ENVIRONMENTAL RESEARCH 5, 451–466 (1972)

A Survey of Respiratory Disease among New York City Postal and Transit Workers

III. Anthropometric, Smoking, Occupational, and Ethnic Variables Affecting the FEV_1 among White Males[1]

JAMES H. STEBBINGS, JR.[2]

Division of Epidemiology, School of Public Health, University of Minnesota, Minneapolis, Minnesota 55455

Received May 5, 1972

Noninhaling of tobacco smoke, and, by implication, use of filter cigarettes, do not appear to reduce the rate of decline of the FEV_1 with age among smokers. The cumulative benefit of not smoking in late adolescence and early adult life appears to be permanently retained despite later smoking.

Workers born in Puerto Rico show distinctly lower values for the FEV_1 than other foreign or native-born workers.

A possible effect of occupational exposure to air pollution among nonsmoking bus drivers was found.

A number of smoking variables beyond current type and amount were shown to contribute to the prediction of the FEV_1. It is suggested that these, and also sitting height and weight, should be accounted for in analyses of air pollution or other environmental effects on pulmonary function. The use of age and standing height in predicting the FEV_1 for mass screening purposes was shown to be only 90% as sensitive as the use of age, standing height, sitting height, and weight.

This third report of the New York City survey of respiratory disease among postal and transit workers presents a detailed analysis of anthropometric, smoking, occupational, ethnic, and certain miscellaneous factors affecting the 1-sec forced expiratory volume among white males aged 30–59. The analysis was undertaken to identify variables to be taken into account in a study of the effects of air pollution on the FEV_1 in these two employed New York City populations (to be published).

The analysis presented here primarily pertains to the question of the circumstances under which body size and tobacco smoking variables, the effects of which on the common pulmonary function measurements are well known, should be decomposed into several variables. The use of standing height as the sole anthropometric variable is certainly warranted in many situations, particularly for calculating an individual's predicted value in clinical medicine. Likewise categorization of individuals by current smoking habits, perhaps with further categoriza-

[1] This study was made possible by USPHS Grants OH-00013 and EC-00233 to the Medical and Health Research Association of New York City. Data analysis was carried out under USPHS Grant OH-00353 at the University of Minnesota.

[2] Address requests for reprints to James H. Stebbings, Jr., Mayo Memorial Bldg., Box 197, University of Minnesota, Minneapolis, Minnesota 55455.

Copyright © 1972 by Academic Press, Inc.
All rights of reproduction in any form reserved.

tion by number of cigarettes smoked, yields results which demonstrate without question the effect of tobacco smoking in reducing pulmonary function. It is arguable, however, that in studies of the subtle effects of air pollution in either the urban or the work environment more accurate corrections for body size and smoking variables are desirable.

In this paper a number of tobacco smoking variables, over and above current smoking categories, are shown to be not only definable and concurrently usable but also to be significantly related to pulmonary function values.

So far as additional anthropometric variables are concerned, the small, but often significant, contribution of sitting height and weight to predicting pulmonary function values is well known; what is investigated here is not their contribution to variance explained, but their effect on the classification of individuals as above or below a screening level for possible disease.

In addition, certain unexpected effects of ethnic background and of marital status on pulmonary function are presented.

The basic design of the New York City survey of postal and transit workers, and the findings with respect to prevalence of respiratory symptoms, were presented by Densen *et al.* in 1967, and an earlier analysis of the ventilatory function test results, emphasizing patterns with respect to race, broad smoking categories, and respiratory symptoms, was published by Densen *et al.* in 1969.

METHODOLOGY

Survey Populations

The original survey was conducted over the period November 1961 through August 1963. The total Post Office population was defined as those carriers and chauffeurs on the employment rosters of stations in Manhattan and the Bronx on the survey dates for the respective stations. The total interviewed numbered 5313, or 87.4% of the defined population. (Thirty-one individuals have since been excluded because of extensive missing data, leaving 5282, or 86.9%).

The Transit population was defined as all those in the payroll titles of motorman, surface line operator, and surface line dispatcher as of February 10, 1962. Of these 6864 (81.3%) were interviewed during the survey period. However, individuals newly employed or who acquired these payroll titles after February 1962 but before time of interview were also surveyed. Of these 427 individuals are included in this analysis, but their response rate is not known.

Further details on the execution of the original survey and the characteristics of nonrespondents are given by Densen *et al.* (1967, 1969).

During the first 2 yr (June 1964–June 1966) of a follow-up study individuals from the defined Transit population who had been missed in the original survey were interviewed when possible. These numbered 683, and increase the response rate in the defined population to 89.4%. These individuals were not included in earlier published analyses (Densen *et al.*, 1967, 1969). The total surveyed Transit population then numbers 7974.

The total number of usable interviews, both postal and transit, is 13,256.

Selection of Individuals for Analysis

For a detailed examination of the variables under consideration it was decided to limit the analysis to white males aged 30–59 in certain smoking categories. In the analysis postal and transit workers are combined into one population, and the occupational difference is taken into consideration statistically. Densen *et al.* (1967, 1969) have found differences in both the FEV_1 and prevalence of respiratory symptoms between whites and nonwhites in these two occupational groups, and inclusion of nonwhites would have undesirably complicated this analysis. The age limits were set for different reasons: the upper limit was set because it became clear during a follow-up survey that retirement, presumably selective for health status, could take place anytime after age 60, and there were a very large number of retirements before age 65. The lower age limit was set because of uncertainty regarding the course of pulmonary function values in the third decade of life (Berglund *et al.*, 1963; Gregg, 1968; Stebbings, 1971c).

Individuals with one or more of the following variables unknown were excluded from the analysis: race, FEV_1, sitting height, standing height, or weight. Of 8577 males of white or unknown race aged 30–59, 134 or 1.56% were excluded from this analysis because of missing information.

It was desired to separate this defined group of white males into nine categories according to a cross-classification of cigarette nonsmokers, exsmokers, and smokers by pipe/cigar (pipe and/or cigar) nonsmokers, exsmokers, and smokers. A cross-classification of types of smokers, rather than separate treatment of cigarette and pipe/cigar smoking, was necessitated by an original coding decision to combine all forms of tobacco when coding quantity of tobacco smoked. Unfortunately the coding did not allow full analysis by lifetime smoking history as only smoking habits at time of interview (or when respondent gave up smoking, in the case of exsmokers) and 10 years previously were coded. Only individuals coded as nonsmokers are known to have never in their lifetime consistently smoked either cigarettes or pipes/cigars. For all other individuals it is assumed that if a form of tobacco (cigarette or pipe/cigar) was not coded for either the current or the 10-yr smoking history, the respondent was a lifetime nonuser of that form of tobacco. This, it was felt, would yield a classification of respondents sufficiently accurate for the purposes of this analysis. The following smoking groups were selected for analysis:

a. *nonsmokers:*
   never smoked cigarettes or pipe/cigars (n = 1,187)
b. *cigarette smokers:*
   currently smoke cigarettes, no history of pipe/cigar smoking (n = 4,086)
c. *pipe/cigar smokers:*
   currently smoke pipe/cigars, no history of cigarette smoking (n = 793)
d. *pipe/cigar/cigarette smokers:*
   currently smoke both pipe/cigars and cigarettes (n = 692)

e. *exsmokers/cigarette:*
exsmokers of cigarettes who have no history of pipe/cigar smoking ($n = 951$)

Four other cross-classifications have been excluded from the analysis because of small numbers and difficulty in analysis. Numbers of individuals in these categories ranged from 51 (exsmokers of pipe/cigars, never smoked cigarettes) to 231 (pipe/cigar smoker, exsmoker of cigarettes), and altogether total 734 individuals, 9.5% of the white males aged 30–59.

Pulmonary Function Measurement

The FEV_1 was measured using a McKesson Vitalor. A trial and two "successful" expirations were obtained and the maximum of the three was chosen for analysis. Standing height was measured with shoes, either with a measuring rod in Transit Authority clinic locations or with a tape measure against a wall in field locations. For sitting height each study participant was seated on a flat-topped stool against a wall and his length from seat to top of head measured against a standard mounted on the wall. All heights were recorded to the next lower quarter inch. Weights, with clothes and shoes but not coats, were measured with balance scales in Transit Authority clinic locations and with spring scales in the field.

Statistical Methods

The analysis is based on standard multiple regression techniques, utilizing the F ratio for testing the significance of the addition of independent variables to the equation. However, two of the techniques used in defining the independent variables are sufficiently obscure to justify a more detailed description.

The first technique is the use of a covariance term Za (or $z \times a$) where a is any continuous variable, age in this example, and z is a dummy variable taking the value 0 or 1 (say, $z = 1$ if respondent is a postal worker, $z = 0$ otherwise). Then, when a model of the form $x = b_0 + b_1z + b_2a + b_3Za$ is fit, the decline of x with age for nonpostal workers is given by b_2 and for postal workers by $b_2 + b_3$. In the same way the constant term b_0 gives the constant term for nonpostal workers, while that for postal workers is given by $b_0 + b_1$. A more detailed discussion of this technique is given by Johnston (1963).

The second technique also involves the use of dummy variables, in this case as a means of handling missing information. As an example, to take one used in this paper, consider the variable "amount smoked, ten years previously." Some individuals were not smoking (0 is not a valid code for this situation) while for other individuals the 10-yr smoking history was not available. If the dummy variable NI (=1 if no information, 0 otherwise) and NA (=1 if not applicable, 0 otherwise) are formed, and the continuous variable AS (amount smoked) is defined as the recorded value where a value was recorded, and as the mean of all recorded values where none was recorded (NI or NA = 1), the set of variables AS + NI + NA fully incorporate all available information. The advantage of this technique is that when dealing with numerous quantitative variables, which are on occasion either inapplicable or not available, the tiresome business of redefining the sample size every time a new variable is investigated may be dispensed

with, while the dummy variables may be found to yield useful additional information. This technique is described by Cohen (1968).

RESULTS

Anthropometric Variables

In addition to standing height (H), sitting height (SH), weight (W), and an index of obesity (W/H^2, Khosla and Lowe, 1967) were tested for their contribution to prediction of the FEV_1. The interaction terms between the dummy variable PO (postal worker) and the anthropometric variables and age were also tested.

Table 1 shows the regression coefficients on those variables which are statistically significant for one or more smoking categories. Neither the obesity index nor its interaction term with PO were significant, indicating no contribution of obesity *per se* over that of height and weight separately to predicting the FEV_1. None of the interaction terms of PO with heights or weights were significant, indicating that the postal and transit populations are not heterogenous with regard to the relationship of the FEV_1 to height and weight.

An attempt was made to create an index, similar to the obesity index, relating sitting height to standing height to improve prediction of the FEV_1. A number of possibilities were tested, mostly of the form SH/H^n or $(SH\text{-}0.5\,H)^n$ where n ranges from ½ to 3. The results did not suggest that such an index was of any use whatsoever. Fit of the regression was also not improved by incorporation of log or square transformations of the height variables.

The age variable does interact with employment, unlike the anthropometric variables, at least among current cigarette smokers where the decline of the FEV_1 with age is significantly greater among transit workers than among postal workers.

TABLE 1
REGRESSION COEFFICIENTS ON AGE AND ANTHROPOMETRIC VARIABLES OF THE FEV_1 AMONG WHITE NEW YORK CITY POSTAL AND TRANSIT WORKERS AGED 30–59[a]

	Regression coefficients				
	Nonsmokers	Cigarette smokers	Pipe/cigar smokers	Pipe/cigar/cigarette smokers	Exsmokers cigarette
Age	−.0239***	−.0396***	−.0304***	−.0315***	−.0336***
Height (cm), standing	.0260***	.0227***	.0251***	.0235***	.0209***
Postal employee (PO)[b]	−.0696**	−.4194	.0669	−.1705	−.0162**
PO × age	−.0005	.0094***	−.0029	.0038	−.0015
Height (cm), sitting	.0100	.0130***	.0202**	.0196**	.0171**
Weight (lb)	−.0023***	−.0002	−.0018*	.0007	.0006
Intercept	−0.512	−0.204	−1.110	−1.371	−0.420
N	1187	4086	793	692	951
R	.45	.54	.50	.53	.54

[a] Statistical significance is given by F test for significance of addition of variable to those preceding. Order of addition of variables was predetermined. One, two, or three asterisks indicate probabilities less than .05, .01, and .001, respectively.

[b] A dummy variable (1 if subject is postal worker, 0 otherwise).

TABLE 2

PREDICTED VALUES OF THE FEV_1 FOR WHITE NEW YORK CITY POSTAL AND TRANSIT WORKERS[a,b]

	Postal workers				Transit workers			
Age	30	40	50	60	30	40	50	60
Nonsmokers	3.61	3.37	3.13	2.88	3.70	3.46	3.22	2.98
Current cigarette smokers	3.46	3.16	2.86	2.56	3.60	3.21	2.81	2.41
Pipe/cigar smokers	3.75	3.42	3.09	2.75	3.77	3.47	3.16	2.86
Pipe/cigar/cigarette smokers	3.50	3.22	2.95	2.67	3.56	3.24	2.93	2.61
Exsmokers/cigarette	3.69	3.34	2.99	2.64	3.75	3.42	3.08	2.74

[a] Height 170 cm, sitting height 90 cm, weight 71 kg. Values are in liters.

[b] Based on equations presented, with a reduced number of significant digits, in Table 1.

Among nonsmokers and exsmokers of cigarettes there was a significant constant difference between postal and transit workers, the postal being lower.

The age-specific expected values of the FEV_1 are shown for both the postal and transit populations in Table 2. These predicted values are standardized for height, sitting height, and weight (after Ashford *et al.*, 1968), and may be compared with other values in the literature so standardized. The values in Table 2 may also be safely (error less than .05 L.) compared with those in the literature standardized only on age and standing height (usually age 40, height 170 cm), since the sitting height standard is very close to the mean for this population and the regression coefficients on weight are small.

Also to allow comparison with results from other studies regression equations incorporating only age and standing height are presented below. (The intercept term and the regression coefficient on age for the postal workers are the sums of the corresponding terms for the transit workers and the regression coefficients on the dummy variable PO and its interaction term with age.)

Nonsmokers	Transit	FEV_1 (L.) $= -.0239$ age $+ .0247$ height (cm) $+ .1487$
	Postal	FEV_1 (L.) $= -.0244$ age $+ .0247$ height (cm) $+ .0907$
Cigarette	Transit	FEV_1 (L.) $= -.0397$ age $+ .0268$ height (cm) $+ .2156$
Smokers	Postal	FEV_1 (L.) $= -.0302$ age $+ .0268$ height (cm) $- .2160$
Pipe/cigar	Transit	FEV_1 (L.) $= -.0302$ age $+ .0288$ height (cm) $- .3033$
Smokers	Postal	FEV_1 (L.) $= -.0325$ age $+ .0288$ height (cm) $- .2653$
Pipe/cigars	Transit	FEV_1 (L.) $= -.0321$ age $+ .0314$ height (cm) $- .8440$
Cigarette	Postal	FEV_1 (L.) $= -.0273$ age $+ .0314$ height (cm) $- 1.0687$
Exsmokers	Transit	FEV_1 (L.) $= -.0334$ age $+ .0284$ height (cm) $- .1024$
Cigarette	Postal	FEV_1 (L.) $= -.0354$ age $+ .0284$ height (cm) $- .1054$

SMOKING VARIABLES

Table 3 shows the significance levels and percentage of variance explained for those smoking variables found to be statistically significant in one or more categories of smokers.

The effect of inhaling was tested ahead of the other smoking variables on the

TABLE 3

SIGNIFICANCE OF SMOKING VARIABLES IN PREDICTING THE FEV[1] AMONG WHITE NEW YORK CITY POSTAL AND TRANSIT WORKERS AGED 30–59[a]

	Cigarette smokers		Pipe/cigar smokers		Pipe/cigar/cigarette smokers		Exsmokers/cigarette	
	% Variance explained	$F(1,4072)$	% Variance explained	$F(1,779)$	% Variance explained	$F(1,678)$	% Variance explained	$F(1,937)$
Does not inhale[b,c]	.0013	7.99**	.0001	0.07	.0073	7.11**	.0006	0.82
Amount smoked[c]	.0133	78.84***	.0002	0.17	.0043	4.23*	.0062	8.40**
Amount smoked[c] × age	.0009	5.27*	.0035	3.74	.0029	2.81	.0020	2.67
Age started smoking	.0017	10.04**	.0011	1.19	.0037	3.57	.0000	0.00
Amount smoked 10 yr before	.0006	3.59	.0033	3.52	.0002	0.18	.0043	5.82*
Nonsmoker 10 yr before	.0015	9.13**	.0052	5.52*	.0005	0.48	.0051	6.94**
10-Year history unknown	.0015	8.73**	.0000	0.00	.0032	3.16	.0009	1.23

[a] Statistical significance is given by F test for significance of addition of variable to those preceding. Order of addition is predetermined. One, two, or three asterisks indicate probabilities less than .05, .01 and .001, respectively.

[b] Except "does inhale" for pipe/cigar smokers.

[c] At time of interview, except exsmokers at time of cessation of smoking.

grounds that whether or not smoke was deliberately inhaled was logically prior to the amount of tobacco consumed. Among cigarette smokers, including those who also smoke pipes and cigars, noninhalers do show significantly ($p < .01$) higher FEV_1 values. The relatively low contributions to the percentage of the variance explained are explained by the low fraction of noninhalers (5.2% among cigarette smokers, 19.4% among pipe/cigar/cigarette smokers). The magnitude of the effect is difficult to estimate; using the regression equations on which Table 3 is based, the estimated effects are +.014 L. for cigarette smokers, and +.095 L. for pipe/cigar/cigarette smokers. Collinearity with other smoking variables clearly affects these regression coefficients: if noninhaling is the only smoking variable included, the estimated effects are +.099 and +.129 liters, respectively. However, the correlation of noninhaling with the other smoking variables is so high that if noninhaling is added to the regression after the other variables shown in Table 3, it is not found to be significant.

The effects described yield the constant difference between inhalers and noninhalers across all age and amount smoked categories. The interaction terms of inhaling with age and amount smoked were investigated among the current cigarette smokers, but neither of these two terms yielded significant results; in particular, the interaction term with age showed no suggestion of any effect. This implies that the rate of decline of the FEV_1 with age or amount smoked in noninhalers is not less than the decline in inhalers.

Of particular interest is the finding that the 15.8% of pipe/cigar smokers who reported that they did inhale showed absolutely no statistically significant reduction in the FEV_1 relative to other pipe/cigar smokers (Table 3). The effect estimated by the regression coefficient is only .01 L., a completely insignificant effect. An investigation of the interaction terms of inhaling with age or amount smoked also showed not even a suggestion of an effect of inhaling on decline of the FEV_1 with age or amount smoked. The same findings hold among exsmokers of cigarettes, but those may be explained as the result of self-selection for cessation of smoking.

Table 3 indicates that the amount of tobacco smoked is the most important smoking variable among cigarette smokers, and that increasing consumption does accelerate the rate of decline with age. Ignoring the effect of correlations among the smoking variables on the reliability of the regression coefficients, and assuming no change in tobacco consumption over 10 yr, one may estimate the effect on the FEV_1 of a one pack difference in cigarette consumption per day among cigarette smokers as $\pm .085$ L. at age 30 and $\pm .214$ L. at age 60 (given by $.0904 \times$ amount $- .0043 \times$ amount $\times$ age $- .0460 \times$ amount smoked 10 yr before).

Pipe/cigar/cigarette smokers and exsmokers of cigarettes show more or less the same effects of amount smoked as do cigarette smokers.

Although one might expect amounts smoked by pipe/cigar smokers to be less reliable than for cigarette smokers, it is still surprising that the amount smoked has no significant effect on the FEV_1, although the interaction term with age is of borderline significance and the regression coefficient does indicate an increased rate of decline with age in the heavier smokers. The mixture of pipe and cigar tobacco in computing amounts, and the less reliable recording of amounts in the

questionnaire make it hazardous to assign a numerical value to this increased rate of decline with age.

Attention should be called to the fact that, except among exsmokers, the amount smoked 10 yr before did not add significantly to the current amount smoked, although the results are very near the .05 level of significance. This reflects the great stability of smoking habits in most individuals rather than indicating that a respondent's past smoking habits have no effect on his current pulmonary function values.

The age a respondent started smoking cigarettes had an effect on his FEV_1. In this population the mean age for beginning smoking (cigarettes only) was 18.2 yr with a standard deviation of 4.1; despite this narrow range, and a low estimated effect of +.0042 L./yr on the FEV_1, the effect on the FEV_1 in later life was highly significant ($p < .01$).

Of even greater interest is that the 2.6% of cigarette smokers who were not smoking 10 yr before (of whom approximately half were lifetime nonsmokers to that point) had FEV_1's significantly ($p < .01$) higher than predicted. The effect was estimated at +.153 liters, and is in addition to the effect of age began smoking. (If this effect is deleted, the effect of age began smoking is increased to .0061 L./yr.) This implies that the benefits of not smoking in the early adult years are not lost if the smoking habit is acquired.

The variable "10 year history unknown" in Table 3 was necessarily included to complete the set of 10 yr smoking history variables; results are not generalizable to other studies.

The use of filter cigarettes was also looked at, but unfortunately the habit was recent in this population. Approximately 50% of cigarette smokers smoked filter cigarettes at the time of interview, while only 10% had smoked them for 10 yr. This relatively brief exposure did not yield any suggestive or significant results when the effects of filter variables (filters, filters × amount, filters × age) were investigated.

Among exsmokers of cigarettes the effect of age quit smoking and number of years since last smoked were investigated but were not found to yield statistically significant results after the variables in Table 3 had been accounted for; however those coded as not smoking 10 yr before showed a significantly ($p < .01$) higher, by .083 L., FEV_1 than other exsmokers.

The relationship of reasons for giving up smoking among the exsmokers of cigarettes to the FEV_1 was also investigated. Dummy variables were assigned to categories "doctor's orders," "symptoms," "distaste," and "no information," and effects of −.04, −.06, +.03, and −.02 liters, respectively, were found (these are relative to values for all other exsmokers who gave financial or other reasons for giving up). While the set of effects is statistically significant, different recollected reasons for giving up smoking are certainly not associated with major differences in pulmonary impairment.

Miscellaneous Variables

Table 4 shows the effect of several miscellaneous variables on the FEV_1 in the five smoking categories. While the marginal effects of these variables on the re-

TABLE 4
THE EFFECT OF MISCELLANEOUS VARIABLES ON THE FEV_1 AMONG NEW YORK CITY POSTAL AND TRANSIT WORKERS AGED 30–59[a]

	Regression Coefficients[b]				
	Nonsmokers	Cigarette smokers	Pipe/cigar smokers	Pipe/cigar cigarette	Exsmokers/ cigarette
Foreign born[c]					
Puerto Rican	−.172	−.126	−.236	−.751 **	−.059
Irish	−.129	+.083	−.161	−.086	−.084
Italian	−.137	−.025	+.036	+.031	+.047
Other foreign born	−.045	−.016	−.033	+.132	−.010
Marital status[d]					
Single	+.166 **	−.003	+.016	+.157	+.003 *
Divorced, widowed, separated	+.075	+.011	−.028	+.067	−.302
Transit Authority occupation[e]					
Dispatcher	+.173 **	−.053	−.014	+.014	−.193 *
Motorman	+.138	−.032	+.012	+.065	+.050
Sleeps with window open[f]					
Except in cold weather	+.036	−.013	+.049	−.128 *	−.048
Never	−.014	+.046	−.026	+.177	+.073

[a] Significance at the .05, and .01 levels given by one or two asterisks, respectively, for set of variables bordered by vertical line.

[b] All variables are dummy variables; hence the regression coefficients are net effects (in liters) relative to the base category. Equations included anthropometric and smoking variables shown in Tables 1 and 3.

[c] Effects are relative to native born.

[d] Effects are relative to married.

[e] Effects are relative to Transit Authority bus drivers.

[f] Effects are relative to those who always sleep with window open.

gression with age was investigated (except for the foreign born categories, known to have atypical age distributions), none of the interaction terms of age with these variables was significant. The effects can therefore be presented as constants, making them easier to interpret.

Foreign birth among nonsmokers is associated with slightly lower FEV_1 values in all ethnic groups. The effect holds across smoking categories only for Puerto Ricans who show consistently lower FEV_1 values than the native born; the Irish are lower than the native born in all smoking categories except cigarette smokers, but the effect is not so strong as among the Puerto Ricans. The Italian and "other foreign born" groups do not show strong or consistent patterns among the smoking groups.

While marital status is statistically significant among both nonsmokers and exsmokers, no meaningful pattern emerges from the data.

Transit Authority occupation, dispatchers and motormen compared to surface line operators, is also statistically significant only among nonsmokers and exsmokers. Both dispatchers and motormen show distinctly higher FEV_1 values than surface line operators among nonsmokers; the superiority of motormen over surface line operators is retained, though much reduced, in all smoking categories

except current cigarette smokers. The dispatchers, a much smaller group of men, show no such consistent pattern.

The variable "sleeps with window open" is nonsignificant except among pipe/cigar/cigarette smokers; this is best presumed to be a chance finding since similar effects are found neither among cigarette smokers nor among pipe/cigar smokers.

DISCUSSION

Anthropometric Variables

Stebbings (1971a) reviewed the literature assessing the contribution of sitting height and weight to the prediction of the FEV_1 after the effect of standing height had been accounted for. While several investigators had found these variables to be of little or no importance, Ashford and co-workers (1965, 1968) in a massive study had found that sitting height and weight did contribute significantly to the prediction of the FEV_1 in many groups of miners. Stebbings (1971a) found that weight contributed significantly to the prediction of the FEV_1

Ferris and Stoudt (1971) published a study of the effect of a large number of anthropometric variables on pulmonary function measurements; they concluded "that age and standing height are still the best predictors for the simple tests of pulmonary function in a white population," although they suggested that for interethnic comparisons other measurements might also be useful.

The results presented in this paper support Ashford and co-workers in finding a significant effect of sitting height and weight beyond the effect of standing height. These relatively small but statistically significant contributions, however, may be of little practical importance: this seems to be the general consensus of investigators.

It would seem, however, that the "importance" should depend on the use to which the prediction of the FEV_1 is to be put. One may wish to predict the FEV_1 for at least three reasons: (1) to make clinical judgements, (2) to correct for the effects of age and body size so that one may investigate the effect of environmental influences, and (3) in mass screening of the population for respiratory impairment.

The use of variables other than age and standing height would seem to be an unnecessary refinement in clinical practice. However, environmental influences such as residential or occupational exposure to air pollution, or influences associated with lower social class, are relatively weak, of the same order of magnitude as the effects of sitting height and weight whether measured in absolute values or by the proportion of the variance contributed. (The effect of cigarette smoking is sufficiently strong as to be detectable under almost any circumstances, and is not under discussion here.) When searching for the effects of environmental influences it would seem to be desirable, as the conservative practice, to account for the effects of sitting height and weight when the data collected allows it. This does not involve any large expenditure of labor in analysis; there is also minimal effort involved in collecting the data unless one is doing home interviewing.

In mass screening the question of the usefulness of sitting height and weight becomes complicated. Here the question becomes: what is the effect of using sitting height and weight on the assignment of individuals to one side or another

of the screening level? This cannot be predicted without making unwarranted assumptions; it has to be measured.

Such an analysis was carried out. The prediction equations estimated from the nonsmoker group were applied to each of the five smoking categories. Screening levels from −.75 L. to −1.5 L. in increments of 0.5 L. were looked at; at each screening level those below the screening level when the equation incorporating standing and sitting height, weight, and age was used were considered truly impaired, and those below the screening level when the equations used incorporated only standing height and age were considered to be screened positive. Results from all five smoking groups combined, screened both at the −1.0 L. ("probably impaired") and −1.5 L. ("impaired") levels are presented in Table 5.

Although the use of age and standing height alone yields a specificity of almost 100%, the sensitivity is only 90% at the −1.0 L. level and 93% at the −1.5 L. level. Table 5 does demonstrate that misclassified respondents, the false positives and false negatives, are a respectable fraction, about 15%, of the screening yield. Since the misclassified individuals are borderline cases anyway, the importance of the misclassifications is open to question: the answer depends on the costs and benefits ascribed to each of the four categories (Blumberg, 1957), and that is too broad a question to be discussed here except to note that the marginal cost of measuring sitting height and weight in the screening situation is minimal.

The results presented here may or may not adequately represent other mass screening situations; however, the principle that the decision to include or exclude variables from the prediction equation should be based not on the amount of population variability explained but on their effect upon the classification of individuals across the screening level seems generally valid.

This principal should hold not only for sitting height and weight, which we use as approximations of body size or lung volume, but also to other variables such as chest breadth and hand grip strength which may be of use in predicting the

TABLE 5
EFFICIENCY OF AGE AND STANDING HEIGHT ALONE IN PREDICTING THE FEV_1 IN MASS SCREENING[a]

	Respiratory impairment							
	FEV_1 1.0 + L. below predicted				FEV_1 1.5 + L. below predicted			
	Yes		No		Yes		No	
	No.	%	No.	%	No.	%	No.	%
Screened								
Positive	425	5.5	21	0.3	98	1.3	8	0.1
Negative	47	0.6	7216	93.6	7	0.1	7596	98.5
Total	472	6.1	7237	93.9	105	1.4	7604	98.6

[a] For five smoking categories combined. Respiratory impairment is defined by deviation below value predicted from age, standing height, sitting height, and weight. Screening equation includes only age and standing height. Both equations estimated from nonsmoker group.

FEV_1 (Stebbings, 1971a) but are particularly likely to be of use in predicting flow measurements (Ferris and Stoudt, 1971).

Smoking Variables

Tables 1 and 2 showed the expected lower values of the FEV_1 among smokers and the greater rates of decline with age. The conclusions of interest here, however, are more specific ones concerning the effects of detailed differences in smoking habits on the FEV_1.

The effect of inhaling was found to be significant among cigarette smokers, but not pipe/cigar smokers, if tested for significance ahead of the other smoking variables. However there was not the slightest evidence of an accelerated rate of decline in pulmonary function with age among noninhalers, nor was inhaling significant when tested after amount smoked and age began smoking. One might conclude that whether or not one knowingly inhales tobacco smoke has no effect on the amount of permanent damage done to the lungs.

There are several possible explanations for this. One is that interview data on inhaling is worthless: the finding that noninhaling has a highly significant constant effect when tested before the other smoking variables is evidence against that hypothesis. A second explanation is that noninhalers make up such a small fraction of persons smoking that the effects are undetectable: however, 19.4% of pipe/cigar/cigarette smokers were noninhalers, although only 5.2% of cigarette smokers were. In addition 15.8% of pipe/cigar smokers were inhalers. These large fractions in two smoking categories, the large absolute numbers of persons in the analysis, and the consistency of negative results among these three smoking categories suggest that the negative findings are real.

There is no question that noninhalers, assuming full accuracy of their responses, respire in a more highly polluted personal atmosphere than nonsmokers. If it can be assumed, and the assumption is of uncertain validity, that noninhalers do not inhale as much smoke as those persons who admit to inhaling, one may hypothesize that the permanent and cumulative effect of tobacco smoke on pulmonary function is not determined by the quantity of smoke inhaled but by the average daily fraction of time during which smoke concentrations in inhaled air are above some minimum level.

One would then expect that more efficient filtering of cigarette smoke would not help to reduce the decline of pulmonary function with age in smokers.

Ferris *et al.* (1971) did not find that use of filter tip cigarettes affected the prevalence of chronic nonspecific respiratory disease (as defined by symptoms and the FEV_1/FVC ratio) among current cigarette smokers. Comstock *et al.* (1970) in a 5-yr longitudinal study found that while smokers of filter cigarettes had slightly higher FEV_1's (.15 L. on the first round, .04 L. on the second) the change over time, $-.40$ L. for filter users and -0.29 for nonfilter users, did not suggest that filter cigarette smokers do not suffer impairment of pulmonary function as rapidly as nonfilter cigarette smokers. Frequency of cough and phlegm, however, was reduced among smokers of filter cigarettes.

Ferris *et al.* (1971) did find, however, that noninhalers, of both sexes and who

smoked at least 15 cigarettes per day, had less chronic nonspecific respiratory disease (as defined by symptoms and the FEV_1/FVC ratio).

In summary, one can say that, while there is some evidence in the literature for a beneficial effect of noninhaling and use of filter cigarettes on chronic bronchitis, there is no evidence that impairment of pulmonary function is retarded. The data presented in this paper suggests that noninhaling (and, indirectly, use of filter cigarettes) does not reduce the rate of decline of pulmonary function.

Also of particular interest, although not unexpected, is the strongly significant effect of the age at which an individual started smoking cigarettes. Two variables, "age began smoking" and "nonsmoker 10 years before" both yielded consistent findings of higher FEV_1 values in those who began smoking later in life or who stopped smoking for a period. This implies that a smoker does not catch up with other smokers in his age group, but retains permanently the advantages accrued during his earlier nonsmoking years.

The last question concerning the smoking history to be considered is the value of including in an analysis more smoking variables than type and amount of tobacco currently consumed. Ignoring the effect of inhaling, the inclusion of other smoking variables (amount × age, age began, amount 10 yr before) increased the percentage of variance explained by 32% among cigarette smokers and by 97% among pipe/cigar/cigarette smokers after accounting for the effect of additional degrees of freedom.

These effects are of the same order of magnitude as other environmental effects of interest to epidemiologists, such as the effects of social class and urban residence (Stebbings, 1971b,c). It would appear desirable, or at least the conservative practice, to account for these additional smoking factors in any analysis of the effects of variables beyond tobacco smoking. At least when dealing with continuous variables such as pulmonary function data these additional smoking variables are easily definable and are readily incorporated into the analysis.

Miscellaneous Variables

These variables (foreign birth, marital status, Transit Authority occupation, and sleeping with window open) yielded some significant findings, but none of great interest.

The strongest finding was the consistently low FEV_1 values among the Puerto Rican born; whether this is due to racial admixture of Negroes (who usually have lower FEV_1's and FVC's; Damon, 1966) is not known and cannot be determined from the data. Puerto Ricans were coded as whites irrespective of skin color. Except among cigarette smokers the Irish also showed consistently low FEV_1 values, although not so low as the Puerto Rican.

Although marital status variables were significant in two groups, no consistent pattern emerged so little can be said other than that there are clearly social variables affecting pulmonary function variables which we have yet to define and measure.

The Transit Authority occupations were tested to see if there was a suggestion of an influence of occupational air pollution exposure. From measurements made in 1965–1966, bus drivers are thought to be exposed to higher levels of total res-

pirable particulates, including organic and sulfate components, than are motormen (O'Connor, 1967). Exposure to NO and aldehydes is also higher, and possibly exposure to NO_2 and SO_2. Nonsmoking bus drivers were, in fact, found to have significantly lower FEV_1's.

Limitations in the air pollution data, namely the small number of routes sampled and the differing daily durations of sampling of bus routes and subway lines, make it difficult to assign numerical values to the differences in air pollution exposure: some conclusions are possible, however.

Measured from a sampling van following bus routes, total respirable particulates (by Aerotec sampler) ranged from 130 μg per cubic meter in Queens to 247 in Manhattan. Sampling from a conductor's compartment on the elevated Flushing (Queens) subway line, on the other hand, yielded an annual mean of 89 while the annual mean for the underground Broadway-7th Avenue line (Manhattan) was 135. The greatest differences were in aldehyde levels where the maximum seasonal (fall) level recorded on the subway lines was 1.36 parts per hundred million on the Flushing line; the minimum level reported on a bus route was 1.92 pphm in the Bronx.

What may be concluded is that respirable particulate levels in the subways are not higher than on the streets, as was at first suspected, and that bus drivers are exposed to higher levels of those pollutants produced by New York's heavy automobile and truck traffic. This last conclusion is obviously reasonable even without the limited supporting data.

While the results presented here suggest that at least among the nonsmokers such an effect exists, such an air pollution effect cannot be stated with certainty. Although the interaction terms of these occupational variables with age were not significant, the estimated coefficients indicated a greater rate of decline with age among dispatchers and motormen than among bus drivers; this is not consistent with the hypothesized effect of air pollution. However, it could be argued that, since dispatchers and motormen, especially the older ones, usually began as bus drivers, the older motormen and dispatchers had more exposure as bus drivers and hence lower values than individuals who became motormen at an earlier age. Without more detailed information concerning employment histories, and also about the process, in the past, of selecting bus drivers for motorman duties, it would be foolhardy to conclude with certainty that an effect of occupational exposure to air pollution exists. Densen *et al.* (1967) found no difference in the prevalence of symptoms among the different categories of Transit Authority workers. Nevertheless, the statistically highly significant results presented here for nonsmokers remain suggestive.

The question "do you sleep with your window open or closed" has appeared in quite a number of questionnaires in the past, though the responses have never, to this investigator's knowledge, been analysed and published. Out of perverse curiosity the variable was investigated here and found not to be of any use, a result which is unlikely to occasion great suprise or consternation among investigators of chronic respiratory disease. I hope these results might even allay some small guilt feelings among those who are compulsive about analysing all collected data!

ACKNOWLEDGMENTS

The author wishes to thank Dr. Paul M. Densen and Mrs. Ellen Jones for allowing and encouraging him to continue analysis of this data.

REFERENCES

ASHFORD, J. R., AND BROWN, S. (1965). The pulmonary ventilatory function of coal miners in the United Kingdom. National Coal Board, Pneumoconiosis Field Research, mimeographed.

ASHFORD, J. R., BROWN, S., MORGAN, D. C., AND RAE, S. (1968). The pulmonary ventilatory function of coal miners in the United Kingdom. *Amer. Rev. Resp. Dis.* **97**, 810–826.

BERGLUND, E., BIRATH, G., BJURE, J., GRIMBY, G., KJELLMER, I., SANDQUIST, L., AND SODERHOLM, B. (1963). Spirometric studies in normal subjects. I. Forced expirograms in subjects between 7 and 70 years of age. *Acta Med. Scand.* **173**, 185–192.

BLUMBERG, M. S. (1957). Evaluating health screening procedures. *Operations Res.* **5**, 351–360.

COHEN, J. (1968). Multiple regression as a general data-analytic system. *Psych. Bull.* **70**, 426–443.

COMSTOCK, G. W., BROWNLOW, W. J., STONE, R. W., AND SARTWELL, P. E. (1970). Cigarette smoking and changes in respiratory findings. *Arch. Environ. Health* **21**, 50–57.

DAMON, A. (1966). Negro–White differences in pulmonary function (vital capacity, timed vital capacity, and expiratory flow rate). *Human Biol.* **38**, 380–393.

DENSEN, P. M., JONES, E. W., BASS, H. E., AND BREUER, J. (1967). A survey of respiratory disease among New York City postal and transit workers. I. Prevalence of symptoms. *Environ. Res.* **1**, 265–286.

DENSEN, P. M., JONES, E. W., BASS, H. E., BREUER, J., AND REED, E. (1969). A survey of respiratory disease among New York City postal and transit workers. II. Ventilatory function test results. *Environ. Res.* **2**, 277–296.

FERRIS, JR., B. G., HIGGINS, I. T. T., HIGGINS, M. W., PETERS, J. M., VAN GANSE, W. F., AND GOLDMAN, M. D. (1971). Chronic nonspecific respiratory disease, Berlin, New Hampshire 1961–1967. A cross-sectional study. *Amer. Rev. Resp. Dis.* **104**, 232–244.

FERRIS, JR., B. G., AND STOUDT, H. W. (1971). Correlation of anthropometry and simple tests of pulmonary function. *Arch. Environ. Health* **22**, 672–676.

GREGG, I. (1968). A study of the causes of progressive airways obstruction in chronic bronchitis. *In* "Current Research in Chronic Respiratory Disease. Proceedings Eleventh Aspen Emphysema Conference, Aspen, Colorado, June 12–15, 1968." Public Health Service, Arlington, Virginia.

JOHNSTON, J. (1963). "Econometric Methods." McGraw-Hill Book Company, Inc., New York.

KHOSLA, T., AND LOWE, C. R. (1967). Indices of obesity derived from body weight and height. *Brit. J. Prev. Soc. Med.* **21**, 122–128.

O'CONNOR, J. R. (1967). Levels of selected atmospheric contaminants encountered by transit workers in New York City. United States Public Health Service, Cincinnati, mimeographed.

STEBBINGS, JR., J. H. (1971a). Chronic respiratory disease among nonsmokers in Hagerstown, Maryland. II. Problems in the estimation of pulmonary function values in epidemiological surveys. *Environ. Res.* 4, 163–192.

STEBBINGS, JR., J. H. (1971b). Chronic respiratory disease among nonsmokers in Hagerstown, Maryland. III. Social class and chronic respiratory disease. *Environ. Res.* 4, 213–232.

STEBBINGS, JR., J. H. (1971c). Chronic respiratory disease among nonsmokers in Hagerstown, Maryland. IV. Effects of urban residence on pulmonary function values. *Environ. Res.* 4, 283–304.