

Lung function prediction equations derived from healthy South African gold miners

Eva Hnizdo, Gavin Churchyard, Rob Dowdeswel

Abstract

Objectives—To estimate lung function prediction equations and to identify appropriate normal reference values for the population of about 250 000 of South African gold miners.

Methods—Data from a lung function screening programme conducted at a large South African gold mine from 1994 to 1998 were used to estimate the lung function prediction equations. The most reliable period of lung function testing was identified in a previous study of a temporal pattern in reliability, and lung function tests from this period were used. Miners with a history of pulmonary tuberculosis or with radiological abnormalities were excluded from the study. The prediction equations were estimated cross sectionally on 15 772 black and 2752 white miners, and published reference equations that fitted most closely the observed data were identified.

Results—The estimated prediction equations for forced vital capacity (FVC) are as follows: for black men, $FVC(l) = -2.901 - 0.025 \times \text{age} + 4.655 \times \text{height}$; and for white men, $FVC(l) = -4.407 - 0.036 \times \text{age} + 5.940 \times \text{height}$. For forced expiratory volume in one second (FEV_1) these equations are: for black men, $FEV_1(l) = -1.654 - 0.30 \times \text{age} + 3.665 \times \text{height}$; and for white men, $FEV_1(l) = -2.341 - 0.038 \times \text{age} + 4.314 \times \text{height}$. Units are years for age and metres for height. Knudson's and the European Community of Coal and Steel (ECCS) reference values provided the closest fit to the data on lung function of white miners, but the lower limits of normal from the ECCS equations were the closest to the observed one sided lower 95% confidence intervals (95% CIs). For black miners, reference equations that fitted best were derived by Louw *et al* on asymptomatic black South African men unexposed to occupational dust. There were significant differences between the two groups of miners in the estimated height adjusted mean lung function values for a 40 year old 1.7 m tall man (220 ml (5.2%) for FVC and 110 ml (3.2%) for FEV_1); white men had higher FVC and FEV_1 , but lower FEV_1/FVC ratio. The ECCS reference values scaled by a conversion factor of 0.93 for the FVC and 0.95 for the FEV_1 provided close fits to the data for black miners, but the rate of decline with age was higher than that in the observed data. None of the linear equations provided a good fit for the 20–29 and more than 55 years old age categories.

Conclusion—The ECCS and Knudson equations provided the best fit to the data for white miners, whereas the equations by Louw *et al* estimated on asymptomatic black South African bank workers provided the best fit to the data for black miners. The ECCS reference values scaled by a factor of 0.93 for FVC and by 0.95 for FEV_1 provided close fits, but the rate of decline with age was higher than that in the data for black miners.

(Occup Environ Med 2000;57:698–705)

Keywords: silica dust; miners; pulmonary function reference equations

The availability of appropriate normal reference values is fundamental to a workplace lung function monitoring programme. Good reference values allow for an early detection of accelerated loss of pulmonary function, for assessment of fitness to work in certain job categories, to wear respirators, and for the diagnosis of a compensational disease, among other purposes.¹

The South African gold mining industry currently employs about 250 000 miners. Because of exposure to dust with a high concentration of crystalline silica, the miners are at risk of developing respiratory diseases—such as silicosis, chronic obstructive lung disease, and pulmonary tuberculosis. Also, there is an extensive population of ex-miners who may be eligible for compensation due to existing lung diseases. Legislation for screening of pulmonary function in all miners has been introduced in South Africa only recently.¹ However, there is uncertainty as to the most appropriate reference values for assessing the miners. The European Community for Coal and Steel (ECCS) reference values² have been suggested for general use,³ but their suitability, particularly for black miners, has never been evaluated. The goal of this study was to identify reference equations appropriate for the large population of gold miners who are predominantly black, but with a significant white subpopulation. The approach adopted was to estimate prediction curves for lung function measurement obtained on a healthy subset of current black and white gold miners and to evaluate how well published reference equations^{4,5} compare with the estimated prediction curves.

Material and methods

STUDY POPULATION

Miners from a large South African gold mining company who had spirometry performed routinely at entry into the industry (initial

National Centre for Occupational Health, Johannesburg 2000, South Africa
E Hnizdo

Arum Health Research, Welkom, South Africa
G Churchyard

Precious Metal Refiners, PO Box 331, Kroonsdal 0350, South Africa
R Dowdeswel

Correspondence to:
Dr Eva Hnizdo, National Institute for Occupational Safety and Health, 1095 Willowdale Rd, MS 2800, Morgantown, WV 26505, USA
exhb@cdc.gov

Accepted 25 May 2000

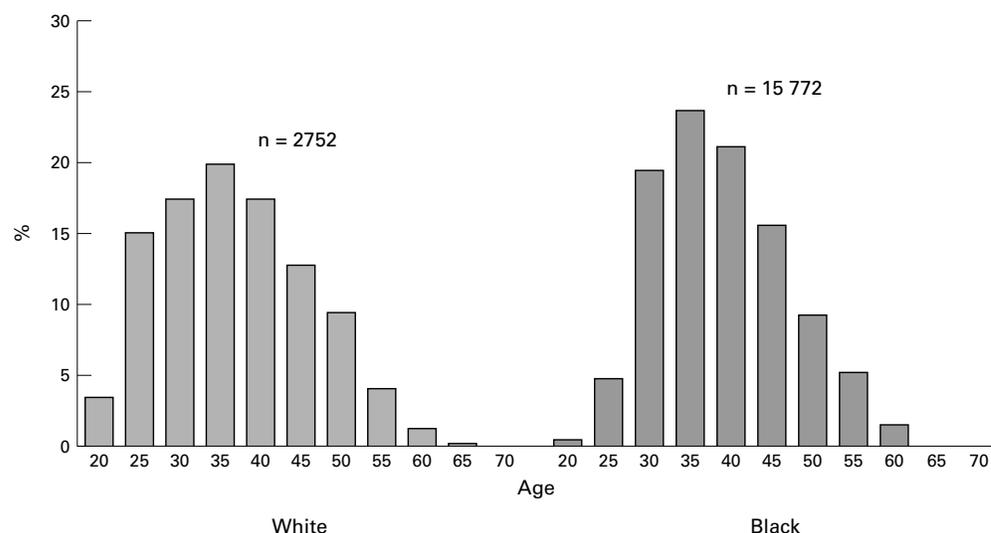


Figure 1 Age distribution for white and black miners.

examination), periodically at 3 year intervals (periodic examination), and on leaving the company (exit examination) comprised the study population. The lung function monitoring programme started in May 1994 and served a population of 71 515 miners in 1994, which decreased to 43 359 miners by 1998.⁶

SPIROMETRY MEASUREMENTS

Maximal forced expiratory manoeuvres were recorded with a Hans Rudolph pneumotachograph (Flowscan, Electromedical Systems). The system software required and validated calibration with a 3 litre syringe and allowed keyboard entry of barometric pressure and ambient temperature for the correction to body temperature, pressure, and saturation (BTPS) conditions. Calibration was done 3–4 times a day. During testing, flow versus volume tracings were displayed. A minimum of three acceptable and reproducible forced expiratory manoeuvres are obtained according to the standards recommended by the American Thoracic Society (ATS).⁷ All testing was done by nursing personnel with a college diploma specialisation in spirometry testing and trained in the techniques of performing spirometry to ATS standards. Height was measured to the nearest centimetre (without shoes). Data recorded for each test included the date of test, date of birth, height, weight, forced vital capacity (FVC), forced expiratory volume in one second (FEV_1), and forced expiratory flow at 25%–75% of forced vital capacity ($FEF_{25-75\%}$).

RELIABILITY OF THE LUNG FUNCTION DATA

During the period May 1994 to March 1998, 113 120 spirometers were recorded in a computerised database. In a previous study⁶ we have examined reliability of the spirometers over time, and in the present study we used data from the most reliable period from January 1995 to August 1996 for which the average coefficient of reliability, G , was 0.93 (the size of random error of measurement was 7%). In total, 45 053 tests were done during the reliable period.⁶ Of these, 36 777 (31 108 on black, and 3270 on white miners) were on miners for whom a complete history of pulmonary tuberculosis and chest radiology was available. The remaining tests were from contract workers with incomplete information. From the 31 108 tests on black miners, we excluded 3905 subjects with a history of pulmonary tuberculosis and 1485 subjects with radiological changes (pulmonary tuberculosis, pneumoconiosis, pneumonia, cardiomegaly, pleural changes, etc; there was an overlap between cases). (All cases of pulmonary tuberculosis were computerised from 1979 to 1998. The yearly radiological screening was done on miniature 100×100 mm chest radiographs, and subjects with abnormalities were recorded cross sectionally from 1994 to 1998.) Of the remaining 26 024 tests, we excluded 71 miners whose age was less than 20 years or more than 63 years, and 255 miners (0.98%) and 84 subjects whose lung function or height measurements were outside the 99.98% confidence

Table 1 The observed means for lung function, the prediction curves, and predicted values for black and white miners

Lung function	Observed values		Regression coefficient (SE)			R^2 (%)	SEE	Prediction, age 40, height 1.7 m
	Mean	SE	Intercept	Age	Height			
Black miners n=15772, age 38.7 (SE 0.06), height 1.71 m (SE 0.005):								
FVC (l)	4.10	0.005	-2.901 (0.120)	-0.025 (0.001)	4.655 (0.068)	33.5	0.528	4.02
FEV_1 (l)	3.44	0.005	-1.654 (0.109)	-0.030 (0.001)	3.665 (0.061)	35.5	0.477	3.38
FEF (l/s)	3.96	0.010	1.363 (0.274)	-0.056 (0.001)	2.780 (0.154)	14.6	1.200	3.86
$FEV_1\%$	84.1	0.055	102.56 (1.511)	-0.237 (0.007)	-5.437 (0.848)	7.24	6.627	83.8
White miners n=2752, age 36.9 (SE 0.06), height 1.80 m (SE 0.001):								
FVC (l)	4.96	0.016	-4.407 (0.339)	-0.036 (0.001)	5.940 (0.182)	44.1	0.547	4.24
FEV_1 (l)	4.04	0.014	-2.341 (0.300)	-0.038 (0.001)	4.314 (0.161)	43.4	0.484	3.49
FEF (l/s)	4.14	0.024	0.824 (0.640)	-0.050 (0.002)	2.866 (0.344)	16.8	1.039	3.70
$FEV_1\%$	81.7	0.129	107.68 (3.601)	-0.172 (0.014)	-10.932 (1.94)	5.8	6.555	82.2

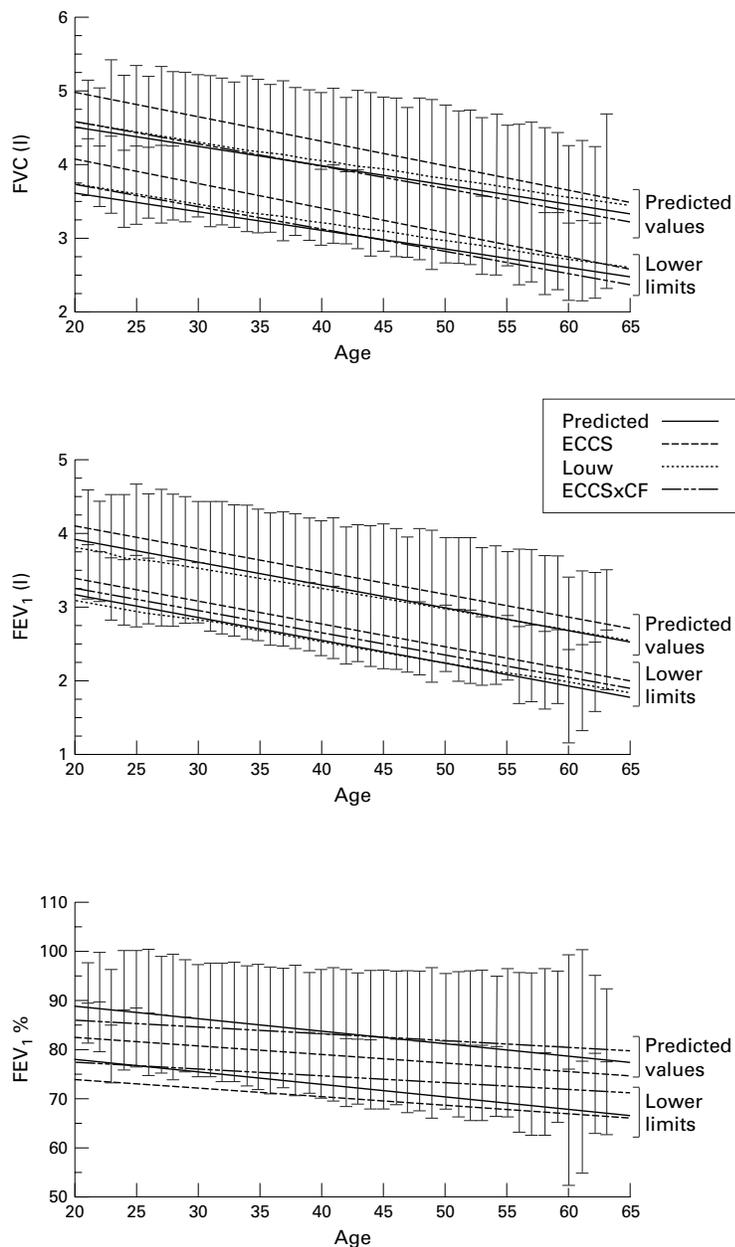


Figure 2 Height standardised mean lung function and 95% CI plotted against age, for black miners. Overlaid are predicted and reference curves and their respective lower limits of normal, as well as the ECCS curves scaled by a conversion factor of 0.93 for FVC and 0.95 for FEV₁.

interval, respectively. Of the remaining 25 614 tests, 1374 were initial, 15 772 periodic, 8267 exit, and 201 were done for other reasons. Of the 3270 tests done on white miners, we excluded two tests because of a history of pulmonary tuberculosis, 20 tests because of radiological changes, and 39 (1.2%) tests were excluded because lung function or height measurements were outside the 99.98% CI. Of the remaining 3248 tests, 158 were initial, 2752 periodic, 174 exit, and 59 were done for other reasons.

In a preliminary analysis, the initial examinations showed a slightly decreased predicted curve for FEV₁. For a 40 year old and 1.7 m tall black man, the decrease was 50 ml. As the number of periodic examinations was large, we

used these exclusively as they would provide the most representative sample of current miners. Measurements from both the initial and exit examinations may be potentially lower because of a learning effect and a selection process, respectively. Finally, the prediction equations were estimated on periodic lung function tests from 15 772 black and 2752 white miners. Because periodic tests are done at 3 yearly intervals, each miner had one test only.

STATISTICAL ANALYSIS

The prediction equations were estimated separately for black and white miners from the linear regression model⁸

$$Y = b_0 + b_1 \times \text{age} + b_2 \times \text{height} + b_3 \times \text{weight} + b_4 \times \text{age} \times \text{height} + b_5 \times \text{age}^2 \quad (1)$$

where Y is the predicted lung function value for a miner of a given age and height, and b_i, i=0,1,...,5 are the regression coefficients. Only the variables that contributed significantly to the R² statistic were retained in the final model. The equations of the final model as well as the published reference equations^{4,5} were then used to predict lung function values of 1.7 m tall men of different ages. The resulting curves were compared with the mean observed lung function values standardised to the average height of black miners of 1.7 m. To allow the comparison, the lung function measurements were standardised to the height of 1.7 m as lung function measurement × 1.7²/ht². The standardisation for height allowed for comparison of lung function measurements from populations with different mean height (fig 5). Selection of the best fitting reference equations was done by a visual inspection with this method. As most of the published curves did not fit well, a visual inspection was sufficient to identify the best fitting curves.

In the next step, we compared the percentage of miners that fell below the lower limit of normal (LL) for the different reference equations. The LL refers to the one sided lower 95% CI calculated as LL = P_(age,1.7) - 1.645 × SEE, where P_(age,1.7) is the age and height specific predicted value, and SEE is the standard error of the estimate of the regression line.^{4,9} For a homogeneous population and a normally distributed variable, it is expected that 5% of the subjects will have lung function values below the predicted LL. For a reference equation to be applicable to a specific healthy population, the percentage below the reference LL should also be around 5%, otherwise too many or too few subjects will be categorised as abnormal.

In the data on miners there were known systematic effects that could have increased the percentage of miners that fell below the LL. These were age below 30 years or above 55 years (see the observed data points in figs 2 and 3) and changes in lung function due to smoking and exposure to dust. These changes usually start to become evident pathologically at around 35–40 years of age—for example, emphysema, pneumoconiosis. Thus, to minimise the systematic effects of age and the adverse effects of environmental exposures, we

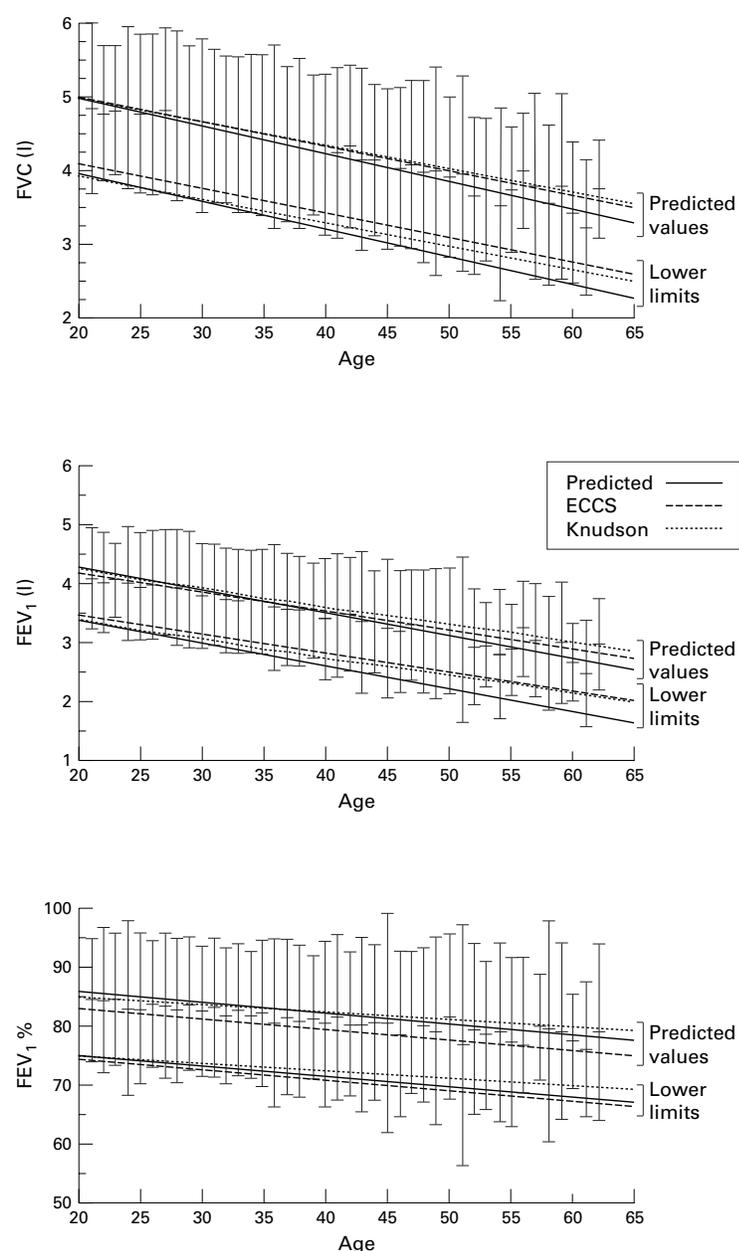


Figure 3 Height standardised mean lung function and 95% CI plotted against age, for white miners. Overlaid are predicted and reference curves and their respective lower limits of normal.

used a cross section of 30–35 year olds to compare the reference equations for the percentage of miners that are below the LL. The number of miners in the 30–35 age category was large (4503 black and 620 white miners) and the frequency distribution was close to normal. The percentage that were below the LL was derived from the standardised normal deviate, z , calculated as the difference between the observed mean and the LL of normal derived for each equation, divided by the observed SD, and the corresponding cumulative probability distribution table values subtracted from one.¹⁰

The percentage predicted values calculated as the ratio between observed and predicted lung function, which corresponded with a one sided 95% CI and the 5th percentile, were also determined for 10 year age categories. The 5th

percentile of a distribution of lung function value is the lung function value such that 5% of subjects have lower values and 95% of the subjects have higher values. The 5th percentile is estimated by a non-parametric method from order statistics.

Results

Figure 1 shows the age distribution of the miners in the study. Table 1 gives the mean lung function and the estimated prediction equations (regression coefficients (SEs)). Because most of the subjects were 30–45 years of age, the contributions of the age² and height \times age interaction to the variation explained by the model were small (the additional R^2 of 0.0002 and 0.0004, respectively, for FEV₁). Thus, the final prediction equations included age and height terms only.

For the black miners, figure 2 shows the observed height standardised lung function means for FVC, FEV₁, and FEV₁%, and the two sided 95% CI, plotted against age. The predicted curve and the best fitting reference curves, and their respective LLs are shown for each lung function test. The reference curve of Louw *et al*⁶ (measurements done by a vitalograph) provided the best fit to the data, whereas the ECCS curves were too high. The ECCS reference values multiplied by a conversion factor (CF) of 0.93 for FVC and 0.95 for FEV₁ were close to the predicted curves. For the white miners, figure 3 shows the observed mean values and the reference equations of the ECCS and Knudson *et al*⁶ which provided the best fit to the data.

Figure 2 indicates that the percentage of black miners that falls below the LL of the ECCS curve is large for younger age groups, but decreases for older age groups. However, a somewhat different trend is found for white miners (figure 3), where the observed means start to decline after 40 years of age more rapidly than the ECCS and Knudson's predictions. To determine the percentage of subjects that were below the LLs, we used the cross section of 30–35 year old miners, because for these miners the observed means corresponded closely with the reference curves (figs 2 and 3) and this subgroup is least likely to be affected by systematic effects already discussed. The percentages that were below the LLs for the different reference curves for the 30–35 year olds are shown in figure 4. Figure 4 shows the frequency distribution for the height standardised FVC, FEV₁, and FEV₁%. Superimposed over each observed frequency distribution is a normal curve derived from the observed mean (SDs). The normal curves derived from the means predicted by the equations from the ECCS and Louw *et al* and from their respective SEE values are also shown in figure 4. (The SEE for the curve of Louw *et al* was available only for models in which the ratio of sitting over standing height was fitted. As the SEE was similar to the models which used standing height only, we used this SEE value.) The percentages below the LLs are shown in table 2. Table 2 presents, for the 30–35 year olds, the descriptive statistics for the observed lung function and the LLs, and the percentage of

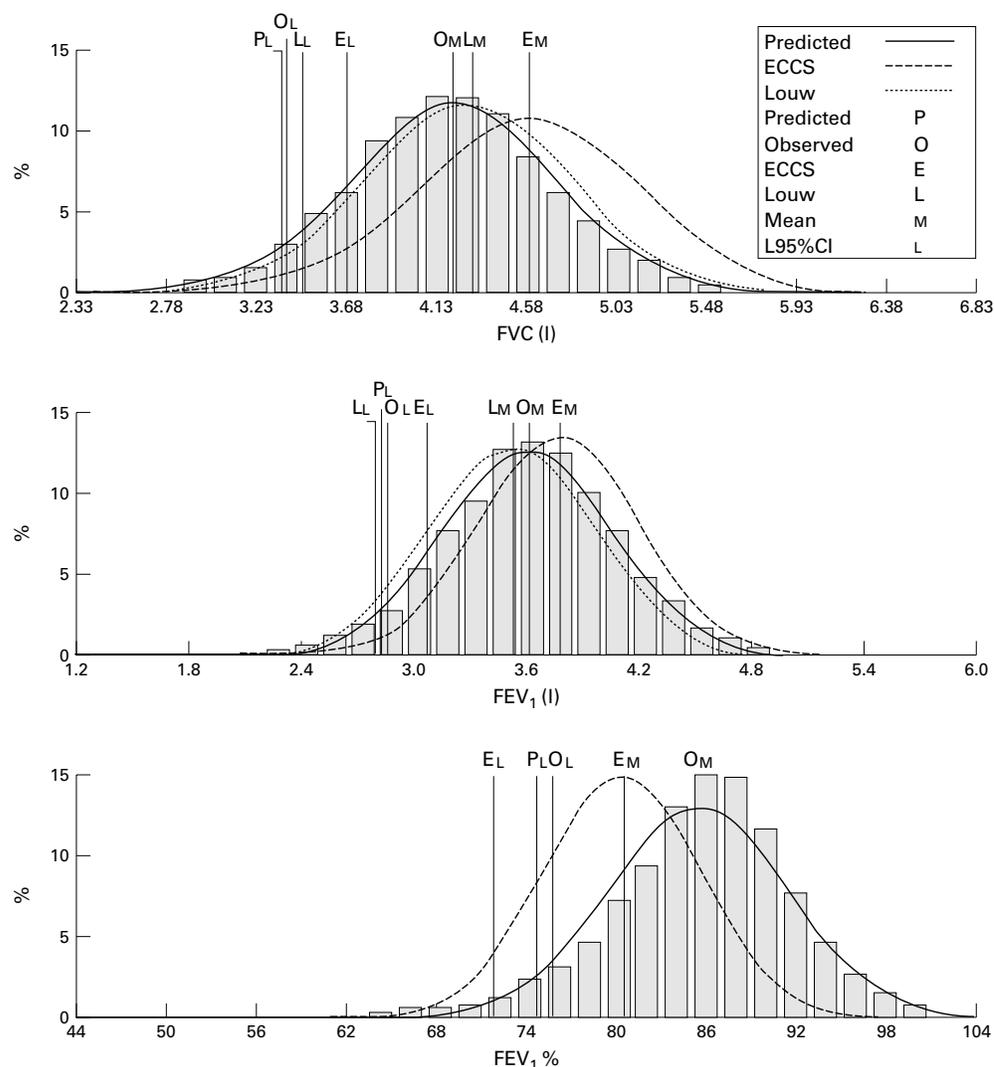


Figure 4 The frequency distribution for height standardised lung function for the cross section of 30–35 year old black miners.

subjects below the LLs, for each equation. Louw *et al*⁶ did not provide a reference equation for the FEV₁%.

The first part of table 3 presents descriptive statistics for the percentage predicted lung function values. The second part of table 3 presents the percentage predicted values that were derived from our prediction equation, the equations of Louw *et al*, and ECCS equations, and that corresponded with the 5th percentile and with the one sided lower 95% CI for 10 year age categories, for black miners only.

Discussion

The objective of the study was to identify reference equations appropriate for the large population of South African gold miners who are predominantly black, but with a notable white subpopulation. The question also arises whether a common reference equation can be used for all the miners irrespective of race, bearing in mind that the reference curves could serve various purposes—such as assessment of fitness to work in certain job categories, diagnosis of compensable disease

Table 2 Descriptive statistics for height standardised lung function for 30–35 year old workers, and the lower limits of normal, and % rejected for the predicted and selected reference equations

Lung function	Mean	SD	Skewness	Lower 95% CI	Lower limit of normal; % rejected for the different equations		
Black miners, n=4503, age=32.6 (SD 1.7), height=1.7:					Predicted	Louw	ECCS
FVC (l)	4.21	0.51	0.14	3.37	3.34; 4.4%	3.44; 6.6%	3.68; 14.9%
FEV ₁ (l)	3.60	0.46	-0.03	2.84	2.81; 4.3%	2.78; 3.8%	3.07; 12.5%
FEV ₁ %	85.6	6.11	-0.74	75.5	74.7; 3.8%	NA	71.7; 1.2%
White miners, n=620, age=32.7 (SD 1.7), height=1.8:					Predicted	Knudson	ECCS
FVC (l)	4.54	0.53	0.15	3.66	3.62; 4.09%	3.54; 3.0%	3.68; 5.3%
FEV ₁ (l)	3.74	0.45	-0.01	3.01	2.96; 4.9%	2.98; 4.6%	3.06; 6.6%
FEV ₁ %	82.5	5.8	-0.27	73.0	72.7; 4.6%	73.0; 5.1%	71.7; 3.1%

Table 3 The % predicted corresponding with the 5th percentile and the one sided 95% CI for different equations, black miners only

Age	n	Mean	SD	Skewness	Predicted equation		Louw's equation		ECCS equation	
					5th Percentile	95% CI	5th Percentile	95% CI	5th Percentile	95% CI
FVC:										
<30	1692	99.2	11.8	0.07	80.6	79.7	78.7	78.0	72.6	72.0
30–39	7430	100.6	12.4	0.17	80.9	80.2	78.9	78.2	73.8	73.0
40–49	4947	100.8	13.7	0.15	79.5	78.3	77.2	76.1	73.4	72.1
≥50	1704	99.9	14.6	0.07	76.1	75.8	73.6	73.3	71.7	70.8
FEV ₁ :										
<30	1692	99.6	12.3	-0.22	80.4	79.4	82.5	81.7	76.7	75.7
30–39	7430	100.1	13.1	-0.01	78.9	78.5	80.7	80.3	75.2	74.9
40–49	4947	99.9	15.0	-0.17	75.3	75.2	76.3	76.3	71.8	71.7
≥50	1704	99.9	16.8	0.23	71.3	72.4	71.2	72.7	68.0	68.9
FEV ₁ %:										
<30	1692	100.5	7.0	-0.68	88.5	89.0	NA	NA	94.6	95.0
30–39	7430	99.9	7.3	-0.74	87.3	87.9	NA	NA	92.6	93.3
40–49	4947	99.8	8.5	-0.94	84.9	85.9	NA	NA	89.3	90.6
≥50	1704	100.6	9.7	0.97	82.5	84.7	NA	NA	86.3	88.6

Table 4 Comparison between studies of occupational groups done on black men in South Africa at higher altitude

First author	Mean age (y)	Mean height (cm)	Altitude	Men, age 38 y, height 171 cm		Type of sample
				FEV ₁ (95% CI)	FVC (95% CI)	
White <i>et al</i> ²	36.1	167.3	1500	4.12 (3.99 to 4.25)	3.46 (3.36 to 3.56)	Textile workers
Louw <i>et al</i> ⁶	41.1	169.7	1700	4.21 (4.14 to 4.28)	3.44 (3.57 to 3.51)	Bank workers, autolink
Hessel and Sluis-Cremer ¹⁴	41.6	171.4	400	4.13 (—)	3.29 (—)	Vermiculite miners
Coetzee, cited in ¹²	33.1	169.4	1700	4.30 (4.25 to 4.35)	3.55 (3.50 to 3.60)	Asbestos workers
Mokoete <i>et al</i> ³	42.0	168.7	1700	4.42 (—)	3.47 (—)	University workers
Fox, cited in ¹²	34.4	171.7	1200	4.20 (4.16 to 4.24)	3.52 (3.46 to 3.58)	Copper or nickel miners
Our study	38.7	170.8	1200	4.12 (4.12 to 4.12)	3.47 (3.47 to 3.47)	Gold miners

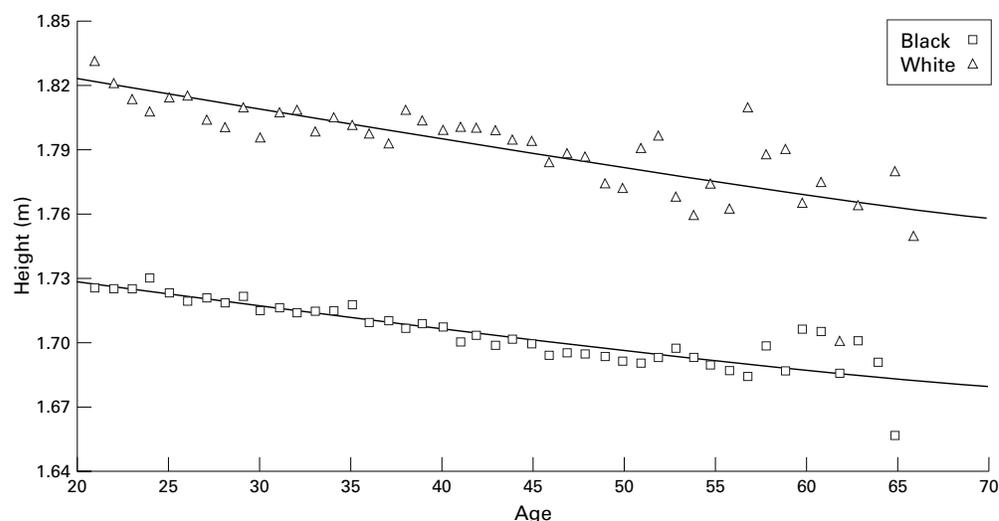


Figure 5 Relation between height and age for black and white miners.

or an early detection of accelerated loss of lung function.

We estimated prediction equations from the data obtained in a lung function screening programme and compared these with published reference values. There were large differences in the observed mean lung function values between white and black miners. These were 860 ml (17.3%) for FVC, 630 ml (14.9%) for FEV₁, -180 ml (-6.8%) for FEF_{25-75%} and -2.38% (-2.9%) for FEV₁% (table 1). However, after adjustment for age and height, the differences for a 40 year old 1.7m tall man decreased substantially to 220 ml (5.2%) for FVC, 110 ml (3.2%) for FEV₁, -160 ml (-4.3%) for FEF_{25-75%} and -1.6% (-2.0%) for FEV₁%. The results show that height is responsible for a large percentage of the observed difference, however, the remaining

differences after adjustment for age and height were still significant.

Of the reference equations published by ATS,^{4,5} the ECCS and Knudson equations corresponded most closely with the prediction curves for white miners (fig 3). However, the observed data showed a steeper decline from about 35 years of age. This may be due to the effect of smoking and exposure to dust previously documented in white gold miners.¹¹ The observed FEV₁% values were higher, however, than the ECCS values. The percentage of miners that were below the LLs of the ECCS and Knudson equations were similar and varied around 5%. For the 30–35 year olds, the percentage of subjects below the LL for the ECCS equation was 5.3% for FVC and 6.6% for FEV₁, whereas for the Knudson's equation it was 3.0% and 4.6%, respectively (table 4).

For black miners, the reference equations that fitted best were those derived for black South African men by Louw *et al*⁶ with a vitalograph (fig 2). The ECCS reference equations are commonly used in South Africa for all subjects, either with an ethnic conversion factor of 0.88 for black men (as recommended by ATS⁴), or without a conversion factor. The data from the present study show that the ECCS reference values are much higher than the observed lung function values, the difference being almost 500 ml at 20 years of age for FVC (figure 2). The percentages below the LLs for equations of Louw and the ECCS, shown in figure 3 and table 2, are also substantially different. For the Louw equation, the percentage rejected was 6.6% for FVC and 3.8% for FEV₁, whereas for the ECCS equation the corresponding values were 14.9% and 12.5%. The ECCS reference values scaled by a factor of 0.93 for FVC and 0.95 for FEV₁ provided the closest fit to the observed data and the predicted curves (fig 1). The scaling down by 8% for FVC and 5% for FEV₁, was higher than the differences found for the age and height adjusted lung function means which were 5.2% for FVC and 3.2% for FEV₁. This reflects the fact that white miners had lower observed values than the ECCS curves because of a steeper decline with age.

The differences in the regression coefficient for age in table 1 show a steeper decline with age in white miners, especially for FVC. The predicted curves for FVC and FEV₁, and the LLs were parallel with the reference curves of Louw *et al* derived from asymptomatic, healthy non-smokers (fig 2), but the white workers had a steeper slope than the ECCS curve. One of the reasons for these differences may be that fewer black miners smoke, and if they do so, then they smoke fewer cigarettes (five cigarettes a day) than white miners (over 20 cigarettes a day). It is interesting to note that for FVC, the curve of Louw *et al* as well as the predicted curve came closer to the ECCS curve (fig 2) with increasing age. This reflects the steeper slope with age for the ECCS curves. The reason for the steeper slope for the ECCS curve is not clear.

The results also show that black miners have higher FEV₁%, but slightly higher decline in FEV₁%, compared with the ECCS curve (fig 2). For 30–35 year old black miners, the observed mean FEV₁ % was 85.5%, the observed one sided 95% CI was 75.5%, and the LL derived from all data was 74.7% (table 2). By comparison, the predicted FEV₁% mean from the ECCS equation was 80.4% and the LL derived from the ECCS equation was 71.7%. For 30–35 year old white miners, the mean observed FEV₁% was 82.5% and the observed one sided lower 95% CI was 73.0% (table 2).

When compared with other published studies, the prediction equations for black gold miners agree most closely with those found in other studies of South African black men occupationally exposed to dust (table 4).^{12,13} For FVC, the predicted values for a 38 year old and 1.7 m tall male gold miner were similar to

those found for textile workers,¹² bank workers,⁵ and other industrial workers.^{12–14} This result suggests that the high ratio found in this study may be due to a real effect. However, the data in table 4 also suggest that the groups exposed to dust may have lower FEV₁ values than the unexposed subjects, and that reference values estimated on healthy South African black men not exposed to dust should preferably be used for the miners.

By comparison with our study and other studies of black workers, a most recent study of black university workers from Johannesburg, South Africa,¹³ found similar FEV₁ values, but the FVC values were systematically higher than those predicted by our and other studies; the estimated increase in FVC was substantial and ranged from 120 ml at 20 years to 200 ml at 60 years of age. The explanation provided for the higher FVC was the changing socioeconomic status of black workers in South Africa. The socioeconomic factors associated with FVC in the study were job category and education. The socioeconomic factors were not associated with FEV₁. Smoking was not found to be a risk factor for FEV₁ or FVC. The effect of socioeconomic factors is also apparent from a secular trend found in FVC measured in different studies of urban black workers done from the 1970s to the 1990s, which ranged from 3.72–3.98 to 4.24–4.42, respectively.¹³ In our study, we found a strong relation between height and age for black and white miners (fig 5), although some of it was due to aging, there was an apparent cohort effect. Changing socioeconomic status, which includes nutrition from birth that results in increased height, stronger build, and lower respiratory morbidity, must play an important part on the respiratory variables of the black South African population¹⁵ whose socioeconomic status is generally improving with urbanisation and education. To take cognisance of that trend, estimation of reference values should be repeated every decade or so in accordance with the recommendations made by the ATS.

Based on these results and given the current circumstances, it seems that the available options are to use the ECCS equations on all miners, with a scaling factor for the black miners, or to use the ECCS equations on white miners and the equations of Louw *et al* on black miners. The curves of Louw *et al* provided an excellent fit to the data, whereas the ECCS curves were too steep with age. From an epidemiological view, the use of the ECCS reference equations without any scaling for all the miners is not recommended. Although the difference of 8% for FVC represents only about 300 ml, which seems to be insignificant in terms of a man's variability in lung function, the difference becomes important if applied systematically to many miners for selections—such as job placement and compensation. Future developments in reference values should account for the curvature in the observed data with age (figs 2 and 3) and the changing socioeconomic status of the black population.

These statistical considerations assume that the distribution of lung function measurements is normal. The skewness statistics in table 2 indicate that the distribution for FVC has a tail to the right, the distribution for FEV₁ is normal, and the distribution for FEV₁% has a tail to the left. This pattern is similar for black and white miners. Because of the skewness and variability in the distribution of the percentage predicted, the 5th percentile and 80% predicted were suggested as preferable criteria for the LL,⁹ and are often used in clinical practice.

The use of 80% predicted as a criterion for abnormal is not recommended, for adults by the ATS.⁴ Although some studies have shown that 80% of predicted is close to the fifth percentile,⁹ other shortcomings inherent in the use of a fixed value of percentage predicted are that shorter, older subjects are more readily classified as abnormal, whereas taller, younger adult subjects are more likely to be erroneously classified as normal.⁴ In South Africa, it has been suggested that the grading of lung function impairment for compensation purposes should be based on percentage of predicted derived from the ECCS reference values.³ The percentage predicted categories of impairment for FEV₁ and FVC were as follows: normal $\geq 80\%$, 79%–65% for mildly impaired, 64%–51% for moderately impaired, and $< 51\%$ for severely impaired. Here, we examined the percentage predicted values corresponding with the 5th percentile and the lower 95% CIs, for different reference equations, according to age categories.

For FVC and FEV₁, the lower 95% CIs (or the LLs) for the percentage predicted corresponded with 80% for values derived from the prediction equation and the equation of Louw *et al*, but was lower for the ECCS equation (table 3). The corresponding percentage predicted based on the prediction equation started to decrease below 80% at 50 years of age for FVC and at 40 years of age for FEV₁. For FEV₁%, the corresponding percentage predicted started at around 89% for the youngest age category, and decreased to around 83% for the oldest age category. There was no important difference between the LL calculated as the 5th percentile or the 95% CI. The percentage predicted values corresponding to the LL depend on how well the reference equation fits the actual data, but even for the best fitting linear curve, the values were age related. Because subjects older than 50 years of age have a steeper decline in lung function than the predicted values based on a linear regression curve, the 80% criterion might not be appropriate for the older subjects.

Figure 5 shows strong relations between height and age, which is likely to be due to a cohort effect. Correlation between predictor variables (age and height) violates the linear regression assumption of independently distributed explanatory variables and could lead to biased estimates for age and height, in particular an overestimated effect of age. In such a case, the variable for age may not be suitable for predicting longitudinal loss of lung function with years of age.

In summary, of the reviewed reference equations,⁵ the ECCS and Knudson equations best fitted the data for white miners. Although the two reference curves were similar with respect to the percentage of miners below the LL, the LLs derived from the ECCS equations were slightly better in terms of rejecting 5% of the subjects. For black miners, the reference equation of Louw *et al* estimated on healthy South African men fitted best, whereas the ECCS reference equations overestimated the percentage of subjects that were below the LL by 10% for FVC and 8% for FEV₁. The ECCS reference values scaled by a factor of 0.93 for FVC and 0.95 for FEV₁ provided a reasonably good fit to the data from black miners, however, the slope for FVC was too steep. The ECCS equations, with a conversion factor 0.88 for the black miners, are available on most spirometers sold in South Africa. However, the results of the present study show that the correction factor of 0.88 is too low and also that the ECCS curve for FVC is too steep to fit well.

Because of the strong correlation between height and age, the estimated regression variables for age may not be suitable for estimating the longitudinal loss due to years of age. The steep trend between height and age indicate that there is a strong cohort effect, and this supports the ATS recommendations that estimation of reference values should be repeated every decade or so.

We thank the mines for letting us use the lung function screening data and Ms Tanusha Singh for her computing support. The study received partial support from the Safety in Mining Research Advisory Committee (SIMRAC).

- 1 Occupational Health and Safety Act, Act 85 of 1993, with Regulations. Johannesburg: Butterworth, 1996.
- 2 Quanjer PhH, Tammeling GJ, Cotes JE, *et al*. Lung volumes and forced ventilatory flows. *Eur Respir J* 1993;6(suppl 16): 5–40.
- 3 White N, Ehrlich R, Rees D. A guide to spirometry as applied to occupational health. *S Afr Med J* 1996;86:807–13.
- 4 American Thoracic Society. Lung function testing: selection of reference values and interpretative strategies. *Am Rev Respir Dis* 1991;144:1202–18.
- 5 Louw SJ, Goldin JG, Joubert G. Spirometry of healthy adult South African men. *S Afr Med J* 1996;86:814–19.
- 6 Hnizdo E, Churchyard G, Dowdeswell R. Assessment of reliability of a pulmonary function screening programme in the South African gold mines. *Am J Respir Crit Care Med* 1999;160:2006–11.
- 7 American Thoracic Society. Standardization of spirometry. *Am J Respir Crit Care Med* 1995;152:1107–36.
- 8 Vollmer WM. Reconciling cross-sectional with longitudinal observations on annual decline. *Occup Med* 1993;8:339–51.
- 9 Knudson RJ, Lebowitz MD, Holberg CJ, *et al*. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis* 1983;127:725–34.
- 10 Snedecor GW, Cochran WG. *Statistical methods*, 6th ed. Iowa, USA: Iowa State University Press, 1967.
- 11 Hnizdo E. Combined effect of silica dust and tobacco smoking on mortality from chronic obstructive lung disease in gold miners. *Br J Ind Med* 1990;47:656–64.
- 12 White NW, Hanley JH, Laloo UG, *et al*. Review and analysis of variation between spirometric values reported in 29 studies of healthy African adults. *Am J Respir Crit Care Med* 1994;150:348–55.
- 13 Mokoetle KE, de Beer M, Becklake MR. A respiratory survey in a black Johannesburg workforce. *Thorax* 1994;49: 340–6.
- 14 Hessel PA, Sluis-Cremer GK. Prediction equations for lung function in black industrial workers at Palabora Mining Company. *S Afr Med J* 1989;76:548–9.
- 15 Myers JE. Differential ethnic standards for lung functions, or one standard for all. *S Afr Med J* 1984;65:768–72.