Reliability of a population survey tool for measuring perceived health problems: a study of patients with osteoarthrosis

SONJA M. HUNT, S. P. McKENNA AND JAN WILLIAMS
From the Department of Community Health, Queen's Medical Centre, University of Nottingham

SUMMARY A study is described which tested the reliability of the Nottingham Health Profile, a standardised measure of perceived health status suitable for use with both large groups and individual patients. Patients with osteoarthrosis, uncomplicated by other diagnosed conditions, were given the questionnaire to fill in on two separate occasions four weeks apart. Analysis of data showed that all sections of the profile achieved a high level of reliability with coefficients of correlation ranging from 0.77 to 0.85 on part I and from 0.44 to 0.86 on part II. Patients reported many problems in the areas of physical mobility, pain, sleep, and energy, but rather fewer in the areas of social isolation and emotional reactions. With the exception of family relationships, most respondents reported a high level of disruption of their daily activities. Patients were very consistent in the number and type of problems which they affirmed.

The Nottingham Health Profile is a two-part, self-administered questionnaire designed to provide a simple and standardised measure of perceived health problems, physical, social, and mental. It was developed in order to supply a tool for the survey of populations, but has also proved to be of value in clinical situations as a valid means of assessing the wider impact of illness on the daily life of patients.

Part I consists of 38 statements with weighted scores in the areas of pain, physical mobility, sleep, energy, social isolation, and emotional reactions, to which the respondent is required to answer yes or no. The development of this part of the profile has been described in previous papers. Part II is designed to give a general estimate of those areas of social function perceived to be affected by the health problems of the individual. This part contains a single statement on each of the following areas: paid employment, jobs around the home, social life, sex life, family relationships, hobbies/interests, and holidays.

RELIABILITY
One of the requirements of a standardised measure is that it should be reliable. Scores should show consistency over time and there should be some numerical estimate of the degree of consistency exhibited by the instrument.

The reliability of a set of scores can logically be defined as the proportion of the measured variance which is the 'true' variance, that is, the genuine value of whatever is being measured. The remainder will be error variance composed of elements which, it is usually assumed, will occur independently and at random. In practice this assumption may be unfounded, for example, where the initial administration of a test sensitises the respondents to the content and affects subsequent administrations of that test.

There are three well-established methods for estimating reliability:

1. The method of internal consistency where items on an instrument are divided into two equivalent parts—the split half technique—and the correlations between the scores on each part are computed.

2. Alternate forms, where two instruments which have been developed in parallel and which measure the same attribute are administered and the scores on one form are correlated with the scores on the other.

3. The test re-test technique, where the same test is administered on two separate occasions to the same group of individuals and the correlation between the two sets of scores is computed.

The first method requires that the items in the instrument should be homogeneous with respect to the attribute being measured and the second method requires that a parallel form of the instrument should exist. Neither of these requirements are fulfilled in the case of the Nottingham Health Profile. The third method, test re-test, is thus the preferred one.
The test re-test technique is concerned with whether or not the instrument will give the same or similar scores if administered to the same person on more than one occasion. This technique requires that the subsequent administration of the instrument should replicate the features of the original administration of the questionnaire and that it should be re-administered after a suitable time interval. In the case of health measurement it may be predicted that the longer the interval between administrations, the greater will be the chance that 'real' changes have taken place in the respondent's health or in the information the respondent has about health matters. In this case the reliability of the instrument would be underestimated. Conversely, if the interval between administrations is too short, respondents may recall the answers that they gave the first time and try to replicate them. In this case the reliability of the instrument will be overestimated.

There have been few studies reporting the reliability of mailed questionnaires concerned with health status. A national health survey in the United States of America, in which 64 patients were tested and re-tested after two weeks, suggested that illnesses are reported with a high degree of reliability, but that symptoms are not.8 Meltzer and Hochstim, using a mailed questionnaire sent to 1530 people on two occasions a week apart, found that responses had a 96% concordance rate overall. Chronic conditions were reported more reliably than acute complaints and negative answers were more consistent than affirmative ones.7

Since the Nottingham Health Profile yields a high proportion of negative answers when given to a 'normal' population, an investigation of its reliability would need to be carried out on a population which could be expected to give a high proportion of affirmative responses in order to avoid overestimating its reliability. However, to avoid the underestimation of reliability, the sample would need to be such that the condition of the respondents would not be expected to change significantly over a short period of time.

For these reasons, patients suffering from osteoarthritis were selected for study, specifically those awaiting hip replacement operations. Osteoarthritis is a degenerative bone condition leading to a chronic state in which, while there may be daily fluctuations, there is little change over a period of weeks, especially when the condition has become severe enough to warrant arthroplasty.

Because of the limitations it imposes on independence and mobility, osteoarthritis has implications for the whole lifestyle of the patient. Surveys of arthritic patients have described the physical, social, and emotional concomitants of the disease in terms of pain, disability, restriction of daily activities and travel, frustration, social isolation, and lack of energy.8 It has been found that the frequency of such complaints is consistent over the period of a month although showing minor variations from day to day.10

Method

The co-operation of eight orthopaedic surgeons in the Nottingham area was obtained and access gained to the lists of patients awaiting arthroplasty of the hip. Patients were selected using the following criteria—that they had no known additional conditions or complications which might affect their health status and that they would not be likely to be called for their operation during the period of the study.

A list of the selected patients was sent to the consultant concerned so that he could verify that the patients fulfilled these criteria and that they would not be likely to be distressed by the study.

The number of subjects obtained in this way was 73. Each person was sent the Nottingham Health Profile, a covering letter, and a prepaid reply envelope. All subjects who responded were sent a second questionnaire four weeks later. The period of four weeks was chosen in order to reduce overestimation of reliability due to memory effects. Patients were assured of confidentiality and it was emphasised that their responses would in no way affect their position on the waiting list.

Results

A response rate of 88% was achieved for the first questionnaire and 90% of those who responded on the first occasion returned the second questionnaire. Only the 58 respondents who completed both profiles are included in the data analysis. The sample consisted of 22 men and 36 women.

There were two respondents in the age group 40–49, 10 in the age group 50–59, 25 in the age group 60–69, and 20 aged 70 or over. One respondent failed to indicate age group.

As expected, the number of affirmative responses was very high, especially in part I on the sections referring to physical mobility, pain, energy, and sleep, and in part II on statements referring to work around the house, social life, hobbies, and sex. These findings are shown in Tables 1 and 2.

Analysis of Data

Part I of the profile was tested for reliability using the Spearman correlation coefficient. For part II, which yields only categorical data, the most suitable
measures of the extent of association between the first and second sets of responses appeared to be the contingency coefficient. However, this statistic has the limitation of not being able to attain unity, even if the variables are perfectly correlated, since the upper limit is a function of the number of categories into which the data are divided. In view of this limitation, part II data were analysed using Cramer's modification of the contingency coefficient which allows C to attain unity. Tables 3 and 4 show the coefficients together with their significance levels.

Discussion

The correlations between the two sets of scores obtained from the initial and subsequent administrations of the Nottingham Health Profile were found to be significantly and satisfactorily high. The use of patients with osteoarthritis ensured a high number of affirmative responses to each statement, thus avoiding a spurious reliability, as a consequence of many negative responses. The fact that respondents perceived themselves to have so many problems suggests that attempts to remember initial responses and replicate them would have been rendered difficult. The time lapse of four weeks between administrations of the profile further reduced the chances of overestimating the reliability of the instrument. The level of reliability obtained compares favourably with other established measures of health status such as the Sickness Impact Profile and the McMaster Health Index Questionnaire and with widely used intelligence tests like the Wechsler Adult Intelligence Scale and the Stanford-Binet test.

Derived correlation coefficients are, strictly, applicable only to the population from which the sample was drawn. Thus we can now say with confidence that the Nottingham Health Profile has proved to be reliable as a measure of the health problems experienced by patients with osteoarthritis. It does seem likely that, in view of previous studies to establish validity, similar levels of reliability would be obtained from other populations as well. Studies currently in progress on pregnant women, surgical patients, and fracture victims, where 'healthy' control groups are also completing the profile on two separate occasions, should allow further tests of reliability.

This research was supported by a grant from the Social Science Research Council No. HR 6157/1.

Reprints from Dr. J. McEwen, Queen's Medical Centre, University of Nottingham, Clifton Boulevard, Nottingham NG7 2UH.

References

Sonja M. Hunt, S. P. McKenna, and Jan Williams


