press and public are more likely to blame the health care worker even though another route is much more likely.

Pell and colleagues criticise as inadequate the current recommendation that health care workers whose lifestyle is likely to put them at risk of HIV should consult an occupational health physician. But what lies between that an compulsory testing of all health care workers? There is no satisfactory course of action that would not discriminate against certain groups.

The risk from health care workers is very low. So why not do nothing? Five years ago the media and the public might have accepted this approach. This is no longer possible, however, and as a result sensational publicity will continue to surround health care workers who are found to be HIV positive, causing additional distress to them and their relatives and undue anxiety to the patients they have treated. The departments of health must accept responsibility for this unfortunate situation.

Statistics Notes

The use of transformation when comparing two means

J Martin Bland, Douglas G Altman

This is the 19th in a series of occasional notes on medical statistics

The usual statistical technique used to compare the means of two groups is a confidence interval or significance test based on the *t* distribution. For this we must assume that the data are samples from normal distributions with the same variance. Table 1 shows the biceps skinfold measurements for 20 patients with Crohn's disease and nine patients with coeliac disease.

Table 1—*Biceps skinfold thickness (mm) in two groups of patients*

Crohn's disease				Coeliac disease	
1·8	2·8	4·2	6·2	1·8	3·8
2·2	3·2	4·4	6·6	2·0	4·2
2·4	3·6	4·8	7·0	2·0	5·4
2·5	3·8	5·6	10·0	2·0	7·6
2·8	4·0	6·0	10·4	3·0	
Mean=4·72 SD=2·42				Mean=3·53 SD=1·96	

The data have been put into order of magnitude, and it is fairly obvious that the distribution is skewed and far from normal. When, as here, the assumption of normality is wrong we can often transform the data to another scale where the assumption of normality is reasonable. The transformation which achieves a normal distribution should also give us similar variances.[1] Table 2 shows the results of analyses using the square root, logarithmic, and reciprocal transformations. The log transformation gives the most similar variances and so gives the most valid test of significance. It also gives a reasonable approximation to a normal distribution.

Table 2—*Biceps skinfold thickness compared for two groups of patients, using different transformations*

Transformation	Two sample *t* test, 27 df: *t*	P	95% Confidence interval for difference on transformed scale	Variance ratio, larger/smaller
None, raw data	1·28	0·21	−0·71 mm to 3·07 mm	1·52
Square root	1·38	0·18	−0·140 to 0·714	1·16
Logarithm	1·48	0·15	−0·114 to 0·706	1·10
Reciprocal	−1·65	0·11	−0·203 to 0·022	1·63

Confidence intervals for transformed data are more difficult to interpret, however. Unlike the case of a single sample,[2] the confidence limits for the difference between means cannot be transformed back to the original scale. If we try to do this the square root and reciprocal limits give ludicrous results. The lower limit for the square root transformation is negative. If we square this we get a positive lower limit and the confidence interval does not contain zero, even though the difference is not significant. If the observed difference were exactly zero the confidence limits would be equal in magnitude but opposite in sign. Transforming back by squaring would make them equal. For the reciprocal transformation the upper limit is very small (0·022) and transforming back by taking the reciprocal again gives 45·5. There is no way that the difference between mean skinfold in these two groups could be 45·5 mm. Thus the confidence interval for a difference cannot be interpreted on the untransformed scale for these transformations.

Only the log transformation gives interpretable (and thus useful) results after we transform back. Using the antilog transformation, we get a confidence interval of 0·89 to 2·03, but these are not limits for the difference in millimetres. How could they be, for they do not contain zero, yet the difference is not significant? They are in fact the 95% confidence limits for the ratio of the geometric mean[2] for patients with Crohn's disease to the geometric mean for patients with coeliac disease. If there were no difference the expected value of this ratio would be 1, not 0, and so lie within the limits. This procedure works because when we take the difference between the logarithms of the two geometric means we get the logarithm of their ratio, not of their difference.[3] We thus have the logarithm of a pure number and we antilog this to give the dimensionless ratio of the two geometric means. The logarithmic transformation is strongly preferable to other transformations for this reason. Fortunately, for medical measurements it often achieves the desired effect.

Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE
J Martin Bland, *professor of medical statistics*

ICRF Medical Statistics Group, Centre for Statistics in Medicine, Institute of Health Sciences, PO Box 777, Oxford OX3 7LF
Douglas G Altman, *head*

Correspondence to: Professor Bland.

BMJ 1996;312:1153

1 Bland JM, Altman DG. Transforming data. *BMJ* 1996;312:770.
2 Bland JM, Altman DG. Transformations, means, and confidence intervals. *BMJ* 1996;312:1079.
3 Bland JM, Altman DG. Logarithms. *BMJ* 1996;312:700.