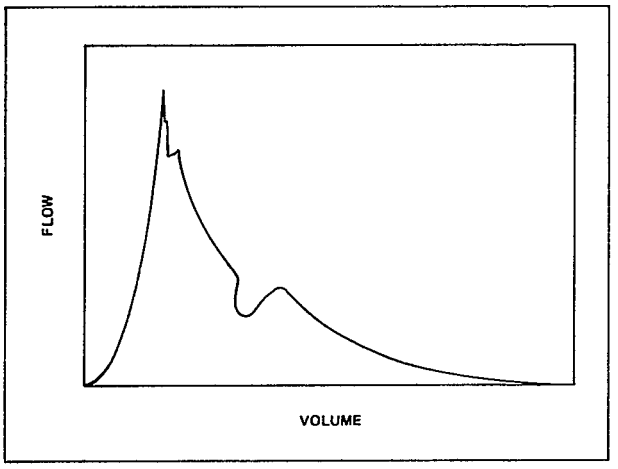


Figure 9. Sample oscilloscope tracing demonstrating the hesitation artifact procedural error

Table A. Procedural error codes and their definitions

Code	Definition
0	No violations occurred.
1	Onset of volume curve occurred less than 150 ms after the beginning of the record (short baseline).
2	End of trial (EOT) was not identified in the 9.19-second record (premature termination by recorder).
3	A volume increment of less than 4 percent between 0.5 second and 1 second after onset of the curve, or an increment between 1 and 2 seconds less than 4 percent (midtrial premature termination by subject), occurred.
4	A negative flow occurred followed by post-EOT positive flows in excess of 50 ml per second over any 0.50-second interval following EOP (inhalation artifact).
5	Peak flow was greater than 3 standard deviation units above subject's predicted peak flow (Venturi artifact).
6	Computed FVC was less than 0.2 l (invalid trial).
7	Post-peak flow but pre-EOT signal showed a marked decrease (25 percent of peak flow) in flow for a time interval of 0.1 second or more and was followed by a marked increase (25 percent of peak flow) in flow (hesitation artifact).
8	The 0.50 second of a trial after EOT had a slope in excess of 50 ml/second (premature termination at end of trial by subject).

ity cannot be demonstrated within that test set, the five-trials test sequence is repeated after the subject has rested.[4,9]

The need for technician monitoring and surveillance.—Uniformity of testing procedures was achieved in the NHANES by the use of appropriate operational procedures, care in the selection and training of technicians, and periodic retraining.

Because data collection in NHANES I extended over a 5-year period, problems of drift in technique were anticipated.[6] This drift was overcome in part by a surveillance program in which spirometry data obtained by each technician were periodically reviewed for trends in procedural and reliability errors. From this information, corrective actions were taken to reduce the continued collection of technically unsatisfactory data. This procedure was accomplished by directly observing the technician as he or she performed the testing in order to identify possible errors in technique. One aspect of this

on-site surveillance was to compare the instructions given to the subjects with the standard instructions shown in appendix II. Another aspect was the on-site review of the subjects' tracings to determine whether the technicians could make accurate judgments from the record and were able to correctly observe the flow-volume loop.

INSTRUMENTATION

The instrumentation used in the NHANES program to acquire and store the spirometry signals in a format suitable for computer analysis comprised an electronic spirometer, a storage X-Y oscilloscope to display the flow-volume curve for monitoring purposes, a single-channel linear strip chart recorder to provide a permanent record of the volume signals, and a data acquisition unit to encode, convert, and record on digital tape the spirometry volume signals. Figure 10 is a schematic representation of the system.

Spirometer and Support Electronics

Spirometry examinations were performed on an Ohio Medical Instruments Corporation model 800 electronic spirometer. This spirometer differs from the more widely used volume displacement "wet" system in that it consists of a dry metal cylinder containing a plastic-faced piston. A silastic rolling membrane forms an air-tight seal between the piston and the cylinder. The piston connecting rod is attached to a low-voltage potentiometer, which varies a fixed voltage signal in a linear manner proportional to the piston displacement.

Expired air from the forced expiratory breathing maneuver flows down the connecting hose, into the spirometer, and displaces the piston, causing the output of a signal from the potentiometer. This signal is transferred to the

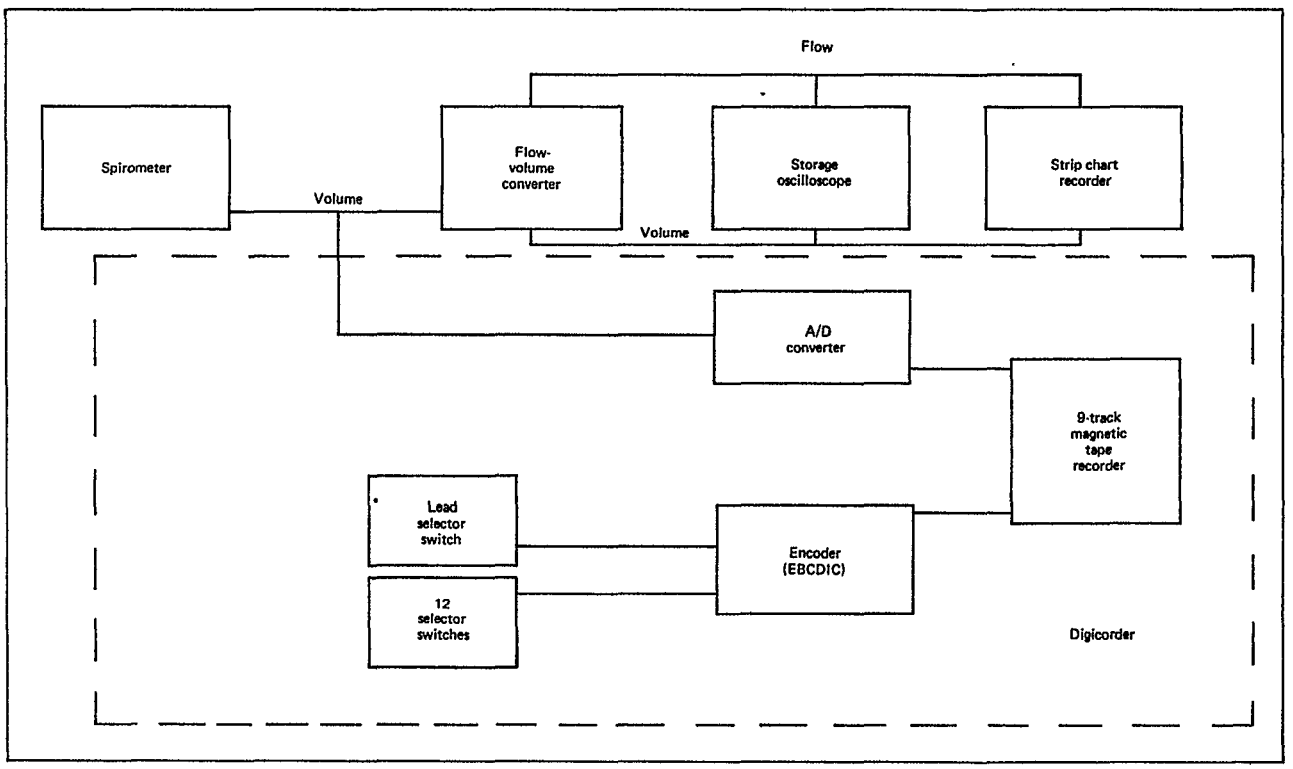


Figure 10. Schematic of the NHANES spirometry system

flow-volume converter where the signal is filtered, amplified, and is also differentiated to generate the flow signal. The outputs from the flow-volume converter are two signals of varying voltages, one of which is directly proportional to the amount of piston displacement (volume signal) and the other directly proportional to the rate of piston displacement (flow signal).

Calibrators

The spirometer is calibrated by means of an internal volume pump operated by a small electric motor that drives a single-lobed cam through a gear reduction train. When the calibrator yoke is attached to the connecting rod of the spirometer piston, and when the electric motor is engaged, the cam rider pushes the piston back and forth, causing the in-and-out movement of a known volume of room air at known flow rates. When the output volume signal is recorded on a paper tracing, as shown in figure 11, the known air movement is represented by a graphic sinusoidal signal, with the trough-to-peak distance representing the volume of air displaced. For example, if the calibrator movement causes 5,000 milliliters (ml) of air to flow in and out of the spirometer, the trough-to-peak distance on any paper recording of the volume signal will represent that 5,000 ml. The volume displacement shown in figure 11 is representative of a midrange calibration.

The Ohio spirometer electronics are preset to convert 1 ml of volume to 1 millivolt (mV); therefore, the trough-to-peak signal will be recorded on the magnetic tape as a difference of 5,000 mV from the baseline voltage. Any variation from the 5,000-mV calibration signal thus indicates either a change in the calibrator or a change in the volume-to-voltage ratio. For preliminary data processing purposes, any such observed change was assumed to be caused by the latter. For example, if the mean trough-to-peak difference was found to be 4,950 ml, a calibration factor of 1.01 was used for the subsequent spirometric analysis. Likewise, a mean difference of 5,050 ml would produce a calibration factor of 0.99.

Finally, a manual check of the program-computed standard deviation is performed, and the coefficient of variation is computed. It is assumed that a coefficient of variation greater than 3 percent indicates a daily variation great enough to warrant the use of different calibration factors for different periods of testing. Specifically, if the coefficient of variation is greater than 3 percent (that is, ±150 ml on a 5,000-ml calibration), hardware maintenance is performed and the affected data set is manually divided into smaller batches until the variation is less than 3 percent; a different calibration factor is computed manually for each batch (from the list of trough-to-peak differences printed out by the program). Spirometric analysis is performed separately for each batch. If the coefficient of variation is within the 3-percent limit, the data on that tape are considered to be a single batch.

Before conducting a spirometry test on a subject, the technician electronically calibrates the spirometer through the use of the signal generation capability of the flow-volume converter. This calibration involves switching the volume-calibration switch from its normal "operate" position to the zero position. The technician then switches between this position and a +5,000 mV (d.c.) position. This switching back and forth between 0 and +5,000 mV activates transmission of a signal that displays graphically as a square wave function, with the bottom step representing 0 mV and the top step representing +5,000 mV.

The difference between electronic and pneumatic calibrations follows. A pneumatic calibration involves all parts of the spirometry data collection system—the mechanical action of the spirometer piston, the mechanical and electronic action of the potentiometer, the amplifier and filtering circuits of the flow-volume converter, the analog-to-digital (A/D) converter and filtering circuits of the data acquisition unit, and the nine-track tape recording device. Conversely, the electronic calibration only tests the electronic portions of the instrumentation system. Thus an evaluation of the pneumatic calibration signal is an evaluation of the accuracy of the entire data collection system, whereas the evaluation of the electronic calibration signal assists the examiner in locating the source of any variation. For example, should the pneumatic variation deviate beyond certain preset limits on a given magnetic

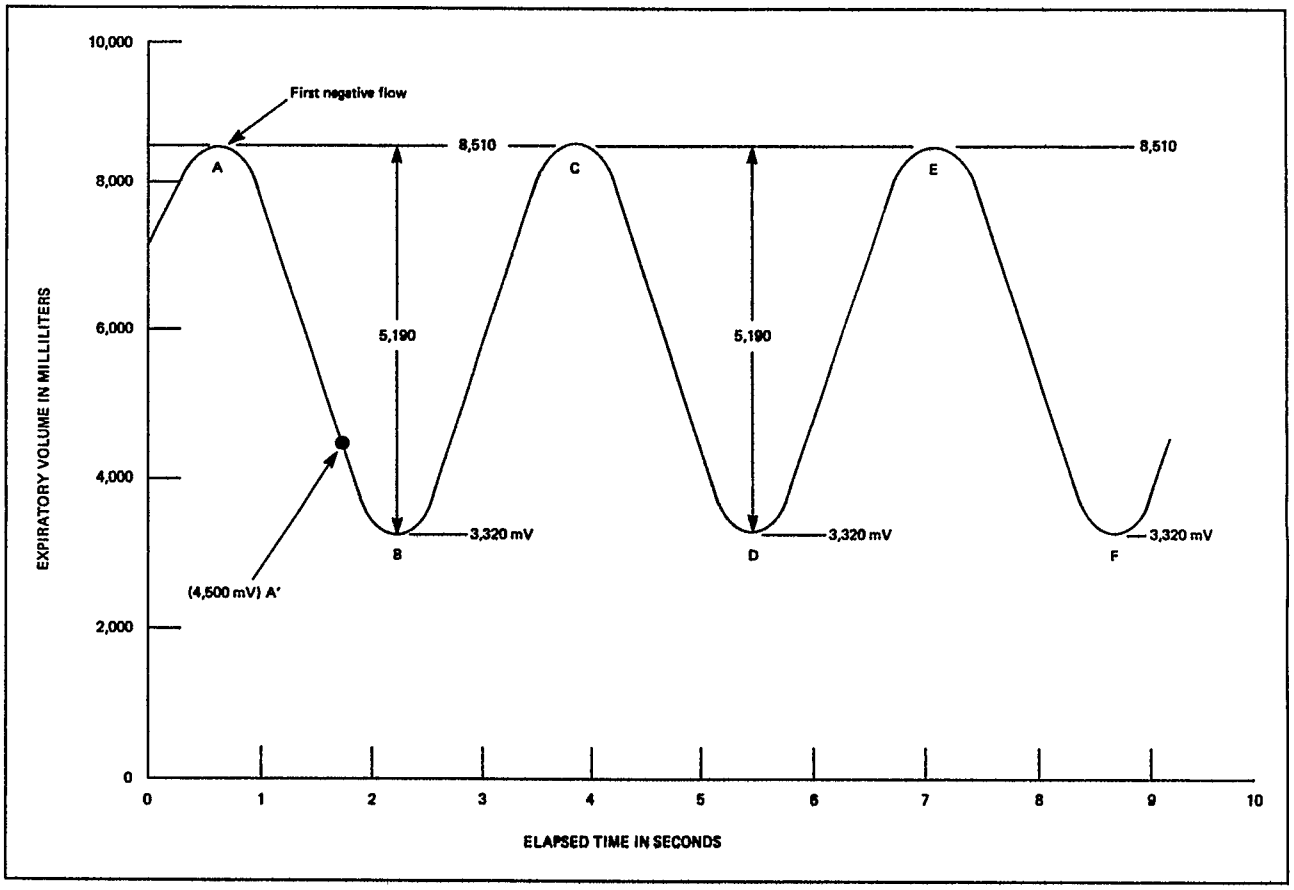


Figure 11. Spirometer data—calibration sine wave

tape, the troubleshooter would first examine the electronic calibration output for that tape. If this check revealed a consistent step-function difference of 5,000 mV, it could be assumed that the electronic portion of the system was functioning normally and that the problem lay in the pneumatic system (i.e., the spirometer itself). Conversely, should the electronic calibration show a significant step-function difference from 5,000 mV, it could be assumed that the spirometer was functioning normally and that the problem emanated from the electronic circuits somewhere in the line after the spirometer. The pneumatic calibration procedure used in NHANES I (1971-75) was not that which was recommended by the American Thoracic Society (ATS) in its Snowbird Standardization Project.[2] NHANES II (1976-80) practice did, however, follow that procedure.

Data Acquisition System

Spirometry data are recorded on a Beckman Digicorder Model No. DRS-1000 digital tape acquisition system. This unit encodes each signal with a series of pulses entered by thumb switches that the computer program identifies as the recording location, subject identification number, age, sex, race, height, technician code, barometric pressure in millimeters of mercury, and temperature in degrees Celsius. The computer uses temperature and pressure to develop a BTPS correction factor that adjusts volume from ambient temperature and pressure saturated with water vapor (ATPS) to body temperature and pressure saturated with water vapor (BTPS). A record of the machine identification number is also encoded. This encoding permits each spirogram to be traced to the machine it was

recorded on and a code (lead) number that indicates to the computer whether the signal was a calibration or an FES. Because this 14-lead data acquisition system is also used to collect a 12-lead electrocardiogram on the same subject, unique lead numbers are assigned to the spirometry examination. Data are recorded on a nine-track digital tape after conversion from analog form via an A/D converter. The tape is processed directly by a digital computer at a later date.

SPIROMETRY DATA ANALYSIS: PROGRAM DESCRIPTION

The analog spirometric signal is converted to digital data and encoded by the Digicorder and then recorded on a digital magnetic tape. Each individual data record (subject trial, calibration, etc.) consists of 18 digits representing the header and identification information, followed by 4,599 data points representing voltages (the spirometer volume curve). All data are recorded at a rate of 500 samples per second. As described below, the number of data points is reduced by computer processing to 100 samples per second, and each resulting data point has a signal resolution of approximately 2 ml of volume.

The automated computation of spirometer trial parameters is performed in three stages. In the first stage the Digicorder data tape is unpacked (reformatted) and the calibration factor (which corrects voltage-to-volume ratios) to be applied to the volume data is computed. In the second stage the flow data are computed from the volume data, the calibration and BTPS factors are applied to the volume and flow data, and the baseline is computed and removed. The third stage is the computation of spirometric parameters from the corrected data.

Calibration Factor Computation (First Stage)

A calibration data record is recognized by two conditions in the 18-digit header. The number 14 must be found in the channel lead indicator (digits 5 and 6) and at least nine 9's must be found in digits 7 through 18. The actual calibration is a sinusoidal wave with the trough-to-peak voltage difference corresponding to a 5-l volume. By computing the average trough-to-peak voltage, a ratio is formed (the calibration factor), which is later used to scale all the volume data.

When a calibration data record is found, the sinusoidal wave data are first reduced from 500 samples per second to 100 samples per second by a five-point average:

$$\bar{V}_n = (V_m + V_{m+1} + V_{m+2} + V_{m+3} + V_{m+4})/5,$$

where

$m = 5(n - 1) + 1$ and
n = the number of the averaged data point, 1-919.

This averaging reduces the data from 4,599 points per record (trial) to 919 points per record. Once the data have been averaged, the first differences (flows) are computed by the relation

$$d\bar{V}_n = \bar{V}_{n+1} - \bar{V}_n.$$

Because the sinusoidal curve may begin with a positive or a negative flow, a starting point must be determined. This point is located by observing the first negative flow with a voltage of less than 4.5 volts (V) (i.e., a point from which to start looking for the first trough). If no such point is found, the record is ignored and the next record is read. When the starting point is located, a search is made for the first positive flow. From this positive flow, the next 175 volume data points are retained (1.75 seconds; because the sine wave period is 3 seconds, this time will contain a minimum and maximum voltage). The trough-to-peak difference is computed and the next cycle is checked, beginning with a positive flow (minimum).

The trough-to-peak differences are computed cycle by cycle and record by record until the end of the data is reached. If no pneumatic calibration signals were found on the data tape, processing is terminated. If one legitimate calibration is found, the calibration factor is computed (cal = 5.0 l/D volts, where D is the average trough-to-peak voltage difference).

Figure 11 shows a typical calibration sine wave. At A, the first negative flow is encountered; however, the voltage level is above the 4,500-mV threshold. At A′, the first negative flow with a voltage less than 4,500 mV is found. The search for the next positive flow proceeds to B, where the minimum threshold of 3,320 mV is recorded. The search then continues to C, where the maximum of 8,510 mV is found. The trough-to-peak difference (B to C) is computed as 5,190 mV. A like difference is computed between D and E, and the average is 5,190 mV or 5.190 V. Thus the calibration factor is

$$\text{cal} = \frac{5.0 \text{ liters}}{5.190 \text{ volts}} = 0.9634 \text{ liters/volt.}$$

Figure 12 shows the data taken from an actual calibration trial where both the trough and peak (B and C) were examined for stability of the signal. As shown in the figure, 9 data points were recorded at 3,330, which preceded the trough, and 9 similar data points were recorded, which followed the trough; moreover, the trough voltage of 3,320 mV appeared as a continuous string of 36 samples. The figure also shows a similar stability at the peak end.

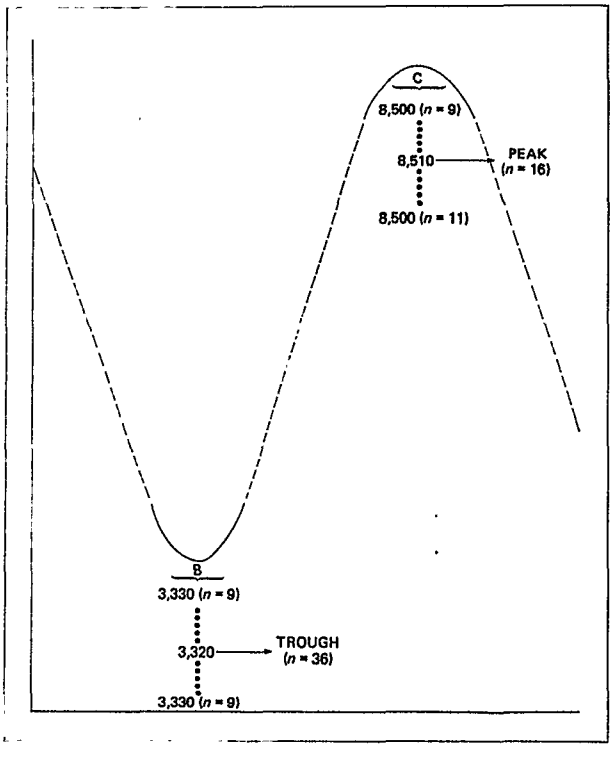


Figure 12. Digital data array from a calibration curve showing trough (B) and peak (C)

Volume and Flow-Rate Signal Data Corrections (Second Stage)

The corrections applied to the spirometric trial volume and flow rate data consist of a calibration factor, a BTPS factor, a reduction of the data sample rate from 500 samples per second to 100 samples per second, and the subtraction of the baseline (from the volume data).

BTPS correction.—The BTPS factor is computed by using the following formula:

$$\text{BTPS} = \frac{(\text{BP} - \text{PH}_2\text{O})}{\text{BP} - 47.067} \times \frac{(310.16)}{(t_k)}$$

where

BP = the barometric pressure (obtained from the header data)

t_k = the spirometer temperature in degrees Kelvin (derived from the header data)

PH_2O = a temperature-dependent water vapor pressure.

A combined correction factor is then computed by multiplying the BTPS and calibration factors.

Sample rate reduction.—The sample rate for the 4,599 volume data points (voltages) is reduced from 500 to 100 samples per second by the five-point averaging technique applied to the calibration data. The volume data are then converted from centivolts (cV) to liters, and the combined correction factor is applied to the averaged volume data. Finally, the flow rates are computed by taking the first differences of the volume data as was done with the calibration curve.

Baseline removal.—Because a volume of zero liters is not generally represented by a zero-voltage signal from the spirometer, a baseline must be determined and removed from the volume data. This baseline is defined as the average value of the points preceding the estimated beginning of the trial.

A flow threshold of 1 l per second, plus a noise tolerance, is used to estimate the beginning of a trial. (The noise tolerance is defined as 3σ, where σ is the standard deviation of all baseline data being processed and is determined as a separate computation.) The flow threshold is based on the minimum step size in the spirometer signal—which is 1 cV or 10 mV. With a combined correction factor of 1, this method would convert 10 ml at 500 samples per second. When the sample rate reduction is performed (five-point averaging), the minimum step size would be reduced to 2 ml. At a sample rate of 100 samples per second, a volume change of 2 ml would produce a flow rate of 0.2 l per second. The 1-l-per-second threshold allows for a small deviation of the baseline above the 0.2-l-per-second minimum step size. The noise tolerance is used to increase the size of the flow threshold if the baseline data are noisy. Therefore, the first flow rate to exceed the flow threshold marks the end of the baseline. (This initial estimate of zero time is refined during the third stage). The baseline digits are then averaged with the weighted average technique at the net volume point, $Avg_n = (Avg_{n-1} + Vol_n)/2$. Once the average baseline volume has been determined, it is subtracted from all volume data to remove the recorder bias. If the baseline is less than 15 points long (150 milliseconds (ms) worth of data), the trial is rejected and the data quality code is set to 1, indicating a short baseline.

Spirometer Trial Parameter Computation (Third Stage)

Peak flow determination.—After the volume and flow curves have been corrected, the flow data are searched and the largest value is recorded as the peak flow, with the corresponding volume. Predicted peaks are computed on the basis of the following formulas:

Predicted peak flow for males $= -1.0028 + (0.0474 \times \text{age}) + (0.2150 \times \text{height})$

Predicted peak flow for females $= -0.5532 + (-0.0331 \times \text{age}) + (0.1493 \times \text{height})$,

where peak flow is in liters per second, age in years, and height in inches. The data quality code is set to 5 (Venturi artifact) if the observed peak exceeds the predicted peak by at least 3.1σ (where: σ male = 1.9585, and σ female = 1.3321).

Zero time.—Once the peak flow has been determined, the zero-time (beginning of trial) estimate can be refined (figure 13). This method for determining zero time therefore replaces the initial estimate derived in the second stage. Using a triangular method for the flow curve, the zero time is corrected by the following equation:

$$t_0 = t_{peak} - 2\frac{V\,peak}{F\,peak},$$

where

t_0 = zero time
t_{peak} = time of peak (referenced to first guess zero time)
V_{peak} = volume at peak flow
F_{peak} = the peak flow rate

with the additional constraint that the corrected t_0 not precede the first guess zero time.

End-of-trial determination and FVC calculation.—After the beginning of trial is determined, the end of the trial (EOT) must be found. The EOT is found by a two-step process. First, the volume data are searched for a plateau. The

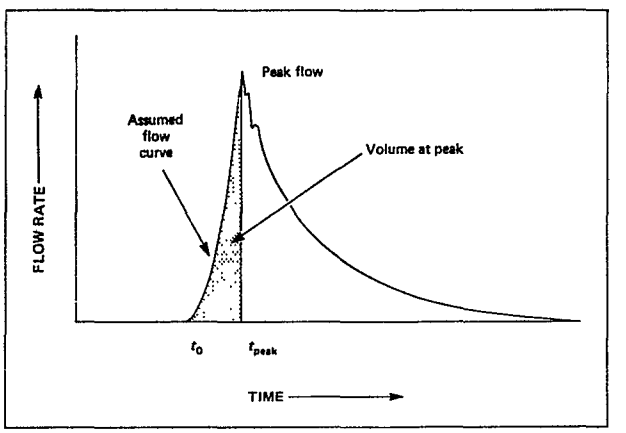


Figure 13. Schematic of flow-volume curve showing the relation of zero time to time of peak flow

plateau is said to be reached when, starting with the zero-time volume and comparing at every 10th point (every 0.1 second), the volume has not increased from the previous 0.1-second point. When a plateau is found, the time of the earliest point is recorded as the first guess of EOT, and the corresponding volume is recorded as the first-guess FVC. If a 10-point plateau is found, a search is made from this first-guess EOT to the end of the volume data (919 data points or 9.19 seconds) for the maximum volume. Current recommendations[2] are for a minimum signal duration of 10 seconds; however, design of this system preceded the development of these recommendations and the current system in use conforms to the 10-second duration of signal. If no volume is found larger than the first-guess FVC, the previously recorded FVC and EOT are used, and the data quality code is set to zero. If a larger volume is found, a search is made of the flow data between the first-guess FVC and the maximum volume. If a negative flow rate is encountered between the first-guess FVC and the higher maximum volume, the EOT is defined as the point just prior to the negative flow and the corresponding value is recorded as FVC. A negative flow is defined as any 0.01-second flow rate less than zero when the baseline σ is zero; when σ is not equal to zero, negative flow is equal to the noise tolerance (or minus 3σ). If no intervening negative flows are found, the maximum value is defined as the FVC, and the corresponding time is recorded as EOT.

A hesitation artifact (procedural error 7 in table A) is reported if a marked decrease in flow occurs after the peak flow for a 0.1-second interval, which is followed by a marked increase in flow. If EOT is found after the 10-point plateau, a check must still be made for premature termination at the end of trial. This check is done by examining the average flow rate during the 0.50-second period preceding EOT. If the average flow rate exceeds 50 ml per second, the data quality code is set to 8 (premature termination at EOT) and no further processing is performed on that trial. If the EOT is at 9.19 seconds and the average flow rate exceeds 50 ml per second, quality control code 2 (premature termination by recorder) is set and no further processing of the trial is performed. If no premature termination is found, the entire trial between zero time and EOT is searched for inhalation artifacts (negative flow rates greater than noise tolerance). If any are found, the data quality flag is set to 4 (inhalation artifact), but the processing continues on that trial.

Calculation of other parameters and quality control checks.—Once the beginning and ending of the trial have been defined and the peak flow and FVC have been determined, the other trial parameters can be computed. The forced expiratory flow rates at 25, 50, and 75 percent of FVC ($FEF_{25\%}$, $FEF_{50\%}$, and $FEF_{75\%}$) are computed from the FVC. The volume data are then searched for a forced expiratory volume of at least 0.2 l. If none is found, the trial is declared invalid (procedural error 6 in table A) and no further processing is carried out. If a volume exceeding 0.2 l is found, the corresponding time is found by linearly interpolating between that volume and the previous one. The same procedure is followed to determine the time for an FEV of 1.2 l. The $FEF_{200\text{-}1,200\ ml}$[c] is then computed as:

$$FEF_{200\text{-}1,200\ ml} = \frac{(V_{1.2} - V_{0.2})}{(t_{1.2} - t_{0.2})}$$

In a like manner, the times are determined for FEV's at 1, 2, 3, 4, 5, and 6 l. Any volumes that are not reached have their corresponding times set to 99.99. The flow rates are also recorded at the times of the various FEV values. Finally, any FEV's that exceed the $FVC_{75\%}$ have their corresponding flow rates set to 99.99.

The times and flows for 25, 50 and 75 percent of FVC are determined by locating the first volume that exceeds that value and recording the corresponding time and flow rates. The $FEF_{25\text{-}75\%}$ measurement (maximum midexpiratory flow rate (MMEF)) is then computed as:

$$FEF_{25\text{-}75\%} = \frac{FVC_{75\%} - FVC_{25\%}}{t_{75\%} - t_{25\%}}$$

[c] Forced expiratory flow rate between 200 and 1,200 ml on the volume curve ($FEF_{200\text{-}1,200}$), formerly known as the maximum expiratory flow rate (MEFR).

SUBJECT NUMBER 12-345 SEX MALE AGE=41.YEARS HEIGHT=70.INCHES WEIGHT=178.LBS. TECH. NO. 2

TRIAL	VOLUME(L)	TIME(SEC)	FLOW(L/S)	TIME(SEC)	VOLUME(L)	FLOW(L/S)	TIME(SEC)	VOLUME(L)	FLOW(L/S)
	0.2	.02	10.31	1/4	2.32	6.02	PEAK .03	.31	12.46
	1.0	.09	10.74	1/2	3.38	3.22	.10	1.43	8.81
	1.2	.11	10.74	3/4	3.93	1.29	.50	3.47	2.58
	2.0	.21	6.66	1.0	4.23	99.99	1.0	4.26	99.99
	3.0	.40	3.65	2.0	4.75	99.99	2.0	4.76	99.99
1	4.0	.80	99.99	3.0	4.97	99.99	3.0	4.97	99.99
	5.0	3.18	99.99	4.0	5.06	99.99	4.0	5.06	99.99
	6.0	99.99	99.99						
	.25FVC	.11	9.24	- ZERO TIME= 1.54			END TIME= 5.04		TIME OF FVC= 3.50
	.50FVC	.29	4.73	- FVC= 5.06 MEFR= 11.48			MMEF= 4.43		RELIABILITY CODE(S)= 8 7
	.75FVC	.67	1.50	-					
	0.2	.01	11.82	1/4	2.28	5.16	PEAK .03	.40	11.82
	1.0	.09	10.31	1/2	3.26	3.22	.10	1.42	8.38
	1.2	.11	9.45	3/4	3.88	1.72	.50	3.35	2.79
	2.0	.21	6.66	1.0	4.22	99.99	1.0	4.25	99.99
	3.0	.43	3.44	2.0	4.75	99.99	2.0	4.75	99.99
2	4.0	.83	99.99	3.0	4.95	99.99	3.0	4.95	99.99
	5.0	3.32	99.99	4.0	5.06	99.99	4.0	5.06	99.99
	6.0	99.99	99.99						
	.25FVC	.11	8.38	- ZERO TIME= 1.06			END TIME= 4.76		TIME OF FVC= 3.70
	.50FVC	.31	4.08	- FVC= 5.06 MEFR= 10.45			MMEF= 4.15		RELIABILITY CODE(S)= 8 7
	.75FVC	.71	1.93	-					
	0.2	.01	11.82	1/4	2.21	6.02	PEAK .04	.50	12.25
	1.0	.09	9.02	1/2	3.27	3.01	.10	1.46	7.52
	1.2	.11	9.02	3/4	3.90	2.15	.50	3.39	2.79
	2.0	.22	6.45	1.0	4.26	99.99	1.0	4.30	99.99
	3.0	.43	3.44	2.0	4.81	99.99	2.0	4.82	99.99
3	4.0	.80	99.99	3.0	5.03	99.99	3.0	5.04	99.99
	5.0	2.84	99.99	4.0	5.06	99.99	4.0	5.06	99.99
	6.0	99.99	99.99						
	.25FVC	.12	7.95	- ZERO TIME= .88			END TIME= 3.98		TIME OF FVC= 3.10
	.50FVC	.31	4.51	- FVC= 5.06 MEFR= 10.27			MMEF= 4.27		RELIABILITY CODE(S)= 8 7
	.75FVC	.70	1.93	-					
	0.2	.01	12.25	1/4	2.28	5.37	PEAK .04	.52	12.68
	1.0	.09	9.67	1/2	3.30	3.44	.10	1.51	8.81
	1.2	.10	8.81	3/4	3.93	1.93	.50	3.41	2.79
	2.0	.21	6.88	1.0	4.31	99.99	1.0	4.35	99.99
	3.0	.41	4.08	2.0	4.83	99.99	2.0	4.83	99.99
4	4.0	.79	99.99	3.0	5.03	99.99	3.0	5.04	99.99
	5.0	2.83	99.99	4.0	5.06	99.99	4.0	5.06	99.99
	6.0	99.99	99.99						
	.25FVC	.11	8.38	- ZERO TIME= .93			END TIME= 4.03		TIME OF FVC= 3.10
	.50FVC	.30	4.73	- FVC= 5.06 MEFR= 10.89			MMEF= 4.27		RELIABILITY CODE(S)= 8 7
	.75FVC	.69	1.93	-					
	0.2	.02	11.82	1/4	2.14	6.02	PEAK .03	.33	12.89
	1.0	.10	8.81	1/2	3.22	3.22	.10	1.33	7.52
	1.2	.11	8.59	3/4	3.88	1.72	.50	3.32	3.22
	2.0	.23	6.45	1.0	4.27	99.99	1.0	4.30	99.99
	3.0	.44	3.22	2.0	4.81	99.99	2.0	4.83	99.99
5	4.0	.82	99.99	3.0	4.96	99.99	3.0	4.96	99.99
	5.0	99.99	99.99	4.0	4.96	99.99	4.0	4.96	99.99
	6.0	99.99	99.99						
	.25FVC	.12	8.59	- ZERO TIME= 1.36			END TIME= 4.16		TIME OF FVC= 2.80
	.50FVC	.31	5.37	- FVC= 4.96 MEFR= 10.89			MMEF= 4.44		RELIABILITY CODE(S)= 8 7
	.75FVC	.68	2.36	-					

Figure 14. Data output sheet for a normal subject showing a set of 5 spirograms

SUBJECT NUMBER 12-345 SEX MALE AGE=41.YEARS HEIGHT=70.INCHES WEIGHT=178.LBS. TECH. NO. 2

TRIAL	VOLUME(L)	TIME(SEC)	FLOW(L/S)	TIME(SEC)	VOLUME(L)	FLOW(L/S)	TIME(SEC)	VOLUME(L)	FLOW(L/S)
	0.2	.07	1.72	1/4	.55	1.72	PEAK .06	.18	2.79
	1.0	.57	1.50	1/2	.92	.86	.10	.39	1.72
	1.2	.75	.86	3/4	1.20	.95	.50	1.00	.86
	2.0	2.41	99.99	1.0	1.40	.86	1.0	1.44	.86
	3.0	99.99	99.99	2.0	1.90	99.99	2.0	1.91	99.99
1	4.0	99.99	99.99	3.0	2.12	99.99	3.0	2.13	99.99
	5.0	99.99	99.99	4.0	2.17	99.99	4.0	2.17	99.99
	6.0	99.99	99.99						
	.25FVC	.25	1.72	- ZERO TIME= 3.47		END TIME= 6.87	TIME OF FVC= 3.40		
	.50FVC	.64	1.07	- FVC= 2.17 MEFR= 1.46		MMEF= .97	RELIABILITY CODE(S)= 8 7		
	.75FVC	1.36	.43	-					
	0.2	.08	1.72	1/4	.51	1.93	PEAK .05	.14	2.79
	1.0	.62	1.07	1/2	.86	1.50	.10	.34	1.07
	1.2	.82	.86	3/4	1.14	.86	.50	.92	1.50
	2.0	2.53	99.99	1.0	1.33	1.07	1.0	1.37	.86
	3.0	99.99	99.99	2.0	1.85	99.99	2.0	1.86	99.99
2	4.0	99.99	99.99	3.0	2.10	99.99	3.0	2.11	99.99
	5.0	99.99	99.99	4.0	2.16	99.99	4.0	2.16	99.99
	6.0	99.99	99.99						
	.25FVC	.26	1.72	- ZERO TIME= 2.01		END TIME= 5.41	TIME OF FVC= 3.40		
	.50FVC	.69	.86	- FVC= 2.16 MEFR= 1.34		MMEF= .87	RELIABILITY CODE(S)= 8 7		
	.75FVC	1.50	.64	-					
	0.2	.07	1.72	1/4	.52	1.93	PEAK .05	.15	2.79
	1.0	.63	.86	1/2	.85	1.50	.10	.35	1.93
	1.2	.83	.86	3/4	1.13	1.07	.50	.91	1.50
	2.0	2.38	99.99	1.0	1.34	1.07	1.0	1.38	.21
	3.0	99.99	99.99	2.0	1.86	99.99	2.0	1.89	99.99
3	4.0	99.99	99.99	3.0	2.13	99.99	3.0	2.14	99.99
	5.0	99.99	99.99	4.0	2.18	99.99	4.0	2.18	99.99
	6.0	99.99	99.99						
	.25FVC	.26	1.07	- ZERO TIME= 1.47		END TIME= 4.77	TIME OF FVC= 3.30		
	.50FVC	.72	.86	- FVC= 2.18 MEFR= 1.31		MMEF= .89	RELIABILITY CODE(S)= 8 7		
	.75FVC	1.48	.43	-					
	0.2	.08	1.93	1/4	.52	1.72	PEAK .02	.08	3.44
	1.0	.61	.86	1/2	.87	1.50	.10	.29	2.15
	1.2	.80	.86	3/4	1.16	.21	.50	.90	.86
	2.0	2.35	99.99	1.0	1.36	.86	1.0	1.38	.64
	3.0	99.99	99.99	2.0	1.88	99.99	2.0	1.89	99.99
4	4.0	99.99	99.99	3.0	2.15	99.99	3.0	2.15	99.99
	5.0	99.99	99.99	4.0	2.28	99.99	4.0	2.28	99.99
	6.0	99.99	99.99						
	.25FVC	.28	1.50	- ZERO TIME= 1.61		END TIME- 5.81	TIME OF FVC= 4.20		
	.50FVC	.75	.21	- FVC= 2.31 MEFR= 1.38		MMEF= .84	RELIABILITY CODE(S)= 8 7		
	.75FVC	1.65	.21	-					
	0.2	.08	1.72	1/4	.72	.86	PEAK .31	.79	2.58
	1.0	.51	.86	1/2	1.00	.86	.10	.90	1.50
	1.2	.71	.86	3/4	1.24	.86	.50	1.29	.86
	2.0	2.27	99.99	1.0	1.42	.43	1.0	1.60	.64
	3.0	99.99	99.99	2.0	1.92	99.99	2.0	2.01	99.99
5	4.0	99.99	99.99	3.0	2.17	99.99	3.0	2.18	99.99
	5.0	99.99	99.99	4.0	2.18	99.99	4.0	2.18	99.99
	6.0	99.99	99.99						
	.25FVC	.13	1.50	- ZERO TIME= 3.06		END TIME= 6.16	TIME OF FVC= 3.10		
	.50FVC	.60	.86	- FVC= 2.18 MEFR= 1.11		MMEF= .87	RELIABILITY CODE(S)= 8 7		
	.75FVC	1.37	.64	-					

Figure 15. Data output sheet for an abnormal subject showing a set of 5 spirograms

The volumes and flow rates are recorded at the following time points: 0.25, 0.50, 0.75, 1.00, 2.00, 3.00, and 4.00 seconds from zero time. For the times 1, 2, 3, and 4 seconds, if the corresponding volume is less than $FVC_{75\%}$, the corresponding flows are set to 99.99. Using the peak time as the reference time, volumes and flows are recorded similarly for 0.1, 0.5, 1.0, 2.0, 3.0, and 4.0 seconds after peak flow. Again, whenever the volume is greater than $FVC_{75\%}$, the corresponding flow rate is set to 99.99.

A final data quality check is performed using the FEV's at 0.5, 1.0, and 2.0 seconds from zero time ($FEV_{0.5}$, FEV_1, FEV_2). If FEV_1 is not at least 4 percent larger than $FEV_{0.5}$, or if $FEV_{2.0}$ is not at least 4 percent larger than $FEV_{1.0}$, the data quality code is set to 3 (midtrial premature termination artifact).

Figures 14 and 15 are examples of five trials of normal and abnormal subject data, respectively.

SPIROMETRY DATA ANALYSIS: VALIDITY OF ALTERNATIVE ALGORITHMS

Methodology

Five separate analyses were performed to evaluate the accuracy, consistency, and validity of the logic selected for calculating five spirometric parameters: zero time, EOT, and the three most commonly used ventilatory parameters (FVC, $FEV_{1.0}$, and $FEF_{25-75\%}$). The calculation of the latter three parameters depends directly on the determination of the first two, and in this section the ventilatory parameters are used to evaluate the performance of various algorithms for those determinations, both in the presence and absence of electronic or physical noise in the volume signal. The algorithms that were chosen as best, and the subsequent calculation of FVC, $FEV_{1.0}$, and $FEF_{25-75\%}$, are described in the previous section entitled "Spirometer Trial Parameter Computation (Third Stage)".

The first analysis consisted of comparing calculations obtained on 19 trials; the comparisons for each of the five parameters were based on three independent measurements:

1. Computer methods as described in the previous section
2. Manual calculations by technician no. 1
3. Manual calculations by technician no. 2.

Differences between the computer-derived parameters and each of the manually derived values were obtained, as well as differences between the values obtained by the two technicians on the 19 trials. The manual calculations were obtained from curves plotted from the same digital data that were introduced into the computer program and were adjusted by the correction factor after subtracting the baseline and averaging to 100 four-figure (nearest milliliter) digits each second. Figure 16 shows 1 l and 1 second measuring 11.4 and 16.8 millimeters (mm), respectively. This figure also shows volume and time-base sensitivities reasonably close to recommended minimums[2] of 10 and 20 mm, respectively.

The second analysis was a comparison of computer-derived parameters in which several alternative algorithms were used. As indicated in table B,[f] four computer methods were developed for determining zero time and four were developed for EOT identification. The initial methods for determining zero time and EOT referred to in the preceding section, entitled "Spirometer Trial Parameter Computation (Third Stage)," constitute methods 1 in table B; moreover, the definitive zero time and EOT time given in that section are methods 2 and 3, respectively. The extrapolation method (method 3) for zero time differs only slightly from the triangular method (method 2) in that the former assumes that flow from time zero to the time when peak flow occurs is equal to the peak flow rate, whereas the latter method assumes that flow averages one-half of peak flow during this short time interval and that the flow increases in

[f]Tables 1-17 showing the results of these analyses are grouped at the end of this report, preceding appendix I.

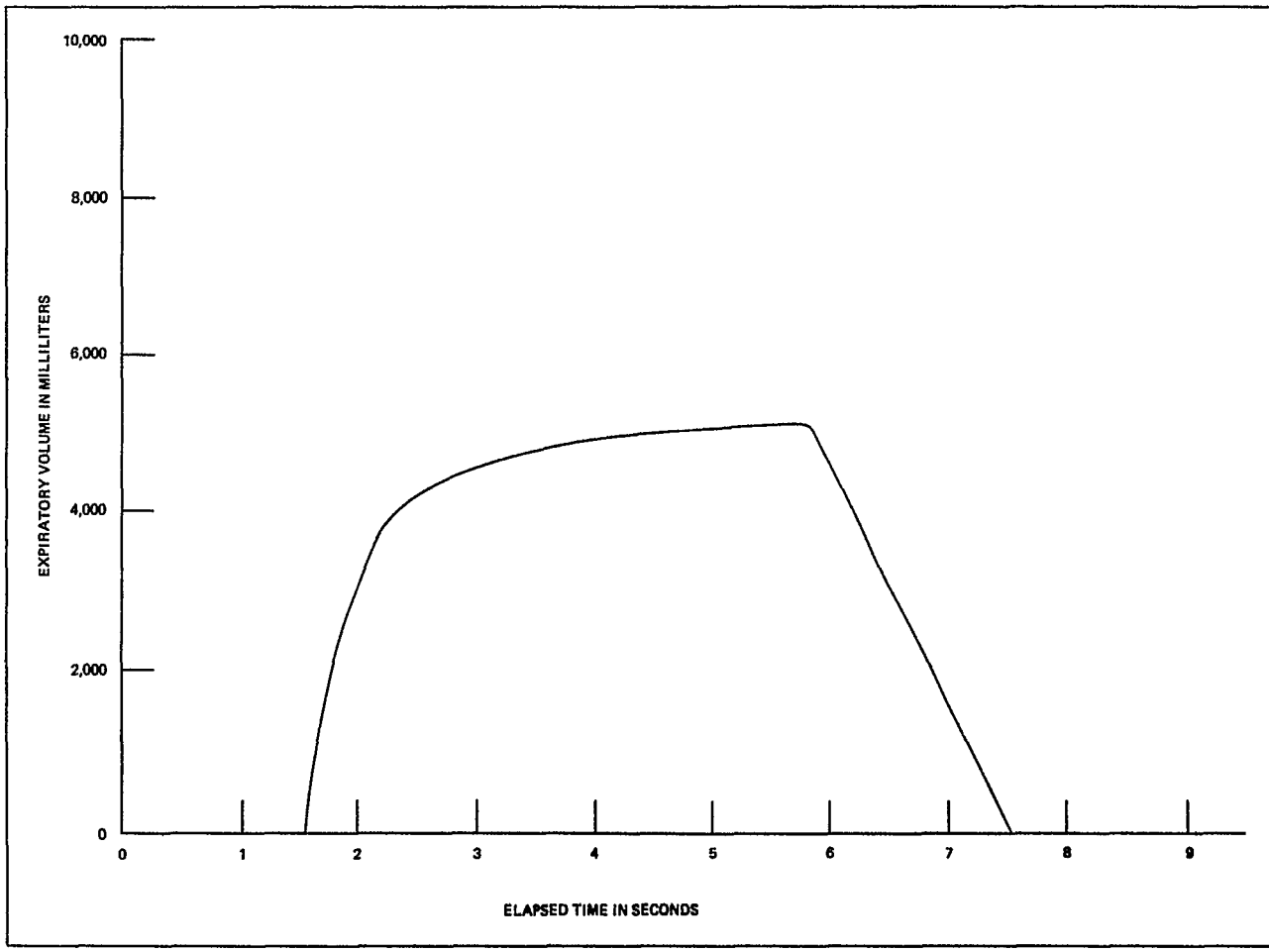


Figure 16. Example of a normal time-volume spirogram

a linear manner between zero time and time when peak flow occurs. Both methods are relatively easy to use in graphic analysis of spirograms, as well as in computer analysis. The extrapolation method for zero time has been recommended previously,[13-15] and the volume threshold (method 4) for determining zero time also has been considered previously.[13] The selection of 30 ml as the threshold (see table B) was based on the assumption that one can read graphic records within 0.5 mm with reasonable accuracy and that for most spirograms this increment would be no less than 30 ml when amplitude is converted from millimeters to milliliters. Thus zero-time comparisons include both algorithms given in the above-mentioned section plus two similar methods that have been examined before. The methods to determine EOT time also consisted of two methods described in the above-mentioned section plus method 2 (slope threshold method) previously recommended[1] for determining EOT time and method 4 (maximum volume method), which disregards any negative or zero-flow events from zero time to EOT time.

The third analysis was a comparison of these algorithms when a noise signal at each of two levels of amplitude was superimposed on the 19 trials. The first noise signal was a sine wave with an average amplitude of 2 cV (0.02 V) superimposed on the original 500-samples-per-second digitized volume signal. The sine wavelength was 0.017 seconds (60 hertz (Hz)). This superimposed noise did not significantly increase the

Table B. Methods for zero-time and end-of-test determinations

Method number and name	Curve used	Critical value(s)	Method description
ZERO TIME (t_0)			
Method 1, flow threshold	Flow	Flow $\geq$ (1 l per second + t), where t is noise tolerance (liters per second)	Search flows from beginning of data (baseline) every 0.01 second. Record t_0 as time when criteria are met: percent of flow $\geq 1 + t$ in liters per second.
Method 2, triangular (triangular extrapolation)	Flow and volume	Flow peak (largest value on flow curve) Volume at peak t_{peak}	Determine peak flow from flow data every 0.01 second; compute t_0 based on following formula: $t_0 = t_{peak} - 2\frac{V\ peak}{F\ peak}$, where flow peak is largest value on flow curve, V peak is corresponding volume, and t_{peak} is corresponding time to nearest 0.01 second.
Method 3, extrapolation (rectangular extrapolation)	Flow and volume	Flow peak Volume peak t_{peak}	Determine peak flow from flow data every 0.01 second, and compute t_0 based on following formula: $t_0 = t_{peak} - \frac{V\ peak}{F\ peak}$, where flow peak is largest value on flow curve, V peak is corresponding volume, and t_{peak} is corresponding time to nearest 0.01 second.
Method 4, volume threshold	Volume	Volume $\geq$ 30 ml	Search volume every 0.01 second for criteria to be met. Record t_0 as first volume to equal or exceed 30 ml.
END-OF-TIME (t_{EOT})			
Method 1, 10-point plateau	Volume	$(Vol_{i+10} - vol_i) \leq 0$	Starting with t_0, compare volumes every 0.1 second (10th point). Record t_{EOT} as time of vol_i when $(vol_{i+10} - vol_i) \leq 0$.
Method 2, slope threshold	Volume	$(Vol_{i+50} - vol_i) \leq 25$ ml	Starting with t_0, compare volumes every 0.01 second with an interval of 0.5 second (50th point). Record t_{EOT} as time (corresponding to vol_i) when $(vol_{i+50} - vol_i) \leq 25$ ml.
Method 3, negative flow	Flow	Flow ($-t$) liters per second where t is noise tolerance (liters per second)	Starting at t_0, compare flows every 0.01 second. Record t_{EOT} as time corresponding to $flow_{i-1}$, when $flow_i < -t$.
Method 4, maximum volume	Volume	Maximum (volume)	Starting at t_0, compare volumes every 0.01 second. Record t_{EOT} as time corresponding to vol_{max}.

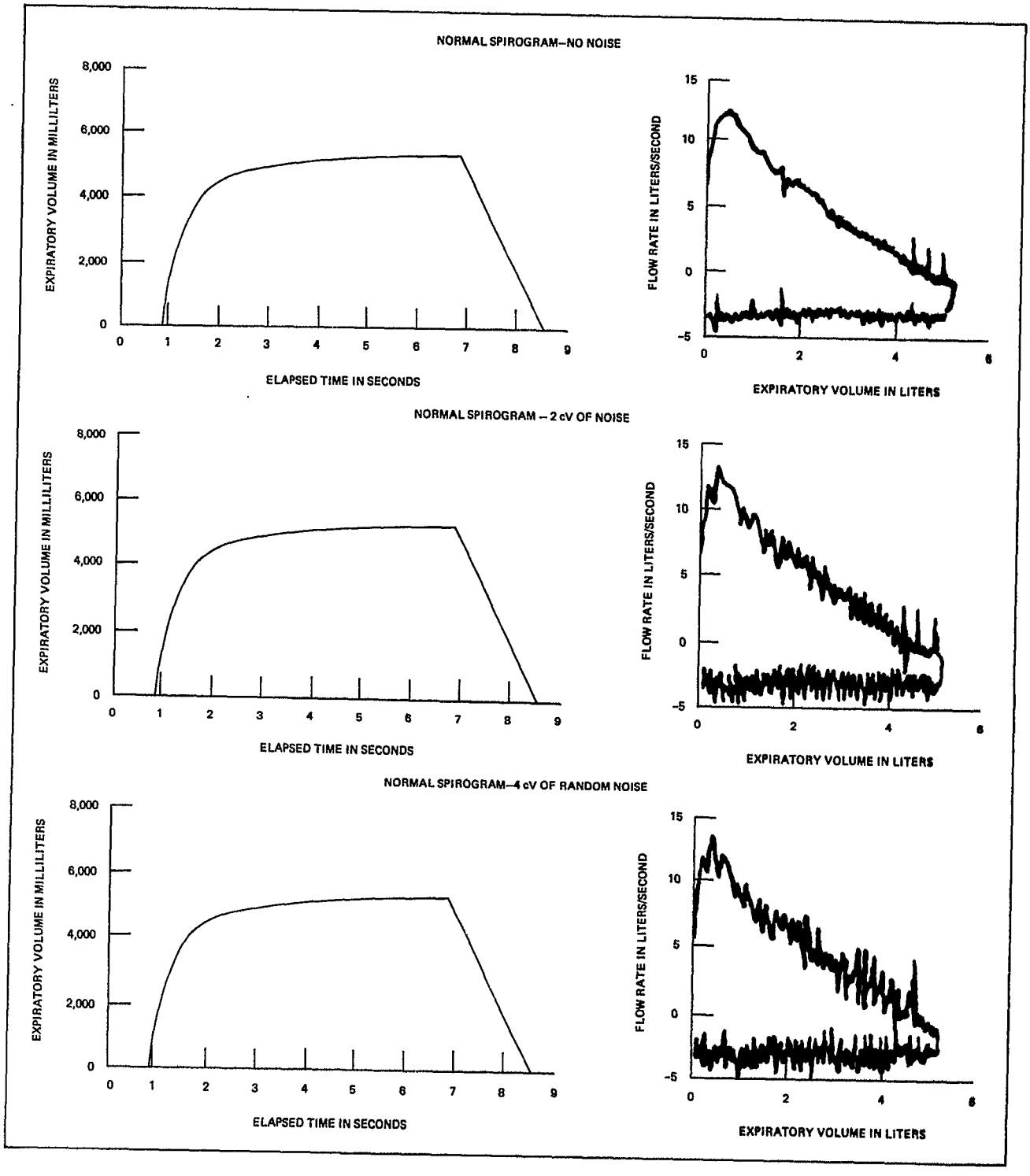


Figure 17. Normal spirogram (time-volume and flow-volume curves) with noise signals superimposed

calculations for either the flow threshold for zero time (method 1) or EOT time (method 4); however, the noise tolerance (t) was equal to 1 l per second when the second noise signal was superimposed—a 4-cV amplitude random sine wave with the same wavelength. Thus the flow threshold for t_0 was defined as 4 l per second with the 4-cV noise and the cutoff for a significant negative flow was −3 l per second with the 4-cV noise when using method 3 for t_{EOT}. Figure 17 shows the results of two levels of noise on a spirometry signal displayed as a time-volume curve and as a flow-volume curve. Note the increased visual sensitivity of the flow-volume representation to detect noise.

Because the 19 trials were obtained from four well-trained subjects with normal ventilatory parameters, the second and third analytic routines were also applied to a set of abnormal spirograms that were derived from the 19 normal spirograms by computer manipulation of the time and volume variables. Therefore, the fourth analysis was a comparison of algorithms for abnormal spirograms with superimposed noise signals. Figure 18 shows an example of a time-volume and flow-volume representation of an abnormal spirogram with superimposed noise.

The electronic spirometry system was carefully adjusted to present as noise-free a signal as possible. This entire baseline signal was examined by visually reviewing a computer listing of all 19 trials; when no noise signals were observed, it was concluded that the spirometry signal was noise free. This conclusion that the signal was relatively noise free should be understood in terms of the necessary 2-ml threshold for a trigger of one bit evident on the computer listing of the 500-signals-per-second data. (This one-bit trigger was described in the section entitled "Baseline Removal.") The computer listing showed a constant 1,050 digital voltage reading for all 44,000 digits examined at 500 samples per second over a baseline signal totaling almost 30 seconds. All four persons used in the tests were symptom free, and each entered into a brief training session on forced expiratory maneuver. Each of the four persons performed the maneuver correctly and no trials were rejected because of subject error or poor motivation. With an excellent or relatively noise-free signal and carefully trained and apparently healthy subjects, one would expect to find few quality control flags and computed values similar to manual values, especially when using rigorous algorithms for zero time and EOT. Alternative algorithms might be more suitable for "abnormal" subjects tested on equipment with a random-noise problem. For example,

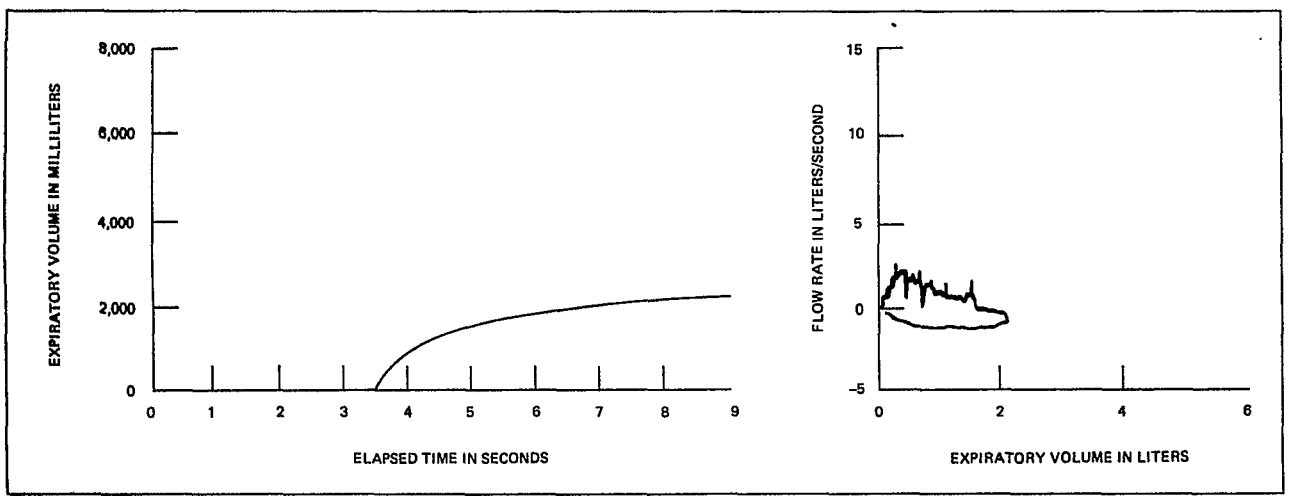


Figure 18. Abnormal spirogram (time-volume and flow-volume curves) with noise signals superimposed

when calculating zero time, a 1,000-ml-per-second flow threshold (method 1) was used initially. Zero time was defined as the interval preceding the first nonzero flow by examining each 0.01-second interval. This determination implies that a volume difference of 10 ml or more adjusted for BTPS would be the first nonzero volume when comparing successive volume signals (5 volume signals averaged to 100 signals per second). When a noise signal is superimposed on a noise-free signal, this flow threshold might be expected to be too rigorous, and an extrapolation or volume method might be preferred. Thus four alternatives, including the triangular method (method 2) described previously, were compared to determine zero time and the impact of noise signals. Similarly, various end-of-trial alternatives were compared to a rigorous negative flow algorithm (method 3), as indicated in the following section. The four end-of-trial alternatives included in this analysis include the two approaches described in the section entitled "End-of-trial determination and FVC calculation."

Zero-Time and $FEV_{1.0}$ Calculations

Zero time was initially identified as the first 0.01-second signal achieving or exceeding a 1-l-per-second flow. For manual readings, the technicians were asked to identify the first detectable departure from baseline. The technicians then measured the 1-second time and the volume at this specific elapsed time. Both time and volume measurements were attempted to the nearest 0.1 ml, recognizing that the fractions of a millimeter were only rough estimates. The computer labeled the zero-time signal as the 0.01 second that was previous to the flow threshold signal and identified the $FEV_{1.0}$ volume as the 100th baseline-corrected digital signal after zero time.

For the first trial (figure 16), the computer calculated a value of 1.50 seconds for zero time, indicating that the first significant departure from baseline occurred with digitized flow signal 151, located on the 100-sample-per-second array that had been derived by averaging the previous 500-sample-per-second array. Technician no. 1 examined the visual display of the same 100-sample-per-second digits and estimated that the first significant departure from baseline had occurred at a point measured as 1.52 seconds, or 25.5 mm of baseline from the initiation of a signal on the plot. Technician no. 2 estimated this point at 1.56 seconds, or 26.2 mm of baseline. The paired differences of zero time for this first trial were calculated as follows:

- Computer minus technician no. 1 = 0.02 seconds
- Computer minus technician no. 2 = 0.06 seconds
- Technician no. 1 minus technician no. 2 = 0.04 seconds

The average paired differences and standard deviations of the 19 differences for the above comparisons are given in table 1. The conclusion was that the two technicians had sets of zero times the means of which were not statistically different when compared as paired differences; however, technician no. 2 identified a mean zero time at an elapsed time significantly later compared with the computer-calculated mean zero time.

The $FEV_{1.0}$ manual calculation began from the manually identified zero time, and then the volume was estimated at 1 second from this point. If the manual readings were slightly biased to late zero-time estimates compared with the computer, the expectation would be that manual $FEV_{1.0}$ values would be slightly greater than the computer values. The mean computer $FEV_{1.0}$ value for 19 trials was 3.974 l (standard deviation = 0.471), whereas technician no. 1 and technician no. 2 had measured mean values of 4.002 and 3.977 l, respectively. These paired mean differences were not significantly different. Although the manual calculations by one technician disagreed slightly with the zero time obtained by the computer program, both technicians were in remarkable agreement (with an average difference of less than 30 ml) with the computer-calculated $FEV_{1.0}$ for 19 trials, and both technicians agreed on $FEV_{1.0}$ within a mean of 30 ml. Note that 30 ml is considerably more than one bit (2 ml) when the computer is

used to calculate the $FEV_{1.0}$, but only about one-third of a millimeter when volume is measured from the visual display.

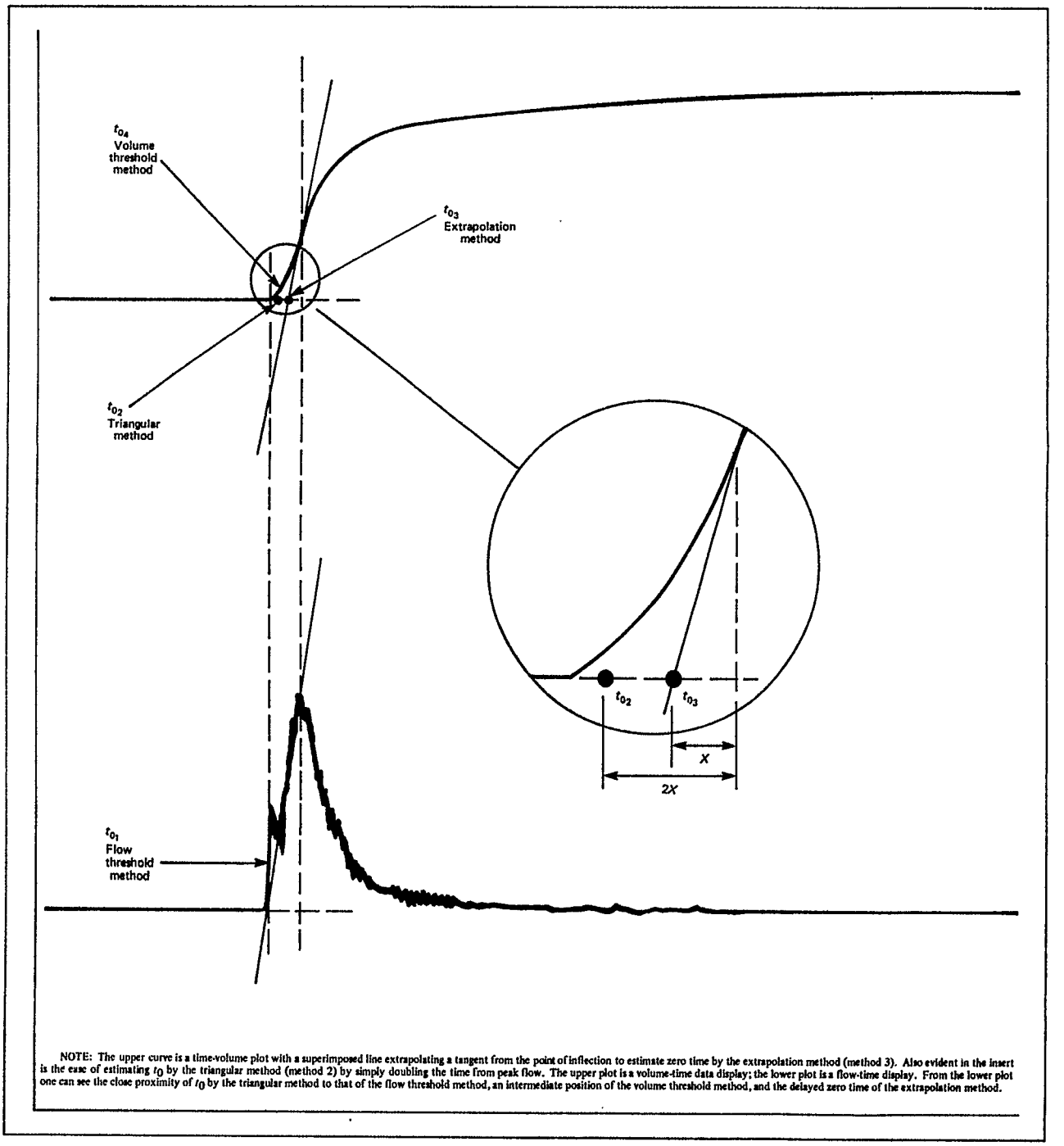
The second test compared the computer-derived values for $FEV_{1.0}$ when three alternative methods for determining zero time were used (see table B for description of methods). As just indicated for trial no. 1, the zero-time value for the flow threshold, as defined above, was 1.50 seconds of elapsed time of baseline signal; if the triangular method had been used, it would have placed zero time at 1.52 seconds. Moreover, by the extrapolation method, zero time would have been at 1.54 seconds, and by the 30-ml volume threshold method it would have been 1.53 seconds. Table 2 presents mean differences of zero-time values for all 19 trials. Thus the extrapolation method (method 3) resulted in a significant shift toward later times, but no significant differences were found for the triangular and the volume threshold methods when compared with the threshold method. It was predicted that the extrapolation method would result in excessive mean $FEV_{1.0}$ values for the 19 trials when compared with the other three methods. Table 3 presents the mean $FEV_{1.0}$ results. The paired mean $FEV_{1.0}$ differences and standard deviations were calculated; significant differences from the flow threshold method were obtained for both the volume threshold ($p < 0.01$) and the extrapolation ($p < 0.05$) methods, as indicated in table 4. This first comparison of methods indicates that (1) zero time and $FEV_{1.0}$ are in remarkable agreement for the flow threshold method and the triangular method as well as in agreement with manual calculations, (2) the volume threshold is of intermediate agreement with the flow threshold method, and (3) the extrapolation method would appear to yield a sizable bias for trial estimates of zero-time and $FEV_{1.0}$ testing for subjects with normal spirograms and a noise-free spirometry system. Figure 19 illustrates each of the four zero-time computations, showing a representative delay in zero time that results from the extrapolation method. As indicated from data in table 4 this delay results in an average error in $FEV_{1.0}$ in the range of 1-2 percent.

After the four methods were compared by using the same 19 signals, a third set of comparisons with a noise signal superimposed on the test signal was performed. For trial no. 1, table 5 gives the shifts in $FEV_{1.0}$ that occurred with a 2-cV amplitude sine wave (slight noise), and with a 4-cV random-amplitude sine wave (greater noise). No significant shift occurred in the mean $FEV_{1.0}$ for 19 trials with increasing noise when the flow threshold method was used; that is, the no-noise $FEV_{1.0}$ was 3.974, the slight noise $FEV_{1.0}$ was 3.971, and the greater noise $FEV_{1.0}$ was 3.979. These differences were considered negligible when the 3.974 value for the noise-free trial was used as the basis for comparison; the extrapolation method (method 3) showed significant effects on $FEV_{1.0}$, as table 6 indicates. Methods 1, 2, and 4 were reasonably similar to one another when the mean-paired differences were used as a basis for comparison. A close examination of the spirometry curves showed that the noise signal often increased the magnitude of peak flow and delayed its occurrence; thus the extrapolation method would yield increased $FEV_{1.0}$ values.

The fourth test was an attempt to have the computer create abnormal curves from the 19 trials by mathematically reducing the volumes at any one time point by first multiplying the volume by 0.5 and then extending the time required for each expired volume by multiplying the time scale by 3.0, as figure 20 shows. For example, the $FEV_{1.0}$ for the first trial was reduced as follows:

- Normal—Trial no. 1 computer-calculated $FEV_{1.0} = 4.19$ l
- Abnormal—Trial no. 1 computer-calculated $FEV_{1.0} = 1.33$ l

The mean of 18 of the 19 values from the transformed trials was 1.251 l. (By performing this mathematical transformation, the baseline was also extended in time; in the process, trial no. 7 was rejected because the baseline was too short. Analysis of 18 abnormal trials was undertaken, however, because the loss of trial no. 7



NOTE: The upper curve is a time-volume plot with a superimposed line extrapolating a tangent from the point of inflection to estimate zero time by the extrapolation method (method 3). Also evident in the insert is the ease of estimating t_0 by the triangular method (method 2) by simply doubling the time from peak flow. The upper plot is a volume-time data display; the lower plot is a flow-time display. From the lower plot one can see the close proximity of t_0 by the triangular method to that of the flow threshold method, an intermediate position of the volume threshold method, and the delayed zero time of the extrapolation method.

Figure 19. Manual calculation of t_0 showing a representation of the 4 alternative methods

would have no bearing on our inferences.) The zero-time and $FEV_{1.0}$ calculations used for these 18 trials again were based on the flow threshold method, and the outcome was compared with the other methods. Table 7 gives the three mean paired differences and standard deviation for zero time and $FEV_{1.0}$ by using the results of the noise-free signal as a basis for comparison. Noteworthy differences, in the range of an error of 3 percent for $FEV_{1.0}$, were

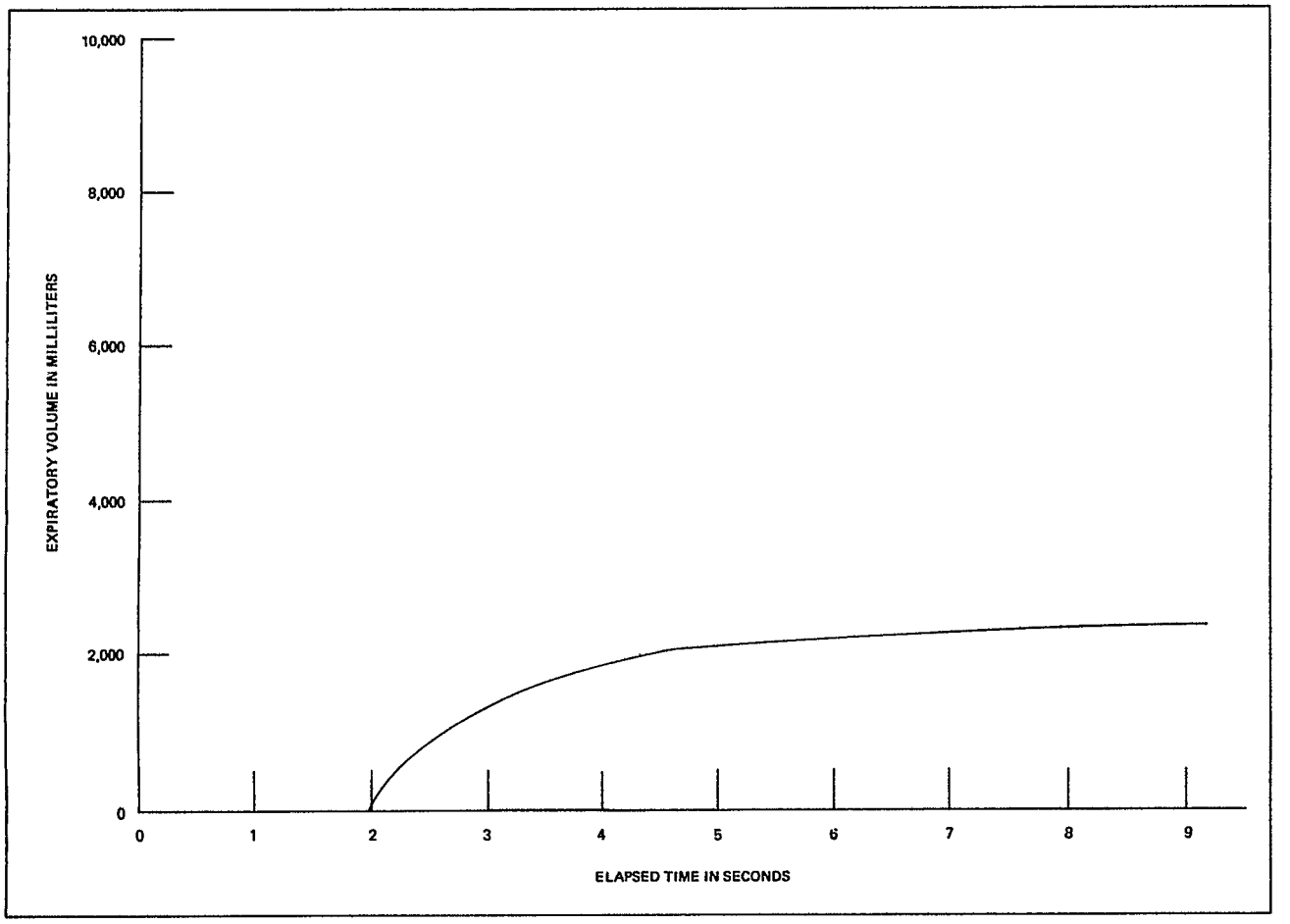


Figure 20. Example of an abnormal time-volume spirogram

found when the flow threshold results were compared with those obtained by direct extrapolation from peak flow (method 3); however, the triangular method of extrapolation from peak flow (method 2) yielded results similar to the flow threshold results. The volume threshold method (method 4) was compared with the flow threshold method and produced paired differences that were statistically significant but were not in the error range as noted above for the extrapolation method.

In the fifth and final test, the introduction of 2- and 4-cV sine wave random noise had no consistent effect on zero time compared with the noise-free flow threshold method. Table 8 presents the results. The shift in zero time was a consistent delay in the t_0 from that obtained on the noise-free signal when the flow threshold method was used as the index for these abnormal spirograms. The extrapolation method (method 3) continued to show the greatest paired difference when compared with the flow threshold index; moreover, at both noise levels the volume threshold method yielded a smaller mean difference for zero time when compared with the triangular method of extrapolation from peak flow. The flow threshold method showed an interesting zero time shift with noise, as follows:

- A′—Flow threshold method with noise-free signal minus flow threshold method with 2-cV noise = −0.076.
- A″—Flow threshold method with noise-free signal minus flow threshold method with 4-cV noise = +0.046.

Table 9 shows how the shift in zero time obtained by various alternative methods and superimposed noise signals combine to yield errors in $FEV_{1.0}$, by using again the mean $FEV_{1.0}$ from 18 abnormal spirograms obtained by the flow threshold method and a noise-free signal as the index for comparison. The paired differences are expressed in table 9 both in liters and as a percentage of the index mean $FEV_{1.0}$. The only method that showed an error in excess of 5 percent was the extrapolation method (method 3), which overestimated $FEV_{1.0}$ in the range of 8-9 percent. The conclusion was that although normal spirograms analyzed by the extrapolation method would result in an error of 1-2 percent, much greater errors are found when abnormal spirograms with or without noise are analyzed with this method. The only method that showed an average error in $FEV_{1.0}$ less than 3 percent was the volume threshold method (method 4). The flow threshold method applied to abnormal spirograms showed a much greater mean $FEV_{1.0}$ change with noise (−38 and +46 ml for the two noise signals given in table 9) compared with the data given previously for normal spirograms and the same two noise signals (−3 and +15 ml). By comparing the results in table 6 with those in table 9, the greater impact of noise on the $FEV_{1.0}$ of abnormal spirograms is evident for all four methods. In table 6, all except the extrapolation method yielded errors less than 1 percent of the mean noise-free $FEV_{1.0}$ by the flow threshold method, although the errors shown in table 9 are 8-9 percent for the extrapolation method compared with approximately 2-4 percent for the other three methods.

In both the normal and abnormal spirograms in the presence of 4 cV of noise, the volume method (4) is affected least since it more closely approximates the threshold reference measurement (method 1) than methods 2 or 3. A close examination of the spirometry data reveals that the additional sensitivity achieved by using the triangular method is jeopardized by the need for a clear peak flow signal. This problem is severely increased in the presence of abnormal data where flow rates are relatively low and peak flow rates are not well defined. (Usually the flow curve demonstrates several points in close proximity to each other in the area around the peak flow.) In many cases, noise caused flow points following the peak flow to be elevated to a greater amplitude than the true peak flow rate, which caused the time of peak flow rate to occur later. This development generally resulted in the zero-time point occurring to the right of the threshold reference point. Method 4, which does not rely on any flow rate data to identify zero time, was relatively unaffected by the high noise levels. This finding would suggest that in a relatively noise-free system (< 2 cV of noise), method 2 is superior to all other methods, although in a system with considerable noise (4 cV or more) method 4 would be the preferred technique. Computerized spirometry data acquisition systems with noise levels similar to that simulated in this study generally should not be used to collect data because measurement accuracy would be severely compromised.

End-of-Trial, FVC, and $FEF_{25-75\%}$ Calculations

The same five tests were performed for the other end of the spirogram and the volume of FVC. Because fractions of FVC are used as the first steps to determine $FEF_{25-75\%}$ volume points, the calculation of FVC is critical to the logic for $FEF_{25-75\%}$.

The mean value for the initial estimate (10-point method—method 1) of end-of-trial time and the mean FVC were both considerably different from the mean values for end of trial and FVC obtained by the more rigorous negative flow method (method 3).[g] This difference is shown in table 10. A significant difference between these two methods is also evident for $FEF_{25-75\%}$. A smaller FVC due to early termination would be expected to result in a much higher $FEF_{25-75\%}$ because of a shift of this slope toward the left where flows tend to be greater on normal spirograms (table 10). The two technicians were asked to identify the FVC plateau and mark the beginning of this plateau as the end of test. By using this approach, the two technicians next calculated FVC and $FEF_{25-75\%}$ for each of the 19 trials. The end-of-trial time

[g]Refer to table B for a description of methods.

and the FVC values were in remarkable agreement (see table 11); however, technician no. 2 had a mean $FEF_{25-75\%}$ that was significantly higher than that of technician no. 1. The data collected by each technician were initially compared with the 19 values obtained by the more rigorous negative flow method that was applied for each of the three measurements. The mean paired difference results are given in table 12. Technician no. 1 reported $FEF_{25-75\%}$ measurements that were significantly lower than those obtained by the computer algorithm consisting of the negative flow method; otherwise, no noteworthy differences were observed.

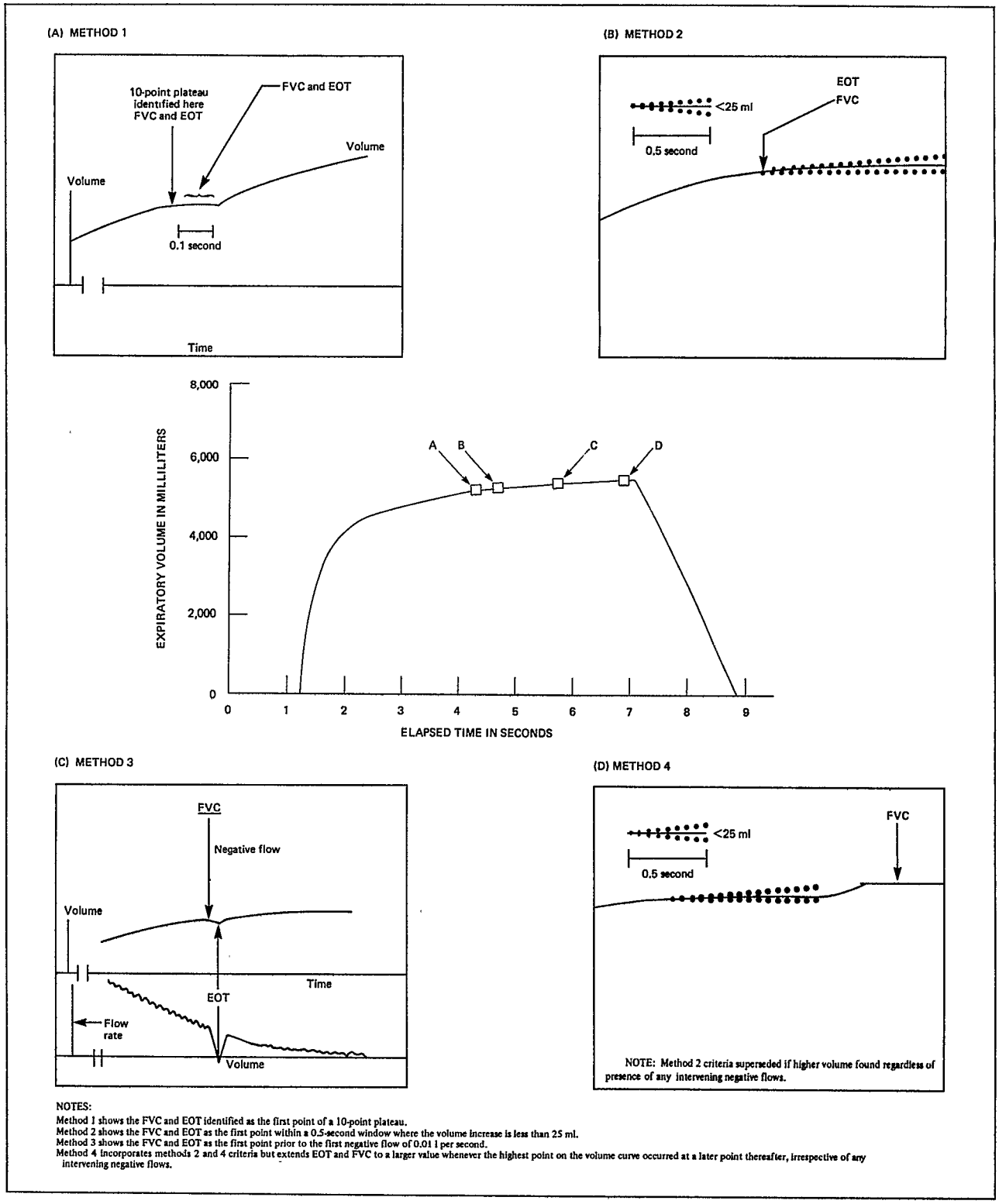
Two alternative methods to determine the EOT were next compared with the negative flow and the 10-point plateau methods. The first alternative was to consider the EOT to be the end point of the first 0.5-second interval where the average flow over this 0.5-second interval was equal to or less than 50 ml per second (method 2); the second alternative (method 4) was the same as the first, but extended the EOT to a larger value whenever the maximal FVC value occurred at a later point (irrespective of any intervening negative flows). Figure 21 illustrates each of the four EOT computations.

Data previously presented on the 19 trials in table 13 show the means and standard deviations for the four methods plus the results obtained by an average of the two manual measurements for each trial. The calculations by the two technicians agreed with the results from both the negative flow and the maximal volume methods, but were very different from the results obtained with the other two methods. It was concluded that a choice between method 3 and method 4 would be difficult. This conclusion is supported by comparisons of these two methods with the average manual measurements. Table 14 provides differences of less than 1 percent for FVC and approximately 1 percent for $FEF_{25-75\%}$. This comparison suggested that the means of the average of manual readings are approximately midway between the results of the two computer methods, perhaps somewhat closer to the maximum volume method. The mean FVC by the technicians was 15 ml more than that determined by the negative flow method and 12 ml less than that from the maximum volume method; in addition, the mean $FEF_{25-75\%}$ obtained by the two technicians was 0.054 l per second higher (+1.4 percent) for the negative flow method, but only 0.018 l per second lower for the maximum volume method. These differences between method 3 and method 4 and the average manual readings were small and not statistically significant. In a second analysis, the paired differences from the two methods and the data from each technician were compared separately (table 15). The second comparison also showed no statistically significant differences except for $FEF_{25-75\%}$, where three of the four differences were significant ($p < 0.05$), the only exception being the 0.080-l-per-second difference of the negative flow method.

The next analysis compared the four end-of-trial methods as related to the effect of the two random noise signals that were superimposed on the normal spirogram. Table 16 gives these results for EOT, FVC, and $FEF_{25-75\%}$ by comparing the mean paired differences for the noise signal obtained by the negative flow method with the other three methods. The effect on the three parameters of the 2-cV noise was to shorten the signal, reduce the FVC by a premature EOT, and increase the $FEF_{25-75\%}$. For the 4-cV random noise signal, the trend was to prolong the EOT time, increase the FVC, and decrease the $FEF_{25-75\%}$ for the negative flow and maximum volume methods; but the trends were inconsistent.

For the final analysis, abnormal spirograms were created and noise signals were superimposed, as described previously. Only 15 of the 19 trials could be interpreted at the tail end of the curve for all noise levels because of the lack of a plateau at the end of the trial when noise was superimposed. Table 17 shows the mean paired differences.

Again, the 2-cV noise signals produced consistent effects of reducing EOT and FVC and of increasing $FEF_{25-75\%}$; similarly, the 4-cV signal prominently showed relatively small reductions in FVC and $FEF_{25-75\%}$ when the negative flow and maximum volume methods were used. These two methods most closely simulated the technicians' measurements; the changes in the FVC, when noise was added to the signal, were



NOTES:
Method 1 shows the FVC and EOT identified as the first point of a 10-point plateau.
Method 2 shows the FVC and EOT as the first point within a 0.5-second window where the volume increase is less than 25 ml.
Method 3 shows the FVC and EOT as the first point prior to the first negative flow of 0.01 l per second.
Method 4 incorporates methods 2 and 4 criteria but extends EOT and FVC to a larger value whenever the highest point on the volume curve occurred at a later point thereafter, irrespective of any intervening negative flows.

Figure 21. Manual calculation of end of trial showing the 4 alternative methods

all less than 3 percent when compared with the noise-free computed values; and for $FEF_{25-75\%}$ the changes were under 5 percent.

The two methods appeared to have a reasonable accuracy when compared with technician-measured EOT time, FVC, and $FEF_{25-75\%}$. In examining the graphs of individual trials where the results of the negative flow method differed from those of the maximum volume method, locations on the spirographic curve were observed where an instantaneous negative flow could have occurred, and in many of these instances the technician may have noted a termination hesitation. The recognition of instantaneous negative flows is beyond the resolution of the human eye; however, if a negative flow continues over a duration of about 0.25 seconds, such an event would likely be identified as an end of test. In an effort to test this hypothesis a fifth method for EOT was developed, which identified the first point in time followed by an average negative flow of 0.25-second duration. The analysis is not given in this report; however, it showed no appreciable accuracy advantage with abnormal spirograms in the presence of noise. Another possible method not yet tested that may yield a more accurate estimate of relatively low FVC and $FEF_{25-75\%}$ in the presence of noise would be a protracted negative flow of 0.25 second. Therefore, either of two methods (negative flow or maximum volume) are recommended with the reservation that neither are very accurate under conditions of a high noise-signal ratio. The slope threshold method recommended in this report is different from recently published recommendations.[1,2] In this report the spirometry signal was 9.19 seconds in duration rather than the recommended minimum of 10 seconds; however, conclusions were not affected because all trials of these "normal" persons were terminated within the 9.19-second signal duration when the slope threshold was used. One subject had a terminal flow pattern suggesting that if the signal duration would have been extended to 10 or more seconds, methods 3 and 4 may have produced FVC values even larger than reported herein and the differences with method 2 would have been even greater.

SUMMARY AND CONCLUSIONS

Because $FEV_{1.0}$ and FVC are the most important ventilatory parameters, three important considerations were: (1) that these two parameters be estimated accurately by computer methods when compared with manual calculations; (2) that the measurements be reproducible with the superimposition of noise in the signal; and (3) that the sequence of computational steps proceed in a logical way, including the determination of significant digits used in the adjustment for baseline signals, BTPS conditions, and calibration information.

With a logical computation of spirometry signals, $FEV_{1.0}$ depends most on the accuracy and consistency of zero time, and FVC depends most on the accuracy and consistency of the EOT time. For zero time, any of three methods (methods 1, 2, and 4) appeared satisfactory; however, the extrapolation method (method 3) led to excessive $FEV_{1.0}$ values even without superimposed noise or when subjects had abnormal or low values. Moreover, when the latter conditions did pertain, the error of method 3 was shown to be in an 8-9-percent range. A defense of the use of the triangular method (method 2) as the definitive method used in the NHANES program can be made because some subjects have hesitation patterns on forced expiration as do certain spirometers with high inertia; both would seem to require some form of linear extrapolation from a peak flow. Method 2 was shown to be essentially equivalent to methods 1 and 4, and the triangular extrapolation can be as easily adapted to manual spirometry as the unacceptable peak flow back extrapolation method. One can calculate method 2 zero time by doubling the time from peak flow to method 3 zero time as shown in figure 19. Figure 22 provides a summary analysis of the comparative zero-time algorithms used in determining the $FEV_{1.0}$ in both normal and abnormal spirometric data, with and without noise superimposed on the signal. The means and standard errors of paired differences from the recommended algorithm (triangular, method 2) are presented as the index in this figure rather than method 1.

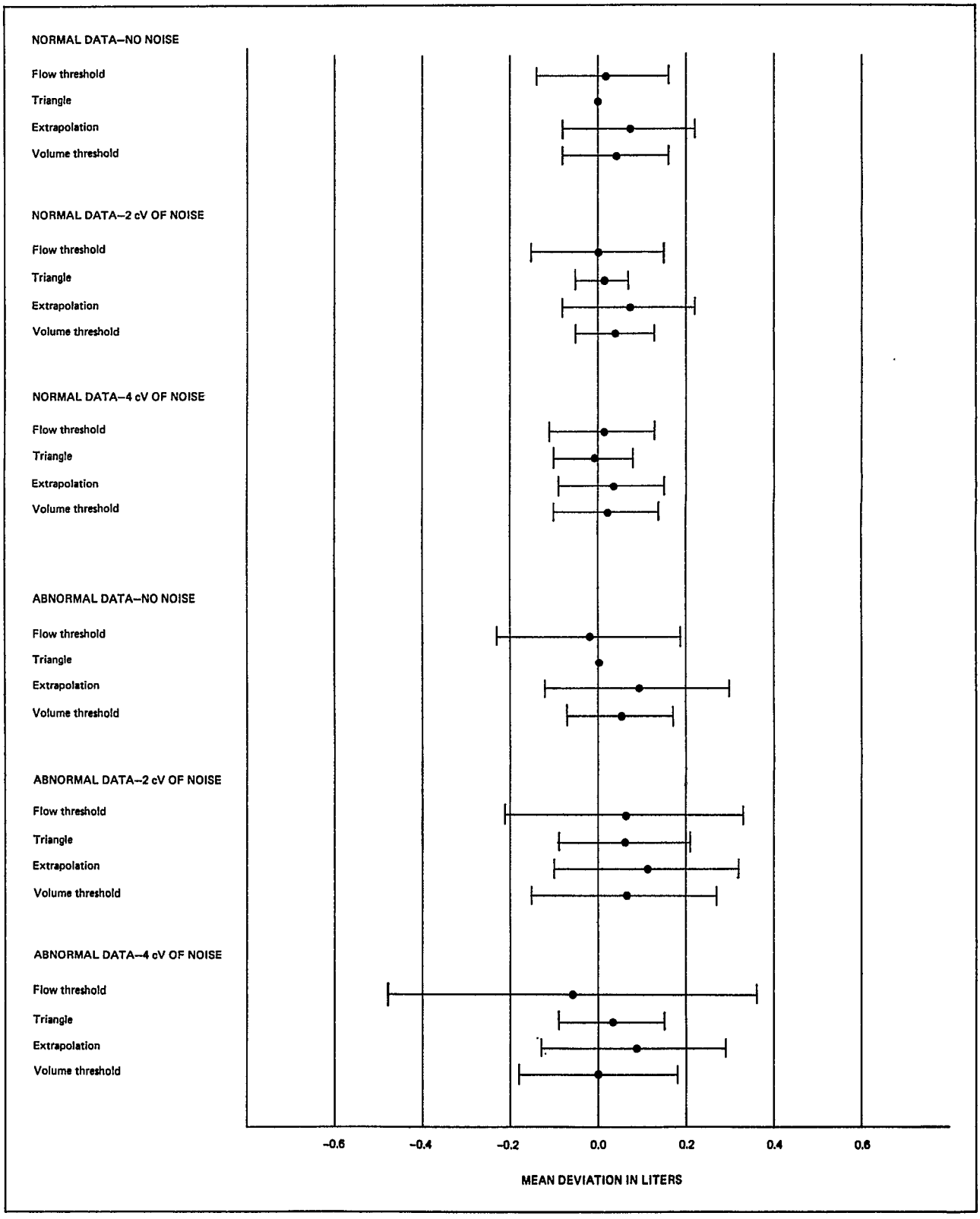


Figure 22. $FEV_{1.0}$ analysis—means and standard deviations (3σ) of paired differences from triangular extrapolation (method 2) measurements

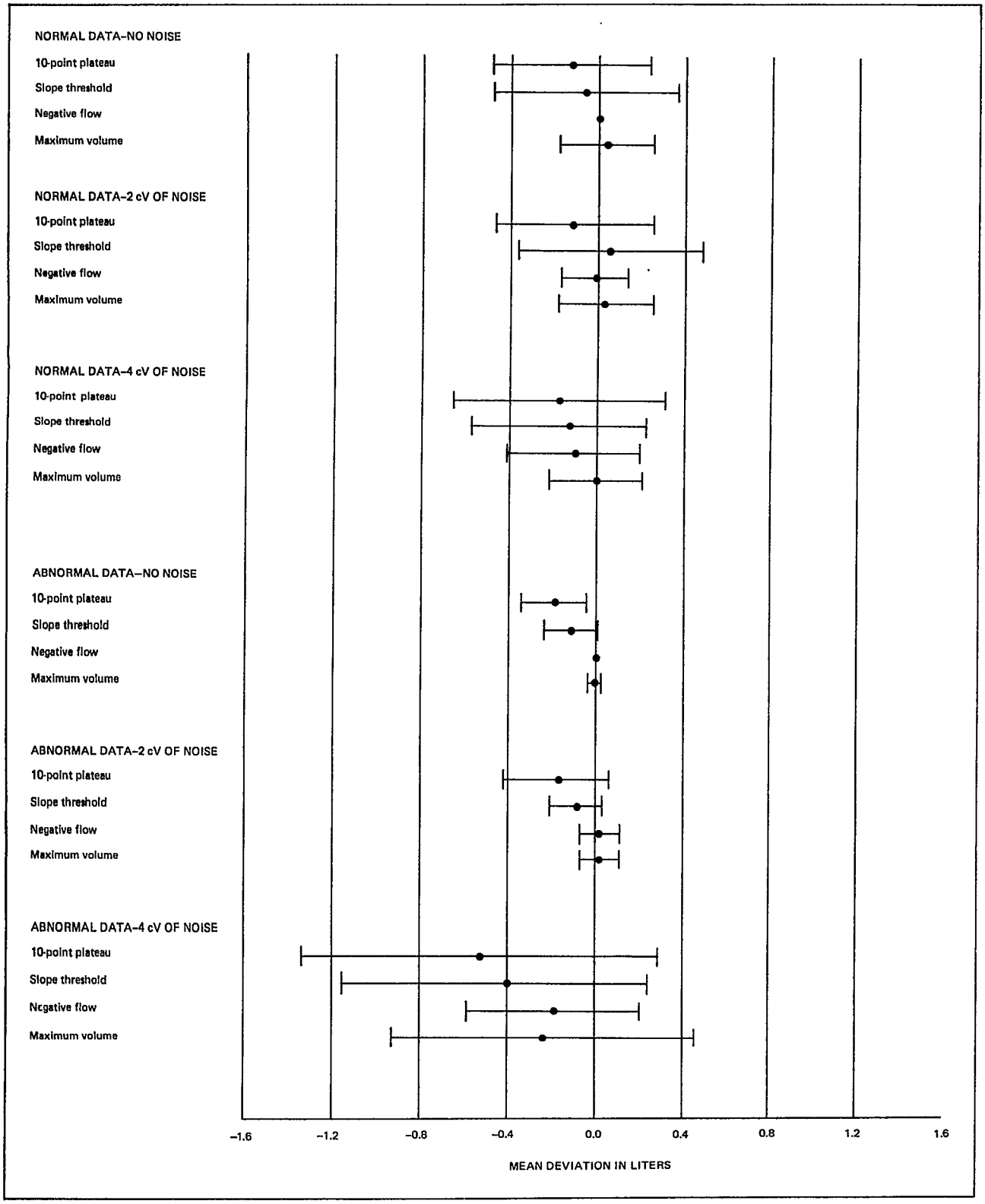


Figure 23. FVC analysis—means and standard deviations (3σ) of paired differences from negative flow (method 3) measurements

Similarly, the slope threshold method (method 2) for determining the end of trial was shown to produce an unacceptable bias in the FVC by lowering FVC by more than 50 ml compared with the negative flow method (method 3) or the maximum volume method (method 4). When either of the latter methods was examined, the FVC accuracy was within 3 percent of manual readings, even with a modest superimposed noise and with abnormal trials with FVC values in the range of 2 l. Figure 23 provides a summary analysis of the comparative EOT algorithms for both normal and abnormal data, with and without noise present on the signal. The means and standard errors of paired differences from the recommended algorithm (negative flow, method 3) are presented.

The procedures used to collect spirometric data during NHANES I represent the state-of-the-art at that time; further improvements have been made for NHANES II, reflecting advances in available instrumentation, such as recording of flow and volume data on sensitive strip-chart recorders. In the future, further refinements are anticipated by use of on-line computerized techniques that judge data quality and reliability and by use of terminal displays showing data trends to assist the attending technician with his or her acceptance decisions.

The computer program criteria described using the recommended algorithms for determination of zero time (method 2) and end-of-test detection (method 4) represent the state-of-the-art to date, superseding those recommendations given in the current NHLBI standards.[1] (These two algorithms, as described in the section entitled "End of trial determination and FVC calculation," have been implemented for current analyses.) More work is required on the development of quality control criteria to preclude the acceptance of questionable data and on development of algorithms to determine the best trial. Regarding the latter issue, this program applies the criteria specified by the ATS recommendations. However, more sensitive procedures must be explored, such as those suggested by Discher and Palmer where a 10-parameter procedure was used to determine the optimal total curve.[17]

———— ooo ————

REFERENCES

[1]Epidemiology standardization project: III. Recommended standardized procedures for pulmonary testing. *Amer. Rev. Resp. Dis.* 118:55, 1978.

[2]Report on Snowbird Work Shop on Standardization of Spirometry. *Amer. Thor. Soc.* 3:20, 1977.

[3]National Center for Health Statistics: Methodologic problems in children's spirometry. *Vital and Health Statistics.* Series 2-No. 72. DHEW Pub. No. (PHS) 78-1346. Public Health Service. Washington. U.S. Government Printing Office, Nov. 1977.

[4]Rosner, S. W., Palmer, A., Ward, S. A., Abraham, S., and Caceres, C. A.: Clinical spirometry using computer techniques. *Amer. Rev. Resp. Dis.* 94:187, 1966.

[5]Ayers, W. R., Ward, D. A., Weihrer, A. et al: Description of a computer program of the forced expiratory spirogram. Part II. Validation. *Comp. Biomed. Res.* 2:220, 1969.

[6]National Center for Health Statistics: Quality control and measurement of non-sampling error in the Health Interview Survey. *Vital and Health Statistics.* Series 2-No. 54. DHEW Pub. No. (HSM) 73-1328. Health Services and Mental Health Administration. Washington. U.S. Government Printing Office, Mar. 1973.

[7]National Tuberculosis and Respiratory Disease Association: *Chronic Respiratory Disease Screening Manual.* New York. National Tuberculosis and Respiratory Disease Association, 1968.

[8]National Center for Health Statistics: Forced vital capacity of children 6-11 years, United States. *Vital and Health Statistics.* Series 11-No. 164. DHEW Pub. No. (PHS) 78-1651. Public Health Service. Washington. U.S. Government Printing Office, Feb. 1978.

[9]Discher, D. P., Massey, F. J., and Hallett, W. Y.: Quality evaluation and control methods in computer assisted screening. *Arch. Environ. Health* 19:323, 1969.

[10]Rosner, S. W., Abraham, S., and Caceres, C. A.: Observer variation in spirometry. *Dis. Chest* 48:265, 1965.

[11]National Center for Health Statistics: *HANES Examination Staff Procedures Manual for the Health and Nutrition Examination Survey, 1971-1973.* Part 15a. Public Health Service. Washington. U.S. Government Printing Service, June 1972.

[12]Dayman, H. G.: Flow control during forced expiration. *Amer. Rev. Resp. Dis.* 95:887, 1967.

[13]Smith, A. A., and Gaensler, E. A.: Timing of forced expiratory volume in one second. *Amer. Rev. Resp. Dis.* 112:882, 1975.

[14]Kanner, R. E. and Morris, A. H.: *Clinical Pulmonary Function Testing.* Salt Lake City. Intermountain Thoracic Society, 1975. p. 1.

[15]Knudson, R. J., Slatin, R. C., Lebowitz, M. D., and Burrows, B.: The maximal expiratory flow-volume curve. Normal standards, variability, and effects of age. *Amer. Rev. Resp. Dis.* 113:587, 1976.

[16]Dischер, D. P., and Palmer, A.: Development of a new motivational spirometer—Rationale for hardware and software. *J. Occ. Med.* 14:9, Sept. 1972.

LIST OF DETAILED TABLES

Table 1. Mean, standard deviation, and significance level for paired differences between zero-time methods on 19 trials for method 1 and manual method, by 2 technicians

Pairs	Paired differences		p
	$\bar{X}_d$	s_d	
	Seconds		
Computer and technician #1	−0.007	0.037	NS
Computer and technician #2	−0.020	0.030	<0.05
Technician #1 and technician #2	−0.013	0.030	NS

NOTES: $\bar{X}_d$ = mean paired difference; s_d = standard deviation of difference; p = significance level; NS = not significant.

Table 2. Mean, mean paired difference, standard deviation of differences, and significance level for zero-time measurements, by 4 methods of computation on 19 trials

[Zero time in seconds]

Zero-time method and mean	Alternative zero-time method and mean	Paired differences		
		$\bar{X}_d$	s_d	p
1 (0.926)	2 (0.917)	−0.009	0.051	NS
1 (0.926)	3 (0.984)	+0.058	0.037	<0.05
1 (0.926)	4 (0.950)	+0.024	0.011	NS

NOTES: $\bar{X}_d$ = mean paired difference; s_d = standard deviation of difference; p = significant level; NS = not significant.

Table 3. Mean forced expiratory volume at 1 second ($FEV_{1.0}$) by 19 trials, by 4 zero-time methods

Zero-time method	Mean $FEV_{1.0}$
	Liters
1	3.974
2	3.966
3	4.034
4	4.002

Table 4. Mean paired difference, standard deviation of differences, and significance level for forced expiratory volume at 1 second ($FEV_{1.0}$), by zero-time method pairs

Zero-time method pairs	Paired differences		
	$\bar{X}_d$	s_d	p
	Liters		
1 and 2	−0.008	0.055	NS
1 and 3	+0.060	0.037	<0.05
1 and 4	+0.028	0.014	<0.01

NOTES: $\bar{X}_d$ = mean paired difference; s_d = standard deviation of difference; p = significance level; NS = not significant.

Table 5. Mean forced expiratory volume at 1 second ($FEV_{1.0}$), by 3 levels of noise and zero-time method

[1 trial]

Zero-time method	No noise	Slight noise (2 cV)	Greater noise (4 cV)
	Mean $FEV_{1.0}$ (liters)		
1	4.19	4.19	4.21
2	4.21	4.19	4.16
3	4.23	4.24	4.22
4	4.22	4.22	4.21

Table 6. Mean paired difference for forced expiratory volume at 1 second ($FEV_{1.0}$), by 3 levels of noise and zero-time method pairs

Zero-time method pairs	No noise	Slight noise (2 cV)	Greater noise (4 cV)
	Mean paired difference (liters)		
1 and 2	−0.008	+0.008	−0.022
1 and 3	[1]+0.060	[1]+0.063	+0.027
1 and 4	+0.028	+0.027	+0.008

[1] $p<0.05$.

Table 7. Mean paired difference, standard deviation of differences, and significance level for zero time and forced expiratory volume at 1 second ($FEV_{1.0}$), by zero-time method pairs on 18 abnormal spirograms

Zero-time method pairs	Zero-time paired differences			$FEV_{1.0}$ paired differences		
	$\bar{X}_d$	s_d	p	$\bar{X}_d$	s_d	p
	Seconds			Liters		
1 and 2	0.028	0.094	NS	0.022	0.073	NS
1 and 3	0.179	0.173	<0.05	0.177	0.114	<0.05
1 and 4	0.042	0.042	<0.05	0.033	0.037	<0.05

NOTES: $\bar{X}_d$ = mean paired difference; s_d = standard deviation of difference; p = significance level; NS = not significant.

Table 8. Mean paired difference between zero-time methods, by 3 levels of noise for 18 abnormal spirograms

Zero-time method pairs	No noise	Slight noise (2 cV)	Greater noise (4 cV)
	Mean paired difference (seconds)		
1 and 2	0.028	0.037	0.177
1 and 3	0.179	0.127	0.265
1 and 4	0.042	0.001	0.099

Table 9. Mean paired difference and percent difference for forced expiratory volume at 1 second ($FEV_{1.0}$) between zero-time methods, by 3 levels of noise for 18 abnormal spirograms

Zero-time method pairs	No noise		Slight noise (2 cV)		Greater noise (4 cV)	
	$\bar{X}_d$	Percent difference	$\bar{X}_d$	Percent difference	$\bar{X}_d$	Percent difference
	Liters		Liters		Liters	
1 and 2	0.022	1.76	0.049	3.92	0.056	4.48
1 and 3	0.177	9.35	0.106	8.47	0.105	8.39
1 and 4	0.033	2.64	0.029	2.32	0.021	1.68
A'	...	...	−0.038	3.04	...	...
A''	...	...	...	...	0.046	3.68

NOTES: A' = Flow threshold method with noise-free signal minus flow threshold method with 2 cV of noise; A'' = flow threshold method with noise-free signal minus flow threshold method with 4 cV of noise.

Table 10. Mean and standard deviation for end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$) with mean paired difference, standard deviation of differences, and significance level, by 2 methods of detecting the end of trial

Parameter	Method 1		Method 3		Paired differences		
	$\overline{X}$	s_x	$\overline{X}$	s_x	$\overline{X}_d$	s_d	p
	Seconds						
End-of-trial time	4.164	0.608	6.024	1.689	−1.860	1.830	<0.01
	Liters						
FVC	4.787	0.474	4.939	0.436	−0.152	0.124	<0.01
	Liters per second						
$FEF_{25-75\%}$	4.287	0.719	3.977	0.905	+0.310	0.269	<0.01

NOTES: $\overline{X}$ = mean; s_x = standard deviation; $\overline{X}_d$ = mean paired difference; s_d = standard deviation of difference; p = significance level; NS = not significant.

Table 11. Mean and standard deviation for end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$) with mean paired difference, standard deviation of difference, and significance level, by 2 technicians

Parameter	Technician no. 1		Technician no. 2		Paired difference		
	$\overline{X}$	s_x	$\overline{X}$	s_x	$\overline{X}_d$	s_d	p
	Seconds						
End-of-trial time	6.258	1.627	6.244	1.693	0.014	0.219	NS
	Liters						
FVC	4.957	0.426	4.951	0.435	0.006	0.025	NS
	Liters per second						
$FEF_{25-75\%}$	3.788	0.860	4.057	0.987	−0.269	0.433	<0.01

NOTES: $\overline{X}$ = mean; s_x = standard deviation; $\overline{X}_d$ = mean paired difference; s_d = standard deviation of difference; p = significance level; NS = not significant.

Table 12. Mean paired difference, standard deviation of differences, and significance level for end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$) between manual method by 2 technicians and method 3 of end-of-trial criteria

Parameter	Paired differences between technician no. 1 and method 3			Paired differences between technician no. 2 and method 3		
	$\bar{X}_d$	s_d	p	$\bar{X}_d$	s_d	p
	Seconds			Seconds		
End-of-trial time	0.234	1.040	NS	0.220	1.034	NS
	Liters			Liters		
FVC	0.018	0.070	NS	0.012	0.072	NS
	Liters per second			Liters per second		
$FEF_{25-75\%}$	−0.189	0.321	<0.01	0.080	0.324	NS

NOTES: $\bar{X}_d$ = mean paired difference; s_d = standard deviation of difference; p = significance level; NS = not significant.

Table 13. Mean and standard deviation for end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$), by 5 methods of detecting the end of trial

Parameter	Method 1		Method 2		Method 3		Method 4		Average of manual readings	
	$\bar{X}$	s_x	$\bar{X}$	s_x	$\bar{X}$	s_x	$\bar{X}$	s_x	$\bar{X}$	s_x
	Seconds									
End-of-trial time	4.164	0.608	5.036	1.077	6.024	1.689	6.342	1.611	6.251	1.657
	Liters									
FVC	4.787	0.474	4.878	0.482	4.939	0.436	4.966	0.444	4.954	0.430
	Liters per second									
$FEF_{25-75\%}$	4.287	0.719	4.103	0.758	3.977	0.905	3.905	0.888	3.923	0.900

NOTES: $\bar{X}$ = mean; s_x = standard deviation.

Table 14. Percentage differences in end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$) between 4 methods of detecting the end of trial and the average manual reading

Parameter	Method 1	Method 2	Method 3	Method 4
	Percent			
End-of-trial time	−33.4	−19.4	−3.6	+1.5
FVC	−3.4	−1.5	−0.3	+0.2
$FEF_{25-75\%}$	+9.3	+4.6	+1.4	−0.5

Table 15. Mean paired difference and standard deviation of differences for end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$) between manual methods by 2 technicians and methods 3 and 4

Parameter	Paired differences between technician no. 1 and:				Paired differences between technician no. 2 and:			
	Method 3		Method 4		Method 3		Method 4	
	$\bar{X}_d$	s_d	$\bar{X}_d$	s_d	$\bar{X}_d$	s_d	$\bar{X}_d$	s_d
	Seconds							
End-of-trial time	0.234	1.040	0.100	0.359	0.220	1.034	−0.086	0.325
	Liters							
FVC	0.018	0.070	0.009	0.021	0.012	0.072	−0.016	0.270
	Liters per second							
$FEF_{25-75\%}$	−0.189	0.321	−0.116	0.222	+0.080	0.324	+0.112	0.270

NOTES: $\bar{X}_d$ = mean paired difference; s_d = standard deviation of difference.

Table 16. Mean paired difference from index or negative flow method without noise for end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$), by 3 levels of noise on 18 normal spirograms and 4 end-of-trial methods

Parameter and method	No noise	Slight noise (2 cV)	Greater noise (4 cV)
End-of-trial time	Mean paired difference (seconds)		
3	...	−1.740	+0.476
4	+0.320	−0.943	+0.473
2	−0.943	−0.943	−1.427
1	−1.818	−1.762	−1.532
FVC	Mean paired difference (liters)		
3	...	−0.142	−0.030
4	+0.022	−0.057	+0.001
2	−0.061	−0.061	−0.115
1	−0.304	−0.150	−0.163
$FEF_{25-75\%}$	Mean paired difference (liters per second)		
3	...	+0.352	−0.127
4	−0.073	+0.112	−0.125
2	+0.137	+0.124	+0.726
1	−0.310	+0.112	−0.125

Table 17. Mean paired difference from index or negative flow method without noise for end-of-trial time, forced vital capacity (FVC), and forced expiratory flow rate between 25 and 75 percent of the FVC ($FEF_{25-75\%}$), by 3 levels of noise on 15 abnormal spirograms and 4 end-of-trial methods

Parameter and method	No noise	Slight noise (2 cV)	Greater noise (4 cV)
End-of-trial time	Mean paired difference (seconds)		
3	...	−3.050	+0.155
4	0.000	−2.310	+0.155
2	−2.318	−2.318	−2.446
1	−3.151	−3.064	−3.120
FVC	Mean paired difference (liters)		
3	...	−0.144	−0.050
4	0.000	−0.080	−0.050
2	+0.555	−0.083	−0.214
1	−0.370	−0.153	−0.765
$FEF_{25-75\%}$	Mean paired difference (liters per second)		
3	...	+0.134	−0.033
4	0.000	+0.089	−0.033
2	+0.074	+0.093	+0.079
1	0.000	+0.138	+0.100

APPENDIXES

CONTENTS

APPENDIX FIGURE

APPENDIX I

GLOSSARY

Trial.—The term "trial" is used to specify a single effort to perform the forced expiratory spirogram (FES) maneuver. A test set comprises five trials; however, a spirometric examination may comprise one to three test sets. The FES always has three parts as shown in figure I and the accompanying explanation.

- Phase I—effort-dependent flow, entailing peak flow for the breath.
- Phase II—constant deceleration of volume exhaled (critical flow).
- Phase III—terminal leakage flow.

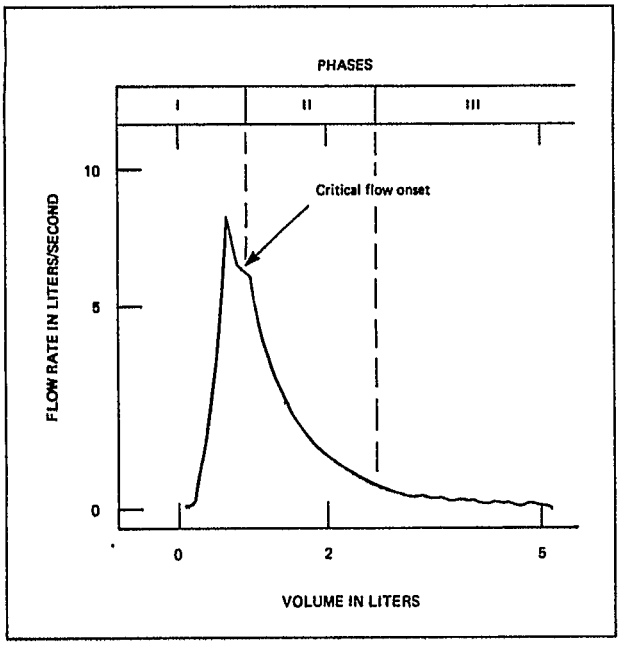


Figure I. Three phases in a forced expiratory spirogram

Procedural error.—A violation of the expected shape (morphology) and completeness of the spirometric flow-volume signal is known as a "procedural error." If artifacts are observed, the indication is that one or more of the following procedural errors has occurred during the test: an *inhalation or hesitation* during the performance of the test, a less-than-expected initial respiratory thrust, an absence of the characteristic decay portions of the flow and volume due to a premature termination of the expiratory effort before reaching residual volume, or flow or volume values higher than clinically possible, perhaps due to a Venturi error.

Detection of these artifacts is cause to challenge the validity of the test effort. For example, if an inhalation artifact is observed, the test has no clinical value and should be discarded. In contrast, if a Venturi phenomenon is observed but the flow and volume history meet the reproducibility criteria on another trial, the criteria for reliable data have been fulfilled and the test should be declared valid (in this example, the higher values observed may be because a person is exceptionally fit, e.g., an athlete).

Best trial.—The best trial refers to a particular trial, within a set of five, that clearly demonstrates the optimum presence of the two variables that are of the most interest (i.e., thrust and sustained expiratory effort). Thrust, as measured by the flow rate, is judged through the visual observation of the flow-volume curve velocities. Sustained expiratory thrust therefore results in the presence of higher flow rates throughout the spirogram, in contrast with other trials in which less thrust was applied.

Variations in thrust can be reflected by overall reductions in the peak flow rate (PFR), the

forced expiratory flows at 25 percent, 50 percent, and 75 percent of the vital capacity, the $FEF_{25-75\%}$, or in all of these. In contrast, sustained expiratory effort is judged by observation of the FVC signal on the volume (horizontal) axis of the oscilloscope display. Because all flow rates are volume dependent (except for the PFR), all flow rates must be judged in the context of the largest FVC. The best trial is the one that exhibits the highest flow rates in the presence of the largest FVC (and in the absence of obvious procedural errors).

Biologic variation often causes some variation of these values from trial to trial, such as the presence of the highest observable flow rates on a trial in which the FVC is slightly lower than in another trial, or conversely, a trial where a larger FVC contains slightly lower flow rates. During an examination, the technicians must use their best judgment to identify the best trial, taking into consideration the phenomenon of biologic variation; when the technician conscientiously applies reliability criteria (described in the following section), this problem will resolve itself.

This "eyeball" type of judgment is validated and supplemented during data processing by a digital computer at a later time when each flow and volume parameter in the spirometric curve is accurately measured and compared with like measures in other curves in the trial sequence. In this way, a consistent method of identifying the best and most reliable data is used, which reduces the probability of unacceptable data quality.

Reliability.—Reproducible spirographic tracings are assumed to represent the very best efforts of the examinee. Reliability is determined by comparing the flow and volume characteristics of the best and second-best spirograms in a test set of five trials by determining the reproducibility of the two trials. If limits of reproducibility are violated, the best trial is declared not reliable, and the test sequence is repeated.

———— ooo ————

APPENDIX II

GENERAL SPIROMETRIC TEST PROCEDURE USED BY NHANES

Instructions for Older Children and Adults[11]

Examinees follow four simple steps during the test procedure:

- Inhaling as deeply as possible (examinee maximally inhales from room air).
- Holding in all the air while placing the lips tightly around the mouthpiece, being careful to keep the tongue under the mouthpiece.
- Blasting out of all the air.
- Maintaining the expiratory effort until all the air is out.

These directives, that is, a deep inspiration, proper placement of the mouthpiece, and the forceful exhalation of air into the tube, are demonstrated to the examinee by the attending technician. The technician overemphasizes the last step by doubling up in an effort to squeeze out air by full exhalation.

Before commencing the first trial, the examinee should have the mouthpiece firmly seated in the hose with both hands around the hose an inch or so from the mouthpiece. The standard instructions for the test procedure are as follows:

- "Take in a deep breath of air, as far as you can go."
- "Hold in the air and place your lips *tightly* around the mouthpiece." (Move the hose toward the subject so he or she will insert mouthpiece; when the mouthpiece is in place, press the recorder button.)
- "Keep blowing!!! Keep blowing!!! Get it all out!!! Keep blowing!!!" By monitoring the oscilloscope, the technician can detect when the volume is not increasing and can judge when 1 or 2 seconds of effort at full expiration have transpired.

For subsequent trial instructions, the technician must repeat the last four standard instructions or use a special instruction for solving a misinterpretation based on the observations that he or she has made. A complete test set requires the recording of five spirograms.

Instructions for Children 6-10 Years of Age[11]

In the current NHANES and in previous surveys, children 6-10 years of age were tested. The following spirometry instructions were applied to the less behaviorally mature children; this group included most of the children 6 and 7 years of age (except for a few poised children with mature behavior), about half of the children 8 years of age, and the least mature of the children 9 and 10 years of age.

Technicians were instructed to proceed slowly, to be patient with the children, to use few words, and to demonstrate mostly by vivid, clear example. A typical instruction situation was

> "We want to see how your lungs work. We will measure how much air you can take in (accompanied by a strenuous demonstration) and how much you can blow out—as hard and as fast as you can."

At this point, the technician asked the child if he or she could blow up a balloon quickly. The

response usually gave the technician an indication of how ready the child was to perform a satisfactory spirogram. The instruction continued with "Let me show you":

> "Take in as much air as you possibly can. Like this."
>
> "Hold it all in until I tell you to blow."
>
> "Lips closed around the tube like this."
>
> "When I say 'blow,' blow out all of your air as fast and hard as you can. Keep blowing out hard until I tell you to stop."

The child was then given one or two practice trials in a relaxed way. If still unsure of the child's understanding and ability to perform, the technician repeated the entire demonstration, stressing anything that the child did not do perfectly (i.e., reemphasizing by use of exaggerated demonstration rather than words how to perform correctly).

A complete spirometry test set also comprised five trials. If further instruction between trials was necessary, it was vivid, pertinent, and brief. If no satisfactory data were obtained during the test, a decision was made either to have the child come back later during the examination session if the technician held hope for better trials, or mark the test as unsatisfactory if the technician believed that a satisfactory test could not be obtained because the child was frightened, too confused, or did not understand.

———— ooo ————

*U.S. GOVERNMENT PRINTING OFFICE: 1980 341-161/24 1-3

VITAL AND HEALTH STATISTICS Series

Series 1. Programs and Collection Procedures.—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions and data collection methods used and include definitions and other material necessary for understanding the data.

Series 2. Data Evaluation and Methods Research.—Studies of new statistical methodology including experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory.

Series 3. Analytical Studies.—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

Series 4. Documents and Committee Reports.—Final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.

Series 10. Data From the Health Interview Survey.—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, all based on data collected in a continuing national household interview survey.

Series 11. Data From the Health Examination Survey and the Health and Nutrition Examination Survey.—Data from direct examination, testing, and measurement of national samples of the civilian noninstitutionalized population provide the basis for two types of reports: (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

Series 12. Data From the Institutionalized Population Surveys.—Discontinued effective 1975. Future reports from these surveys will be in Series 13.

Series 13. Data on Health Resources Utilization.—Statistics on the utilization of health manpower and facilities providing long-term care, ambulatory care, hospital care, and family planning services.

Series 14. Data on Health Resources: Manpower and Facilities.—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

Series 20. Data on Mortality.—Various statistics on mortality other than as included in regular annual or monthly reports. Special analyses by cause of death, age, and other demographic variables; geographic and time series analyses; and statistics on characteristics of deaths not available from the vital records based on sample surveys of those records.

Series 21. Data on Natality, Marriage, and Divorce.—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports. Special analyses by demographic variables; geographic and time series analyses; studies of fertility; and statistics on characteristics of births not available from the vital records based on sample surveys of those records.

Series 22. Data From the National Mortality and Natality Surveys.—Discontinued effective 1975. Future reports from these sample surveys based on vital records will be included in Series 20 and 21, respectively.

Series 23. Data From the National Survey of Family Growth.—Statistics on fertility, family formation and dissolution, family planning, and related maternal and infant health topics derived from a biennial survey of a nationwide probability sample of ever-married women 15-44 years of age.

For a list of titles of reports published in these series, write to:

Scientific and Technical Information Branch
National Center for Health Statistics
Public Health Service
Hyattsville, Md. 20782

DHHS Publication No. (PHS) 81-1360
Series 2-No. 86

NCHS
U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Office of Health Research, Statistics, and Technology
National Center for Health Statistics
3700 East-West Highway
Hyattsville, Maryland 20782

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

For publications in the *Vital and Health Statistics* Series call 301-436-NCHS

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF H.H.S.
HHS 396

THIRD CLASS