The Use of Pulmonary Function Testing and Questionnaires as Epidemiologic Tools in the Study of Occupational Lung Disease*

Brian A. Boehlecke, M.D. and
James A. Merchant, M.D., Dr. P.H.

Screening has been defined as "the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly to sort out apparently well persons who probably have a disease from those who probably do not."[1] Criteria for tests useful for screening or other epidemiologic studies differ somewhat from those applicable to laboratory investigations. The equipment and procedure must be relatively simple to prevent malfunction and delays in a field setting. To obtain high rates of participation, tests must not only be safe, but also relatively free of discomfort for subjects. Test results should not be influenced by subject cooperation or observer bias, although this is never completely true for any test. Spirometry and standardized questionnaires meet these criteria quite well and have found wide use as epidemiologic tools in studying occupational lung disease. Two other attributes of screening tests are important: reproducibility or precision and validity, or the ability to correctly identify persons with the disease or condition being sought. These will be discussed in detail for spirometry and respiratory questionnaires.

Reproducibility

The reproducibility of spirometric test results is determined by two factors: (1) measurement error, caused by equipment limitations and the observer who carries out the test and measures or calculates the results, and (2) true biologic variation. Measurement errors can be reduced by setting accuracy requirements for the equipment, standardizing procedures and calculations, and requiring frequent calibration. This has recently been done for spirometry by the American Thoracic Society and the Department of Labor.[2,3,20] Highlights of the minimum equipment requirements and recommendations for test performance and calculations are given in Table 1. Thirteen of 18 currently available spirometers met the recommended accuracy requirements for measurement of the forced expiratory volume in the first second (FEV_1) when tested on a laboratory simulator.[4] Spirometric tests performed to fulfill medical surveillance requirements of an Occupational Safety and Health Administration (OSHA) standard, such as that for occupational exposure to cotton dust, now must be carried out by a technician who has attended a National Institute for Occupational Safety and Health (NIOSH) approved training course in pulmonary function testing (Table 2) to become proficient in a standardized technique for administering pulmonary function tests. When multiple tests on an individual are to be compared, biologic variation can be minimized by maintaining conditions during testing as nearly identical as possible (*eg*, same time of day). This is also important if results from two groups of subjects are to be compared. Even when spirometry is

Table 1—*Highlights of "Standardization of Spirometry" (Statement by the American Thoracic Society[2])*

Equipment (minimum requirements)
7 liter volume
Accumulate volume for 10 seconds
Volume accuracy ±3% of reading or 50 ml, whichever is greater
Flow range 0-12 L/sec
Resistance <1.5 cmH_2O/L/sec at 12 L/sec flow
Flow accuracy ±5% of reading or 0.2 L/sec, whichever is greater
Test Performance and Measurement
Minimum of 3 acceptable maneuvers, (no leaks, obstruction of mouthpiece, cough, Valsalva, or early termination)
Start point defined by back extrapolation or equivalent method
Back extrapolated volume <10% of FVC or 100 ml, whichever is greater
Best 2 of 3 acceptable curves vary for FVC by no more than ±5% of reading or ±100 ml, whichever is greater
Best FVC and FEV_1 used regardless of whether they occur on same curve

Table 2—*Highlights of NIOSH Approved Pulmonary Function Training Courses*

1. Course Director: M.D., nurse or health professional with advanced degree in pulmonary physiology or related field with training and experience with equipment and procedures.
2. Minimum of 1 instructor and spirometer system including calibrating syringe per 4 students.
3. Minimum of 16 hours of instruction with at least: a. 4 hours lecture and/or A-V material b. 8 hours small group practical instruction c. 2 hours evaluation and testing of each student's skills in spirometry testing
4. Content to include: a. Basic physiology b. Instrumentation requirements including calibration c. Performance of testing d. Measurement of tracings and calculations e. Discussion of data quality including sources of error, corrective action, reproducibility f. Supervised use of equipment
5. Application for approval to: Director, Division of Training and Manpower Development, NIOSH Robert A. Taft Laboratories 4676 Columbia Parkway Cincinnati, OH 45226

*National Institute for Occupational Safety and Health, Appalachian Laboratory for Occupational Safety and Health, Morgantown, West Virginia.
Reprint requests: Dr. Merchant, 944 Chestnut Ridge Road, Morgantown, West Virginia 26505

DOW 06818

Table 3—*Mean Intrasubject Coefficient of Variation*

Interval		FVC (%)	FEV_1 (%)	$\dot{V}_{max50}$ (%)	$\dot{V}_{max75}$ (%)
Days*		2.9	2.9	5.1	8.2
10 weeks†		5	6	11	15
1 year‡	Smokers	4	10	19	
	Non-smokers	3	9	16	

*Reference 5
†Reference 6
‡Reference 7

performed properly on accurate equipment under optimal conditions, the results vary on repeated testing. Table 3 contains the coefficient of variation (standard deviation divided by mean value and expressed as a percentage) for the forced vital capacity (FVC), forced expiratory volume in the first second (FEV_1), and the maximum expiratory flow rate after 50 percent and 75 percent of the FVC has been exhaled ($\dot{V}_{max50}$ and $\dot{V}_{max75}$, respectively) on repeat measurements in the same individual. Clearly the FVC and FEV_1 are the most reproducible. No significant difference in the reproducibility of these measurements between smokers and nonsmokers was present.

Reproducibility is especially important when the change in a measurement over a short period of time is of interest, *eg*, the change in pulmonary function over a workshift. Cochrane and co-workers[4] have stated that in short-term situations where the subject is used as his own control, changes of 6 percent of FEV_1 are significant, but at least a 19 percent change in $\dot{V}_{max75}$ is needed "to provide a similar degree of confidence." Reproducibility must also be considered in longitudinal studies. Figure 1 shows two theoretical patterns for change in FEV_1 with age in three individuals. The variability depicted in the bottom panel results from both errors in estimation of the true FEV_1 for each individual at the two ages and true differences in the rate of decline of FEV_1 between individuals. If the follow-up period is relatively short, estimation errors are of major importance.[6] Note that a cross-sectional study can provide an accurate estimate of the mean rate of change in FEV_1 with age in either situation, but cannot provide an accurate estimate of the variation between individuals or within an individual over time. The latter makes prediction of long-term changes from short periods of observation difficult. Figure 2 shows the FEV_1 of an asymptomatic smoker without other exposure to respiratory hazards, repeatedly measured in our laboratory over four years. During the first year, the decline was over 700 ml, yet without change in smoking habits, the overall rate of decline for the four years was approximately 80 ml per year. The inaccuracy of using the decline in the first year to predict long-term results is obvious.

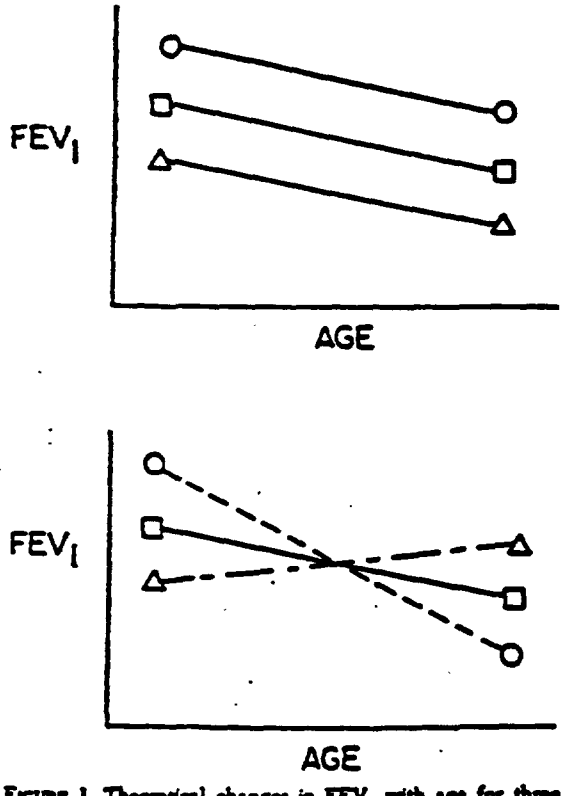

FIGURE 1. Theoretical changes in FEV_1 with age for three individuals. *Top:* Uniform rate of change with age. *Bottom:* Variable rate of change with same mean change as in top figure.

Responses to questionnaires may be used to define a condition (Monday chest tightness for byssinosis), to provide information on the presence of factors confounding the relationship between pulmonary function and the exposure being studied (*eg*, smoking and previous exposures), or to estimate total cumulative occupational exposure to an agent when quantitative measurements for an individual are not available. Interviewers must be well trained in a standardized method of administering respiratory questionnaires, if reproducible and valid responses are to be obtained. Table 4 lists some features of interviewer training.

The reporting of smoking and occupational history, although showing some variability, is generally quite

Table 4—*Training Interviewers for the Use of Respiratory Questionnaires*

1. Questionnaire and instructions should be studied in detail and difficulties discussed with an experienced interviewer or investigator
2. Interviewer should apply the questionnaire to 10 or more subjects who have some respiratory symptoms (*eg*, pulmonary clinic patients)
3. Practice interviews should be witnessed by an experienced colleague or preferably tape-recorded for correction and clarification later
4. The importance of following *exactly* the printed wording of the questionnaire must be stressed
5. Acceptable explanatory statements should be included in the instruction manual and strictly followed when needed

IST0055550

DOW 06819

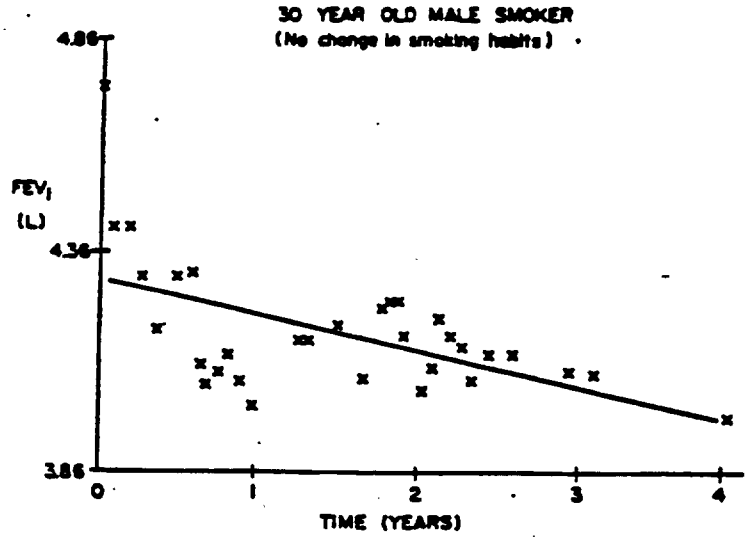

FIGURE 2. Actual measured FEV_1 in 30-year-old male smoker with no change in smoking habits or exposures over four years. × indicates measured value.

reproducible,[9] as shown in Table 5. However, average cigarette consumption is reported to be less reproducible than current consumption. This may result in unreliable estimates of "pack-years" in older subjects.[10] The reported prevalence of symptoms may be quite variable. Table 6 shows that the symptoms of cough and phlegm increased significantly over a one-year period during which the workers were exposed to an educational program concerning the risks of asbestos. Schilling and co-workers[11] compared the prevalence of byssinosis in 183 male cotton workers as recorded by two observers. Table 7 shows that complete agreement was obtained in only 76 percent of cases, but agreement on the presence or absence of disease was achieved for 93 percent of the workers. These studies point out that symptom prevalence is variable on repeat measurement even when recorded by trained interviewers using standardized questionnaires.

Table 5—*Reliability of Smoking and Occupational Histories*

	Mean Reported		
	Survey 1	Survey 2*	R
Smoking (pack-years)	23.3	24.9	.81
Years employed as pipe coverer	9.0	10.3	.95
Years asbestos exposure	10.8	11.3	.84

Adapted from reference 9
*Survey 2 performed 1 year after survey 1
R = correlation coefficient

Table 6—*Reliability of Questionnaire Responses*

	Prevalence (%) On Clinical Questionnaire		
Symptom	Survey 1	Survey 2	Agreement
Dyspnea	1.7	6.7	.93
Cough	33.3	51.7	.65*
Phlegm	36.7	56.7	.63*
Wheeze	3.4	13.4	.83
History of pneumonia	25.0	20.0	.82
Sinus trouble	26.7	25.0	.82

Adapted from reference 9
*$P < .05$

Table 7—*Prevalence of Byssinosis: Comparison of Gradings By Two Observers in 183 Male Cotton Workers Ages 40-59*

			Observer A			
			Normal	Grade I	Grade II	Total
			79	67	37	183
Observer B	Normal	78	72	6	0	
	Grade I	70	6	47	17	
	Grade II	35	1	14	20	
	Total	183				

Complete reproducibility in 72+47+20 = 139 (76%)

Agreement on presence or absence of byssinosis in 139+14+17 = 170 (93%)

Normal = no evidence of byssinosis

Grade I = chest tightness on Mondays only

Grade II = chest tightness on Mondays and other working days (reference 11)

Table 8—*Predictive Value of Positive Test*

% Sensitivity	% Specificity	Prevalence	Predictive Value (%)
95	95	50/100	95
95	95	5/100	50
95	95	1/100	16
95	95	1/1,000	2
60	98	1/1,000	3

DOW 06820

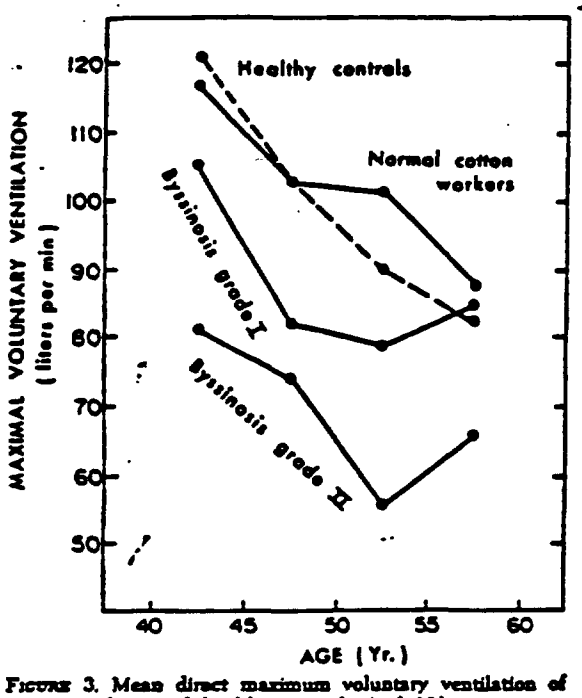

FIGURE 3. Mean direct maximum voluntary ventilation of cotton workers and healthy controls (ref 13).

VALIDITY

The usefulness of a screening test lies in its ability to distinguish those who probably have the disease in question from those who probably do not. The probability that a person with a positive screening test result actually has the disease sought is called the predictive value of a positive test.[12] It is important to recognize that this probability depends not only on the validity of the test as measured by its sensitivity and specificity, but also on the prevalence of the disease in the screened population (Table 8). Thus, if a disease is rare, many false positives will occur, even with a highly specific test, before a true positive is found.

To determine the validity of a screening test, some other measure must be used to define true positives and true negatives. The validity of the grading of byssinosis by questionnaire responses has been evaluated by comparing pulmonary function test results for workers classified in the different grades.[13,14] Figure 3 shows that for groups of workers, pulmonary function is progressively reduced below normal as byssinosis grade increases. Workers with grade 2 byssinosis by questionnaire response not only have lower baseline values for FEV_1 than grade 0 or grade 1/2 workers, but also show persistent decrements in FEV_1 over the workshift throughout the work week (Fig 4). Thus, the grading of byssinosis by symptoms reported on a standardized questionnaire appears to separate groups of workers with objective differences in pulmonary function. Such validation against an external standard should be done whenever possible.

In large surveys the use of a separate measure to define true positives is usually not possible for spirometry, since no other pulmonary function tests are done. Rather, an individual's results are compared with a predicted normal value derived from a pulmonary function survey of a group of presumably healthy persons. Factors which affect the predicted value include the population from which the reference group was chosen, the criteria for normality used to select individuals for inclusion, and the methodology for testing and calculating results. The discussion which follows draws heavily on the work of Miller and Thornton.[15]

PREDICTED NORMAL VALUES FOR SPIROMETRY

The predicted normal values derived from a reference group depend upon the distribution within that group of age, race and ethnic group, socioeconomic class, occupation, urban-rural residence, and other fac-

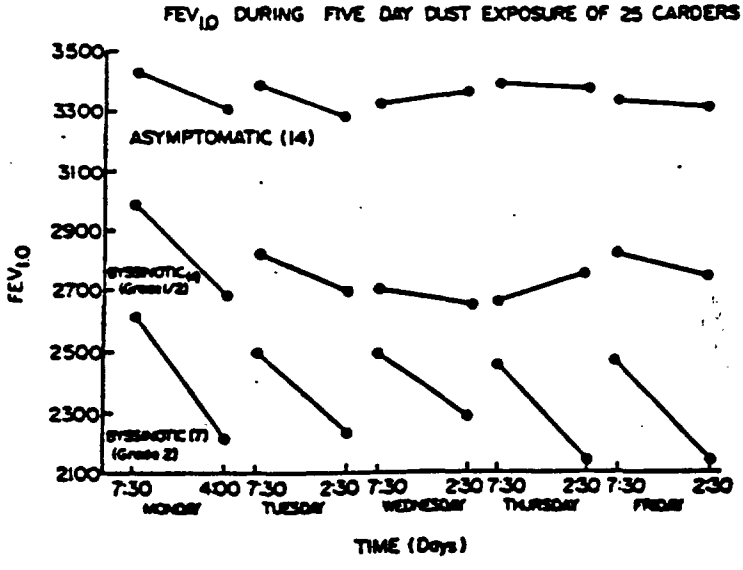

FIGURE 4. Mean FEV_1 of carders with differing grades of byssinosis during five days of dust exposure (ref 14).

IST0055552

DOW 06821

Table 9—*Pulmonary Function In Black And White Coal Miners Matched For Age, Height, And Smoking History*

Mean	White	Black	Difference
Age (yrs)	34.7	34.9	–
Height (in)	69.9	70.0	–
FEV_1 (L)	4.16	3.58	13.9%
FVC (L)	5.20	4.48	14.4%
TLC (L)	7.10	6.01	15.3%

reference 16

tors. These characteristics must be considered when evaluating the applicability of the predicted values from a given reference group to the current study group. Several studies have documented a difference for FVC and FEV_1 between healthy white and black men of the same age and standing height. Table 9 shows results from the study of Lapp and co-workers[16] who found the FEV_1 of black subjects to be approximately 14 percent lower than white subjects. Although simple adjustment of predicted values obtained from studies of white reference groups for use in blacks is not completely satisfactory, use of such a scaling factor is preferable to ignoring the difference until adequate prediction formulae for blacks are available. Schoenberg and co-workers[17] have produced regression equations for spirometric values for black subjects, but the number of adult men was relatively small (120), and the equations contain the ln (weight) and $(age)^2$ as variables, thus making computations somewhat cumbersome. Rossiter and co-workers[18] suggest reducing the predicted values of FEV_1 and FVC from white reference groups by 13.2 percent when they are applied to blacks. Since the FVC and FEV_1 appear to be reduced approximately proportionately, no adjustment for the FEV_1/FVC ratio is necessary. Lanese and co-workers[19] found similar differences between white and black textile workers.

The common assumption that the use of age in the prediction formulae derived from a reference group completely controls for any differences in age distribution from the study group is not totally accurate. This is discussed under considerations of the "normal range" given below. The other factors mentioned are more difficult to evaluate and have usually been ignored in deriving normal values.

Criteria used for selection of "healthy" individuals for inclusion in the reference group in various studies have varied considerably. While most modern studies exclude subjects with respiratory symptoms, past history of severe or chronic respiratory or cardiac disease, and smokers, only recently have the effects of occupational exposure been considered. Very few studies have required chest radiographs or physical examinations for inclusion. Although theoretically one might wish to obtain a reference group as totally free from respiratory system impairment as possible, the more selective the criteria for inclusion, the more likely that any comparison of this group with an unscreened group from the general population will show a statistically "significant" prevalence of abnormalities in the unscreened group. The practical significance of such a difference may be difficult to determine.

Table 10—*Influence of Number of Maneuvers Performed on The Measured FEV_1*

	Best FEV_1
Maneuver 1	29%
2	19%
3	17%
4	16%
5	19%

Best FEV_1 on maneuvers 4 or 5 exceeded best on maneuvers 1-3 by:

≥ 5% in 8% of subjects

≥10% in 2% of subjects

Adapted from reference 20

Methodologic factors also influence predicted values. Equipment and test procedures have already been discussed. The optimum number of maneuvers performed has been debated citing a "learning effect" for the first few maneuvers followed by a "fatigue effect." Table 10 is adapted from the work of Knudson and associates[20] who found that the largest FEV_1 occurred in the first three trials in 65 percent of over 3,000 subjects tested in a general population survey. When the largest FEV_1 occurred on the 4th or 5th trial, it exceeded the largest from the first three maneuvers by more than 5 percent in only 8 percent of subjects and by more than 10 percent in only 2 percent of subjects. Tager and co-workers[21] found the largest FEV_1 in the 4th or 5th trial in 31 percent of subjects and found the average single maximal FEV_1 to be 105 ml larger than the mean of the three largest values. These differences are not necessarily negligible when large groups are studied and small absolute differences may be statistically significant. Nevertheless, the recommendation of the National Heart, Lung, and Blood Institute Epidemiology Standardization Project[22] is that "three acceptable" maneuvers be obtained. Whatever practice is followed, it is essential that the same procedure be used for the study and the reference or "control" groups.

It is also important to obtain an accurate measurement of the maximum FVC. Falsely low values for FVC will produce artificially high FEV_1/FVC ratios and very significantly increased flow rates near the end of expiration. The flow rate measured at a given percentage of the falsely low FVC is actually occurring at a higher lung volume than the flow rate measured at the same percentage of the true FVC. Hankinson and Petersen[23] have shown that at least 10 seconds of expiratory time must be recorded if 94 percent of subjects with true FEV_1/FVC ratio less than .70 are to be correctly classified (Table 11). The predicted values

DOW 06822

Table 11—*Effect of Spirometer Cut-off Time On Observed FEV_1/FVC Ratio*

Cut-off Time (sec)	% of 203 Subjects Correctly Classified With $FEV_1/FVC < .70$
5	57
6	67
7	81
8	86
9	93
10	94
11	97
12	97

reference 23

from three recent surveys for FVC, FEV_1, and FEV_1/FVC for a 50-year-old man who is 70" tall are given in Table 12. Morris and co-workers[24,25] used the Kory technique[26] for measuring FEV_1. This method has been shown to result in values which are an average of 179 ml lower than the back extrapolation technique.[27] The FVC is unaffected by this method of measurement and agrees with that of Petersen and co-workers[28] who used a rolling seal spirometer and a method equivalent to back extrapolation to measure the FVC and FEV_1. Knudson and associates[29] used a pneumotachograph to measure flow and then integrated to obtain volume. Their value for FEV_1 agrees with that of Petersen et al, but the FVC value is lower, resulting in a higher predicted FEV_1/FVC ratio. If the FVC is falsely reduced, predicted values for flow rates near the end of expiration will also be too high. The interpretation of these flow rates is made difficult by this sensitivity to errors in measurement of the FVC.

Table 12—*Predicted Values For Male Aged 50 Years and 70" Tall*

	FVC (L)	FEV_1 (L)	FEV_1/FVC (%)
Petersen et al[28]	4.88	3.70	76*
Morris et al[24,25]	4.87	3.58	73
Knudson et al[29]	4.65	3.70	81

*Obtained by taking the ratio of predicted FEV_1 to predicted FVC

Table 13—*Prevalence of "Abnormal" Spirometry in A Healthy Population*

Males Age (yrs)	FVC (%)	FEV_1 (%)	FEF_{25-75} (%)
≤39	5.5	4.2	17.2
40-49	3.3	6.6	17.6
50-59	7.8	9.1	19.7
≥60	7.1	11.9	28.5

Calculations from data of Morris[24] with abnormal FVC or FEV_1 defined as less than or equal to 79% of mean predicted and $FEF_{25-75} \leq 75\%$ of mean predicted.[29]

Because healthy individuals of the same sex, race, age, and height exhibit considerable variation in spirometric results, classifying a subject as "abnormal" depends upon some arbitrarily defined lower limit of normal. The most common clinical practice is to set the lower limit of normal at some percentage of the mean predicted value, often 80 percent, and to label all values below this as "abnormal." This practice is not completely satisfactory. As discussed previously, the prediction equations derived from a given reference group are never completely applicable to any other group. Also, because a given percent predicted deviates less from the mean value for small mean values than for large ones, more abnormalities will always be found among older and short subjects when this method for defining "abnormal" is used. Miller and co-workers[30] have calculated the prevalence of "abnormal" spirometric values among the participants in the survey of healthy subjects by Morris.[24] Table 13 shows the results for men where "abnormal" is defined as a value ≤79 percent of mean predicted for the FVC and FEV_1 and ≤75 percent of mean predicted for the $FEF_{25-75\%}$. The prevalence of "abnormal" values in this healthy group increases with age so that 12 percent of men over 60 years would have an "abnormally low" FEV_1 and 29 percent an "abnormal" $FEF_{25-75\%}$ by this definition.

Use of a statistical definition of "abnormal" is one means to avoid the problem just discussed. By setting the lower limit of normal at the mean predicted value less some multiple of the standard error of the estimate (SEE) of the regression line, the number of healthy persons arbitrarily called abnormal is predetermined. A lower limit of normal of the mean value −1.64 SEE will place 5 percent of the group below normal; if two SEE's are used, only 2.5 percent of the group will be below the limit. This method may be more or less stringent than using 80 percent of mean predicted value depending on the SEE. Figure 5 shows a comparison of the mean −1.64 SEE and

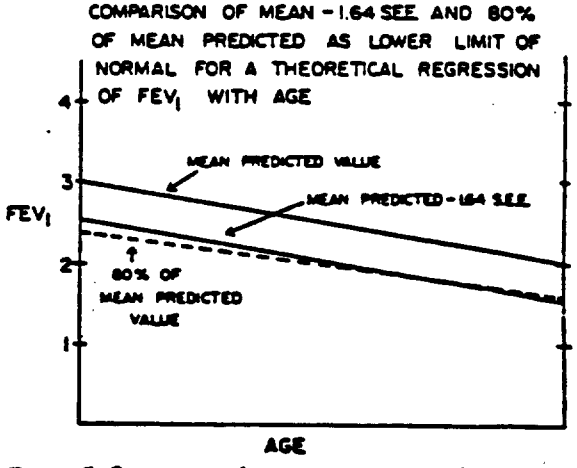

FIGURE 5. Comparison of mean −1.64 SEE and 80 percent of mean for defining lower limit of normal for a theoretical regression line of FEV_1 and age (ref 29).

STOD0556

DOW 06823

80 percent of the mean for a theoretical regression equation relating age and FEV_1. At younger ages the statistical definition is more stringent, while the opposite holds true for the older ages. Although use of a statistical definition of the normal range alleviates the problem of increasing prevalence of "abnormality" for those with smaller predicted values, it does assume that the SEE is the same for any point along the regression line. This may not be true, especially if the reference group has proportionately fewer members at the extremes of age. Also, some variables do not exhibit a Gaussian (normal) distribution about the mean value. If the distribution is skewed or the SEE large, the mean value less 2 SEE may include zero, thus invalidating this definition of the normal range.[20]

When the reference group is large enough, the most useful method for individual comparisons may be to determine the percentage of observations in the reference group above given values of the variable in question (percentile ranks). Results for an individual can then be classified by the percentile rank in which they would fall in the reference group. Since the probability of being truly "abnormal" increases as the percentile rank decreases, the physician does not need to arbitrarily interpret one level as the "lower limit of normal."

Comparison of Pulmonary Function Between Two Groups

Population studies may be undertaken to determine the prevalence of disease, to elucidate the natural history of a disease, or to establish an association of disease with a given agent. In occupational lung disease epidemiology, the objective is often to determine a dose-effect relationship between an occupational exposure and reduction in pulmonary function. Proper interpretation requires not only defining "abnormal" for individuals, but also detection and interpretation of any differences between the study group as a whole and a properly chosen reference ("control") group. When comparing the pulmonary function of the two groups, the basis should be the distribution of observed values —not the percentages of observations classified as "abnormal." Miller and Thornton[15] point out that two groups may have significantly different percentages of observations below any arbitrary lower limit of normal, yet not differ in the mean value of observations (Fig 6). Contrariwise, two groups may have the same percentage observations below a lower limit of normal but markedly different distributions (Fig 7).

The most powerful techniques for comparing study and reference groups are the analysis of covariance and matching procedures. Matching requires a large pool of subjects in the reference group and the knowledge of which factors are important. Analysis of covariance allows direct estimation of the difference between groups while adjusting for the effect of confounding variables. Certain assumptions (*eg*, linearity of response) are made in the analysis, and care must be taken that true differences between the groups are not obscured by limitations of the technique. Statistical consultation is necessary when attempting to use either of these techniques.

Before beginning a comparison of this type, the investigator must also select what magnitude of difference in function he considers important to detect. This determines the sample size he must use. The overall significance of a difference in pulmonary function between two groups depends on the basic question the study is attempting to answer and should not necessarily be equated with either "statistical significance" or "clinical significance." In some situations, small absolute differences between groups may have important implications while in others statistically significant differences may not.

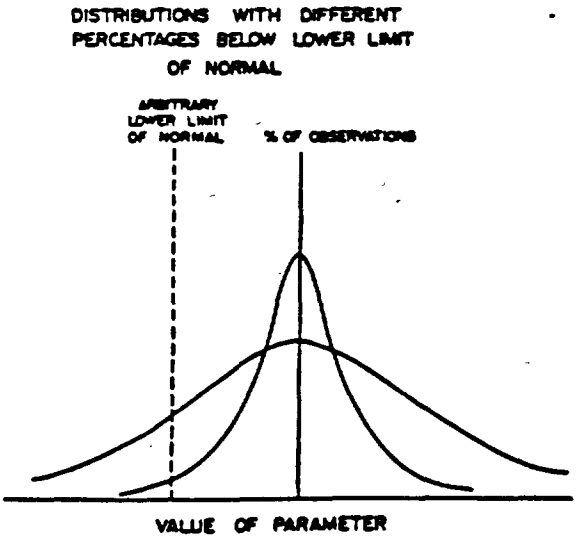

Figure 6. Distributions with markedly different percentages of observations below arbitrary lower limit of normal but not significant difference in mean value (adapted from ref 15).

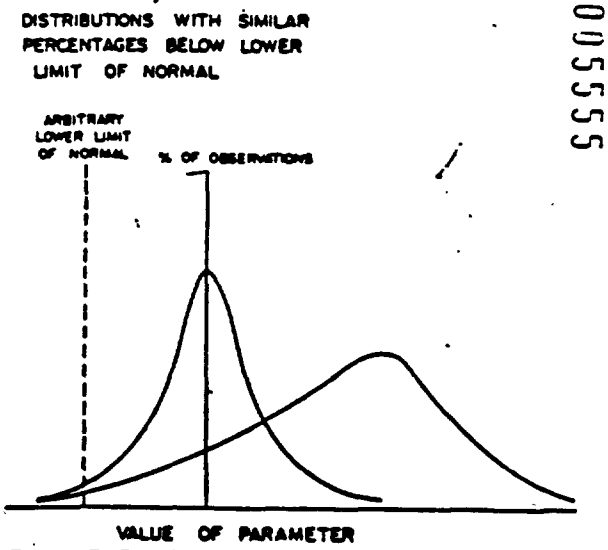

Figure 7. Distributions with similar percentages of observations below arbitrary lower limit of normal but markedly different mean values (adapted from ref 15).

ST0055555

DOW 06824

The presence of smokers and ex-smokers in a study group presents a difficult problem. Although the effect of smoking on pulmonary function can be quantitated to some extent for large groups,[21] the impact on an individual is so variable that no useful predicted values for pulmonary function for smokers are available. Fletcher and co-workers[12] state that "if the whole of the population were exposed to (say) 15 cigarettes/day for 30 years, the resultant FEV losses would have a skew distribution with many quite small loses, but a 'tail' of clinically significant loses." Thus, simply matching two groups for percentage of smokers and ex-smokers or even pack-years of smoking does not necessarily control for the effect of smoking on pulmonary function in producing differences between the groups.

Summary

Pulmonary function testing and questionnaires are valuable tools in epidemiologic studies of occupational lung disease. Accurate equipment and standardized methodology are vital to obtain reproducible responses. For spirometry, the FVC and FEV_1 show the least intrasubject variability and on questionnaires, occupational and smoking history are more reproducible than symptoms. The limitations of any method used to define a lower limit of normal should be kept in mind and, whenever possible, groups should be compared by use of the distribution of observations in the two groups—not just the prevalence of "abnormal" findings.

References

1 Commission on Chronic Illness: Chronic illness in the United States, Vol 1, Cambridge: Harvard University Press, 1957
2 Standardization of Spirometry—Statement of the American Thoracic Society. Am Rev Respir Dis 1979; 119: 831-38
3a 29 CFR 1910.1043 Appendix D, Occupational exposure to cotton dust, Final Mandatory Safety and Health Standards.
b 20 CFR 718.103 Appendix B, Standards for determining coal miners' total disability or death due to pneumoconiosis.
4 Gardner R, Hankinson J, West B. Evaluating commercially available spirometers. Am Rev Respir Dis 1980; 121:73-82
5 Cochrane G, Prieto F, Clark T. Intrasubject variability of maximal expiratory flow volume curve. Thorax 1977; 32:171-76
6 McCarthy D, Craig D, Cherniak R. Intraindividual variability in maximal expiratory flow-volume and closing volume in asymptomatic subjects. Am Rev Respir Dis 1975; 112:407-12
7 Tattersall S, Benson M, Hunter D, Mansell A, Pride N, Fletcher C. The use of tests of peripheral lung function for predicting future disability from airflow obstruction in middle-aged smokers. Am Rev Respir Dis 1978; 118:1035-50
8 Berry G. Longitudinal observations, their usefulness and limitations with special reference to the forced expiratory volume. Bull de Physio-Path Respir 1974; 10:643-55
9 Samet J, Speizer F, Gaensler E. Questionnaire reliability and validity in asbestos exposed workers. Bull de Physio-Path Respir 1978; 14:177-88
10 Eckert HL, Petersen MR, Reger RB, Hahon N. Reliability of smoking histories in standard questionnaire interviews (abstract). Annual Meeting, Professional Association of the U.S. Public Health Service, San Francisco, CA: April 3-6, 1977
11 Schilling R, Hughes J, Dingwall-Fordyce I. Disagreement between observers in an epidemiological study of respiratory disease. Br Med J 1955; 1:65-68
12 Vecchio TJ. Predictive value of a single diagnostic test in unselected populations. N Engl J Med 1966; 274: 1171-73
13 Schilling R, Hughes J, Dingwall-Fordyce I, Gilson J. An epidemiological study of byssinosis among Lancashire cotton workers. Br J Industr Med 1955; 12:217-27
14 Merchant J, Halprin G, Hudson A, Kilburn K, McKenzie W, Bernnanzolin P, Hurst D, Hamilton J, Germino V. Evaluation before and after exposure—the pattern of physiological response to cotton dust. Ann NY Acad Sci 1974; 221:38-43
15 Miller A, Thornton J. The interpretation of spirometric measurements in epidemiologic surveys: Standards and inconsistencies. Presented at the Annual Meeting, ACCP Subcommittee on Criteria for Occupational Lung Disease, Washington, D.C., 1978
16 Lapp NL, Amandus H, Hall R, Morgan WKC. Lung volumes and flow rates in black and white subjects. Thorax 1974; 29:185-88
17 Schoenberg J, Beck G, Bouhuys A. Growth of decay of pulmonary function in healthy blacks and whites. Respir Physiol 1978; 3:367-93
18 Rossiter C, Weill H. Ethnic differences in lung function: Evidence for proportional differences. Int J Epidemiol 1974; 3:55-61
19 Lanese RR, Keller MD, Foley MF, Underwood EH. Differences in pulmonary function tests among whites, blacks, and American Indians in a textile company. J Occup Med 1978; 20:39-44
20 Knudson R, Slatin R, Lebowitz M, Burrows B. The maximal expiratory flow-volume curve normal standards of variability and effects of age. Am Rev Respir Dis 1976; 113:587-600
21 Tager I, Speizer F, Rosner B, Prang G. A comparison between the three largest and three last of five forced expiratory maneuvers in a population study. Am Rev Respir Dis 1976; 114:1201-03
22 Epidemiology Standardization Project (Ferris BC, Principal Investigator). Am Rev Respir Dis 1978; 118: Part 2
23 Hankinson J, Petersen M. Data analysis for spirometry instrumentation standards (abstract). Am Rev Respir Dis 1977; 115:Supplement, 116
24 Morris J, Koski A, Johnson L. Spirometric standards for healthy non-smoking adults. Am Rev Respir Dis 1971; 103:57-67
25 Morris J, Temple W, Koski A. Normal values for the ratio of one-second forced expiratory volume to forced vital capacity. Am Rev Respir Dis 1973; 108:1000-03
26 Kory R, Callahan R, Boren H, Syner J. The Veterans Administration—Army cooperative study of pulmonary function. I. Clinical spirometry in normal men. Am J Med 1961; 30:243-58
27 Smith A, Gaensler E. Timing of forced expiratory volume in one second. Am Rev Respir Dis 1975; 112:883-85
28 Petersen M, Amandus H, Reger R, Lapp N, Morgan

ST0005556

DOW 06825

WKC. Ventilatory capacity in normal coal miners prediction formulae for FEV_1 and FVC. J Occup Med 1973; 15:899-902
29 Sobol B. The early detection of airways obstruction: another perspective. Am J Med 1976; 60:819-24
30 Miller A, Thornton J, Smith H, Morris J. Prevalence of "abnormal" spirometry in a normal reference male population—a reconsideration of the 1971 Oregon survey. Am J Industr Med (in press)
31 Burrows B, Knudson R, Cline M, Lebowitz M. Quantitative relationships between cigarette smoking and ventilatory function. Am Rev Respir Dis 1977; 115: 195-205
32 Fletcher C, Peto R, Tinker C, Speizer F. The natural history of chronic bronchitis and emphysema. Oxford: Oxford University Press: 1976, 129

Worker Monitoring in Byssinosis*

Harold R. Imbus, M.D., Sc.D.†

In late 1970, we began our program of medical surveillance of workers exposed to raw cotton dust. This program followed initial research efforts with Duke University and the North Carolina State Department of Health, which gave insights into the cause and incidence of byssinosis. Accordingly, we embarked on a program that included worker monitoring and environmental control. We describe the following three aspects of worker monitoring: procedures, some of the results achieved, and some of the problems involved.

Procedures

The procedures used in our program are testing, initial follow-up, referral, and notification of results.

Testing

Our testing program, begun in 1970, consists of a forced expirogram, namely, the one-second forced expiratory volume (FEV_1), the forced vital capacity (FVC), and the British Research Council (MRC) Respiratory Questionnaire Modified for Byssinosis. At first, this was a team effort consisting of nurses and myself from the corporate headquarters in addition to the plant nurse who had been trained in spirometry and administration of the questionnaire. Later the program was carried out on a yearly basis by the plant nurse, with results monitored by the corporate headquarters.

Pulmonary function testing is done on a waterless spirometer on the first day of the work week, before the worker goes to his assigned place of work. The worker returns between 4½ to 6 hours into the work shift, and the spirometry is repeated. Pulmonary function testing before and after the worker has been exposed to cotton dust is compared to determine if there has been a significant decrease. Currently, the Occupational Safety and Health Administration (OSHA) defines a 5 to 10 percent decrease as significant. The graphs are read by the plant nurse and recorded on a form that also contains the results of the questionnaire. These results are then processed by the computer, which records each employee's FEV_1, FVC percentage of predicted, and any decrement of FEV_1 during the work shift. Also, the computer records the frequency with which the worker complains of Monday tightness under the Schilling Classification System. Under Schilling's system, the frequency of these complaints places a worker in a byssinosis grade of ½, 1, or 2 based on symptoms only.

Initial Follow-up

If the nurse suspects unsatisfactory results owing to technical problems, such as unsatisfactory effort, or any other reason for obtaining a poor pulmonary function test result, she will repeat the test at another date before sending the results in on a computer form. Once the results are obtained for each individual, there is a follow-up for certain categories of workers. This follow-up is made by either a physician or a nurse who has had special training and wide experience in the program and who comes from the corporate headquarters. The following categories of workers are interviewed: (1) those with an FEV_1 of less than 80 percent of predicted unless they have been previously interviewed and are in a stable condition; (2) those with a shift decrement of their FEV_1 of 5 percent or more; (3) those who have byssinosis symptoms whether or not there is any impairment; (4) those whose predicted FEV_1 has decreased by more than 5 percent since the baseline test, the first test performed.

During this interview, additional questions are asked about symptoms, other diseases, smoking habits, and medical treatment. Another pulmonary function test may be administered on the spot if there is any question about the technical quality of the test results that indicated a problem. Many of the workers who are interviewed are assigned a surveillance schedule requiring them to undergo retesting once every six months. Those who have been interviewed by the corporate staff have the results and findings discussed with them, including abnormally low pulmonary function or reaction to cotton dust, either in the form of a decrement of pulmonary function or the Monday tightness. If the employee appears to be a reactor, he will be advised to wear a respirator or to transfer and will be placed under special surveillance. Smoking habits are also discussed, and the worker is advised that smoking can contribute to the problem.

Referral

We refer to independent specialists, at company expense and on company time, those who have an FEV_1 below 60 percent of predicted, who have an accelerated deterioration in pulmonary function, or who have other problems or complaints not readily explained. This evaluation is done by physicians who have special expertise in chest medicine. This evaluation must include as a minimum a complete history, physical examination, ECG, SMA 12 or SMA 23, urinalysis, posteroanterior and lateral chest x-ray film, spirometry, usually arterial oxygen and carbon dioxide tensions, and any other study that the specialist believes to be important. Since it has been shown that emphysema is not related to cotton dust exposure, the specialist in the future will

*From Burlington Industries, Inc. Greensboro, North Carolina.
†Director of Health and Safety.

DOW 06826

JOHN H Nicholson

References

1 Morgagni JB. The seats and causes of diseases investigated by anatomy. 1768 Translated by B. Alexander. London: Millar, 1769; 1:361 Hafner Publishing.
2 Van Coetsem. De la pneumonie produite par la poussière de coton. Annales de la Médicine Belge et Etrangère (Brussels) 1836; 3:3
3 Greenhow H. The pathology of flax dressers lung. Transactions of the Pathological Society of London. 1869; 20: 48-59
4 Schilling K. Uber die Schadlichen Einwirkungen des Baumwollstaubes auf die Atmungsorgane. Deut Arch f Klin Med. 1925, 146-163. Abstracted in Public Health Bulletin, Washington No. 297, 1947
5 Landis HRM. The relation of organic dust to pneumoconiosis. J Ind Hygiene 1925; 7:1
6 Shaw Dunn J, Sheehan HL. Report on ten autopsies on cotton-mill workers. Appendix III, Report of the Departmental Committee on Dust in Card Rooms in the Cotton Industry. London: HMSO 1932, 68-70
7 Rüttner JR, Spycher MA, Engeler ML. Pulmonary fibrosis induced by cotton fibre inhalation. Pathol et Microbiol 1963; 32:1-14
8 Edwards C, Macartney J, Rooke G, Ward F. The pathology of the lung in byssinotics. Thorax 1975; 30:612-623
9 Takizawa T, Thurlbeck WM. Muscle and mucous gland size in the major bronchi of patients with chronic bronchitis, asthma and asthmatic bronchitis. Am Rev Respir Dis 1971; 104:331-36

Factors Influencing the Interpretation of FEV_1 Declines Across the Working Shift*

Henry Glindmeyer III, D.Eng.; John Diem, Ph.D.; Janet Hughes, Ph.D.; Robert N. Jones, M.D., F.C.C.P.; and Hans Weill, M.D., F.C.C.P.

Guidelines for the evaluation and management of workers exposed to cotton dust have been defined by the Department of Labor in the recently adopted cotton dust standard. The success of this standard is dependent on the precision with which an acute or chronic change in lung function can be detected at the workplace. Although strict requirements have been adopted for test instruments and procedures, the inherent variability of measurements of lung function still pose a significant problem in distinguishing reactors from nonreactors. According to the standard, possible acute reactors are defined as those exhibiting a fall in Monday FEV_1 in excess of 5 percent or 200 ml, whichever is less. Since for some persons the within-subject SD for FEV_1 is in excess of 200 ml, individual variability can cause the cutoff for reactors to be exceeded, yielding both false-positive and false-negative results. To quantify this effect, the individual and group variability of data collected at four cottonseed crushing mills was examined relative to symptoms, smoking status, and work shift.

Materials and Methods

Four cottonseed crushing mills in the southern United States were visited in 1975, 1977, and 1978. Lung function measurements were performed before work and after the subjects had been away from the mill for at least 36 hours and after at least five hours at the workplace. Spirometric tests were conducted on a dry, rolling-seal spirometer with the output displayed on a Hewlett-Packard XYY recorder as volume-time or flow-volume plots of the maximal forced expiratory vital capacity maneuver. The measurements included the 7-sec forced vital capacity (FVC), the forced expiratory volume in .5 and 1 second ($FEV_{.5}$ and FEV_1), and the forced expired flow between 25 and 75 percent of the forced vital capacity (FEF25-75%). The beginning of time for calculating the $FEV_{.5}$ and FEV_1 was obtained by the method of backward extrapolation, and all measurements were corrected to the body temperature and pressure saturated with water vapor (BTPS). Subjects were tested in the standing position, nose clips were used, and at least four maximal expiratory maneuvers were performed at each testing session.

A trained interviewer administered a British MRC questionnaire modified for detection of byssinosis. Responses to the questionnaire were analyzed for the prevalence of byssinosis by the Schilling definition and for chronic bronchitis, atopy, dyspnea, and groups of upper and lower respiratory tract symptoms.

An analysis of variability within subjects for FEV_1 was performed using the four largest preshift and postshift volumes obtained on 192 subjects tested at the four mills in 1977. For each individual, within-subject SD was computed as the square root of the average preshift and postshift FEV_1 variances. This within-subject SD was analyzed in relation to symptoms, smoke status, and work shift. The SD of the preshift-postshift FEV_1 difference for each individual was computed by multiplying the within-subject SD by the square root of $(1/N_1 + 1/N_2)$, where N_1 and N_2 represent the number of preshift and postshift measurements.

Acute changes in lung function in relation to shift were analyzed for all mills and for all visits. Only subjects with acceptable over-the-shift spirometry values were included in analyses. Acceptable spirometry required visually acceptable curves, with the added stipulation that the two largest FVC values differed by less than 3 percent. Acceptable over-the-shift spirometry was obtained at least once for 185 linter area workers at three visits.

Results

The individual within-subject SD for the 192 subjects studied in 1977 ranged from .020 to .428 L, with 74 percent less than .100 L. The average within-subject variability was .102 L. SD across the shift ranged from .014 to .302 L, with .072 L as a representative value.

Figure 1 displays the average FEV_1 within-subject variability between groups answering "yes" or "no" to

*From the Tulane University School of Medicine, New Orleans.

This work was supported by U.S. Public Health Service grant HL-15092 of the National Heart, Lung, and Blood Institute.

Reprint requests: Dr. Glindmeyer, Department of Pulmonary Diseases, 1700 Perdido Street, New Orleans 70112

S-10005558

DOW 06827

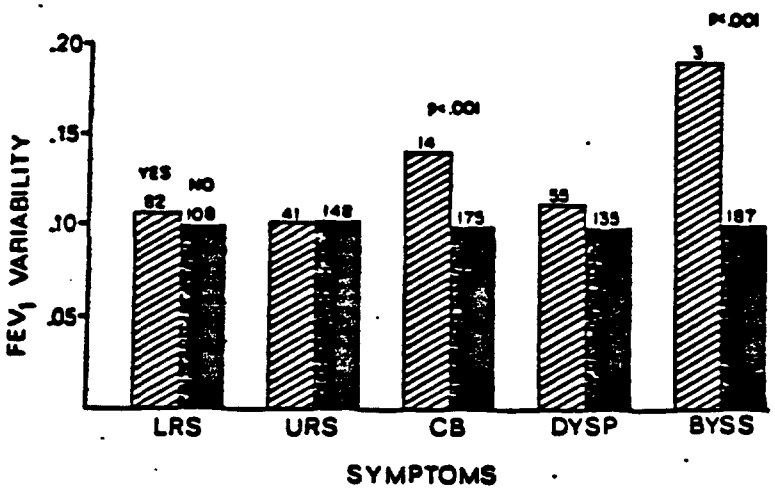

FIGURE 1. Average FEV_1 within-subject variability between groups answering "yes" or "no" to lower respiratory tract symptoms (LRS), upper respiratory tract symptoms (URS), chronic bronchitis (CB), dyspnea (DYSP), and byssinosis (BYSS). Number of subjects included in each category are listed at top of each bar.

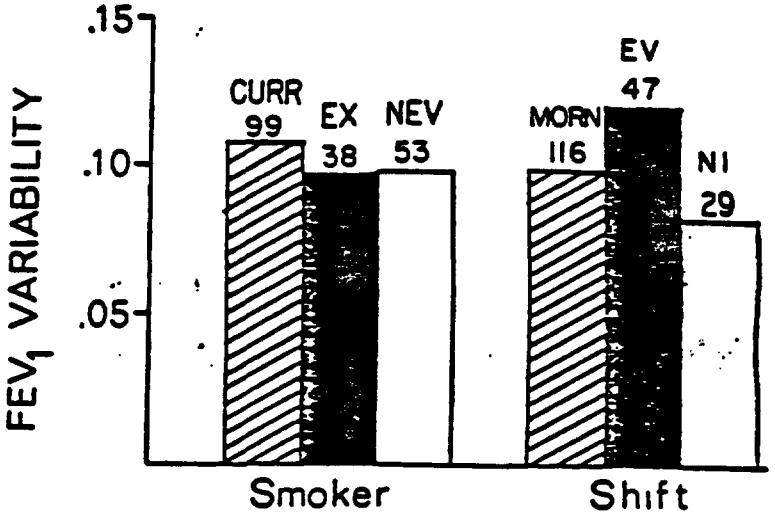

FIGURE 2. Average FEV_1 within-subject variability displayed for current (CURR), exmokers (EX), and never smokers (NEV), and for workers on the morning (MORN), evening (EV), and night (NI) shifts. Number of subjects included in each category are listed at top of each bar.

lower or upper respiratory tract symptom complexes, chronic bronchitis, dyspnea, and byssinosis. The variability of groups answering "yes" to bronchitis and byssinosis (.141 L and .192 L, respectively) was significantly different (P = .001) from the nonbronchitic subjects and the subjects without symptoms of byssinosis (.099 L and .102 L, respectively). No statistical difference in variability was exhibited between groups answering "yes" or "no" to lower or upper respiratory tract symptom complexes or dyspnea.

Figure 2 displays the average within-subject variability in FEV_1 as a function of smoking status and shift. For smoking, there was no significant difference in variability between current, exmokers, and never-smokers. In addition, the morning, evening, and night shifts yielded statistically similar repeatability.

Changes over the working shift were related to time of day for all three visits and for all four mills, even though preshift measurements were not significantly related to shift. For FVC, FEV_1, and FEF25-75%, there was no significant difference between morning and night shifts. However, there was always a significant difference for all of these parameters between evening shift and the morning and night shifts. Table 1 shows the mean change for all visits collapsed for all mills. The significant difference in FEV_1, FVC and FEF25-75% between shifts persisted even though there was no significant difference between shifts in smoking status, symptoms, exposure (all were linter area workers), race, age, years employed in mill, and work area.

DISCUSSION

The ability with which one can detect a meaningful change in FEV_1 is related to shift and is an inverse

Table 1—*Mean Change in Spirometric Values for All Visits to All Mills*

Shift	Mean Time of Preshift Spirometry	N	Mean Change Over All Visits (Linter Area Workers) FVC	FEV_1	FEF25-75%
Morning	7:00 AM	72	−.042	−.051	−.148
Evening	2:30 PM	64	−.083	−.135	−.465
Night	11:00 PM	49	−.031	−.030	−.098

DOW 06828

function of within-subject variability. This variability is related to symptoms of byssinosis and chronic bronchitis. The current practice of comparing individual decline in FEV_1 across the shift to a fixed value, eg, > 200 ml, in order to classify reactors, is misleading, since the across-shift SD of some subjects is in excess of 250 ml. Conversely, very repeatable subjects (SD < 50 ml) should be classified as reactors if their across-shift decline is in excess of 100 ml. In addition, the excessive variability of subjects with symptoms of byssinosis is a possible explanation of the lack of correlation between acute decline and symptoms. The shift effect implies that an adverse change in FEV_1 observed across the day shift or night shift would be significantly lower than an equally meaningful change observed in the evening. Therefore, a more accurate classification of reactors should take into account symptoms, shift, and an assessment of individual variability.

Summary

An analysis of variability within individuals was conducted for FEV_1, obtained from subjects employed in the cottonseed industry. Individual SDs ranged from .020 to .428 L. The within-subject SD of the preshift-postshift difference ranged from .014 to .302 L. When within-subject variability of FEV_1 was analyzed in relation to symptoms, smoking history, and shift, the only statistical difference occurred in the 14 bronchitic subjects (.141 L) relative to the 179 without bronchitis (.099 L), and the three subjects with byssinosis (.192 L) relative to those without symptoms (.102 L). Changes over the working shift were significantly different for evening relative to morning and night shifts, even though there was no significant difference in smoking status, symptoms, exposure, race, age, years employed in the mill, and work area. In addition, baseline measurements were not significantly related to shift. Therefore, the effect of individual variability, symptoms, and shift should be considered if an accurate classification of reactors based on change in FEV_1 across the shift is to be obtained.

ACKNOWLEDGMENT: The authors would like to acknowledge the programming and technical assistance of Daniel Seale, B.S., in the preparation of this manuscript.

The Effect of Mediator Modifying Drugs in Cotton Bract-Induced Bronchospasm*

E. Neil Schachter, M.D., F.C.C.P.,† Steven Brown, B.S.; Eugenia Zuskin, M.D.; Gerald J. Beck, Ph.D.,‡ Marion Buck, Ph.D.; Barbara Kolack, B.S.; and Arend Bouhuys, M.D., Ph.D.§

Byssinosis is a chronic occupational lung disease associated with the inhalation of cotton and other textile dusts.[1-3] In its early phases acute reversible symptoms, such as wheezing, chest tightness, and shortness of breath, accompany reversible changes in pulmonary function.[3] The therapy for this disease remains incompletely defined. However, considerable evidence from both *in vitro* and *in vivo* studies indicates that the mechanism underlying acute airway obstruction in this disease is related to the nonantigenic release of histamine and possibly of other mediators by the action of an airway constricting agent in cotton bracts.[4-6] To test the therapeutic implications of these findings, we studied the effect of mediator modifying drugs on cotton bract-induced bronchospasm in healthy subjects. Two classes of agents were studied, a cromone, disodium cromoglycate, and two antihistamines, an H_1-blocking agent, chlorpheniramine, and cimetidine, an H_2-blocking agent.

*From the Department of Internal Medicine, Yale University School of Medicine, New Haven.
†Associate Professor of Internal Medicine.
‡Assistant Professor of Biostatistics.
§Deceased June 15, 1979 (Professor of Internal Medicine).
Reprint requests: Dr. Schachter, Yale University School of Medicine, 333 Cedar Street, New Haven 06511

Decreases in lung function owing to cotton dust can be reproduced in the laboratory in healthy subjects, never before exposed to textile dust, after they inhale an aerosolized aqueous extract of cotton bract. Most healthy subjects will show some decrease in pulmonary function when exposed to these extracts, but previous studies have shown that the degree of induced bronchospasm to a standardized dose will vary with the individual.[6] These differences permit investigators to classify healthy subjects as responders and nonresponders. The ability of mediator modifying agents to protect against cotton bract extract-induced bronchospasm was studied in healthy responders.

Subjects and Methods

A total of 21 healthy subjects whose ages ranged from 18 to 34 years were selected from a larger group of 31 subjects on the basis of their increased reactivity to cotton bract extract. Nineteen subjects participated in our study with cromolyn sodium, and ten participated in our study with antihistamines. Six individuals from the group smoked, and 13 were then nonsmokers. None of the subjects related a history of asthma or other atopic diseases. Pulmonary function data showed normal lung function for the group as a whole (Table 1).

The study consisted of three protocol days for the disodium cromoglycate study and four protocol days for the antihistamine study. On the first day of each study, the subject's response to cotton bract extract was determined. Baseline pulmonary function was obtained immediately before the challenge with cotton bract and subsequently was measured 30, 60, 90, 120, and 150 minutes after the exposure. This permitted us to group our subjects into responders and nonresponders. The protocol was followed exactly on each subsequent study day, with the exception that the drug or placebo was administered before the chal-

ST0005560

DOW 06829