

- 2 Rawles J. Magnitude of benefit from early thrombolytic treatment in acute myocardial infarction: new evidence from Grampian region early anistreplase trial (GREAT). *BMJ* 1996;312:212-6.
- 3 British Heart Foundation Working Group. Guidelines for the early management of patients with myocardial infarction. *BMJ* 1994;308:767-71.
- 4 Rawles J. On behalf of the GREAT group. Halving of mortality at one year by domiciliary thrombolysis, in Grampian region early anistreplase trial (GREAT). *J Am Coll Cardiol* 1994;23:1-5.
- 5 Rawles J, Adams JN. Implications for the Grampian region early anistreplase trial (GREAT) for pre-hospital coronary care in Scotland. *Health Bulletin* 1994;52:106-17.
- 6 Barradell LB, Goa KL. Alteplase: a pharmacoeconomic evaluation of its use in the management of myocardial infarction. *Pharmacoeconomics* 1995;8:428-59.

(Accepted 12 December 1996)

Statistics notes

Cronbach's alpha

J Martin Bland, Douglas G Altman

This is the 28th in a series of occasional notes on medical statistics

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 1997;314:572

Many quantities of interest in medicine, such as anxiety or degree of handicap, are impossible to measure explicitly. Instead, we ask a series of questions and combine the answers into a single numerical value. Often this is done by simply adding a score from each answer. For example, the mini-HAQ is a measure of impairment developed for patients with cervical myelopathy.¹ This has 10 items (table 1) recording the degree of difficulty experienced in carrying out daily activities. Each item is scored from 1 (no difficulty) to 4 (can't do). The scores on the 10 items are summed to give the mini-HAQ score.

When items are used to form a scale they need to have internal consistency. The items should all measure the same thing, so they should be correlated with one another. A useful coefficient for assessing internal consistency is Cronbach's alpha.² The formula is:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum s_i^2}{s_T^2} \right)$$

where k is the number of items, s_i^2 is the variance of the i th item and s_T^2 is the variance of the total score formed by summing all the items. If the items are not simply added to make the score, but first multiplied by weighting coefficients, we multiply the item by its coefficient before calculating the variance s_i^2 . Clearly, we must have at least two items—that is $k > 1$, or α will be undefined.

The coefficient works because the variance of the sum of a group of independent variables is the sum of their variances. If the variables are positively correlated, the variance of the sum will be increased. If the items making up the score are all identical and so perfectly correlated, all the s_i^2 will be equal and $s_T^2 = k^2 s_i^2$, so that $\sum s_i^2 / s_T^2 = 1/k$ and $\alpha = 1$. On the other hand, if the items

are all independent, then $s_T^2 = \sum s_i^2$ and $\alpha = 0$. Thus α will be 1 if the items are all the same and 0 if none is related to another.

For the mini-HAQ example, the standard deviations of each item and the total score are shown in the table. We have $\sum s_i^2 = 11.16$, $s_T^2 = 77.44$, and $k = 10$. Putting these into the equation, we have

$$\alpha = \frac{10}{9} \times \left(1 - \frac{11.16}{77.44} \right) = 0.95$$

which indicates a high degree of consistency.

For scales which are used as research tools to compare groups, α may be less than in the clinical situation, when the value of the scale for an individual is of interest. For comparing groups, α values of 0.7 to 0.8 are regarded as satisfactory. For the clinical application, much higher values of α are needed. The minimum is 0.90, and $\alpha = 0.95$, as here, is desirable.

In a recent example, McKinley *et al* devised a questionnaire to measure patient satisfaction with calls made by general practitioners out of hours.³ This included eight separate scores, which they interpreted as measuring constructs such as satisfaction with communication and management, satisfaction with doctor's attitude, etc. They quoted α for each score, ranging from 0.61 to 0.88. They conclude that the questionnaire has satisfactory internal validity, as five of the eight scores had $\alpha > 0.7$. In this issue Bosma *et al* report similar values, from 0.67 to 0.84, for assessments of three characteristics of the work environment.⁴

Cronbach's alpha has a direct interpretation. The items in our test are only some of the many possible items which could be used to make the total score. If we were to choose two random samples of k of these possible items, we would have two different scores each made up of k items. The expected correlation between these scores is α .

Table 1 Mini-HAQ scale in 249 severely impaired subjects

Item	Mean score	SD of score s_i
Stand	2.96	1.04
Get out of bed	2.57	1.11
Cut meat	2.91	1.12
Hold cup	2.41	1.06
Walk	2.64	1.04
Climb stairs	3.06	1.04
Wash	3.25	1.01
Use toilet	2.59	1.09
Open a jar	2.86	1.02
Enter/leave car	2.80	1.03
Mini-HAQ	28.06	$s_T = 8.80$

- 1 Casey ATH, Crockard HA, Bland JM, Stevens J, Moskvich R, Ransford AO. Development of a functional scoring system for rheumatoid arthritis patients with cervical myelopathy *Ann Rheum Dis* (in press).
- 2 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-333.
- 3 McKinley RK, Manku-Scott T, Hastings AM, French DP, Baker R. Reliability and validity of a new measure of patient satisfaction with out of hours primary medical care in the United Kingdom: development of a patient questionnaire. *BMJ* 1997;314:193-8.
- 4 Bosma H, Marmot MG, Hemingway H, Nicholson AC, Brunner E, Stansfield SA. Low job control and risk of coronary heart disease in Whitehall II (prospective cohort) study. *BMJ* 1997;314:558-65.