

# Development and validation of a core outcome measure for palliative care: the palliative care outcome scale

Julie Hearn, Irene J Higginson, on behalf of the Palliative Care Core Audit Project Advisory Group

## Abstract

**Objectives**—To develop an outcome measure for patients with advanced cancer and their families which would cover more than either physical symptoms or quality of life related questions. To validate the measure in various specialist and non-specialist palliative care settings throughout the UK.

**Design**—A systematic literature review of measures appropriate for use in palliative care settings was conducted. In conjunction with a multidisciplinary project advisory group, questions were chosen for inclusion into the scale based on whether they measured aspects of physical, psychological, or spiritual domains pertinent to palliative care, and whether similar items had shown to be valid as part of another measure. A staff completed version was developed to facilitate data collection on all patients throughout their care, and a patient completed version was designed to enable the patient to contribute to the assessment of their outcomes when possible. A full validation study was conducted to evaluate construct validity, internal consistency, responsiveness to change over time, and test-retest reliability. Assessments were timed.

**Setting**—Eight centres in England and Scotland providing palliative care, including inpatient care, outpatient care, day care, home care, and primary care.

**Patients**—A total of 450 patients entered care during the study period. Staff collected data routinely on patients in care long enough to be assessed ( $n=337$ ). Of these, 262 were eligible for patient participation; 148 (33%) went on to complete a questionnaire.

**Main measures**—The Palliative Care Outcome Scale (POS), the European Organisation for Research on Cancer Treatment, and the Support Team Assessment Schedule.

**Results**—The POS consists of two almost identical measures, one of which is completed by staff, the other by patients. Agreement between staff and patient ratings was found to be acceptable for eight out of 10 items at the first assessment. The measure demonstrated construct validity (Spearman  $\rho = 0.43$  to  $0.80$ ). Test/re-test reliability was acceptable for seven items. Internal consistency was good (Cron-

bach's  $\alpha = 0.65$  (patients),  $0.70$  (staff)). Change over time was shown, but did not reach statistical significance. The questionnaire did not take more than 10 minutes to complete by staff or patients.

**Conclusion**—The POS has acceptable validity and reliability. It can be used to assess prospectively palliative care for patients with advanced cancer.

(Quality in Health Care 1999;8:219–227)

Keywords: hospice; palliative care; quality of life

Care for patients with advanced cancer can be provided in a wide variety of settings, including specialist palliative services such as inpatient hospices, day care, hospice at home, specialist hospital or home support teams, as well as primary and secondary care.<sup>1</sup> Increasingly, elderly patients with advanced cancer are also being cared for in nursing and residential homes. The specific aim of palliative care is to ensure that the patient has the highest possible quality of life until they die, and to provide support for their nearest family members or carers.<sup>2,3</sup>

Quality of life and outcome measures for patients with advanced illness need to be able to assess the key goals of palliative care. Various outcome measures and systems for evaluating palliative care have been developed in recent years.<sup>4</sup> These include the Edmonton Symptom Assessment Schedule (ESAS),<sup>5</sup> the European Organisation for Research on Cancer Treatment Quality of Life Questionnaire (EORTC QLQ-C30),<sup>6</sup> McGill Quality of Life Questionnaire (MQOL),<sup>7</sup> the Palliative Care Assessment (PACA),<sup>8</sup> and the Support Team Assessment Schedule (STAS).<sup>9</sup> Measures such as these are constantly being amended and updated, and are not consistently revalidated for individual settings.<sup>4</sup> In addition, none of these systematically covers all those domains considered important to palliative care.<sup>2</sup> Patients with advanced cancer require palliative care, where the emphasis is not only on the relief of pain and symptoms but also on the resolution of emotional, social, psychological, and spiritual problems; the provision of information; good communication; and support for the family.<sup>2</sup>

It could be argued that a battery of measures would encompass all aspects of care. Unfortunately, such an approach has failed in the past because the patients were unable to complete lengthy forms.<sup>10</sup> Patients' and staffs' views and assessments normally systematically and

Department of Palliative Care and Policy, King's College School of Medicine and Dentistry and St Christopher's Hospice, New Medical School, Bessemer Road, London SE1 9PJ, UK  
J Hearn, *honorary research fellow*  
I J Higginson, *professor*

Correspondence to:  
Professor Higginson

Accepted 4 October 1999

randomly differ.<sup>11</sup> Patient completed measures are important in evaluating the impact of health care,<sup>12</sup> but due to the very ill nature of the palliative population, patients are often unable to fill in forms (this is discussed in more detail elsewhere<sup>13</sup>). Capturing both views where possible in the palliative care population is therefore important. It also allows us to compare, where collected, patient and staff assessments and thereby improve the accuracy of staff evaluation through feedback and training.

The aim of this study was to develop a core outcome measure for patients with advanced cancer and their families which would cover more than either physical symptoms or quality of life, but remain brief and simple to administer. In the future, additional items could be developed for use in a particular setting, or to assess a particular domain further. This paper describes the validation study that was subsequently conducted in various specialist and non-specialist palliative care settings. In developing the measure, a staff rated and a patient rated component were agreed upon. Hence, this study also aimed to test the level of agreement between patients' self ratings and staff assessments.

## Methods

### SUBJECTS

The validation study included all new referrals to a purposive sample of eight services providing palliative care: St Christopher's Hospice, two of four inpatient wards and all the home care teams; King's College Hospital palliative care team, inpatient hospital service; the Leicestershire Hospice home care team; the Liverpool Marie Curie Centre, inpatient and outpatient services; Peace Hospice, day care service; James Paget Hospital palliative care team, inpatient, outpatient, and community services; Highland Hospice inpatient service; and Suffolk Community palliative care services. Patients were recruited in 1997 over a three month period, or until a pre-defined target number of patients had been accrued, depending on the referral rate to individual study centres. The anticipated sample size was 450 patients recruited into the study, of whom one third would be expected to be able to complete 1 or more questionnaires themselves.<sup>7</sup> At the initial site visit the study was presented to staff in the individual centres as a validation study of a new measure.

### ELIGIBILITY AND CONSENT

Local Research Ethics Committee approval for the study was obtained for all eight centres. Staff assessments were done on all patients in care for long enough before discharge or death, as if an assessment was part of routine data collection at first point of contact. Patients were considered to be eligible for patient participation if they fulfilled all of the following inclusion criteria:

- Had a good command of the English language

- Did not have an impaired mental status, either as previously assessed or as judged by a member of staff
- Were considered to be physically able to complete a questionnaire.

Patients were asked to complete a questionnaire upon entry to the service and at two timepoints thereafter. Questions asked about the patients' wellbeing over the previous three days. The exact timing of successive assessments was dependent on the nature of the service providing care and when they would be expected to have an impact on the outcomes of the patient. For example, patients receiving inpatient care were assessed every three days, and patients attending outpatient clinics or receiving home care were assessed at each contact (usually weekly).

### MEASURES

The Palliative Care Outcome Scale (POS) was developed using data from a review of other outcome measures used, or proposed for use in evaluating the palliative care of patients with advanced cancer.<sup>4</sup> Ten questions were chosen for inclusion based on those reliable and valid questions used in other measures which covered the physical, psychological, and spiritual domains of life within the remit of palliative care. In addition, a space was provided to list "main problems". There is a staff completed version of the questionnaire, and a patient completed version for use by those patients able to fill in the form. Data can therefore be gathered on all patients by staff throughout the course of a patient's care, avoiding the problem of incomplete data if using a solely patient completed outcome measure.<sup>4</sup> Patients can provide, however, additional input to the measurement of their outcomes. The staff questionnaire includes an additional question on performance status.<sup>14</sup>

The questionnaire was devised in collaboration with a multidisciplinary project advisory group, which consisted of representatives from the main palliative care organisations in the UK, plus a patient representative and local researchers involved in a similar project. The original questionnaire was piloted on 25 patients from five of the study centres, and revised accordingly to yield the final version (see appendix).

The measure was designed to provide information at item level and therefore provide an item profile for use in clinical practice, audit, and research.

### VALIDITY AND RELIABILITY

The validity of the measure was tested by evaluating the various components of validity and reliability described below.

#### *Content and consensus validity*

These aspects tested whether the measure could: (1) yield information of clinical relevance to palliative care, (2) cover those domains considered important to this type of care and nothing more, and (3) achieve a consensus of agreement among specialists that (1) and (2) had been met. These three components

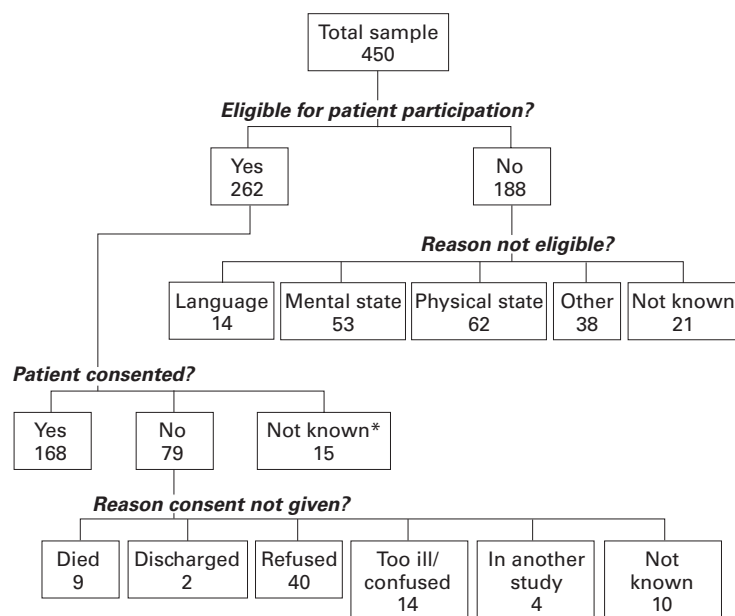


Figure 1 The eligibility and consent of patients entering the study. \*There was no record of consent given on the study centres' registers. Hence, patients may not have given consent or were not asked for their consent by staff.

of validity were determined from an initial literature review, through expert discussion within the project advisory group, and systematic feedback from staff involved in the study. The literature review was presented to the project advisory group and three versions of the questionnaire were devised before the draft pilot version. Staff participating in the study were asked to complete questionnaires before and after the study to assess the acceptability of using such a measure for audit purposes.

#### Face validity: patients' views

Patients attending one day care centre over a three day period towards the end of the study were asked to complete a questionnaire. In-depth interviews were then conducted to explore whether the questionnaire was tiring to the patient, if any questions were confusing, upsetting, not relevant, or felt to be missing, and whether the patient believed the information obtained would be of use to staff. The responses were transcribed and a content analysis conducted.

#### Construct validity

The extent to which the new measure conforms to the results obtained using other established scales was evaluated. There are currently no gold standard measures which could be used for comparison, hence the most established and validated scales were selected from our literature review (one patient completed scale and one staff completed scale).<sup>4</sup> Each centre was asked to select 10 patients at random to complete the EORTC QLQ-C30<sup>6</sup> at the same time as the patient rated POS. Staff were also asked to complete the STAS<sup>9</sup> on 10 patients at random in each centre directly after completing the staff rated POS. As a result of recruitment difficulties, the final numbers of EORTC QLQ-C30 and STAS questionnaires completed was less than expected. Because the

measures were not viewed as gold standards, items were grouped into the subscales where correlation between the measures should be expected: (1) physical symptoms; (2) all problems excluding the quality of life/existential domains; (3) quality of life.

#### Comparison of staff and patient ratings

The agreement between staff and patient scores at all three timepoints was tested by using the matched data for patients who had completed an assessment themselves and who had had a staff assessment completed on them at the same time.

#### Internal consistency

Internal consistency was measured using data from questionnaires completed on entry to the study. Although the measure was designed for use at item level, this was conducted to determine the overall consistency of approach of the scale.

#### Test/re-test reliability

A separate group of patients who we anticipated would be a relatively stable population—that is, who would not “change” substantially over time—were recruited from two palliative day care centres to conduct the test/re-test evaluation. The inclusion criteria were that the patients had been receiving palliative day care for over four weeks and had a prognosis of over one year. Questionnaires were completed by patients and staff on two consecutive visits. Any apparent change in the patient's physical or psychological status was recorded by staff so that this could be accounted for in the analysis.

#### Responsiveness to change

This terminally ill patient population would be expected to change rapidly over time, and therefore between consecutive assessments of newly referred patients. Comparisons were made between assessments at timepoints one and two, and one and three for both the patient completed and staff completed questionnaires.

#### Time to complete

The completion of questionnaires was timed by the member of staff completing the questionnaire or staff member who had provided the patient with a questionnaire. This was done at all three time points for 10% of assessments.

#### STATISTICAL ANALYSIS

Patient and staff ratings were compared using tests for agreement and correlation. Agreement is the ideal requirement for validity, but correlation can indicate whether one rater reflects another's view. There were two tests of agreement: (1) the proportion of cases where the staff member's rating was equal to the patient's rating, or was within +1 or -1 of the score; and (2) Cohen's weighted  $\kappa$  which measures agreement but controls for chance agreement. An acceptable level of  $\kappa$  was set in advance at  $\kappa = 0.3$  or above, as used in the development of the Spitzer quality of life index.<sup>15</sup>  $\kappa$  should be used with caution when data show a highly skewed distribution or when

Table 1 Distribution of the total population, the number eligible for patient participation for each centre, and the sample able and willing to complete at least one questionnaire

Centre	Total number of patients (% of total)	Number eligible (% of centre sample)	Number completing 1 or more questionnaires (% of centre sample)
St Christopher's	178 (40)	88 (49)	56 (32)
Leicestershire Hospice	99 (22)	76 (76)	22 (22)
King's palliative care team	50 (11)	15 (30)	7 (14)
Highland Hospice	35 (8)	9 (26)	9 (26)
James Paget Hospital	34 (8)	25 (74)	18 (53)
Liverpool Marie Curie	28 (6)	25 (89)	20 (71)
Peace Hospice	20 (4)	19 (95)	12 (60)
Suffolk Community	6 (1)	6 (100)	4 (67)
Total	450 (100)	263 (58)	148 (33)

most of the ratings are the same, as this increases the probability of chance agreement. Spearman correlation were calculated to test for association between the staff and patient ratings, and when comparing the POS to the EORTC QLQ-C30 and the STAS measures. Owing to the different scoring systems used, the data from the three scales were transformed using a linear transformation to facilitate comparison. Cronbach's  $\alpha$  was used to test for internal consistency. The Wilcoxon matched

Table 2 Patient demographic characteristics for the total sample (n=450) and those patients who completed at least 1 questionnaire (n=148)

	Number of patients (%)	Number of patients completing 1 or more questionnaires (%)
Age		
<65 years	130 (30)	47 (32)
≥65 years	299 (70)	98 (68)
Sex		
Men	216 (49)	66 (45)
Women	227 (51)	82 (55)
Ethnic origin		
White UK	384 (89.1)	138 (93.8)
Black-Caribbean	18 (4.2)	1 (0.7)
Black-African	4 (0.9)	1 (0.7)
Indian	6 (1.4)	1 (0.7)
Pakistani	4 (0.9)	1 (0.7)
Bangladeshi	1 (0.2)	1 (0.7)
Chinese	1 (0.2)	1 (0.7)
other/not known	13 (3.0)	3 (2.0)
Total:	431	147
Marital status		
Married	210 (50.7)	70 (47.6)
Widowed	116 (28.0)	46 (31.3)
Divorced/separated	28 (6.8)	9 (6.1)
Single	39 (9.4)	17 (11.6)
Partnered	7 (1.7)	2 (1.4)
Other/not known	14 (3.4)	3 (2.0)
Total:	414	147
Diagnosis (ICD-10 categories)		
Digestive organs	124 (28.7)	44 (29.7)
Respiratory tract	86 (19.9)	26 (17.6)
Genitourinary tract*	77 (17.8)	33 (22.3)
Breast	51 (11.8)	20 (13.5)
Lymph/haematopoietic	13 (3.0)	3 (2.0)
Other cancers	70 (16.2)	20 (13.5)
HIV/AIDS	2 (0.5)	0
Stroke, MND	9 (2.0)	2 (1.4)
Total:	432	148
Place of care at entry to study		
Home	303 (70.0)	105 (72.4)
Hospital	121 (27.9)	38 (26.2)
Hospice	3 (0.7)	2 (1.2)
Other	6 (1.4)	0
Total:	433	145
Performance status (ECOG)		
Fully active	23 (7.3)	14 (10.3)
Restricted	65 (20.7)	36 (26.5)
Ambulatory	46 (14.6)	22 (16.2)
Limited ability	131 (41.7)	56 (41.2)
Disabled	49 (15.6)	8 (5.9)
Total:	314	136

\*Cancers of the male and female genitals and the urinary tract were combined to form this category. Not all subtotals add up to the original sample totals due to a small amount of missing data for each variable; missing data are data not recorded and not obtainable from patient records, data which were not known refers to data the member of staff were unable to record or ascertain.

pairs signed rank test was used to test for the difference between observations on the same patient over time. In some instances the percentages of severe scores are presented to illustrate changes in outcomes over time. A probability of  $p < 0.05$  (two tailed) was taken as significant for the above tests.

## Results

### SUBJECT CHARACTERISTICS

A total of 450 patients were recruited. Of these, 262 (58%) were eligible for patient participation, of whom 168 (64%) agreed to complete a patient questionnaire (fig 1).

Staff were unable to complete assessments on a patient if the person died within hours of admission to the service. In total, 337 (75%) of the sample had one or more staff assessment carried out on them, and 148 (33%) were in care long enough to complete one or more patient assessment. Eligibility for completion of a patient questionnaire was lowest in the hospital palliative care team and in the hospice inpatient settings, and highest in day care and general practice (table 1).

Those patients completing one or more questionnaires had higher levels of mobility (as measured by the ECOG score) compared with the total population (table 2). There were fewer patients from ethnic minority groups able to complete 1 or more questionnaires. This may have been the result of the majority of patients having been cared for at King's College Hospital, a setting where few patients were able to complete questionnaires themselves (table 1). There were no other major demographic differences among patients able to complete questionnaires.

### VALIDITY AND RELIABILITY

#### Content and consensus validity

In addition to the input from the project advisory group during the development of the questionnaire, staff doing the validation study reported that the measure was useful as a basic tool for measuring outcomes in palliative care, covering more than simply the physical problems experienced by patients with advanced disease. The extent to which this held true was dependent on how much research and evaluation had previously been done by staff in the individual centres. Those who had never used an outcome measure before felt it gave them the opportunity, under the guise of research, to introduce and discuss what they felt were the more difficult existential domains and social aspects of care with their patients. In one setting there were concerns raised about the question on whether "life was worthwhile". Staff found this question difficult to ask, and believed that patients found it difficult to answer. In two other settings this question was welcomed and found to be helpful in initiating discussions.

#### Face validity

12 patients were interviewed, none of whom found completing the questionnaire tiring or any questions confusing. One patient found the questions relating to family members upsetting. One patient felt the question on wasted

Table 3 Criterion validity: comparison of the POS with the EORTC QLQ-C30 and the STAS

Subscale	EORTC QLQ-C30 v patient POS (n=29) (95% CI)		STAS v staff POS (n=43) (95% CI)	
	Spearman's rho	p Value	Spearman's rho	p Value
Physical symptoms	0.51 (0.18 to 0.74)	0.005	0.80 (0.66 to 0.89)	0.000
All non-quality of life problems	0.53 (0.20 to 0.75)	0.003	0.67 (0.46 to 0.81)	0.000
Quality of life	0.43 (0.08 to 0.69)	0.022	0.51 (0.25 to 0.70)	0.001

time was irrelevant to their own circumstances. None of those interviewed reported feeling that there were any questions missing. In terms of usefulness of the information to staff, nine patients believed it would help staff to know more about how they were currently feeling. Two of those who did not feel it would be of use expanded their answers by saying that their care was well under control by staff anyway.

#### Construct validity

Twenty nine patients completed both a POS and EORTC QLQ-C30 questionnaires, and 43 staff completed a STAS assessment in addition to the staff POS. The correlations ranged from 0.43 to 0.80 (table 3).

#### Comparison of staff and patient ratings

Table 4 details the percentage of answers scored as severe by the patients and staff (scores 3 and 4), plus measures of agreement and correlation between the two people recording information. The level of agreement, as measured by Cohen's  $\kappa$ , is acceptable (greater than 0.3) for eight out of 10 of the items at the first assessment (all except family anxiety and support), and for all items by the second assessment. The correlation coefficients illustrate reasonable correlation between

Table 4 Agreement between scores for staff and patient assessments

Item	No of patients	Patient score (% severe)	Staff score (% severe)	K (weighted)	Spearman correlation	Proportion agreement within 1 score
At first assessment: 145 matched assessments						
Pain	140	24.3	20.0	0.56	0.67	0.87
Other symptoms	140	27.2	26.4	0.43	0.60	0.86
Patient anxiety	140	23.6	30.0	0.37	0.56	0.83
Family anxiety	137	49.6	46.0	0.28	0.37	0.72
Information	135	12.6	13.4	0.390	0.36	0.79
Support	135	10.4	14.1	0.22	0.32	0.79
Life worthwhile	133	13.6	16.5	0.43	0.54	0.82
Self worth	132	15.9	23.5	0.37	0.53	0.82
Wasted time	135	5.9	6.7	0.33	0.32	0.95
Personal affairs	129	7.8	13.2	0.42	0.49	0.96
Second assessment: 97 matched assessments						
Pain	94	8.5	10.7	0.58	0.68	0.91
Other symptoms	94	13.8	17.1	0.49	0.61	0.89
Patient anxiety	93	10.7	20.4	0.42	0.57	0.86
Family anxiety	92	32.6	31.6	0.44	0.59	0.84
Information	94	13.8	14.9	0.56	0.55	0.97
Support	94	13.9	11.7	0.35	0.48	0.74
Life worthwhile	93	6.5	5.4	0.38	0.41	0.87
Self worth	91	11.0	8.8	0.32	0.36	0.86
Wasted time	92	2.2	2.2	0.54	0.59	0.99
Personal affairs	92	5.4	5.4	0.42	0.42	0.98
At third assessment: 66 matched assessments						
Pain	64	9.4	7.8	0.50	0.59	0.94
Other symptoms	64	14.4	7.9	0.32	0.52	0.83
Patient anxiety	64	9.4	12.5	0.28	0.38	0.81
Family anxiety	62	40.3	30.7	0.58	0.69	0.85
Information	62	14.5	9.7	0.39	0.35	0.87
Support	65	15.4	15.4	0.48	0.62	0.83
Life worthwhile	64	7.8	7.8	0.46	0.49	0.89
Self worth	63	12.6	9.5	0.42	0.47	0.84
Wasted time	61	3.3	1.6	0.35	0.44	0.97
Personal affairs	62	6.5	1.6	0.29	0.27	0.95

Table 5 Test/re-test reliability results for additional set of day care patients (n=34)

Item	K	Proportion agreement within one score
Pain	0.29	0.88
Other symptoms	0.10	0.74
Patient anxiety	0.43	0.82
Family anxiety	0.40	0.82
Information	0.44	0.82
Support	0.62	0.85
Life worthwhile	0.37	0.91
Self worth	0.40	0.88
Wasted time	0.53	1.00
Personal affairs	-0.08	0.97

staff and patients, for most of the items at most timepoints. Interestingly, the proportion of scores that are within one score of the ideal situation (when scores are perfectly matched) was very high throughout, indicating that the staff's perception of the patient's situation was close to the patient's report, if not always exact.

#### Test/re-test reliability

The 34 patients included in the test/re-test reliability scored 0 or 1 for most items. Weighted  $\kappa$  for three items—pain, other symptoms, and personal affairs—showed low values, possibly due to the uneven distribution of answers. However, the proportion agreement within one score between assessments for the questionnaire items was very good, ranging from 0.74 to 1 (table 5). Staff noted some degree of change in four patients since their last visit to day care, and these were therefore excluded from the analyses.

#### Internal consistency

The  $\alpha$  reliability coefficients were 0.65 (n=128), and 0.70 (n=308) for the patient rated scale and the staff rated scale, respectively.

#### Responsiveness to change

Figures 2 and 3 give the results of the changes in severe scores over time for patient and staff assessments, respectively. The data presented are matched data for those 66 patients who completed three self assessments and had all three staff assessments completed on them. The figures illustrate that by the second assessment improvements were seen for all but the item "support" rated by the patient, and the item "personal affairs" rated by the staff. The improvements were only statistically significant ( $p < 0.05$ ), however, for the patient rated pain and family anxiety, and the staff rated pain. Some items, particularly "wasted time" and "personal affairs", were rarely severe and there appeared to be some floor effects.

When the data for the total of 248 patients who had two staff assessments were considered, all of the staff rated questionnaire items except for the item support showed statistically significant improvements. The scores at first assessment suggest that those patients able to complete questionnaires at all three timepoints were less ill than those for whom one or possibly two assessments could be made. Of the original 148 patients who completed a first questionnaire, seven out of 10 items were rated

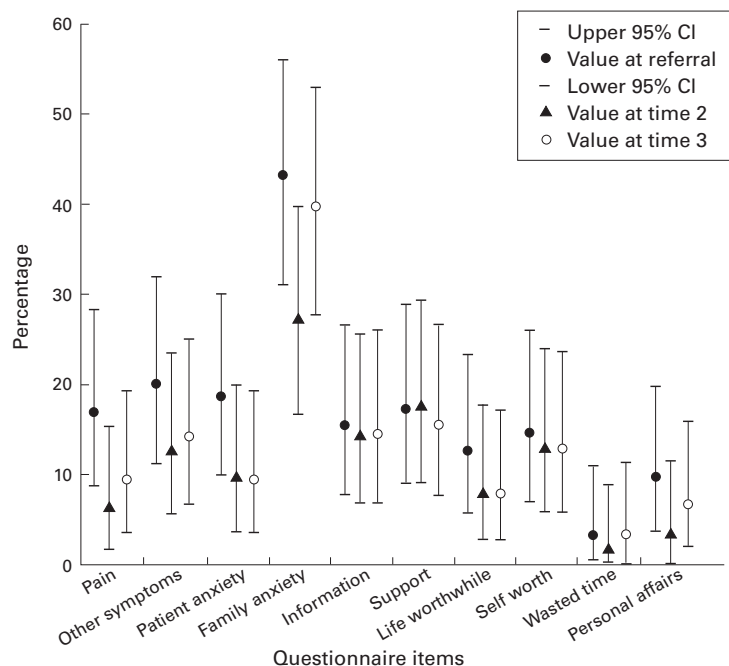


Figure 2 Percentage of scores rated severe (3 or 4) by patients at three timepoints (95% confidence intervals (CI)) for patients completing all three assessments (n=66).

severe (score 3 or 4) by more than 20% of patients. It is possible that those with less problematic symptoms remained in the study for longer.<sup>16</sup>

*Time to complete*

No timed assessment by patient or staff took longer than 10 minutes. Patients took longer than staff, but the time to complete decreased over successive assessments. The mean time to complete a questionnaire at the first assessment was 6.9 minutes (patients), and 5.7 minutes (staff). By the third assessment both patients and staff had a mean time to complete of less than four minutes.

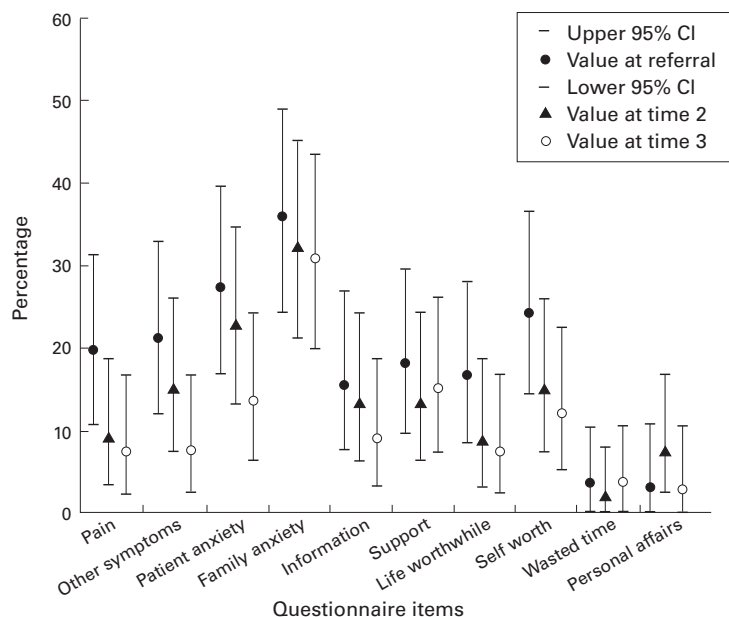


Figure 3 Percentage of scores rated severe (3 or 4) by staff over time (95% confidence intervals (CI)) for patients completing all three assessments (n=66).

*Open sections of the questionnaire*

These were completed by both patients and staff in most instances, although staff provided more information. Main problems related to symptoms, particularly breathlessness, weakness, feeling tired, sick or not eating, or to concerns about finances, future planning, and members of the family. These were reflected in the POS ratings.

**Discussion**

Despite the difficulties inherent in undertaking prospective research in palliative care,<sup>17</sup> this multicentre study provides comprehensive data on the validity and reliability of a new core outcome measure for palliative care. The POS was developed using components of other measures which appear to have worked well, and includes aspects about pain and symptom control, patient and family psychosocial needs, and communication and information.

Except when testing for construct validity against other measures, all data presented illustrate the validity of individual items. We have developed a measure that assesses the key domains important in the provision of palliative care, yet is not intended to provide only a summary score which may obscure the multidimensional nature of problems which an individual faces.<sup>18</sup> However, the internal consistency shown using Cronbach's  $\alpha$  indicated some consistency of POS as a summary scale. The questionnaire also provides room to list the main problems in the previous three days. This open ended question gives the patient the opportunity to emphasise the importance of a problem, or indicate what else is bothering them; information which may assist clinical care. The high concentration of symptom reports suggest that future development of POS should incorporate some more symptom information, notably breathlessness, weakness, and nausea. Symptom extensions are already available for the STAS.<sup>19</sup>

From the limited data available on the acceptability of the measure to patients, it would appear that the measure is acceptable to those patients completing it in day care, the majority of whom believed it would be of clinical use to staff. Note, however, that this group of patients were likely to have had fewer severe problems than the majority of patients who would complete the questionnaire. Nevertheless, this group and the patient representative on the project advisory group provided some initial patient evaluation.

Of the 450 patients entering care during the study period, only one third actually completed one or more patient questionnaires. This number is disappointing but similar to that found in other studies.<sup>7</sup> Those patients willing and able to fill in the questionnaires were drawn from various settings providing care for patients with advanced disease, including inpatient hospice care, inpatient hospital care, hospice outpatient care, hospice day care, and home care. This is a strength of the measure in that it appears that both staff and patients are able to complete it across settings, a factor that has not been shown by any other outcome

measure in palliative care.<sup>4</sup> Eligibility of patients was lowest in hospitals and hospices, and in those patients with lower mobility. Unfortunately, there was insufficient data from each setting to allow re-analysis to verify whether the measure meets all the criteria for validity and reliability in each setting separately. This type of analysis will be the focus of future work.

The study relied upon local personnel to determine whether patients were eligible or not to participate. One third of patients were reported not to have met the pre-set eligibility criteria and therefore only had staff assessments done on them as part of routine practice. Although it was essential to have fixed criteria to try to obtain a representative sample from each centre, the criteria were open to interpretation. Hence, it is possible that some patients were not recruited who could have been, as a result of staff acting as the “gate keepers” of patients. The proportion eligible in each centre varied greatly. Moreover, the process of obtaining consent was felt by some staff to be prohibitive to asking the patient to complete the questionnaire. This factor could be overcome in routine practice by including the questionnaire as part of the initial assessment, enabling more patients to contribute to the evaluation of their care.

The data presented show that the POS has demonstrated validity and reliability in the majority of items for most constructs investigated. Some of the items, notably “wasted time”, “personal affairs”, and “life worthwhile” need review, and there is scope to include more items. Test-retest reliability was shown for most items but was difficult to assess because of (1) difficulty in identifying a stable population and (2) many of the ratings being the same (0—no problem) in this population, which limited our ability to test across a range of scores. We acknowledge that to achieve this validation, multiple significance tests were done. Hence, it is possible that some of the significant values arose due to chance alone. Future work could attempt to expand and replicate these findings.

We believe that the POS has potential as a core outcome measure to assess prospectively the palliative care of patients. Moreover, the conceptual and methodological issues pertinent to outcome measurement in this field have been documented, issues that were recently highlighted as lacking in the presentation of quality of life measures.<sup>18</sup> We hope to develop a training package to accompany the tool, which would help staff to make consistent assessments of patients and their families, and contain suggestions of how to present the forms to patients for completion, based on the experiences of those conducting the study.

There is an inherent difficulty in using self completed measures to assess outcomes in a palliative care population as many patients are too ill to complete them, or die early during care.<sup>7</sup> Hence, this measure fills the need for a tool which has a valid staff completed component, but which can be enhanced by the use of patient input whenever possible.

The use of outcome measures can help to determine whether a method of treatment or a particular intervention package is worthwhile.<sup>12–20</sup> Outcome measures are also valuable in assessing and auditing care.<sup>21</sup> By providing a core measure for use in palliative care to which additional, validated items could be added in the future, we may be able to begin the process of determining which interventions or packages of care work best for patients with a particular set of problems associated with their advanced disease.

We would like to thank the NHS Executive clinical audit unit for providing funding for this study, and in particular Dr Robert Dunlop and the members of the project advisory group for their continued support over the past two years. We would also like to thank the staff in the study centres for their tireless efforts in collecting data, and most importantly, all those patients who gave their time and energy to complete the questionnaires. Our thanks also to Yetunde Adeyina who undertook the indepth interviews and assisted in their analysis.

The Palliative Care Core Audit Project Advisory Group members are: I J Higginson, J Hearn, R Dunlop, B Biswas, A Eastaugh, J Ellershaw, R George, P McDaid, B Miles, K Nobes, H Pickles, H Smith, J Smith, S Thompson, E Wilson.

## Appendix

### The POS staff questionnaire

Patient Name:

Unique Number:

Care Setting:  
Date:

Date of Birth:  
Assessment No:

Please answer the following questions by ticking the box next to the answer which you think most accurately describes how the patient has been feeling. Thank you.

- (1) Over the past 3 days, has the patient been affected by pain?  
0 Not at all, no effect  
1 Slightly—but not bothered to be rid of it  
2 Moderately—pain limits some activity  
3 Severely—activities or concentration markedly affected  
4 Overwhelmingly—unable to think of anything else
- (2) Over the past 3 days, have other symptoms e.g. nausea, coughing or constipation seemed to be affecting how well they feel?  
0 No, not at all  
1 Slightly  
2 Moderately  
3 Severely  
4 Overwhelmingly
- (3) Over the past 3 days, have they been feeling anxious or worried about their illness or treatment?  
0 No, not at all  
1 Occasionally  
2 Sometimes—affects their concentration now and then  
3 Most of the time—often affects their concentration  
4 Patient does not seem to think of anything else—completely preoccupied by worry and anxiety
- (4) Over the past 3 days, have any of their family or friends been anxious or worried about the patient?  
0 No, not at all  
1 Occasionally  
2 Sometimes—it seems to affect their concentration  
3 Most of the time  
4 Yes, they always seem preoccupied with worry
- (5) Over the past 3 days, how much information has been given to the patient and their family or friends?  
0 Full information—patient feels free to ask  
1 Information given but not always understood by patient  
2 Information given to patient on request—patient would have liked more  
3 Very little given and some questions have been avoided  
4 None at all
- (6) Over the past 3 days, has the patient been able to share how they are feeling with family or friends?  
0 Yes, as much as they wanted to  
1 Most of the time  
2 Sometimes  
3 Occasionally  
4 No, not at all with anyone
- (7) Over the past 3 days, do you think they have felt life was worth living?  
0 Yes, all the time  
1 Most of the time  
2 Sometimes  
3 Occasionally  
4 No, not at all
- (8) Over the past 3 days, do you think they have felt good about themselves?  
0 Yes, all the time  
1 Most of the time  
2 Sometimes  
3 Occasionally  
4 No, not at all
- (9) Over the past 3 days, how much time do you feel has been wasted on appointments relating to the healthcare of this patient, e.g. waiting around for transport or repeating tests?  
0 None at all  
2 Up to half a day wasted  
4 More than half a day wasted
- (10) Over the past 3 days, have any practical matters resulting from their illness, either financial or personal, been addressed?  
0 Practical problems have been addressed and their affairs are as up to date as they would wish  
2 Practical problems are in the process of being addressed  
4 Practical problems exist which were not addressed  
0 The patient has had no practical problems
- (11) If any, what have been the patient's main problems in the last 3 days?  
 1.  
 2.
- (12) What is the patient's ECOG scale performance status?  
 (0—fully active; 1—restricted; 2—ambulatory; 3—limited self care; 4—completely disabled)



- 1 Eve A, Smith AM. Palliative care services in Britain and Ireland: update 1991. *Pall Med* 1994;**8**:19–27.
- 2 Standing Medical Advisory Committee. *The principles and provision of palliative care*. London: Joint report of the Standing Medical Advisory Committee and Standing Nursing and Midwifery Advisory Committee, 1992.
- 3 Higginson IJ, McCarthy M. Validity of the support team assessment schedule: do staff's ratings reflect those made by patients or their families? *Pall Med* 1993;**7**:219–28.
- 4 Hearn J, Higginson IJ. Outcome measures in palliative care for advanced cancer patients: a review. *J Public Health Med* 1997;**19**:193–9.
- 5 Bruera E, Kuehn N, Miller MJ, *et al*. The Edmonton symptom assessment system (ESAS): a simple method for the assessment of palliative care patients. *J Pall Care* 1991;**7**:6–9.
- 6 Aaronson NK, Ahmedzai S, Bergman B, *et al*. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;**85**:365–76.
- 7 Cohen SR, Mount BM, Strobel MG, *et al*. The McGill quality of life questionnaire: a measure of quality of life appropriate for people with advanced disease. A preliminary study of validity and acceptability. *Pall Med* 1995;**9**:207–19.
- 8 Ellershaw JE, Peat SJ, Boys LC. Assessing the effectiveness of a hospital palliative care team. *Pall Med* 1995;**9**:145–52.
- 9 Higginson I. A community schedule. In: Higginson I, editor. *Clinical audit in palliative care*. Oxford: Radcliffe Medical Press 1993:34–7.
- 10 McWhinney IR, Bass MJ, Donner A. Evaluation of a palliative care service: problems and pitfalls. *BMJ* 1995;**309**:1340–2.
- 11 Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. *J Clin Epidemiol* 1992;**45**:743–60.
- 12 Kind P, Dolan P, Gudex C, *et al*. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* 1998;**316**:736–41.
- 13 Higginson IJ. Clinical and organisational audit in palliative care. In: Doyle D, Hanks GWC, MacDonald N, editors. *Oxford textbook of palliative medicine*. Second edition. Oxford: Oxford University Press 1998;**2**:67–81.
- 14 Eastern Cooperative Oncology Group. Eastern cooperative oncology group performance status scale. *Am J Clin Oncol* 1982;**5**:649–55.
- 15 Spitzer WO, Dobson AJ, Hall J, *et al*. Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. *J Chronic Dis* 1981;**34**:585–97.
- 16 Axelsson B, Sjoden PO. Quality of life of cancer patients and their spouses in palliative home care. *Pall Med* 1998;**12**:29–39.
- 17 Rinck GC, van den Bos GAM, Kleijnen J, *et al*. Methodologic issues in palliative cancer care: a systematic review. *J Clin Oncol* 1997;**15**:1697–707.
- 18 Muldoon MF, Barger SD, Flory JD, *et al*. What are quality of life measurements measuring? *BMJ* 1998;**316**:542–5.
- 19 Edmonds PM, Stuttaford JM, Penny J, *et al*. Do hospital palliative care teams improve symptom control? Use of a modified STAS as an evaluation tool. *Pall Med* 1998;**12**:345–51.
- 20 Bowling A. *Measuring disease: a review of disease-specific quality of life measurement scales*. Milton Keynes: Open University Press, 1994.
- 21 Higginson IJ. *Clinical audit in palliative care*. Oxford: Radcliffe Medical Press, 1993.

Cartoon used with permission of the artist, Ross Scriven.



"He's an evidence-based parrot. He memorises the Cochrane Library and I take him on ward rounds."