Journal of Epidemiology and Community Health, 1988, **42**, 383–389

Relationship of lung function loss to level of initial function: correcting for measurement error using the reliability coefficient

LES IRWIG,[1]* HENNIE GROENEVELD,[1 2] AND MARGARET BECKLAKE[3]

From the Institute for Biostatistics of the South African Medical Research Council;[1] the Department of Statistics,[2] University of Pretoria, South Africa; and the National Centre for Occupational Health and Department of Community Health,[3] University of Witwatersrand, Johannesburg, South Africa

SUMMARY The regression of lung function change on the initial lung function level is biased when the initial level is measured with random error. Several methods have been proposed to obtain unbiased estimates of regression coefficients in such circumstances. We apply these methods to examine the relationship between lung function loss over 11 years and its initial level in 433 men aged about 20 when first seen. On theoretical and practical grounds the best method is the correction of the regression coefficient using the reliability coefficient. This is defined as the ratio of the error free variance to the variance of the variable measured with error, and is easily estimated as the correlation between repeat measurements of the underlying level. In young men the loss of some lung functions (forced vital capacity [FVC], forced expiratory volume in one second [FEV_1], forced expiratory flow in the middle half of expiration, and the ratio FEV_1/FVC) do not appear to be related to initial level.

In epidemiological research, it is often relevant to determine if a change in measurement of a variable between two occasions is related to the initial level of that variable. Random measurement or observational error in the initial measurement results in a biased estimate of this association. Despite extensive documentation of this phenomenon (regression to the mean) over several decades, [1–4] inferences are often still based on associations which may well be explained by random measurement error, as recently pointed out in the case of blood pressure.[5] The purpose of this paper is to review methods of analysis of change on initial level which attempt to take account of the consequences of random error of measurement and apply them to determine whether lung function loss is related to initial level in young men.

Theoretical considerations

MEASUREMENT ERROR

The linear model of the relationship between change and initial level, assuming there is no measurement error and no confounding, is:

$$L_2 - L_1 = \alpha + \beta_L L_1 + \varepsilon \quad (1)$$

where: L_1 = the error free value or "level" on the first occasion
L_2 = the error free value on a second occasion, eg, ten years later
α = a constant
β_L = the regression coefficient for L_1
ε = the residual term, ie, random deviation from the linear model.

Note that change is measured as the second minus the first measurement. A negative value therefore indicates loss.

In practice, L_i (i = 1 or 2) is usually observed with random error (δ_i) resulting in a measurement M_i, where $M_i = L_i + \delta_i$. The model from equation 1 then becomes:

$$M_2 - M_1 = \alpha + \beta_L M_1 + \varepsilon - \delta_1 + \delta_2 - \beta_L \delta_1 \quad (2)$$

alternatively expressed as:

$$M_2 = \alpha + (1 + \beta_L) M_1 + \varepsilon + \delta_2 - (1 + \beta_L)\delta_1 \quad (3)$$

The error component now includes $(1 + \beta_L)\delta_1$, which is not independent of M_1 and therefore violates the necessary assumptions of regression analysis.[6] The coefficient $1 + \beta_L$ is defined as $\sigma_{L_2L_1}/\sigma_{L_1}^2$, where $\sigma_{L_2L_1}$ is the covariance between L_1 and L_2 and $\sigma_{L_1}^2$ is the variance of L_1. When M_i is used instead of L_i, the use of $\sigma_{M_1M_2}/\sigma_{M_1}^2$ leads to a

* Present address and address for correspondence and reprints: Professor L M Irwig, University of Sydney, Department of Public Health, Building A27, Sydney, New South Wales 2006, Australia

biased estimate $\widehat{1+\beta_L}$. As the random measurement error terms δ_1 and δ_2 are by definition uncorrelated, substituting M_i for L_i does not alter the expected covariance, ie, $\sigma_{M_1M_2} = \sigma_{L_1L_2}$.[7] Because $\sigma_{M_1}^2$ is greater than $\sigma_{L_1}^2$, the expected value of $\widehat{1+\beta_L}$ will be smaller than $1 + \beta_L$. Assuming that L_1 and δ_1 are independent and normally distributed, $\sigma_{M_1}^2$ is equal to $\sigma_{L_1}^2 + \sigma_{\delta_1}^2$. The ratio of the variance of error free values ($\sigma_{L_1}^2$) to that of measurements ($\sigma_{M_1}^2$) is called the reliability or generalisability coefficient (G):[7–12]

$$G = \frac{\sigma_{L_1}^2}{\sigma_{M_1}^2} = \frac{\sigma_{M_1}^2 - \sigma_{\delta_1}^2}{\sigma_{M_1}^2} = 1 - \frac{\sigma_{\delta_1}^2}{\sigma_{M_1}^2} \quad (4)$$

It follows that the expected value of $\widehat{1+\beta_{L_1}}$ is equal to $G(1 + \beta_L)$. The implications for longitudinal data are easily demonstrated by a simple example in which change is not associated with initial level, ie, $\beta_L = 0$ or $1 + \beta_L = 1$. In the presence of measurement error, G is less than 1. Therefore the expected value of $\widehat{1+\beta_L}$ will be less than 1 and hence the expected value of $(\widehat{1+\beta_L}) - 1$ will be less than zero. In other words, random error of measurement induces a spurious negative regression of change on initial level, the phenomenon known as regression to the mean.[1–4]

CORRECTION OF REGRESSION COEFFICIENTS FOR RANDOM ERROR OF MEASUREMENT USING THE RELIABILITY COEFFICIENT (METHOD 1)

The unbiased estimate of $1 + \beta_L$ can be obtained by dividing the calculated regression coefficient by the relevant reliability coefficient,[13] ie:

$$\widehat{1+\beta_L} = \frac{\widehat{1+\beta_L}}{G} \quad (5)$$

The corrected $\hat{\beta}_L$ is then obtained by subtracting 1 from $\widehat{1+\beta_L}$. The regression coefficient for L estimated from samples can be corrected for random error of measurement by the method outlined above using estimates of G ($\hat{G}$). The variance of a corrected regression coefficient is estimated approximately[4] by:

$$\sigma^2_{\hat{\beta}_L} = \sigma^2_{\widehat{1+\beta_L}} = \quad (6)$$

$$\widehat{(1+\beta_L)^2}\left[\frac{\hat{\sigma}^2_{\widehat{1+\beta_L}}}{\widehat{(1+\beta_L)^2}} + 2\left(\frac{1-\hat{G}}{\hat{G}}\right)^2\left(\frac{1}{n-1} + \frac{1}{n_1-1}\right)\right]$$

where n = sample size for calculation of $\widehat{1+\beta_L}$, and n_1 = number of pairs of measurements for calculation of $\hat{G}$.

As n and n_1 approach infinity, the variance of the corrected regression coefficient decreases to $\hat{\sigma}^2_{\widehat{1+\beta_L}}/\hat{G}^2$. Because $\widehat{1+\beta_L} = \widehat{1+\beta_L}/\hat{G}$, the probability of rejecting the null hypothesis ($1 + \beta_L = 0$) will be unaltered after correction if n and n_1 are very large; otherwise the test will be less significant. However, the hypothesis one wishes to test is whether change is independent of initial level, ie, whether $\beta_L = 0$ or $1 + \beta_L = 1$. If the biased estimate of β_L is negative, the corrected estimate will approach zero or become positive and its standard error will increase. If the corrected estimate of β_L is negative, the probability of rejecting the null hypothesis will be decreased.

The reliability coefficient is estimated easily by the correlation between repeat measurements of the same underlying level (ρ).[9] The reason is as follows. Imagine repeat measurements (M_A and M_B) of the same underlying level L. Assuming that δ is independent of L and that δ_A is independent of δ_B, the expected value of the covariance of M_A and M_B will be the same as that of error free measurements L_A and L_B,[7] ie:

$$\sigma_{M_AM_B} = \sigma_{L_AL_B} = \sigma^2_L \quad (7)$$

$$\text{Now} \quad \rho_{M_AM_B} = \frac{\sigma_{M_AM_B}}{\sigma_{M_A}\sigma_{M_B}} = \frac{\sigma^2_L}{\sigma^2_M} = G$$

In the case of lung function measurements concerned with the detection of chronic airways obstruction, repeat measurements should be days or weeks apart so as to include both technical and short term biological components of the random error variance. This is in contrast to much of the published literature on lung function variability, in which random error of measurement is often assessed on only one piece of equipment by one technician and sometimes on only one occasion.[14]

Correction of regression coefficients using the reliability coefficient is based on the assumption that random measurement error is normally distributed and independent of the true level of lung function.[8 9 15] However, correction of regression coefficients is consistent if the error is not normally distributed, though the variance of the regression coefficient may be slightly biased.[15]

OTHER METHODS OF CORRECTING REGRESSION COEFFICIENTS FOR RANDOM MEASUREMENT ERROR

Apart from the use of the reliability coefficient, the literature contains several other methods of correcting for the effect of measurement error in longitudinal data.[16] All of them are based on the principle of substituting the predictor M_1 by some other measure of L_1.

Substituting L_1 predicted from other variables (Method 2)[16] — One can use an independent estimate of initial level, M_{1a}, predicted by a set of instrumental variables. These are defined as variables which are correlated with the predictor of interest but

independent of its random error of measurement. If the error of M_{1a} is δ_{1a}, then:

$$M_2 - M_1 = \alpha + \beta_L M_{1a} + \varepsilon - \delta_1 + \delta_2 - \beta_L \delta_{1a} \quad (8)$$

For example, one can predict some lung functions by height, and use the predicted value as M_{1a} in equation (8). This method is equivalent to using an instrumental variable (H), to estimate the regression coefficient as follows:[17]

$$\widehat{1 + \beta_L} = \frac{\hat{\sigma}^2_{M_2 H}}{\hat{\sigma}^2_{M_1 H}} \quad (9)$$

Even if height is measured with minor error, making the $\beta_L \delta_{1a}$ term in equation (8) negligible, the method has several disadvantages. First, the predicted value excludes the component of variance due to biological variability unrelated to height. This component may have a different relationship to lung function change than that of the height related component. Second, the variance of the predicted M_{1a} will be smaller than that of L_1. This results in a large standard error of the β_L estimate.

Substituting $(M_1 + M_2)/2$ for L_1 (Method 3)[3] — The regression of $M_2 - M_1$ on $(M_1 + M_2)/2$ is a method of longitudinal data analysis which is in use in the lung function literature.[18] The method is based on the following concept. If the variance of random measurement error of a single observation is σ_δ^2, that of the mean of j observations is σ_δ^2/j, assuming independent errors on the two occasions.[8] The reliability coefficient (G) of the mean of j observations increases as j increases:

$$G = \frac{\sigma_L^2}{\sigma_L^2 + \frac{\sigma_\delta^2}{j}} \quad (10)$$

This equation demonstrates the value of multiple measurement of a predictor. However the method described uses the mean of M_1 and M_2 rather than repeat measurements at the time M_1 is measured. The model is then:

$$M_2 - M_1 = \quad (11)$$

$$\alpha + \beta_L \left(\frac{M_1 + M_2}{2} \right) + \varepsilon - \delta_1 + \delta_2 - \beta_L \left(\frac{\delta_1 + \delta_2}{2} \right)$$

The method partly takes account of the association between the predictor and the residual because the variance of the term

$$\beta_L \left(\frac{\delta_1 + \delta_2}{2} \right)$$

will be less than that of $\beta_L \delta_1$ alone. The predictor is also less associated with δ_1 but is now associated with δ_2. However, the most important defect of the model is that it is biologically incorrect. It incorporates a component of change in the $(M_1 + M_2)/2$ term on which β_L is estimated.[16] This results in overestimation of β_L and the possibility of showing a spurious positive association between lung function change and initial level. This is easily demonstrated by analysis of the hypothetical data in table 1. The data are free of random error of measurement and not intended to represent lung function data. The least squares regression equations calculated from this data set are:

$M_2 - M_1 = 22{\cdot}5 + 0{\cdot}00L_1$
(using the error free M_1 as the measure of L_1)
and $M_2 - M_1 = 11{\cdot}05 + 0{\cdot}14L_1$
(using $(M_1 + M_2)/2$ as the proxy of L_1).

Whether this effect is important in lung function data will be explored later in this paper when all of the above methods are applied to a data set.

Table 1 *Hypothetical data for measurements for M_1 and M_2 (measured without error)*

M_1	M_2
40	40
40	45
40	60
40	65
40	80
40	85
100	100
100	105
100	120
100	125
100	140
100	145

Substituting an independent measure of L_1 (Method 4)[19] — One can obtain two measures of L_1 (M_1 and M_{1a}) on occasions close enough to ensure that they both measure the same underlying level, but for which the random error of measurements of M_1 and M_{1a} are independent. The model is described in equation (8). The method solves part of the problem of the association between residual terms and explanatory variables, because δ_1 is independent of M_{1a}. However, $\beta_L \delta_{1a}$ is still associated with M_{1a} and the estimate of β_L obtained from the regression analysis should be corrected using the G of M_{1a}. In contrast to the situation when $M_2 - M_1$ is regressed on M_1, the absolute value of the β_L estimate will *increase* with correction, because $M_2 - M_1$ is obtained independently of the predictor M_{1a}.

Application of methods to longitudinal lung function data

THE DATA SET

All men who intend to work in dusty occupations in South African mines must have a medical examination before starting employment. Examinations are repeated 6 months later and then annually. The sample was assembled from white men who had their first routine examination between 1964 and 1966. Sampling was done by clerical staff on the basis of the availability of lung function laboratory personnel and equipment, and without knowledge of the men's medical status. Spirometric and peak expiratory flow rate measurements were made at the first routine examination 6 months after entry into the industry, and at irregular intervals thereafter. During 1976 and 1977, about 11 years after the initial examination, all men in the original sample were retested if they were still working in the mining industry in the region.

Spirometry was performed on one of four 9 litre Godart water sealed spirometers. Spirometers were calibrated regularly and methods of measurement were very similar to recommendations subsequently published by the American Thoracic Society.[20] Spirometry was performed with the subject seated and using a noseclip. The forced vital capacity (FVC), forced expiratory volume in 1 second (FEV_1, referred to as FEV in this paper) and forced expiratory flow in the middle half of expiration ($FEF_{25-75\%}$, referred to as FEF) were measured on each of three acceptable tracings. Lung function measurements were adjusted to body temperature, atmospheric pressure and saturated water vapour pressure (BTPS). The numerically highest values of FVC, FEV, and FEF were used, irrespective of whether they came from the same or different tracings. The value for FEF was therefore not selected from the tracing with the largest sum of FVC and FEV, as is recommended.[20] The procedure used tends to overestimate FEF by about 5%.[21] After spirometry, peak expiratory flow rate was measured on one of five calibrated Wright Peak Expiratory Flow Meters with the subject standing. The highest of three blows judged acceptable by the technician was recorded as the value for each man. The proportion of the sample excluded because of incomplete data or extreme outlying measurements of change (>4 standard deviations from the mean) was less than 1%.

Only 652 of the 1974 men in the initial sample remained in the mining industry long term, defined as having their final lung function measurement at least 8 years after the initial measurement. The average time between initial and final lung function tests on these men was 11·3 years, with 79% of men having between 10 and 12 years of follow up. Long term longitudinal data analysis was restricted to individuals in the narrow age interval of the greatest biological interest, ie, the 433 men aged between 18·00 and 22·99 years of age. These 433 men with long term follow up represent only about one third of all 18–22 year olds who had initial examinations. Concern over how representative they are is reduced by two findings. First, over half of the men who did not remain in the mining industry long enough to qualify for inclusion had left within two years of their initial examination. Within so short a time it seems unlikely that their leaving was related to the development of respiratory impairment. Second, there were no major differences in the initial age, anthropometric characteristics or lung function measurements of the 433 men who remained in the industry and those who left.

The means of the initial spirometric measurements were toward the upper end of the range expected for 20 year old men from prediction equations[22–24] and the mean peak expiratory flow rate was similar to that expected using the Wright flow meter[25–27] (table 2). Lung function change (LFT change) was calculated as the final minus the initial measurement. Change with a negative sign, as shown in table 2, therefore indicates a loss of function over the 11 years between the initial and final lung function test. The mean change of all lung function measurements was within the range expected over 11 years from other studies.[28–30] The distributions of initial lung function data did not show any major deviations from normality. Lung function change data were also distributed reasonably symmetrically around the mean, though somewhat kurtosed.

Correlation coefficients (product-moment or Spearman) between potential confounders (smoking, the number of years worked in dusty occupations, age, weight change and the time between initial and final examinations) and initial lung function tests all had

Table 2 *Descriptive statistics for lung function measurements*

	Initial		*Final*		*LFT change (final-initial)*	
LFT	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
FVC (litres)	5·57	0·61	5·46	0·68	−0·11	0·42
FEV (litres)	4·68	0·52	4·34	0·60	−0·35	0·42
FEF (litres/s)	5·49	1·23	4·38	1·34	−1·12	1·00
FEV%	84·3	6·4	79·6	6·8	−4·7	5·5
PEFR (litres/min)	575	67	574	73	−2	65

LFT = Lung function test; FVC = forced vital capacity; FEV = forced expiratory volume in 1s; FEF = forced expiratory flow between 25% and 75% of expiration; FEV% = $FEV_1/FVC \times 100$; PEFR = peak expiratory flow rate.

absolute values less than 0·15. Bias in regression coefficients due to confounding was therefore considered unlikely.

RELIABILITY COEFFICIENTS

A subset of data on 331 men with examinations one year apart was used to estimate the reliability coefficient of lung function data. Reliability coefficients for lung function may be underestimated using this method because a 1 year time period between examinations is likely to include some true variation in lung function level. However, lung function variability measured over a few weeks or one year is very similar.[31] Correlelograms have been suggested as a method of estimating reliability coefficients including only that component of variance due to random error of measurement.[10] The results of correlelogram analysis were not dissimilar to those obtained from repeat measurement at 1 year, though heavily dependent on model assumptions.[32] Estimates of the reliability coefficient based on equations (4) and (5) ranged from 0·929 for FVC to 0·786 for FEV% (table 3). Because the random error variance was found to be associated with level (the mean of the two readings) for FEF, $\sqrt{FEF}$ was used in all further analyses so that the necessary assumption of independence could be met.

Table 3 *Reliability coefficient ($\hat{G}$) based on lung function tests one year apart (n = 331)*

*LFT**	$\hat{G}$
FVC	0·929
FEV	0·899
FEF	0·836
$\sqrt{FEF}$	0·847
FEV%	0·786
PEFR	0·818

* See table 2 for description

COMPARISON OF RESULTS OF DIFFERENT METHODS OF CORRECTING REGRESSION COEFFICIENT (table 4 and 5)

All the ordinary least squares (OLS) regression coefficients of lung function change on initial level without any correction for measurement error were significantly negative (table 4). Regression coefficients corrected using the reliability coefficients (Method 1) did not differ significantly from zero, except for PEFR, for which the coefficient remained significantly negative, and FEV%, which was only marginally significant. Standard errors of the regression coefficients were increased by the correction procedure, but did not differ markedly, whether calculated from equation (6) or simply by dividing the standard error of the OLS regression by G (table 4). The regression coefficients using height as an instrumental variable for FVC and FEV (Method 2) were not significantly different from zero (table 4). The predicted variables had less variation than equivalent observations because only 40% of the variance of FVC and 27% of the variance of FEV were explained by height. The standard errors of the regression coefficients of change on predicted initial lung function were therefore larger than those of the corrected coefficients of change on initial values. Predicted values were not calculable for the other lung function tests as there were no appropriate instrumental variables. When change was regressed on the mean of the initial and final lung function results

Table 4 *The regression of change on initial measurement using various methods (n = 433)*

*LFT**		*OLS regression*	*Regression corrected using G (Method 1)*	*Regression on predicted initial LFT (Method 2)*	*Regression on (M1 + M2)/2 (Method 3)*
FVC	β_L	−0·109 (−0·16)†	−0·041	+0·076	+0·129
	SE	0·033	0·036 (0·036)‡	0·052	0·033
	p	0·001	0·3	0·1	<0·001
FEV	β_L	−0·154 (−0·19)	−0·059	−0·032	+0·159
	SE	0·037	0·043 (0·041)	0·072	0·037
	p	<0·001	0·2	0·7	<0·001
$\sqrt{FEF}$	β_L	−0·128 (−0·15)	0·030	NA	+0·242
	SE	0·042	0·053 (0·050)		0·039
	p	0·002	0·6		<0·001
FEV%	β_L	−0·303 (−0·35)	−0·113	NA	+0·076
	SE	0·039	0·056 (0·050)		0·044
	p	<0·001	0·04		0·08
PEFR	β_L	−0·382 (−0·39)	−0·244	NA	+0·103
	SE	0·043	0·055 (0·053)		0·050
	p	<0·001	<0·001		0·04

β_L = Regression coefficient estimate; NA = Not applicable; * See table 2 for description; † Bracketed values in this column are correlations between lung function change and initial measurement; ‡ The first standard error is estimated from equation (6), the bracketed standard error is estimated from SE(OLS)/$\hat{G}$.

(Method 3), the coefficients were positive in all cases and only failed to reach significance for FEV% (table 4).

The regression of lung function change on an independent initial measurement (Method 4) required three separate lung function test results. This analysis was therefore done only on a subset of 221 men who had an additional ("middle") lung function test between their initial and final examination. Change was calculated as the final minus the middle lung function adjusted to the mean time between the two examinations (6·5 years). The results of the regression coefficients of lung function change on the initial lung function (measured independently on average of 4·4 years earlier) are shown in the third column of table 5. Only that for FVC differed significantly from zero. For comparison, the OLS regression coefficients of change on middle measurement, without and with correction for measurement error, are shown in columns 1 and 2 respectively. The results obtained by Methods 1 and 4 were reasonably similar, with the notable exception of PEFR for which change was strongly negatively associated with initial level when Method 1 was used. The effect of correcting Method 4 for random error of measurement in M_{1a} was to marginally increase the absolute magnitude of the coefficients and their standard errors without decreasing their p values (fourth column of table 5).

Discussion

The method we favour for correcting coefficients for random error of measurement is that using the reliability coefficient (Method 1). Apart from its appeal on conceptual grounds, its advantages over the other methods are as follows. (1) Unlike Method 2, it does not require additional data on instrumental variables, which are in fact not always available. Even when they are, the standard errors of the corrected regression coefficients are very large. (2) Unlike Method 3, it is not expected, on theoretical grounds, to give biased results. The coefficients obtained in our data suggest that the bias in Method 3 is of sufficient magnitude to cause concern. (3) Unlike Method 4, which requires repeat initial measurement on the full sample, a reliability coefficient (G) can be estimated from a subset or alternative sample. Furthermore the coefficients from Method 4 are still biased unless further corrected for the measurement error in the initial independent measurement.

Our analysis suggests the absence of any important relationship between change and the initial level of FVC, FEV and FEF in 18–22 year old men. At first sight this appears contrary to the well known "horse racing effect" whereby loss is negatively associated with attained level, ie, the lower one's function, the more one loses.[18] This was first described in a data set on older men and is too large to be explained by the bias inherent in Method 3, which was used to correct for random measurement error in those data.[18] The effect occurs because lung function level in older men is in part determined by loss from the peak level attained in the early 20s until the age at which lung function is first measured. Subsequent loss is likely to be correlated with the loss prior to the first measurement and therefore with the first measurement itself. In our data, the first measurement was approximately at the time the peak level of lung function was attained. The lack of association of loss with level in young men is therefore not incompatible with the notion of the "horse racing" association in older men. However, our data do suggest that those

Table 5 *Regression of change (final–middle lung function tests) on middle and initial lung function tests (Method 4, n = 221)*

*LFT**		*OLS on middle LFT*	*On middle LFT corrected by G (Method 1)*	*On independent initial LFT (Method 4)*	*Method 4 further corrected by G*
FVC	β_L	−0·191	−0·129	−0·115	−0·124
	SE	0·034	0·038†	0·038	0·041†
	p	<0·001	0·001	0·002	0·002
FEV	β_L	−0·232	−0·146	−0·073	−0·081
	SE	0·038	0·044	0·048	0·053
	p	<0·001	0·001	0·1	0·1
$\sqrt{FEF}$	β_L	−0·215	−0·073	−0·026	−0·031
	SE	0·041	0·053	0·052	0·062
	p	<0·001	0·2	0·6	0·6
FEV%	β_L	−0·300	−0·109	−0·072	−0·092
	SE	0·043	0·062	0·048	0·061
	p	<0·001	0·08	0·1	0·1
PEFR	β_L	−0·423	−0·295	−0·062	−0·076
	SE	0·054	0·069	0·068	0·083
	p	<0·001	<0·001	0·4	0·4

LFT = Lung function tests; * See table 2 for explanation; β_L = Regression coefficient estimate; † Standard errors estimated from equation (6).

individuals with a lower peak adult level, because of genetic factors or environmental influences during childhood, are not at risk of a more rapid subsequent age related loss of function in adult life.

The example in this paper illustrates the effect of random measurement error on regression coefficients of change on initial level. The magnitude of the effect is considerable even in the case of lung function, which is less prone to error than many other measurements in medicine. It highlights the need to estimate random error of measurement. For this purpose the reliability coefficient, which indicates the proportion of variance attributable to true variability, is of more value than the more commonly used coefficient of variation. The reliability coefficient, estimated as the correlation between repeat measurements, is of practical value in correcting regression coefficients in many areas of epidemiological research.

We thank: Dr E Baskind, Mr I P Olivier and the staff of the lung function laboratory of the Medical Bureau for Occupational Diseases for doing the lung function tests; Magda de Beer and April Anderson for computing assistance; and Dr Carl Lombard for critical comment on an earlier draft. The late Dr Helen van Doorn was responsible for initiating the collecting of longitudinal lung function data on which the paper is based. This paper was based on a thesis by L M I, accepted for the degree of PhD at the University of the Witwatersrand, 1986.

References

[1] Davis CE. The effect of regression to the mean in epidemiologic and clinical studies. *Am J Epidemiol* 1976; **104:** 493–8.

[2] Healy MJR, Goldstein H. Regression to the mean. *Ann Hum Biol* 1978; **5:** 277–80.

[3] Oldham PD. A note on the analysis of repeated measurements of the same subjects. *J Chronic Dis* 1964; **15:** 969–77.

[4] Svärdsudd K, Blomqvist N. A new method for investigating the relation between change and initial value in longitudinal blood pressure data. *Scand J Soc Med* 1978; **6:** 85–95.

[5] Gill JS, Beevers DG, Zezulka AV, Davies P. Relationship between initial blood pressure and its fall with treatment. *Lancet* 1985; i: 567–9.

[6] Snedecor GW, Cochran WG. *Statistical methods*. 6th ed. Iowa: The Iowa State University Press, 1967: 164–6.

[7] Kupper LL. Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies. *Am J Epidemiol* 1984; **120:** 643–8.

[8] Gardner MJ, Heady JA. Some effects of within-person variability in epidemiological studies. *J Chronic Dis* 1973; **26:** 781–95.

[9] Shepard DS. Reliability of blood pressure measurements: implications for designing and evaluating programs to control hypertension. *J Chronic Dis* 1981; **34:** 191–209.

[10] Werts CW, Linn RL. A general linear model for studying growth. *Psycho Bull* 1970; **73:** 17–22.

[11] Lord FM, Novick MR. *Statistical theories of mental test scores*. Reading, Massachusetts: Addison Wesley, 1968: 198–223.

[12] Goldstein H. Some models for analysing longitudinal data on education attainment. *J R Statist Soc A* 1979; **142:** 407–42.

[13] Blomqvist N. On the relation between change and initial value. *J Am Statist Assoc* 1977; **72:** 746–9.

[14] Becklake MR, Permutt S. Evaluation of tests of lung function for 'screening' for early detection of chronic obstructive lung disease. In: Macklem PT, Permutt S, eds. *The lung in the transition between health and disease*. New York: Marcel Dekker, 1979; 345–87.

[15] Warren RD, White JK, Fuller WA. An error-in-variables analysis of managerial role performance. *J Am Statist Assoc* 1974; **69:** 886–93.

[16] Blomqvist N, Svärdsudd K. A new method for investigating the relation between change and initial value in longitudinal blood pressure data. II. Comparison with other methods. *Scand J Soc Med* 1978; **6:** 125–9.

[17] Maddala GS. *Econometrics*. New York: McGraw-Hill, 1977; 296–300.

[18] Fletcher C, Peto R, Tinker C, Speizer FE. *The natural history of chronic bronchitis and emphysema. An eight-year study of early chronic obstructive lung disease in working men in London*. Oxford: Oxford University Press, 1976.

[19] Wu M, Ware JH, Feinleib M. On the relation between blood pressure change and initial value. *J Chronic Dis* 1980; **33:** 637–44.

[20] American Thoracic Society. ATS Statement—Snowbird workshop and standardization of spirometry. *Am Rev Respir Dis* 1979; **119:** 831–8.

[21] Peslin R, Bohadana A, Hannhart B, Jardin P. Comparison of various methods for reading maximal expiratory flow-volume curves. *Am Rev Respir Dis* 1979; **119:** 271–7.

[22] Burrows B, Cline MG, Knudson RJ, Taussig LM, Lebowitz MD. A descriptive analysis of the growth and decline of the FVC and FEV_1. *Chest* 1983; **83:** 717–24.

[23] Glindmeyer HW III. Predictable confusion. *J Occup Med* 1981; **23:** 845–9.

[24] Knudson RJ, Lebowitz MD, Holberg CG, Burrows B. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis* 1983; **127:** 725–34.

[25] Stebbings JH Jr. Chronic respiratory disease among nonsmokers in Hagerstown, Maryland. II. Problems in the estimation of pulmonary function values in epidemiological surveys. *Environ Res* 1971; **4:** 163–92.

[26] Gregg I, Nunn AJ. Peak expiratory flow in normal subjects. *Br Med J* 1973; **3:** 282–4.

[27] Brooks AGF, Waller RE. Peak flow measurements among visitors to a public health exhibition. *Thorax* 1972; **27:** 557–62.

[28] Beaty TH, Menkes HA, Cohen BH, Newill CA. Risk factors associated with longitudinal change in pulmonary function. *Am Rev Respir Dis* 1984; **129:** 660–7.

[29] Kauffmann, F, Drouet D, Lellouch J, Brille D. Occupational exposure and 12-year spirometric changes among Paris area workers. *Br J Ind Med* 1982; **39:** 221–32.

[30] Krzyzanowski M. Changes of ventilatory capacity in an adult population during a five year period. *Bull Europ Physiopath Resp* 1980; **16:** 155–70.

[31] Tashkin DP, Detels R, Coulson AH, Rokaw SN, Sayre JW. The UCLA population studies of chronic obstructive respiratory disease. II. Determination of reliability and estimation of sensitivity and specificity. *Environ Res* 1979; **20:** 403–24.

[32] Irwig LM. *Correcting for the effect of measurement error on the association between environmental exposure and respiratory impairment. Research Report*. Johannesburg: Institute for the Biostatistics of the South African Medical Research Council, 1986.

Accepted for publication July 1988