

**Original Article**

# Validation and Clinical Application of the German Version of the Palliative Care Outcome Scale

Claudia Bausewein, MD, MSc, Martin Fegg, PhD, Lukas Radbruch, MD, Friedemann Nauck, MD, Silvia von Mackensen, PhD, Gian Domenico Borasio, MD, and Irene J. Higginson, BmedSci, BMBS, FFPHM, PhD

*Interdisciplinary Center for Palliative Medicine (C.B., M.F., G.D.B.), University of Munich, Munich, Germany; Department of Palliative Medicine (L.R.), University of Aachen, Aachen, Germany; Center for Palliative Medicine (F.N.), Malteser Hospital Bonn, Rheinische-Friedrich-Wilhelms University, Bonn, Germany; Institute of Medical Psychology (S.v.M.), Center of Psychosocial Medicine, University Hospital Hamburg-Eppendorf, Hamburg, Germany; and Department of Palliative Care and Policy (I.J.H.), King's College London, London, United Kingdom*

---

**Abstract**

*The Palliative Care Outcome Scale (POS) is a multidimensional instrument covering physical, psychosocial, spiritual, organizational, and practical concerns. This study validated the German version of the POS and used the tool in different palliative care settings in Germany and Austria. Patients and staff were asked to complete the POS three times and evaluate the questionnaire and the translation afterwards. One hundred eighteen patients (44 male, 74 female, mean age 63 years, all suffering from advanced cancer) completed the POS one time, 55 patients two times, and 36 patients three times. Spearman's rho was highly significant for pain, other symptoms, anxiety, and life worthwhile in the first two assessments. The third assessment showed significant correlations for pain, other symptoms, anxiety, and family anxiety. Seventy-seven of 87 patients answered questions regarding the scale and the translation. All questions other than "Over the past 3 days, have you felt good about yourself?" were understandable for patients. Almost half of the staff was undecided whether the tool was reflecting the patients' condition. In contrast, the majority of patients liked it. Thus, the German version of the POS is well accepted by patients and staff and appears to be valid, although there are some areas where the scale would benefit from expansion to more closely capture staff and patient concerns. J Pain Symptom Manage 2005;30:51–62. © 2005 U.S. Cancer Pain Relief Committee. Published by Elsevier Inc. All rights reserved.*

**Key Words**

*Palliative care, outcome measures, quality of life, German*

---

*Address reprint requests to: Claudia Bausewein, MD, MSc, Interdisciplinary Center for Palliative Medicine, Munich University Hospital–Grosshadern, Marchioninistrasse 15, 81377 Munich, Germany.*

*Accepted for publication: January 17, 2005.*

**Introduction**

Hearn and Higginson developed the Palliative Care Outcome Scale (POS) as a multidimensional outcome measure for patients with advanced cancer.<sup>1</sup> The scale covers physical and

psychological symptoms, spiritual considerations, practical concerns, emotional concerns of the patient and family, and psychosocial needs of the patient and family. The POS demonstrated construct validity in comparison to the EORTC QLQ-C30 and the Support Team Assessment Schedule (STAS), and good internal consistency.<sup>1</sup> The POS has been used as a clinical tool, in audit and research, and as a teaching aid.<sup>2</sup>

Palliative care in Germany is a young but growing specialty. Questions about quality assurance or research studies in palliative care are only a recent concern. One of the reasons might be the lack of validated German versions of outcome measures suitable for palliative care. Of the 12 outcome measures Hearn and Higginson identified as satisfactory for palliative care, only the EORTC QLQ-C30 has a validated German version.<sup>3,4</sup>

The overall aim of this study was to provide a valid version of the POS in German. The following study objectives were defined: to translate the POS into German, to use the German POS in different palliative care settings (palliative care unit, hospice, hospital support team, home care setting), to assess the validity and reliability of the German translation of the POS for patients and staff, and to evaluate the patients' and staff's response to the tool.

## Methods

The linguistic validation of the POS consisted of the following phases: 1) backward and forward translation of the POS into German, 2) pilot testing and feasibility testing of the POS, and 3) psychometric analysis of the POS.

Ethical approval for the validation and the use of the POS was gained from the Institutional Review Board from the University of Munich, the University of Cologne, and the Institutional Review Board of the Medical Council Nordrhein. Patients gave signed consent to participate in the study.

## Measures

The POS contains two almost identical measures, each with 11 items, one to be completed by staff, the other one by the patient. It covers physical and psychological symptoms, spiritual considerations, practical concerns, emotional

concerns of the patient and family, and psychosocial needs of the patient and the family. The last question is an open question for the patient and staff regarding the main problems of the patient in recent days. The staff version contains the Eastern Cooperative Oncology Group (ECOG) performance status to reflect the patients' functional status. The answers in all but the last question are scored using Likert scales from 0 to 4 with numerical and descriptive labels. POS scores of the individual questions number 1–10 can be summarized into a total score. The overall score ranges from 0 to 40, while the maximum impairment is 40, in both the staff and patient versions. The English POS showed acceptable agreement between staff and patient rating for 8/10 items, demonstrated construct validity, and had acceptable test/re-test reliability.<sup>1</sup>

## Translation

According to the EORTC recommendations for translation procedures of the EORTC Quality of Life Study Group,<sup>5</sup> the original POS was translated forward and backward by two professional translators. After adaptation of the different versions, a final POS version was agreed on, which was then used in different palliative care settings in Germany and Austria (hospital support team, palliative care unit, hospice, pain clinic, GP practice/home care). Review of the German POS by professionals revealed content and consensus validity.

## Pilot and Feasibility Testing

All new patient referrals to the hospital support team, the palliative care units, and the hospice beginning May 2002 and over a period of 3 months were included in the pilot testing. All patients in the pain clinic and general practice were included during this time. Patients were asked to complete the questionnaire by themselves or with the help of a staff or family member on Days 1–3 after admission and two times after 3 or 4 days each. If patients needed help completing the POS, staff members were asked to answer the POS questions before attending the patient. Patients in a home care or outpatient setting were asked to complete the POS at each visit. Patients were eligible for participation if they had good understanding of the German language, did not have impaired mental capacity (either assessed or

judged by staff), and were physically able to complete the questionnaire. A structured interview was performed with patients after they completed the POS regarding their opinion about the questionnaire and to assess its relevance for the patients' concerns (content validity). A questionnaire was developed to ask staff members who administered the POS their opinion (to assess staff's view of content validity of the POS). We had planned to compare the German POS with the EORTC-QLQ-C30 in order to analyze the construct validity of the new POS, but piloting indicated that completing both questionnaires and the open interview was too much for patients.

### *Psychometric Analysis*

In addition to quantitative analysis of data, staff and patient ratings were compared for agreement and correlation in the statistical analyses (to assess reliability). The five scores of answers were grouped into three categories: Category 1 was used for scores that were rated with 0, Category 2 for scores rated 1 and 2, and Category 3 for severe scores (rated 3 and 4). Cohen's  $\kappa$  coefficient was calculated for proportion of responses for simple agreement.<sup>6</sup> Spearman's rho was calculated to test the association between staff and patients for severe scores (answer Category 3), and for categories of grouped scores. Proportion agreement within one score was calculated. Matched pairs were compared for the group of patients and staff with severe codes (answer Category 3) who completed the POS three times (longitudinal view). The Chi-square test was used to test for significant levels of change of severe codes over time. Where the expected values were less than 5, Fisher's exact test was used. The answers to the open question No. 11 ("If any, what has been your/the patient's main problem in the last 3 days?") were content analyzed and categorized in somatic (pain and symptoms), psychological (mood, anxiety, and future), social (family and further care), and spiritual aspects (meaning of life, religion, and spirituality). Three potential problems could be mentioned. The SPSS version 11.0 for Windows statistical package was used to perform data analysis.

### *Results*

From May 2002 until January 2003, a total of 118 patients completed the POS in eight

different palliative care settings and institutions in Germany and Austria (Table 1). The demographic data are shown in Table 2. One hundred eighteen patients and staff completed the POS for the first assessment a median of 2 days (range 0–892) after admission. The second assessment was completed by 62 staff members after 12 days, and by 55 patients after 10 days. Thirty-six patients and 42 staff members were able to complete the POS a third time. Mean time between assessment 1 and 3 was 16 days for staff and 14 days for patients (median 8 days for both).

About one third of the patients (30%) were able to complete the questionnaire without external help at the first assessment. This number increased to 33% at the second assessment and 43% at the third measure point. About another third needed help from relatives at all three measure times (34%, 37%, 33%). Staff help was necessary for 36% of patients at first assessment, 30% at second assessment, and 28% at third assessment.

The patients took longer to complete the POS than did the staff. Patients needed, on average, 11 minutes at the first assessment, and 9 minutes at the second and third assessments; the staff needed 6 minutes at the first assessment, and 4 minutes at the second and third assessments.

### *Agreement of Staff and Patient Assessment for Grouped Codes*

In Table 3, Cohen's  $\kappa$  and Spearman's rho are shown for grouped scores between the three categories (for the answer categories 0, 1–2, 3–4). In the first assessment, Spearman's rho ranged from 0.11 to 0.54. Spearman's correlation was highly significant ( $P < 0.001$ ) for the items "pain" and "other symptoms." All other items showed significant levels of correlation except "family anxiety" and "self worth." Weak agreement was found for all questions in the first assessment (Cohen's  $\kappa$ ). For the second assessment, Spearman's rho was between  $-0.01$  and  $0.5$ . All other items had significant correlations except "support," "self worth," and "personal affairs." Weak agreement existed for all items except "support" ( $\kappa = 0.005$ ) and "personal affairs" ( $\kappa = 0.06$ ). At the third assessment, Spearman's rho was between  $-0.15$  and  $0.48$ . Significant levels of correlation were revealed for "pain," "other symptoms," "anxiety,"

Table 1  
Distribution of Total Population, Number Eligible for Participation, and Study Sample

Institution	Palliative Care Setting	No. Included	Percentage of Study Sample
University Hospital Munich	Hospital support team	31	26
Hospital Barmherzige Brüder Munich	Palliative care unit	15	13
Hospital Harlaching Munich	Palliative care unit	14	12
Munich Hospice	Hospice	9	8
General Practice Regensburg	GP practice/home care	6	5
Malteser Hospital Bonn	Palliative care unit	18	15
University of Cologne	Pain clinic	11	9
Hospital Lainz	Palliative care unit	14	12

and “family anxiety.” Weak agreement was given for all items except “information,” “personal affairs,” and “support.”

#### Agreement of Staff and Patient Assessment for Severe Codes

As in grouped scores, Spearman’s rho ranged from 0.13 to 0.61 (Figs. 1 and 2). “Pain” and “other symptoms” showed highly significant levels of correlation. For “family anxiety,” no correlation was found in the first assessment.

Table 2

#### Patient Demographic Characteristics (n = 118)

	n (%)
Age	
< 60 years	50 (42)
≥ 60 years	68 (57)
Range	27–94 years
Mean	63 years
Sex	
Men	44 (37)
Women	74 (63)
Marital status	
Married	61 (52)
With partner	14 (12)
Widowed	23 (20)
Divorced/separated	6 (5)
Single	12 (10)
Not known	2 (2)
Diagnosis	
Digestive organs	21 (18)
Respiratory tract	22 (19)
Genitourinary tract	30 (25)
Breast	27 (23)
Lymph/hematopoietic	2 (2)
Other cancers	16 (14)
ECOG (first assessment)	
Fully active	2 (2)
Restricted	17 (15)
Ambulatory	25 (22)
Limited ability	44 (39)
Disabled	25 (22)
Missing	5 (4)
Place of care at end of study	
Discharged home	63 (55)
Died	35 (31)
Still under care	16 (14)

Cohen’s  $\kappa$  showed weak agreement for all questions at the first assessment except “family anxiety.” For the second assessment, Spearman’s rho ranged between 0.04 and 0.5. “Pain,” “other symptoms,” and “anxiety” had highly significant correlation levels. “Information,” “life worthwhile,” “self worth,” and “wasted time” correlated significantly in this assessment. Weak agreement was shown for all questions except “family anxiety,” “support,” “life worthwhile,” “self worth” and “personal affairs.” For the third assessment, Spearman’s rho ranged between  $-0.08$  and  $0.61$ . “Pain” and “anxiety” showed highly significant levels of Spearman’s rho; correlation levels for “other symptoms,” “life worthwhile” and “personal affairs” were significant. Weak agreement was shown for “pain,” “other symptoms,” “anxiety,” “family anxiety,” “life worthwhile,” “self worth,” and “wasted time.” Patients and staff did not agree in their assessment concerning “information,” “support,” and “personal affairs.”

#### Proportion Agreement Within One Score

Table 4 describes the proportion agreement within one score, which reflects the proportion of scores that are within one score of the ideal of perfectly matched answers. The proportion agreement is rather high, indicating that in most answers, the staff was close to patients’ perception of their situation.

#### Severe Codes Over Three Assessments

The severe scores (answer Category 3) of the 36 patients who completed the POS three times are shown in Fig. 1. The frequency of severe scores decreased from time 1 to time 2 significantly for “anxiety” (Fisher’s exact test;  $P = 0.046$ ), “family anxiety” (Fisher’s exact test;  $P = 0.035$ ), “support” (Fisher’s exact test;  $P = 0.035$ ), and “self-worth” (Fisher’s exact test

Table 3  
Agreement Between Grouped Scores of Staff and Patient Assessments

		No. of Patients	Kappa for Grouped Scores (3 Categories)	Spearman's Correlation	Significance Level of Spearman's Correlation
<b>First Assessment</b>					
1	Pain	113	0.39	0.54	<0.001 <sup>a</sup>
2	Other symptoms	114	0.29	0.39	<0.001 <sup>a</sup>
3	Anxiety	113	0.23	0.32	0.001 <sup>b</sup>
4	Family anxiety	106	0.09	0.11	0.269
5	Information	109	0.23	0.24	0.012 <sup>b</sup>
6	Support	107	0.22	0.33	0.001 <sup>b</sup>
7	Life worthwhile	106	0.21	0.28	0.003 <sup>b</sup>
8	Self worth	99	0.21	0.16	0.108
9	Wasted time	104	0.19	0.25	0.010 <sup>b</sup>
10	Personal affairs	99	0.26	0.20	0.047 <sup>b</sup>
<b>Second Assessment</b>					
1	Pain	59	0.24	0.33	0.010 <sup>b</sup>
2	Other symptoms	59	0.31	0.50	<0.001 <sup>a</sup>
3	Anxiety	59	0.35	0.49	<0.001 <sup>a</sup>
4	Family anxiety	58	0.23	0.41	0.001 <sup>b</sup>
5	Information	57	0.36	0.36	0.005 <sup>b</sup>
6	Support	57	0.005	-0.01	0.946
7	Life worthwhile	57	0.35	0.45	<0.001 <sup>a</sup>
8	Self worth	56	0.18	0.22	0.099
9	Wasted time	56	0.39	0.44	0.010 <sup>b</sup>
10	Personal affairs	58	0.06	0.04	0.782
<b>Third Assessment</b>					
1	Pain	36	0.37	0.45	0.006 <sup>b</sup>
2	Other symptoms	36	0.27	0.42	0.010 <sup>b</sup>
3	Anxiety	36	0.32	0.48	0.003 <sup>b</sup>
4	Family anxiety	35	0.25	0.41	0.014 <sup>b</sup>
5	Information	35	NA	0.10	0.558
6	Support	34	-0.02	-0.15	0.399
7	Life worthwhile	34	0.19	0.27	0.118
8	Self worth	33	0.18	0.20	0.257
9	Wasted time	32	0.29	0.33	0.066
10	Personal affairs	34	NA	0.37	0.029

NA = not available.

<sup>a</sup>Significance level <0.001.

<sup>b</sup>Significance level <0.05.

$P = 0.035$ ). Staff scores are shown in Fig. 2. None of the changes in staff scores were significant.

#### POS Sum Scores

The average of sum scores is presented in Table 5. The minimum sum score for patients was 2 and the maximum 31. For staff, the minimum score was 2 and the maximum was 28. The average of sum scores was between 13 and 15 in patients' assessments and between 12 and 15 in staff assessments.

#### Open Question No. 11

The categories found for Question 11 are shown in Table 6. Somatic problems played a major role in all three assessments and were mentioned by staff and patients. Psychological

problems were regularly reported in the first and second assessment, but far less in the third assessment. Social problems were presented in all three assessments, but much less than the somatic and psychological problems.

#### Evaluation of the Questionnaire and Translation by the Patient

Of the 87 patients who completed the POS and answered questions regarding the scale, 77 evaluated the questions and the translation as well. Seven percent liked the questionnaire a lot, 59% liked it, 29% were undecided, and 5% did not like it very much. Nine percent agreed strongly that the questionnaire reflects their actual condition, 69% agreed, 12% were undecided, 7% disagreed, and 3% strongly disagreed. Nine percent dealt with the answers



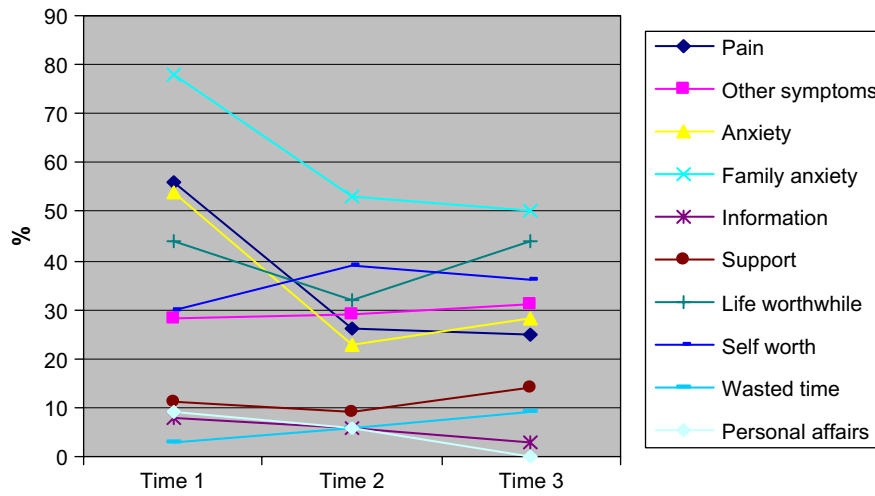


Fig. 1. Severe scores (3 and 4) of patients.

very well, 76% well, 7% were undecided, and 8% dealt with the answers less well.

The following questions were rated as most important by the patients: Q1 (pain,  $n = 23$ ), Q2 (other symptoms,  $n = 10$ ), Q5 (information,  $n = 12$ ), Q7 (life worthwhile,  $n = 5$ ), and Q11 (open question,  $n = 11$ ). Thirteen patients found that all questions were important.

We asked for aspects of quality of life that were not covered by the POS. Forty-five patients (52%) did not miss any aspect. The following aspects were mentioned as missing by single patients: “home and environment,” “religious life,” “sexuality and illness,” “loss of interest,” “loss of hope,” “change in quality of life,” and “living with problems and restrictions.”

Concerning improvement of the questionnaire, 41 patients (47%) indicated that there was nothing that they would improve or did not like. Two patients reported difficulties with the answer categories. One patient thought that the time span of three days was too short, one wanted more questions about everyday life, and one patient complained that there were too many questions.

The last item of the questionnaire’s evaluation concerned the comprehensibility and clearness of the translation of the POS, and the importance of the questions. The results are demonstrated in Table 7.

The following are comments for improvement of the questions or the translation. For

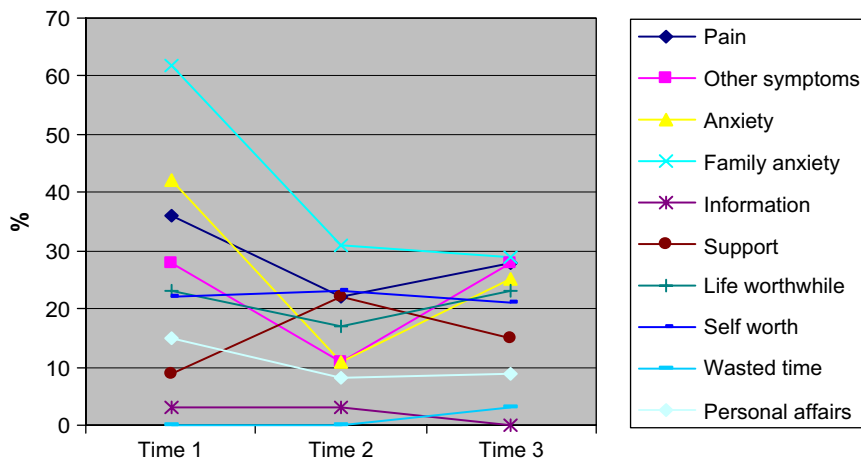


Fig. 2. Severe scores (3 and 4) of staff.

Table 4  
Agreement Between Severe Scores (3/4) of Staff and Patient Assessments

	No. of Patients	Staff (%)	Patients (%)	Kappa for Each Score (5 Categories)	Spearman's Correlation	Significance Level of Spearman's Correlation	Proportion Agreement Within 1 Score
<b>First Assessment</b>							
1 Pain	113	35	46	0.31	0.61	<0.001 <sup>a</sup>	0.87
2 Other symptoms	114	33	35	0.13	0.38	<0.001 <sup>a</sup>	0.78
3 Anxiety	113	44	41	0.18	0.32	0.001 <sup>b</sup>	0.77
4 Family anxiety	106	57	72	0.08	0.13	0.170	0.70
5 Information	109	4	5	0.21	0.24	0.010 <sup>b</sup>	0.83
6 Support	107	12	10	0.18	0.29	0.002 <sup>b</sup>	0.78
7 Life worthwhile	106	24	36	0.13	0.33	0.001 <sup>b</sup>	0.71
8 Self worth	99	22	29	0.14	0.22	0.030 <sup>b</sup>	0.74
9 Wasted time	104	4	10	NA	0.26	0.009 <sup>b</sup>	0.93
10 Personal affairs	99	14	10	0.26	0.20	0.047 <sup>b</sup>	0.87
<b>Second Assessment</b>							
1 Pain	59	20	24	0.21	0.44	<0.001 <sup>a</sup>	0.83
2 Other symptoms	59	16	34	0.14	0.48	<0.001 <sup>a</sup>	0.80
3 Anxiety	59	19	26	0.19	0.50	<0.001 <sup>a</sup>	0.85
4 Family anxiety	58	33	53	0.09	0.35	0.070	0.69
5 Information	57	3	7	0.26	0.35	0.008 <sup>b</sup>	0.86
6 Support	57	20	10	0.02	0.08	0.530	0.74
7 Life worthwhile	57	24	29	0.04	0.31	0.019 <sup>b</sup>	0.81
8 Self worth	56	26	28	0.04	0.31	0.022 <sup>b</sup>	0.80
9 Wasted time	56	3	7	0.39	0.44	0.001 <sup>b</sup>	0.96
10 Personal affairs	58	9	5	0.06	0.04	0.782	0.93
<b>Third Assessment</b>							
1 Pain	36	23	24	0.31	0.58	<0.001 <sup>a</sup>	0.89
2 Other symptoms	36	28	29	0.23	0.47	0.003 <sup>b</sup>	0.75
3 Anxiety	36	28	26	0.34	0.61	<0.001 <sup>a</sup>	0.83
4 Family anxiety	35	26	47	0.29	0.33	0.050	0.66
5 Information	35	0	3	NA	0.09	0.615	0.80
6 Support	34	15	13	-0.03	-0.08	0.631	0.71
7 Life worthwhile	34	29	42	0.12	0.38	0.027 <sup>b</sup>	0.71
8 Self worth	33	22	34	0.18	0.11	0.545	0.67
9 Wasted time	32	2	8	0.29	0.33	0.066	0.97
10 Personal affairs	34	10	0	NA	0.37	0.029 <sup>b</sup>	0.94

NA = not available.

<sup>a</sup>Significance level <0.001.

<sup>b</sup>Significance level <0.05.

Q2, a symptom checklist was suggested. One patient expressed difficulty in how to answer this question, as he was suffering from several symptoms at the same time. Two patients found that the answers for Q5 were too long and one patient thought that the answers were not appropriate for this question. Several patients did not understand the expression "self-esteem," which was used to translate Q8 ("Have you felt

good about yourself"). Q9 was too long for one patient and not appropriate for patients who were in the hospice or the pain clinic.

#### *Evaluation of the Questionnaire and Translation by Staff*

Twenty-two members of the staff, 11 doctors and 11 nurses, expressed their opinion about POS. Forty-one percent of the staff liked the

Table 5  
Sum Scores of Patients and Staff

	1st Assessment mean (min-max)	2nd Assessment mean (min-max)	3rd Assessment mean (min-max)
Patients	15.9 (3-31)	13.5 (2-27)	13.7 (2-28)
Staff	15.4 (2-28)	13.7 (2-24)	11.9 (2-24)

Table 6  
Categories Open Question 11 (n = 80)

	Staff 1st Answer	Patient 1st Answer	Staff 2nd Answer	Patient 2nd Answer	Staff 3rd Answer	Patient 3rd Answer
<b>First assessment</b>						
Somatic	55	46	37	22	13	6
Psychological	7	23	16	11	7	5
Social	18	11	5	4	4	2
<b>Second assessment</b>						
Somatic	30	17	11	4	6	6
Psychological	5	10	6	5	0	0
Social	5	3	3	1	2	1
<b>Third assessment</b>						
Somatic	17	9	10	7	0	2
Psychological	4	6	1	1	0	0
Social	1	1	2	1	0	1

questionnaire, 22% were undecided, and 36% did not like it very much. Thirty-six percent thought that the questionnaire reflected the patients' actual condition, 45% were undecided, and 14% thought it did not. Eighteen percent thought that the patients dealt with the questions well, 45% were undecided, and 36% thought the questions were dealt with less well. Thirty-two percent thought the patients managed the answers well, 36% were undecided, and 27% thought patients dealt with the answers less well. The general comments were:

- POS is good for patients who are well, but difficult for those who are too ill or sick.

- Using POS shortly after admission was difficult because patients are not known well enough.
- Using the same questionnaire for patients and staff was very helpful, as staff members learned more about the differences between patients' and staff members' views.
- It was difficult to use POS in the hospital support team in the first days of admission.
- Some patients found it difficult to assess a time span of three days.

Comments regarding the questions were:

- The most important questions were: pain (Q1), other symptoms (Q2), information (Q5), and open question (Q11).

Table 7  
Evaluation of the Questionnaire by Patients (n = 77)

No.	Question	Question Clear		Question Important	
		Yes	No	Yes	No
1	Over the past 3 days, have you been affected by pain?	99	1	98	2
2	Over the past 3 days, have other symptoms (e.g., nausea, coughing, or constipation) been affecting how you feel?	100	0	93	7
3	Over the past 3 days, have you been feeling anxious or worried about your illness or treatment?	96	4	100	0
4	Over the past 3 days, have any of your family or friends been anxious or worried about you?	97	3	95	5
5	Over the past 3 days, how much information have you and your family or friends been given?	92	8	95	5
6	Over the past 3 days, have you been able to share how you are feeling with your family or friends?	96	4	97	3
7	Over the past 3 days, have you felt that life was worthwhile?	97	3	100	
8	Over the past 3 days, have you felt good about yourself?	92	8	94	6
9	Over the past 3 days, how much time do you feel has been wasted on appointments relating to your healthcare (e.g., waiting around for transport or repeating tests)?	89	11	69	31
10	Over the past 3 days, have any practical matters resulting from your illness, either financial or personal, been addressed?	89	11	89	11
11	If any, what have been your main problems in the last 3 days?	97	3	100	0



- Questions were too long for patients.
- Question about information (Q5) and about wasting time (Q9) were not relevant for patients in the home care setting, the hospice, and the pain clinic.
- Questions concerning quality of life (Q7) and self-worth (Q8) were too complicated for patients and many did not understand them.
- The open question (Q11) was often the beginning for an open communication.

Comments regarding answers were:

- Many answers were too long and required a lot of concentration from the patient.
- Classification of answers of psychosocial questions was difficult for patients.

## Discussion

Different outcome scales in palliative care have been developed in recent years. Hearn and Higginson identified 12 of 41 measures that met the following criteria: suitable for patients with cancer or advanced disease receiving palliative care, more than one domain covered, and suitable for patients with all types of cancer.<sup>3</sup> Hearn and Higginson concluded in their systematic review that none of the measures reviewed systematically covers all domains that are considered important in palliative care. Furthermore, each measure meets some, but not all, criteria for validity, reliability, responsiveness, and appropriateness.<sup>3</sup>

### *Comparison of the Original Version of the POS with the German POS*

As in the original validation study of Hearn and Higginson,<sup>1</sup> the German version of the POS could be used in different palliative care settings, including palliative care units, hospices, pain clinics, and GP practices. Of the known number of patients cared for during the study period, 42% assessed the POS. Although this number is even higher than in the original study, where only about one-third of patients were able to complete the POS,<sup>1</sup> the known number of patients included in this study is much smaller. Only about one-third of the patients were able to complete the POS without external help. The average time to complete the POS was comparable for staff, but patients

needed more time than in the original study. It is not mentioned how many patients in the Hearn and Higginson study<sup>1</sup> were able to complete POS on their own.

### *Summary of the Results*

The first assessment of the POS was planned within three days after admission to palliative care; however, only 60% of patients received the POS within that time. The remainder received the first POS within 20 days after admission. Data are skewed, as many patients who participated in the study were under the care of a hospice, pain clinic, and GP practice for a long time (up to almost a year) before entering the study. Furthermore, the staff may have protected patients to give them time to feel comfortable on the unit before distributing the questionnaire. Consequently, data from the first assessment do not necessarily reflect the quality of care prior to admission as patients might have had improved symptom control and psychosocial support during their inpatient stay already.

Further assessments were completed in much longer time spans than planned. A delay of more than a week was noted, especially between the first and second assessment. Due to deterioration or discharge, the number of patients able to complete the POS for the second and third time declined.

Severe scores provided important information about the patient's status and the areas in greatest need of help. Frequencies of severe scores were of higher clinical value than mean scores.<sup>7</sup> As intended, severe codes for "pain" and "other symptoms" decreased between first and second assessment. Interestingly, a decrease in "anxiety" scores between the first and second assessment was paralleled by an increase in the feeling that life is worthwhile. "Family anxiety" also was reduced significantly. In parallel, the patients felt more support from their relatives. Thus, families seemed to profit at least as much as patients from palliative care in this group.

The POS gave the patients the opportunity to name their main problems during the past days. With one exception, somatic problems were mentioned mainly by patients and staff, through all three assessments. This was observed in Hearn and Higginson's study as well,

and led to the decision to include more symptoms in the POS,<sup>1</sup> which is already implemented in the new St. Christopher's version.<sup>8</sup> In this study, staff rated somatic problems higher than patients, whereas patients mentioned more psychological problems than staff noticed. This is interesting, as palliative care claims to be holistic and to look at psychosocial and spiritual issues as well as physical ones. However, this cannot be taken for granted and requires special training and attitude from all staff members. Because none of the patients reported spiritual problems, it is uncertain whether an additional question on this issue is needed. Currently, several tools exist that try to measure religious and spiritual needs.<sup>9,10</sup> One could also argue that spiritual questions will not be addressed as long as somatic problems such as pain are not sufficiently relieved.

*Reliability.* Cohen's  $\kappa$  is a measure of the proportion of responses where two raters agree and is, therefore, a measure for inter-rater reliability. In Hearn and Higginson's validation study of the POS,<sup>1</sup> agreement could be shown for eight of ten items of the first assessment and for all items of the second assessment. In this study, levels of Cohen's  $\kappa$ , both in grouped and severe scores, differed remarkably from Hearn and Higginson's study. Looking at grouped scores, it was only possible to show weak agreement for all items except family anxiety in the first assessment. In the second and third assessment, weak agreement could only be shown in some of the items. In this respect, reliability was only given for some items of the POS but not yet for the questionnaire.

This raises the question why agreement is so different between the two studies. One could argue that palliative care in the UK is more advanced and staff is more experienced, both in patient care and in using questionnaires such as the POS. However, all institutions participating in this study were already in operation for several years. On the other hand, it could be questioned whether correlation between staff and patients is indeed a measure of reliability, as there are known problems with proxy ratings in palliative populations.<sup>11</sup> An advantage of the missing agreement could be that the data being gained from the POS are even more important because it might give new information about the patient, and help staff to improve

care. Differences between items should then indicate areas for action to support the patient. It also could be argued that only one third of the patients completed the POS without help, and help from relatives or staff may have led to under-reported anxiety and other non-physical problems. Furthermore, experienced staff may use nonverbal communication to identify problems that patients have suppressed (e.g., anxiety). Patients may report pain resulting from psychosocial or spiritual stress, and experienced staff may correctly interpret the underlying construct and score less pain but more anxiety for the patient.

*Validity.* Content and consensus validity for the German POS was revealed by professionals. Content validity was also tested by asking the patients about the importance of the items, missing aspects, and if their actual situation was reflected by the questionnaire. Construct validity could not be tested as originally planned, as it was too much for patients to complete another questionnaire.

Evaluation of the Questionnaire by Patients and Staff: Content and Consensus Validity. This area of the study showed a divergence of the opinion between patients and staff. In the patient group, the majority liked the questionnaire and agreed that the POS reflected their actual situation, while almost one-third of patients were undecided. The staff was much more critical—only half of the staff liked the POS, while one-third did not like it very much. Almost half of the staff members were undecided whether the POS really reflected the patients' actual situation. This discrepancy could be due to several reasons. A responder bias might be that patients want to please their carers and judge the tool more positively than they would do with an independent researcher. As none of the institutions had its own research assistant, the POS and the evaluation always had to be carried out by the staff. This bias is underlined by some comments of staff, who thought that the patients were mainly agreeing to take part in the study to do them a favor. Some staff questioned the time span over three days, which they found difficult for patients to recall. The main criticism regarding the POS was that it was too difficult for very ill patients and that some questions and answers were too long.

The most important questions for patients were those asking about physical symptoms, information, whether life is worthwhile, and the opportunity to name main problems in an open question. Interestingly, the staff supported the importance of the same questions, except the question whether life is worthwhile. They found the symptom questions easier to manage for patients than psychosocial questions. Depending on the setting, some questions were not appropriate, such as wasted time, in the hospice or home care, and should be omitted if the POS is used in these settings.

A similarly controversial picture emerged regarding the answering options given in the POS. The vast majority of patients thought they coped very well or well with the answers. In contrast, the staff thought that many answers were too long, especially for patients who were not able to complete the POS on their own. Some of the answers required a lot of concentration (i.e., if a patient had to listen and then choose the appropriate answer).

Evaluation of the Translation by Patients and Staff. An important question of this study was whether the wording of the translated POS version was clear and understandable. The majority of patients found that all questions were clear and understandable. The only expression that patients questioned related to Question 8, where the words “feeling good about yourself” were translated with a German expression which actually means “positive self-esteem” (“positives Selbstwertgefühl”). The comments of several staff support this critique, as they observed that several patients did not really understand what was meant by “positive self-esteem.” Therefore, this expression might need to be changed and re-evaluated. Interestingly, this question caused patients discomfort in the UK as well, and alternative wordings are being sought.<sup>12</sup>

## Conclusion

In summary, the German version of the POS is well accepted by patients and staff and has satisfactory reliability and validity. The German POS was already used in the “Core Documentation” project<sup>13</sup> to gather data on quality of care and to gain additional information on validity,

reliability, and appropriateness in a representative sample of patients in German hospices and palliative care units.

## Acknowledgments

This paper was based on the Masters dissertation of C.B. within the MSc course for Palliative Care at King’s College London. Special thanks to all colleagues and nurses in the different palliative care institutions in Germany who took part in the study and interviewed the patients, especially Susanne Roller, Elisabeth Albrecht, Michaela Werny, Susanne Vogel, Susanne Steinhauser-Kirsch and Annette Hasenender.

## References

1. Hearn J, Higginson IJ. Development and validation of a core outcome measure for palliative care: The Palliative Care Outcome Scale. Palliative Care Core Audit Project Advisory Group. *Quality in Health Care* 1999;8(4):219–227.
2. Aspinall F, Hughes R, Higginson I, et al. A user’s guide to the Palliative Care Outcome Scale. Palliative Care & Policy Publications, Kings College, London; 2002.
3. Hearn J, Higginson IJ. Outcome measures in palliative care for advanced cancer patients: a review. *J Public Health Med* 1997;19(2):193–199.
4. Aaronson N, Ahmedzai S, Bergman B, et al. The European Organization 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;85:365–376.
5. Cull A, Sprangers M, Bjordal K, Aaronson N. EORTC quality of life study group: translation procedure. Brussels: EORTC, 1998.
6. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use, 2nd ed. Oxford: Oxford University Press, 1995.
7. Higginson IJ, Carr AJ. The clinical utility of quality of life measures. In: Carr AJ, Higginson IJ, Robinson PG, eds. *Quality of life*. London: BMJ Books, 2003:63–78.
8. Chidgey J. Use of the Palliative Care Outcome Scale (POS) at St. Christopher’s Hospice. London: St. Christopher’s Hospice & Department of Palliative Care & Policy, King’s College London, 2000.
9. King M, Speck P, Thomas A. The royal free interview for spiritual and religious beliefs: development and validation of a self-report version. *Psychol Med* 2001;31(6):1015–1023.

10. Puchalski C, Romer A. Taking a spiritual history allows clinicians to understand patients more fully. *J Palliat Med* 2000;3(1):129–137.
11. Sterkenburg CA, King B, Woodward CA. A reliability and validity study of the McMaster Quality of Life Scale (MQLS) for a palliative population. *J Palliat Care* 1996;12(1):18–25.
12. Higginson IJ. Measuring the quality of care in palliative care services: an interview with Irene Higginson, by AL Romer. *Innovations in End-of-Life Care* 2000;2(1). [www.edc.org/lastacts](http://www.edc.org/lastacts). Accessed January 2003.
13. Lindena G, Nauck F, Bausewein C, et al. Qualitätssicherung in der Palliativmedizin—Ergebnisse der Kerndokumentation 1999-2002 (Quality Assessment in Palliative Medicine—Results from the core documentation, 1999–2002). Submitted.