

# Are the modified “simple questions” a valid and reliable measure of health related quality of life after stroke?

Paul Dorman, Martin Dennis, Peter Sandercock, on behalf of the United Kingdom Collaborators in the International Stroke Trial (IST)

## Abstract

**Objectives**—Two “simple questions” were developed as a minimalist measurement tool to assess outcome in large trials and epidemiological studies after stroke. A previous study of their validity had disclosed ambiguities in their wording. In this study, the clarity, validity, and reliability of a modified version of these simple questions were examined. The relation between patients’ responses to these questions and two widely used generic measures of health related quality of life were also studied.

**Methods**—A hospital based stroke register cohort of 152 patients, who were all visited at home by a study nurse, was used to study validity. A cohort of 1753 patients derived from the International Stroke Trial was used to study the relation with measures of quality of life. The sensitivity, specificity, and accuracy with which responses to each question predicted the patients’ outcome measured using standard instruments was assessed. The distribution of scores for the EuroQol and SF-36 was examined for patients classified as dependent, independent, and fully recovered by the combined use of the modified simple questions.

**Results**—The modified “dependency” question had excellent sensitivity (>85%), specificity (>79%), and accuracy (>82%) for identifying dependency after stroke. The “problems” question had good sensitivity (65–88%) and moderate specificity (36–72%) for the detection of problems in a broad range of domains. The combined use of the modified dependency and problems questions provided a valid, simple, and reliable overall indicator of health related quality of life after stroke.

**Conclusions**—The modified simple questions have excellent face validity and good measurement properties for the assessment of outcome after stroke. They are particularly well suited for large epidemiological studies and randomised trials. (J Neurol Neurosurg Psychiatry 2000;69:487-493)

Keywords: stroke; outcome assessment; quality of life

The selection of the optimal measure of outcome after stroke must take account of the purpose of measurement. Complex stroke scales focus on impairments (the neurological signs) which are less important to the patient’s

daily life than disability and handicap (what they can and what they actually do). Although stroke scales provide a relatively objective and sensitive means for assessing outcome, they may be criticised for poor interpretability and little relevance to patients.<sup>1 2</sup> In general they are not appropriate for use in large clinical trials which attempt to determine the role of an intervention in clinical practice. Scales that measure disability, handicap, or quality of life are more relevant in this setting.

Lindley *et al* developed two “simple questions” to assess disability and outcome after stroke for use in large scale trials.<sup>3</sup> One question assessed dependency: “In the last two weeks, did you require help from another person for everyday activities?” and one assessed recovery: “Do you feel that you have made a complete recovery from your stroke?”<sup>3</sup> These questions have been used in two recent randomised controlled trials in patients with stroke.<sup>4 5</sup> The data were analysed in different ways in each of these studies: in the trial of low molecular weight heparin, the dependency question was used to provide an estimate of the proportion of surviving patients disabled at follow up.<sup>4</sup> In the International Stroke Trial, both questions were used together to classify surviving patients into one of three hierarchical levels of outcome (dependent, independent but not recovered, independent and fully recovered).<sup>5</sup> Although this hierarchical classification has intuitive appeal, its relation with other classifications of outcome (for example, health related quality of life or the World Health Organisation (WHO) Classification of Impairments, Disabilities, and Handicaps) has not been established.

The initial assessment by Lindley *et al* of the measurement attributes of the questions suggested they had good validity and acceptable reliability.<sup>3</sup> However, a more recent study highlighted some ambiguities in the wording.<sup>6</sup> Some patients who answered “yes” to the dependency question in fact meant that they needed help from someone, in addition to their normal helper, and not just that they needed help from any person at all.<sup>6</sup> Similarly, other patients answered “yes” to the recovery question when they meant that they had stopped recovering, and not that they had returned to their prestroke state.<sup>6</sup> These ambiguities may have reduced the questions’ validity and reliability and so reduced their power to detect true differences in outcome.

Dennis *et al* proposed a modification to both questions to try and improve their clarity,

Department of Neurology, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne NE1 4LP, UK  
P J Dorman

Department of Clinical Neuroscience, University of Edinburgh, Scotland  
M Dennis  
P Sandercock

Correspondence to:  
Dr Paul Dorman  
pauljdorman@hotmail.com

Received 10 November 1999  
and in final form  
10 April 2000  
Accepted 20 April 2000

validity, and reliability.<sup>6</sup> They removed the clause referring to “another person” from the dependency question to give “Do you need help from anybody with everyday activities?” and reworded the recovery question to give “Has the stroke left you with any problems?”. We examined the validity of these modified questions in the current study. We examined whether the modified dependency question is a valid measure of dependency and whether the combined use of the questions to classify patients into one of three outcome groups (dependent, independent with problems, and recovered (independent without persisting problems)) is a valid measure of overall health related quality of life. We also assessed the test-retest reliability of this classification and the agreement between patients and their proxies (carers) for responses to the modified simple questions, because patients are often unable to complete questionnaires or participate in interviews after a stroke.

## Methods

### SELECTION OF PATIENTS

#### *Lothian Stroke Register series*

We studied the validity and the proxy validity of the modified simple questions in a series of 152 patients from our prospective registry of inpatients and outpatients with first or recurrent stroke. The patients, and their available proxies, were asked the questions as part of a previous study to examine the validity of the EuroQol questionnaire after stroke.<sup>7,8</sup> We have described elsewhere how we identified the patients and their proxies and how we administered the instruments.<sup>7,8</sup> Briefly, all patients were visited by a research nurse at home. The nurse administered the modified simple questions, the EuroQol,<sup>9</sup> the Frenchay activities index,<sup>10</sup> the hospital anxiety and depression scale (HADS),<sup>11</sup> and a visual analogue pain scale as questionnaires to be completed by the patient as far as possible. When patients could not complete the questionnaires by themselves, the nurse administered the questionnaires by interview. The nurse always administered the modified simple questions first, to limit interaction with the subsequent questions. The nurse assessed the Barthel index<sup>12</sup> and the OPCS disability scores<sup>13</sup> by direct questioning at the end of the interview. We asked patients to

ensure that a friend or relative (a proxy) who knew them well was available at the time of the interview. The nurse asked each proxy to independently complete a questionnaire booklet including the modified simple questions on behalf of the patient.

#### *International Stroke Trial series*

We also studied the relation between responses to the modified simple questions and the assessments of health related quality of life, measured with the EuroQol and SF-36, in the cohort of patients included in the randomised comparison of these instruments.<sup>14</sup> We have described the selection of patients, allocation of patients to the questionnaires, and the methods of questionnaire administration in detail elsewhere.<sup>14</sup> Briefly, we randomly allocated 2253 patients to receive either the EuroQol or the SF-36. The study included patients with confirmed or suspected ischaemic stroke who had been enrolled between 2 March 1993 and 31 May 1995 by any of the United Kingdom hospitals participating in the International Stroke Trial.<sup>5</sup> We included all patients who were not known to have died by the time of the survey. We incorporated the EuroQol and SF-36 into booklets that included some additional questions recording the patients' demographic details, their functional outcome after stroke, and whether the patient completed the booklet by themselves. The questionnaire booklets were identical in all respects, other than the nature of the health related quality of life (HRQoL) instrument. The modified simple questions were readministered with the health related quality of life questionnaires for both a study of the reliability,<sup>15</sup> and a study of the relation between the different quality of life assessments (fig 1).<sup>16</sup> We used data from both of these substudies to determine the test-retest reliability of the individual questions and that of the overall classification of outcome using the modified simple questions.

#### STATISTICAL ANALYSIS

For the Lothian Stroke Register cohort, we assessed the concurrent validity of the modified simple questions by calculating the sensitivity, specificity, and accuracy with which the responses to each question predicted whether the patient scored as good or bad (score above or below the cut offs for the appropriate

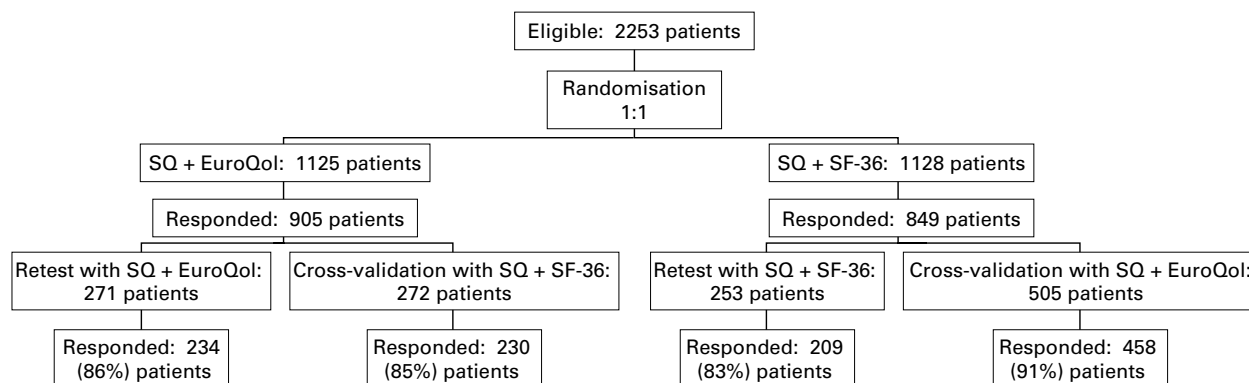


Figure 1 Flow of patients through the study of outcome among United Kingdom cohorts in the International Stroke Trial (IST). SQ = sample quotient; SF-36 = short form 36.

standard instrument which separated good from bad outcome). The cut offs were selected either if they had face validity or if they had been used in previous studies for each subscale.<sup>3 6 17 18</sup>

The modified simple questions were used to classify surviving patients into one of three hierarchical outcome groups: patients who responded with “yes” to the dependency question were classified as dependent, patients who responded with “no” to the dependency question and “yes” to the problems question were classified as independent and patients who responded with “no” to both questions were considered independent and recovered. We investigated the validity of this classification by plotting histograms that showed the distribution of responses to our standard measures for patients who were dependent, independent, and recovered. For both cohorts, we further assessed the concurrent validity of this classification by calculating the median score and interquartile range for each standard instrument for each of these three outcome groups. We used the Terpstra-Jonckheere test to compare the ordering of the distribution of scores for each group.

For the International Stroke Trial cohort, we examined the agreement between patients’ responses to the modified simple questions on test and retest using the  $\kappa$  statistic.<sup>19</sup> We also examined the agreement between patients and their proxies for their classification of the patients’ outcome using the  $\kappa$  statistic for the Lothian Stroke Register cohort.<sup>19</sup>

Data were held in an Access 2.0 (Microsoft Corporation) database and analyzed using the statistical software package SPSS for Windows (Release 7.5).

## Results

### LOTHIAN STROKE REGISTER SERIES

The patients were assessed at a median interval of 72 weeks after the onset of their index stroke (interquartile range: 43 to 104 weeks). Of the 152 patients who participated in this study, 92 were able to complete the questionnaires by themselves without help; the remaining 60 patients could only be assessed by interview.

One hundred and forty seven of the 152 patients (97%) completed both modified simple questions. Three patients did not complete

the modified dependency question and another two patients did not complete the problems question. Fifty four (37%) patients replied “yes” to the dependency question and were classified as dependent; eight (15%) of these 54 dependent patients replied “no” to the question about problems from their stroke (but in fact five of these eight had been rated as having an Oxford handicap score of at least 2 or more before their index stroke).<sup>20 21</sup> The remaining 93 (63%) patients were independent in everyday activities; of these, 52 (56%) reported being left with problems after their stroke (independent, but not recovered), whereas the other 41 (44%) denied any problems (recovered).

### Validity

The sensitivity, specificity, and accuracy of the modified simple questions for the assessment of outcome in the domains of mobility, self care, social functioning, pain, and psychological functioning are reported in table 1. The modified dependency question had excellent sensitivity, specificity, and accuracy for the assessment of activities of daily living as defined by the Barthel index. Predictably, it was less accurate at predicting psychological dysfunction as defined by the HADS. The problems question proved highly sensitive for the detection of problems in all of the domains assessed, but it had low specificity. Its specificity was better (63%) when used to identify patients with problems in any of the unidimensional measures, table 1.

The median score and interquartile range for the standard instruments are shown for groups defined by their responses to both questions (table 2). The median scores with the standard instruments were ordered appropriately and were statistically distinct. The mean utility scores with the EuroQol differed significantly among groups of these patients classified by their responses to the simple questions (table 3). The best outcome in all domains (including overall health related quality of life with the “thermometer”, table 3) was reported by the fully recovered group of patients.

*Validity of the proxy assessments*—The study nurse, using the OPCS disability measure, rated six patients as having significant difficulties in communication. These six patients did

Table 1 Sensitivity, specificity, and accuracy of the modified simple questions in predicting patients’ responses to more complex measures (Lothian Stroke Register cohort)

|                                  | “Do you need help from anybody with everyday activities?” |                 |              | “Has the stroke left you with any problems” |                 |              |
|----------------------------------|---|-----------------|--------------|---|-----------------|--------------|
|                                  | Sensitivity (%)   | Specificity (%) | Accuracy (%) | Sensitivity (%)                             | Specificity (%) | Accuracy (%) |
| OPCS locomotion (>0)*            | 72  | 95              | 84           | 84  | 49              | 65           |
| Barthel index                    |   |                 |              |   |                 |              |
| <20)                             | 85  | 85              | 85           | 87  | 42              | 56           |
| <18)                             | 91  | 79              | 82           | 85  | 38              | 49           |
| Frenchay activities index:       |   |                 |              |   |                 |              |
| >34)                             | 100   | 42              | 49           | 65  | 72              | 71           |
| <22)                             | 67  | 90              | 79           | 77  | 42              | 59           |
| <17)                             | 77  | 88              | 84           | 81  | 41              | 56           |
| Visual analogue pain scale (>0)* | 52  | 73              | 64           | 70  | 36              | 50           |
| HADS depression subscale (<7)*   | 60  | 71              | 67           | 88  | 42              | 56           |
| HADS anxiety subscale (<7)*      | 50  | 70              | 61           | 83  | 46              | 62           |
| Problems on any single measure†  | —   | —               | —            | 75  | 63              | 79           |

\*Low scores indicate better outcome.

†Problems were defined using above cut offs (<17 for Frenchay activities index and <18 for Barthel index).

Table 2 Concurrent validity of the outcome defined by response to modified simple questions (Lothian Stroke Register cohort)

| Standard instrument        | Median score on standard instruments (interquartile range) for groups defined by response to modified simple questions |                    |                  | p Value |
|----------------------------|--|--------------------|------------------|---------|
|                            | Dependent (n=54)   | Independent (n=52) | Recovered (n=41) |         |
| OPCS locomotion score      | 7 (3–9.5)  | 0 (0–3)            | 0 (0–0)          | <0.0001 |
| Barthel index              | 17 (12–20)   | 20 (20–20)         | 20 (20–20)       | <0.0001 |
| Frenchay activities index  | 9 (5–15)   | 26 (21–31)         | 33 (26–36)       | <0.0001 |
| Visual analogue pain scale | 15 (0–50)  | 0 (0–20)           | 0 (0–10)         | <0.0001 |
| HADS depression subscale   | 6 (4–8)  | 4 (2–7)            | 3 (2–4)          | <0.0001 |
| HADS anxiety subscale      | 7 (2–9)  | 4 (2–10)           | 3 (1–5)          | 0.005   |

p Value=significance of hierarchical differences between levels using the Terpstra-Jonckheere test; dependent=needing help from anybody with everyday activities; independent=not needing help from anybody for everyday activities, but persisting problems after stroke; recovered=not needing help from anybody for everyday activities and no persisting problems after stroke.

Table 3 Relation between patients' responses to the modified simple questions and overall health related quality of life assessed by the EuroQol

|                     | Mean score on EuroQol for groups defined by their responses to modified simple questions (95% CI of the mean) |                     |                  | p Value† |
|---------------------|---|---------------------|------------------|----------|
|                     | Dependent   | Independent         | Recovered        |          |
| LSR Series (n=147): |   |                     |                  |          |
| Overall HRQoL*      | 58 (52–64)  | 65 (58–71)          | 77 (74–80)       | <0.0001  |
| EuroQol utility‡    | 0.38 (0.29–0.47)  | 0.74 (0.69 to 0.79) | 0.88 (0.80–0.96) | <0.0001  |
| IST Series (n=867): |   |                     |                  |          |
| Overall HRQoL       | 48 (47–50)  | 67 (64–70)          | 77 (74–80)       | <0.0001  |
| EuroQol utility     | 0.31 (0.29–0.34)  | 0.71 (0.68–0.74)    | 0.88 (0.84–0.92) | <0.0001  |

\*Measured with vertical visual analogue scale which forms part of the EuroQol; †significance of hierarchical differences between levels using the Terpstra-Jonckheere test; ‡utilities were generated from patients' categorical responses to the EuroQol, using the preferences of the general public.

not therefore provide data for the assessment of agreement between the patients and their proxies. We obtained proxy assessments of the patients' outcome with the modified simple questions for 121 (82%) of the 147 patients. Agreement between the patient and their proxy

for the dependency question was very good (agreement=103/121 (85%);  $\kappa=0.70$ , 95% confidence interval (95% CI) 0.57–0.83). Agreement between the patient and their proxy for the problems question was good (agreement=99/121 (82%);  $\kappa=0.58$ , 95% CI 0.42–0.74). Agreement between the patient and their proxy for the overall outcome classification with the modified simple questions was good (agreement = 87/121 (72%);  $\kappa=0.57$ , 95% CI 0.45–0.69). There was an overall trend towards proxies assessing the patients' functioning as worse than that assessed by the patients themselves (21 patients were rated worse by their proxies than by themselves versus 13 patients who were rated better by their proxies than by themselves, sign test  $p>0.05$ ).

#### INTERNATIONAL STROKE TRIAL SERIES Relation between the simple questions and the EuroQol

Figure 2 shows the distribution of responses to the EuroQol questionnaire for the patients classified by their responses to the modified simple questions. As before, each of the groups have distinct responses for all of the domains of the EuroQol and the pattern of responses was very similar to that observed in the other (Lothian Stroke Register) cohort (histogram not shown, but available from authors). For this larger cohort of patients, the mean estimates of the EuroQol utilities were also ordered appropriately and were statistically distinct for each of the three levels (table 3). The mean estimates of overall health related quality of life showed a similar pattern (table 3).

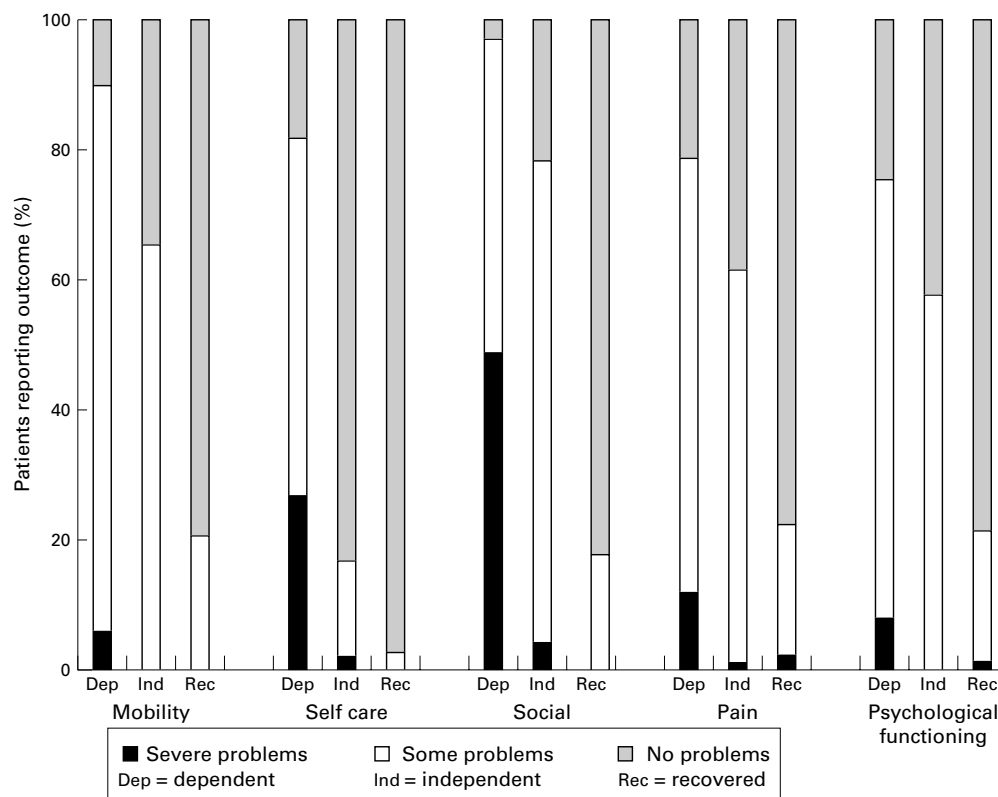


Figure 2 Health related quality of life in each domain of the EuroQol among patients categorised as dependent, independent, or recovered by their response to the modified simple questions (ISTS).

Table 4 Relation between patients' responses to the modified simple questions and health related quality of life assessed by the SF-36 (International Stroke Trial series)\*

| Domain of SF-36         | Median score on domains of SF-36 for groups defined by their responses to modified simple questions (interquartile range) |                     |                   | p value† |
|-------------------------|---|---------------------|-------------------|----------|
|                         | Dependent (n=564)   | Independent (n=137) | Recovered (n=123) |          |
| Physical functioning    | 5 (0-25)  | 50 (30-65)          | 68 (50-85)        | <0.0001  |
| Physical role function  | 0 (0-0)   | 0 (0-25)            | 75 (0-100)        | <0.0001  |
| Social functioning      | 25 (13-50)  | 63 (50-75)          | 88 (63-100)       | <0.0001  |
| Mental health           | 60 (52-68)  | 64 (56-68)          | 64 (60-68)        | 0.01     |
| Emotional role function | 0 (0-25)  | 25 (0-100)          | 100 (25-100)      | <0.0001  |
| Bodily pain             | 41 (22-62)  | 62 (41-100)         | 100 (62-100)      | <0.0001  |
| Vitality                | 30 (15-45)  | 40 (30-55)          | 60 (52-75)        | <0.0001  |
| General health          | 35 (25-52)  | 52 (35-72)          | 72 (55-87)        | <0.0001  |

\*Simple question data missing in 25 patients; †significance of hierarchical differences between levels using the Terpstra-Jonckheere test.

Table 5 Test-retest reliability of the modified "dependence" question (International Stroke Trial series)

|                   | First assessment |             |     |
|-------------------|------------------|-------------|-----|
|                   | Dependent        | Independent |     |
| Second assessment | Dependent        | 679         | 46  |
|                   | Independent      | 37          | 271 |

$\kappa=0.81$  (0.77–0.85).

Table 6 Test-retest reliability of the "problems" question (International Stroke Trial series)

|                   | First assessment |             |     |
|-------------------|------------------|-------------|-----|
|                   | Problems         | No problems |     |
| Second assessment | Problems         | 862         | 25  |
|                   | No problems      | 42          | 154 |

$\kappa=0.78$  (0.73–0.83).

#### Relation between the simple questions and the SF-36

Although the responses to the modified simple questions seemed to be closely related to the patients' health related quality of life as defined by the EuroQol, we then confirmed the relation by relating the simple questions with the SF-36 (table 4). For all but two of the domains (physical role functioning and mental health), the median scores for each domain were ordered appropriately and differed from each other. In the physical role functioning domain, the dependent and independent patients both scored zero out of a possible 100. This provides evidence for a floor effect in this domain of the SF-36. Similarly, the median score in the mental health domain of the SF-36 was the same for patients who were classified as independent or recovered.

#### Test-retest reliability of the modified simple questions

In the study of the reliability of the EuroQol and SF-36 questionnaires, 443 patients returned repeat assessments of outcome.<sup>15</sup> These

Table 7 Test-retest reliability of the overall classification of outcome derived using both modified simple questions (International Stroke Trial series)

|                   | First assessment |             |           |     |
|-------------------|------------------|-------------|-----------|-----|
|                   | Dependent        | Independent | Recovered |     |
| Second assessment | Dependent        | 679         | 38        | 8   |
|                   | Independent      | 27          | 98        | 21  |
|                   | Recovered        | 9           | 11        | 129 |

$\kappa=0.76$  (0.71–0.80).

assessments included responses to the modified simple questions. A further 688 patients also returned repeat assessments of outcome which included the modified simple questions as part of the comparison of the EuroQol and SF-36 questionnaires.<sup>16</sup> Thus, 1131 patients potentially completed the simple questions at test and retest (fig 1). Test-retest reliability was very good for both modified questions and the classification of overall outcome with the modified simple questions (tables 5, 6, and 7). There was a mean delay of 30 days between completion of the initial and repeat questionnaires (SD 12 days).

## Discussion

MODIFIED DEPENDENCY QUESTION ("DO YOU NEED HELP FROM ANYBODY WITH EVERYDAY ACTIVITIES?")

We found that the modified dependency question was a valid measure of dependency in activities of daily living after stroke. It had excellent sensitivity, specificity, and accuracy for identifying dependency after stroke. Although the current study was not primarily designed to directly compare the validity of the modified dependency question with that of the original, it had a greater sensitivity than the original at identifying dependency (by Barthel index) after stroke (85% v 61%); however, its specificity was slightly worse than that of the original question (85% v 96%).<sup>3</sup> The modified dependency question was also highly accurate at identifying poor mobility and social functioning after stroke. It was, as would be predicted, less accurate at identifying patients with psychological dysfunction. These indirect comparisons suggest that, compared with the original, the modified version of the question had concurrent validity that was at least as good as the original and better face validity. We found the modified dependency question also had excellent test-retest reliability ( $\kappa=0.81$ , 95% CI 0.77–0.85); this compares very favourably with the interrater reliability of the original question ( $\kappa=0.51$ , 95% CI 0.36–0.67).<sup>6</sup> This difference may reflect the improved wording of the new question. Alternatively it may just reflect the differences between interrater and test-retest reliability or the method of questionnaire administration. Dennis *et al* employed two raters who administered the simple questions on different occasions by face to face interview,<sup>6</sup> whereas we administered the questions by post.

PROBLEMS QUESTION ("HAS THE STROKE LEFT YOU WITH ANY PROBLEMS?")

We completely reworded the recovery question because patients found the original wording ambiguous and our assessments of its validity also disclosed uncertainty about which aspect of outcome it addressed.<sup>6, 18</sup> The reworded question aims to detect those patients who are left without appreciable problems resulting from the stroke, rather than to identify patients who recover to the point of having no problems at all, whatever their cause. This alternative emphasis, although relatively subjective and non-specific, may be more relevant to both

patients and their families. The reworded question had excellent sensitivity and moderate specificity at detecting problems in all the domains examined (mobility, self care, social functioning, and psychological functioning). Indeed, this analysis may underestimate the sensitivity of the problems question as we would expect some patients to report problems on the standard measures which are not attributable to the stroke. These measurement properties differ clearly from those of the original recovery question and support the validity of the reworded question. The reworded question also had better reliability than the original ( $\kappa=0.78$  (95%CI 0.73–0.83) *v*  $\kappa=0.61$  (95%CI 0.46–0.76)).<sup>6</sup> However, it is not clear whether this improved reliability resulted from better wording or from differences in the study design. It is also not clear from the current data how patients with problems after their stroke would have responded to the question “have you made a complete recovery from your stroke?” (the original version).

#### USE OF BOTH MODIFIED QUESTIONS TOGETHER TO ASSESS HEALTH RELATED QUALITY OF LIFE

The modified simple questions may be used to classify surviving patients into one of three hierarchical outcome groups. About half of the independent patients were left with significant problems as a consequence of the stroke. This illustrates the ceiling effect with simple measures of disability that focus exclusively on activities of daily living—that is, if the dependency question was the only measure of outcome, then other significant limitations (for example, problems with communication, household maintenance, and social and psychological functioning) would not be captured. Additional assessments of health are therefore required in independent patients who might otherwise be considered as having achieved a “good outcome” after their stroke. The statistically distinct profile of scores for this group of patients with the standard instruments support the validity of this hierarchical classification.

The validity of this classification is further supported by its close relation with patients’ responses to the assessments of health related quality of life. The same pattern of responses was found, in both cohorts of patients, for assessments with the EuroQol and in the IST cohort for assessments with the SF-36. Therefore, the combined use of these modified simple questions, to classify patients as dependent, independent, and fully recovered, seems to offer a simpler approach to the assessment of global health related quality of life, than with even the EuroQol. These assessments also had very good test-retest reliability, which was of a similar order of magnitude to that found for the assessments of overall health related quality of life with the EuroQol or general health with the SF-36.<sup>15</sup> Moreover, agreement between patients and their proxies for their responses to the modified simple questions was also moderately good ( $\kappa=0.57$ ). There would be several potential advantages associated with the use of these simple

questions for the assessment of health related quality of life after stroke. Firstly, because of their brevity and simplicity they are likely to place less of a burden on patients, and so may be more feasible than complex measures. Short and simple measures, with a high frequency of response, may reduce non-response bias and enhance the interpretation and generalisability of results. Secondly, analysis of data based on a single classification of outcome by the simple questions would be simpler to both perform and interpret, than analysis of multidimensional data from a health related quality of life instrument. Thirdly, treatment effects described in terms of the proportion of patients reporting a change in response to the simple questions would have more immediate interpretability than a change in numerical score or utility. Finally, a single categorical classification of outcome by the simple questions allows the analysis to include the dead patients (dead, dependent, independent, and recovered)—dead patients could not be easily scored with the EuroQol or SF-36.

However, there are also disadvantages to replacing detailed assessments of health related quality of life with the cruder information provided by these questions. Firstly, the modified simple questions could only be used to provide a broad and global assessment of health related quality of life. Therefore, the opportunity of obtaining information about outcome in specific domains—for instance, psychological functioning—would be lost. This would reduce the opportunity to explain how a treatment improves overall health related quality of life—for example, by improving psychological functioning or by improving mobility. Secondly, the modified simple questions were developed as a disease specific instrument. They do not, therefore, provide outcome information which can be used to compare directly the effects of different treatments in groups of patients with different diagnoses. However, such comparisons could still be performed, albeit indirectly, by using the data from table 3 to map patients’ responses to the simple questions into health related quality of life utilities with the EuroQol.

#### SIMPLE VERSUS MORE COMPLEX METHODS OF MEASUREMENT

Simple measures of outcome, such as the modified simple questions, may be criticised for lacking measurement precision because they categorise stroke survivors into only three groups: dependent, independent with residual problems, and independent with no problems. This would be a reasonable point of view if the aim of the questions was to make a precise measure of outcome in an individual patient. However, these assessments were designed to estimate the difference in outcome between large groups, each comprising thousands of patients. Individual patients differ so greatly from each other that a rather crude yet valid measure can provide a precise and unbiased estimate of treatment effects in such large groups of patients.<sup>22 23</sup> Furthermore, detailed

measures can be surprisingly unreliable if studies only include a few subjects, as the effect of random variation is likely to be larger than the effect of the treatment.

Simple categorical data, of the type generated by the modified simple questions, can convey a surprisingly large amount of information. For example, if an imaginary and really accurate measurement scale has a possible range of 0 to 100, a measure which can identify those with a score of 0–50 and those scoring 51–100 would reduce the variance of non-measurement by about 75%. In other words, three quarters of the information which would have been obtained using the full 100 point scale (at great effort, time, and cost) can be obtained by a simple dichotomy. An equal three way split reduces the variance by 89% (Peto R, personal communication). Finally, more recent data has also demonstrated that patients respond significantly more often to simple measures of outcome versus longer more complex measures.<sup>14</sup>

### Conclusions

Patients completed these modified simple questions either by interview or questionnaire in the presence of the study nurse. Although the study nurse did not routinely ask patients about their interpretation of either question, no significant ambiguities were noted. These modifications therefore seem to have improved both the face and content validity as well as the reliability of these simple questions. The current study has provided strong support for the concurrent validity of the modified questions when used either individually or together to classify patients into one of three hierarchical groups which are relevant to the patients' overall health related quality of life. Proxies were able to provide a moderately accurate and unbiased assessment of the patients' outcome after stroke with these modified simple questions. Therefore, the "modified" versions of the simple questions seem to be valid, and although they have not been formally compared, they seem preferable to the original. The modified simple questions are a simple, pragmatic, and clinically relevant measure of outcome after stroke. They have a wide range of potential applications, from large formal epidemiological studies (and trials) to local audit studies.

PD was supported by a United Kingdom Medical Research Council Training Fellowship. PS was also supported by grants from the United Kingdom Medical Research Council. The study was supported by a grant from Glaxo-Wellcome plc. We would like to thank all the patients, their families, and carers for their keen participation. We are also grateful to Fiona Waddell for her collaboration.

- van Gijn J. Measurement of outcome in stroke prevention trials. *Cerebrovasc Dis* 1992;2:23–34.
- van Gijn J, Warlow C. Down with stroke scales. *Cerebrovasc Dis* 1992;2:244–6.
- Lindley RJ, Waddell F, Livingstone M, et al. Can simple questions assess outcome after stroke? *Cerebrovasc Dis* 1994;4:314–24.
- Kay R, Wong K, Yu Y, et al. Low-molecular-weight heparin for the treatment of acute ischemic stroke. *N Engl J Med* 1995;333:1588–93.
- International Stroke Trial Collaborative Group. The international stroke trial (IST): a randomised trial of aspirin, subcutaneous heparin, both or neither among 19 435 patients with acute ischaemic stroke. *Lancet* 1997;349:1569–82.
- Dennis M, Wellwood I, O'Rourke S, et al. How reliable are simple questions in assessing outcome after stroke? *Cerebrovasc Dis* 1997;7:19–21.
- Dorman PJ, Waddell FM, Slattery J, et al. Is the EuroQol a valid measure of health related quality of life after stroke? *Stroke* 1997;28:1876–82.
- Dorman PJ, Waddell FM, Slattery JM, et al. Are proxy assessments of health status after stroke with the EuroQol questionnaire feasible, accurate and unbiased? *Stroke* 1997;28:1883–7.
- The EuroQol Group. EuroQol: a new facility for the measurement of health related quality of life. *Health Policy* 1990;16:199–208.
- Wade DT, Legh-Smith J, Hewer RL. Social activities after stroke: measurement and natural history using the Frenchay activities index. *International Rehabilitation Medicine* 1985;7:176–81.
- Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361–70.
- Mahoney F, Barthel D. Functional evaluation: the Barthel index. *Maryland Medical Journal* 1965;14:61–5.
- Wellwood I, Dennis M, Warlow CP. A comparison of the Barthel index and the OPCS disability instrument used to measure outcome after acute stroke. *Age Ageing* 1995;24:54–7.
- Dorman PJ, Slattery JM, Farrell B, et al. A randomised comparison of the EuroQol and SF-36 after stroke. *BMJ* 1997;315:461.
- Dorman PJ, Slattery JM, Farrell B, et al. A qualitative comparison of the reliability of health status assessments with the EuroQol and SF-36 after stroke. *Stroke* 1998;29:63–8.
- Dorman PJ, Dennis MS, Sandercock PAG, et al. How do scores on the EuroQol relate to scores on the SF-36 after stroke? *Stroke* 1999;30:2146–51.
- O'Rourke SJ, MacHale S, Signorini D, et al. Detecting psychiatric morbidity after stroke: comparison of the GHQ and the HAD scale. *Stroke* 1998;29:980–5.
- Dennis M, Wellwood I, Warlow C. Are simple questions a valid measure of outcome after stroke? *Cerebrovasc Dis* 1997;7:22–7.
- Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ* 1992;304:1491–4.
- Bamford JM, Sandercock PAG, Warlow CP, et al. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1989;20:828.
- Rankin J. Cerebral vascular accidents in patients over the age of 60. II: prognosis. *Scott Med J* 1957;2:200–15.
- Peto R. Monitoring cancer patients in clinical trials need not be precise. In: Symington T, Williams AE, McVie JG, eds. *Cancer: assessment and monitoring: 10th Pfizer International Symposium*. Edinburgh: Churchill Livingstone, 1980:377–81.
- Tukey JW. The future of data analysis. *Annals of Mathematical Statistics* 1962;33:13–14.