Forced expiratory volume in one second: why dichotomise?

Jørgen Vestbo, Kim M Knudsen, Finn V Rasmussen

Numerous epidemiological surveys have tried to evaluate the effects of exposures relevant to pulmonary disease.¹ A number of studies have analysed forced expiratory volume in one second (FEV₁) using linear regression analysis, thus enabling adjustment for age, height, and sex as well as potential confounding variables. Within the last few years, studies analysing dichotomous FEV₁ variables with different cutoff points in multivariate regression models have been published.^{2,3}

As there seems to be no obvious reason for dichotomising the continuous FEV₁, we have compared the results of the two different approaches to analysis of FEV₁.

Material and methods

Our study is based on a cross sectional survey of 1404 men carried out in 1974.⁴ In the present study we have used information on 1045 men representing three occupational categories: cement workers (277 men), other blue collar workers (574 men), and white collar workers (194 men). All the men had spent at least half of their working life in the occupational category in which they were placed.⁵ As well as occupational categories, smoking category at the time of interview was considered to be of interest. Men were classified as never smokers (less than one cigarette a day for one year; 88 men), ex-smokers (with no limit to the duration of the abstinence period; 204), light smokers (less than 10 g of tobacco a day; 246 men), and heavy smokers (10 g of tobacco or more a day; 507 men). Spirometry was performed using a Godart Expirograph type 16 000 (Bilthoven,

Holland). All indices were reported at body temperature, pressure, and saturation as mean values of three measurements. In this study the following indices of spirometry were investigated. The continuous variable FEV₁ (FEV₁-cont); The dichotomous FEV₁ variables (FEV₁-65, FEV₁-70, FEV₁-75, FEV₁-80, and FEV₁-85). For the calculation of predicted values we used the equation from a Danish reference population.⁶ Dichotomous FEV₁ variables were constructed using 65%, 70%, 75%, 80%, and 85% of predicted as dividing values respectively.

Smoking habits were entered as a factor with four levels in the multivariate models as three indicator variables: ex-smokers *v* never smokers, light *v* never smokers, and heavy *v* never smokers. Likewise, occupation was entered having three levels as two indicator variables: cement workers *v* white collar workers and blue collar workers *v* white collar workers. The continuous FEV₁ measurement, FEV₁-cont, was used as the dependent variable in a linear regression analysis including age and height as well as smoking habits and occupation:

$$\text{FEV}_1\text{-cont} = \text{intercept} + \beta_a \times \text{age} + \beta_h \times \text{height} + \beta_{s1} \times \text{ex-smoker} + \beta_{s2} \times \text{light smoker} + \beta_{s3} \times \text{heavy smoker} + \beta_{o1} \times \text{cement worker} + \beta_{o2} \times \text{blue collar worker}$$

The five dichotomous indices of spirometry were analysed using a multivariate logistic regression model with the dichotomous indices as the dependent variable. Separate analyses were made for each of the dichotomous indices. With *p* denoting the risk of having a lower FEV₁ value than the % of the predicted, the model equation is:

$$\log(p/(1-p)) = \text{intercept} + \beta_{s1} \times \text{ex-smoker} + \beta_{s11} \times \text{light smoker} + \beta_{s111} \times \text{heavy smoker} + \beta_{o1} \times \text{cement worker} + \beta_{o11} \times \text{blue collar worker}$$

Results

A total of 994 men (95%) performed an acceptable spirometry; 130 men had an FEV₁ < 65%. For 70%, 75%, 80%, and 85% the numbers were 162, 220, 311, and 426 men respectively. Table 1 shows the effects of the selected variables of interest on the continuous FEV₁. The effects are expressed as devi-

Medical Department P and the Chest Clinic, Bispebjerg Hospital, DK-2400 Copenhagen NV, and Clinic of Occupational Health, Rigshospitalet, DK-2200 Copenhagen N, Denmark

J Vestbo

Statistical Research Unit, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

K M Knudsen

Medical Department B, Hillerød Hospital, DK-3400 Hillerød, Denmark

F V Rasmussen

Table 1 Determinants of FEV₁ estimated using multiple linear regression analysis (n = 994)

Variable	FEV ₁ -cont Estimated regression coefficient	(95% CI)	p Value
Intercept	3.61 l		
Age (per 10 years)	-0.41 l	(-0.47 to -0.35)	<0.0001
Ex-smokers*	-0.19 l	(-0.35 to -0.03)	0.02
Light smoking*	-0.27 l	(-0.43 to -0.11)	0.001
Heavy smoking*	-0.38 l	(-0.53 to -0.23)	<0.0001
Cement workers†	-0.22 l	(-0.34 to -0.10)	0.0002
Blue collar workers†	-0.22 l	(-0.33 to -0.11)	0.0001
Height	0.032 l/cm	(0.026 to 0.038)	<0.0001

*By contrast with never smokers. †By contrast with white collar workers.

ations from the intercept, which can be read as the expected FEV₁ for a 50 year old 170 cm tall non-smoking white collar worker. As an illustration, men classified as cement workers had in average 220 ml less in FEV₁ than white collar workers (95% confidence interval (95% CI) 100–340 ml). Table 2 shows the results of multivariate logistic regression analyses for each of the five dichotomous FEV₁ variables. Age and height were not included as independent variables because adjustment was already made using predicted values. For all dichotomous variables the determinants are expressed as odds ratios (ORs). An OR of 1.00 indicates no increased risk of the outcome under study. As an example, men classified as light smokers had roughly twofold increased probability (OR = 2.0) of having an FEV₁ < 70% of predicted values in comparison with never smokers.

From table 1 it can be seen that the index of regression coefficients for heavy smokers/cement workers was 1.7. For the dichotomous variables in table 2, a comparable index of log ORs for heavy smokers/cement workers varied from 1.4 for FEV₁–70 to 2.5 for FEV₁–75.

Discussion

Our comparison of different analyses of FEV₁ did not give particularly surprising results. All analyses were qualitatively fairly consistent—that is, heavy smoking appeared to have a greater impact than either light or ex-smoking and the effects of cement work and other blue collar work was, in no analysis, greater than that of heavy smoking. Quantitatively, however, differences appeared. The “relative importance” of smoking and occupation varied depending on the dividing point used for dichotomising the FEV₁ variable. Also, associations in the logistic regression analyses were weaker than in the linear regression analysis. This was due to the reduction in information in the process of dichotomising a linear outcome variable. Thus, our results indicate that comparison of surveys reporting different indices of spirometry and using different statistical models must be made with great caution. In our study we used % of predicted to define “abnormal”. This choice is in no way self evident.^{7,8} We chose % of predicted as this is both widely used and understood. It seems to us, however, that no matter how “disease” is defined, the reasonableness of a division in respiratory disease of yes or no using one particular dividing point is questionable.

Large epidemiological studies on respiratory dis-

Table 2 Determinants of the five dichotomous FEV₁ variables (n = 994) estimated using multiple logistic regression analyses

Variable	FEV ₁ –65	FEV ₁ –70	FEV ₁ –75	FEV ₁ –80	FEV ₁ –85
Ex-smokers:*					
OR	1.9	1.9	2.1	1.8	1.7
p Value	0.16	0.13	0.06	0.07	0.07
95% CI	0.7–4.8	0.8–4.5	0.9–4.8	0.9–3.5	0.9–3.1
Light smokers:*					
OR	1.8	2.0	2.5	1.9	2.2
p Value	0.19	0.09	0.01	0.04	0.005
95% CI	0.7–4.5	0.9–4.7	1.1–5.6	1.0–3.7	1.2–4.0
Heavy smokers:*					
OR	2.1	2.4	3.2	2.8	3.4
p Value	0.07	0.02	0.0006	0.0004	<0.0001
95% CI	0.9–5.1	1.1–5.3	1.5–6.9	1.5–5.1	2.0–5.8
Cement workers:†					
OR	1.6	1.9	1.6	1.7	2.1
p Value	0.15	0.02	0.06	0.01	0.0004
95% CI	0.8–2.9	1.1–3.4	1.0–2.6	1.1–2.7	1.4–3.1
Other blue collar workers:†					
OR	1.8	1.9	1.7	1.6	1.8
p Value	0.03	0.01	0.01	0.01	0.001
95% CI	1.0–3.2	1.1–3.2	1.1–2.7	1.1–2.4	1.2–2.5

*By contrast with never smokers. †By contrast with white collar workers.

ease tend to concentrate on a single or a few spirometric parameters. This is done mainly for the purpose of clarification, for the purpose of testing many exposures or combinations of exposures, and because of limitations on length of manuscripts submitted to medical journals. With a possible bias of such journals towards mostly publishing positive studies,⁹ investigators may be tempted to select results that seem most "interesting" in terms of p values and ORs. For both editors and readers of medical journals this can be difficult to discern; therefore, papers reporting indices other than the continuous FEV₁ ought to state the reason for their choice and preferably mention how replacement of their parameter with the continuous FEV₁ variable would influence their results.

This study was supported by grants from The Danish Medical Research Council (12-5877) and The National Union against Lung Diseases.

Requests for reprints to: Jørgen Vestbo, Medical Department P, Bispebjerg Hospital, Bispebjerg Bakke 23, DK-2400 Copenhagen NV, Denmark.

- 1 Hensley MJ, Saunders NA, eds. *Clinical epidemiology of chronic obstructive pulmonary disease*. New York: Marcel Dekker, 1989.
- 2 Korn RJ, Dockery DW, Speizer FE, Ware JH, Ferris BG Jr. Occupational exposures and chronic respiratory symptoms. A population-based study. *Am Rev Respir Dis* 1987;136:298-304.
- 3 Marine WM, Gurr D, Jacobsen M. Clinically important respiratory effects of dust exposure and smoking in British coal miners. *Am Rev Respir Dis* 1988;137:106-12.
- 4 Rasmussen FV, Borchsenius L, Holstein B, Sølvsteen P. Lung function and long-term exposure to cement dust. *Scandinavian Journal of Respiratory Diseases* 1977;58:252-64.
- 5 Vestbo J, Knudsen KM, Rasmussen FV. Should we continue using questionnaires on breathlessness in epidemiologic surveys? *Am Rev Respir Dis* 1988;137:1114-8.
- 6 Danish Society of Pulmonary Medicine. *Spirometry, a recommendation*. Odense: AiO Tryk as, 1986. (In Danish.)
- 7 Harber P, Tockman M. Defining "disease" in epidemiologic studies of pulmonary function: percent of predicted or difference from predicted? *Bull Europ Physiopath Respir* 1982;18:819-28.
- 8 Kauffmann F. Means or cut-offs? *Bull Europ Physiopath Respir* 1983;19:73-5.
- 9 Fletcher RH, Fletcher SW, Wagner EH. *Clinical epidemiology*. Baltimore: Williams & Wilkins, 1982.

Accepted 7 October 1991