available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.ejconline.com

Assessing health-related quality of life in palliative care: Comparing patient and physician assessments

Morten Aa. Petersen[a,*], *Henrik Larsen*[a], *Lise Pedersen*[a], *Nan Sonne*[a], *Mogens Groenvold*[a,b]

[a]*The Research Unit, Department of Palliative Medicine, Bispebjerg Hospital, Bispebjerg bakke 23, Opgang 20d, DK-2400 Copenhagen NV, Denmark*

[b]*Institute of Public Health, University of Copenhagen, Copenhagen, Denmark*

ARTICLE INFO

Article history:
Received 16 December 2005
Received in revised form
4 January 2006
Accepted 6 January 2006
Available online 18 April 2006

Keywords:
Agreement
Questionnaires
Palliative care
Patients
Physicians
Proxy
Quality of life

ABSTRACT

It is often difficult to recruit patients for palliative care studies and severe attrition must be expected resulting in biased findings. This may be avoided if equivalent information could be obtained from sources other than the patients. Therefore, we investigated whether physician assessments can be used to evaluate the patients' health-related quality of life (HRQOL). Patient and physician assessments of the patients' HRQOL were obtained once a week for up to 13 weeks using EORTC QLQ-C30 items. The agreement between patients and physicians at first contact ($N = 115$) and for the following 13 weeks combined (total $N = 263$) was investigated. Significant differences between patient and physician assessments were observed for all HRQOL domains assessed. Physicians reported patients to have fewer problems/symptoms than patients did for all HRQOL domains except for physical and social functioning. The agreement between patients and physicians was poor. Using physician assessments may bias findings and cannot be recommended as a substitute for patient self-assessment in palliative care.

1. Introduction

Palliative care aims at providing the best possible care to improve the quality of life of patients and their families.[1] To achieve this goal, it is necessary to evaluate the effect of the care. However, it is often difficult to recruit patients for palliative care studies and severe attrition due to deterioration must be expected.[2,3] This biases results and limits the broad applicability of findings.[4] If equivalent information could be obtained from other sources than the patients, this would be preferable in descriptive studies in palliative settings. This could potentially improve the generalizability and the range of descriptive studies ethically and practically feasible in palliative care. Such alternative sources or 'proxies' could be physicians, nurses, or significant others.

There is general agreement that the patient is the most appropriate source of information regarding his/her quality of life. Therefore, proxy assessments should only be used if in agreement with the patients' own ratings. Several studies have compared patient and proxy ratings in general. A review from 1992 concluded that the concordance between patient and proxy ratings was far from optimal.[5] The 'up-date' from 2002 was more positive, concluding that proxy ratings are reasonably accurate and substantial discrepancies occur in a minority of cases only.[6] However, it was noted that proxies tended to overestimate the patients' problems and symptoms.[6] A number of studies have compared the responses of terminally ill/palliative care patients with proxy responses[7–16] (see Ref. [17] for a review). Except for one, these studies were small (less than 50 participants[8,9,13,15]) or of

* *Corresponding author:* Tel.: +45 3531 2025; fax: +45 3531 2071.
E-mail address: map01@bbh.hosp.dk (M.Aa. Petersen).
0959-8049/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2006.01.032

moderate size (less than 80 participants[10–12,14,16]). The agreement ranged from poor to good agreement across the studies with a tendency to better agreement for concrete and overt symptoms, such as impaired physical functioning, vomiting, and dyspnoea, while poorer agreement often was observed for subjective aspects such as emotional functioning, feelings, and pain.

The majority of proxy studies in palliative care have evaluated the use of family caretakers/significant others as proxies. Some significant others may be too emotionally affected by their relative's condition to be able and willing to participate as proxies. The health care professionals, e.g. the physicians, could in principle provide information about all patients regardless of the patients' conditions. Ratings of the patients could be incorporated into the physicians' daily rounds. We have only identified two studies comparing responses from the professional staff and patients,[7,16] and one of these studies evaluated the agreement on one overall quality of life measure only.[16] Both studies found relatively poor agreement and recommended that staff ratings be avoided or used with caution. In light of the clear advantages of using staff ratings in palliative care, further evidence is needed to determine whether the agreement is sufficient or staff ratings should be avoided. Therefore, we initiated a study investigating whether ratings from physicians in a palliative care unit could replace patient ratings. The aim was to evaluate the agreement between physicians' and patients' ratings of different aspects of the patients' health-related quality of life.

2. Patients and methods

2.1. Patients

From June 1998 to August 2003, a longitudinal study evaluating the symptomatology of palliative care patients using patient self-assessment questionnaires was carried out in the Department of Palliative Medicine, Bispebjerg Hospital, Copenhagen, Denmark. Inclusion criteria were admittance to the department as inpatient, outpatient, or home-care, Danish speaking, age at least 18 years, and informed consent.[18] Home-care patients are patients receiving palliative care from the department in their own home or in a few cases in a nursing home. The current study of proxy assessments was implemented by expanding the evaluation study with physician assessments in the period from June 2001 to August 2003. In this period, 353 eligible patients were referred to the palliative department.

2.2. Patient self-assessments

Patients participating in the evaluation study received a set of questionnaires on the day of first contact with the department (T0). The questionnaires were to be filled in on this day or the day after. The patient received identical questionnaires once a week for the next 13 weeks (T1–T13) or as long as the patient was able and willing to participate. The questionnaire booklet given to the patients included a number of self-assessment questionnaires.[18] For the proxy study, only the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30, version 3.0)[19] was used.

The EORTC QLQ-C30 is one of the most widely used cancer-specific quality of life questionnaires.[20] It consists of 30 items. Twenty-four of the items form nine scales – six functional scales and three symptom scales – and six items are single-item symptom measures. The scale scores are constructed by summation of the item responses followed by transformation to a 0–100 scale.[21] For functional scales, a higher score reflects better functioning whereas a high score on a symptom measure reflects a high intensity of symptoms.

2.3. Physician assessments

All patients referred to the department in the study period should in principle be evaluated by a physician. The 'proxy questionnaire' filled in by the physicians included 20 of the items from the QLQ-C30: items 3–5 from the physical functioning (PF) scale, items 12 and 18 from the fatigue (FA) scale, items 21 and 24 from the emotional functioning (EF) scale, the items of the cognitive functioning (CF), social functioning (SF), nausea/vomiting (NV), and pain (PA) scales, and the single item measures on dyspnoea, insomnia, appetite loss, constipation, and diarrhoea (items 8, 11, 13, 16, and 17). There were two reasons for including only 20 of the 30 QLQ-C30 items in the proxy questionnaire. Firstly, when the study was initiated the EORTC Quality of Life Group was in the process of developing a shortened version of the EORTC QLQ-C30 for use in palliative care.[22] While the final version was not yet available, some results were ready. Secondly, the physicians were already burdened by having to complete a 10-page questionnaire for each patient each week. Therefore, we decided to omit 10 items: eight were expected to be omitted from the palliative care questionnaire and two items were judged so subjective that there was little idea in asking for the physician assessments. The agreement between patient and physician assessments was evaluated across the seven EORTC scales and five EORTC single item measures in the proxy questionnaire.

To be included in the analyses, a patient and a physician assessment should be filled in with one day apart at the most. This restriction was used to minimise the risk that possible differences between patient and physician ratings were caused by changes in the patient's conditions between the two assessments.

Notice that each physician may have evaluated more than one patient. In principle, this may introduce bias when using standard statistical methods. However, investigations of this (details omitted) showed that adjustment for the multiple assessments per physicians did not affect the findings. Therefore, to simplify the analyses we ignored that we had multiple assessments from some physicians in the analyses presented.

2.4. Analyses

2.4.1. Comparison of participants and non-participants

The participating patients were compared to the non-participants with regard to: age, gender, cancer site (breast, colorectal, gynaecological, lung, or others), Karnofsky Performance

Status (KPS),[23] affiliation to the department (inpatient, outpatient, home care, or nursing home), residual lifetime (days from first contact to death), Mini-Mental State Examination (MMSE) score,[24] social group (1–5, group 1 being the most affluent),[25] marital status (married, never married, divorced, or widow), partner (living with partner/alone), having children (yes/no), and the physician reported scores on the QLQ-C30 scales and single item measures. Wilcoxon rank-sum test, log-rank test (residual lifetime), and Fisher's Exact test (2×2-tables) were used. Variables showing a significant difference ($P < 0.05$) between participants and non-participants were included as explanatory variables in a multiple logistic regression analysis with participation (yes/no) as outcome.

2.4.2. *Agreement between patients and physicians*

To evaluate the agreement between patients and physicians at the patient's first contact with the department, T0, we calculated the mean difference and the mean of the absolute patient–physician differences (i.e. the differences ignoring the direction) for each of the scales and single item measures. The mean difference may be viewed as a measure of validity and the absolute mean as a measure of reliability of the physician assessments at the group level. Using Wilcoxon signed rank test, we tested whether the patient and physician assessments were significantly different. To evaluate the agreement at the individual level, we calculated the percent exact agreements and the percent agreements within one score level. Further, we calculated the Spearman correlation and the weighted kappa coefficient between the patient and physician assessments since these are very commonly used as measures of agreement in proxy studies[6] and thereby enable comparisons with most proxy studies. Finally, to investigate whether the agreement depended on the patient's level of symptoms/functioning, we plotted the mean of the physician scores for each patient reported score level.

As shown in Table 1, few patients participated during the whole study period. This is by no means surprising in a palliative care setting: only 25% of the patients were alive 13 weeks after referral to the department. Because of the large attrition, we chose to focus primarily on the assessments at T0. The following time-points, T1 through T13, were used for verification of the findings for T0. To get an overall evaluation of the agreement at T1–T13, we combined the patient–physician pairs from the 13 time-points and then analysed these as one sample. That is, a patient with patient and physician responses from all 13 time points would contribute with 13 observations, while a patient with pairs from T1 to T3 would contribute with three. This approach was inspired by the study of Stephens and colleagues.[26] For the combined sample of patient–physician pairs from T1 to T13, we repeated the analyses conducted for T0 except we did not test whether patient and physician assessments were significantly different since usual tests are not valid with multiple assessments per patient.

All analyses were carried out using the SAS statistical software package v. 8.02.

3. Results

3.1. *Participants*

In all, 115 patients met the criteria for participation at T0. The 238 non-participants were excluded for the following reasons: 210 did not have a patient assessment at T0, 6 did not have a physician assessment, and for 22 patients either the patient or the physician assessments were not filled in within the timeframe allowed. Clinical and socio-demographic data for the 115 participants are shown in Table 2.

3.2. *Comparison of participants and non-participants*

Participation was significantly higher among inpatients than for the other patient groups, the participants were judged by the physicians to have significantly more severe nausea/vomiting, and significantly poorer emotional functioning than non-participants (Table 2). Otherwise, there were no significant differences between participants and non-participants. The logistic regression analysis with participation as outcome and affiliation, EF, and NV as explanatory variables showed that only affiliation to the department was a significant independent predictor of participation ($P < 0.001$) indicating higher participation among inpatients. Comparisons of patient–physician agreement for inpatients, outpatients, and home care patients did not indicate any significant differences in agreement across the affiliations (results not shown).

3.3. *Analyses of agreement*

The results of the analyses of agreement between patients and physicians at T0 are summarised in Table 3. The mean differences between the patient and physician assessments ranged from −15 (dyspnoea) to 18 (social functioning) on a 0–100 scale. The closest average agreements were observed for fatigue (−7.0) and constipation (−8.7). The average mean difference (ignoring the signs) across the 12 domains was 12. The patient and physician assessments were significantly different for all 12 domains (all P-values <0.01). Except for physical and social functioning, the physicians judged the patients to have fewer

Table 1 – Number of patients and physicians filling in the questionnaire and the number of eligible patient–physician pairs[a] at each of the 14 time points, T0–T13

	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13
Patients	143	121	95	74	53	45	45	40	39	28	26	24	21	20
Physicians	274	149	104	85	65	53	42	40	41	30	32	20	21	19
Pairs[a]	115	68	40	39	26	18	16	15	13	5	10	5	5	3

a Eligible pairs: both patient and physician answers to at least one item at the given time point answered at no more than one day apart.

Table 2 – Mean scores/distributions for participants and non-participants and tests of no differences between the two groups

Variable	Participants (N = 115)		Non-participants (N = 238)		P-value[b]
	N	Mean/percent[a]	N	Mean/percent[a]	
Age (mean)	115	62.7	238	64.5	0.476
Gender					
Male	50	43.5%	109	45.8%	0.733
Female	65	56.5%	129	54.2%	
Diagnosis					
Breast cancer	21	18.8%	30	14.6%	0.707
Colorectal cancer	12	10.7%	27	13.1%	
Gynaecological cancer	10	8.9%	16	7.8%	
Head and neck	9	8.0%	10	4.9%	
Lung cancer	20	17.9%	40	19.4%	
Other cancer sites	49	43.8%	83	40.3%	
Affiliation					
Inpatient	72	62.6%	98	41.2%	**0.002**
Outpatient	13	11.3%	36	15.1%	
Home care patients	39	25.2%	94	39.5%	
Nursing home	1	0.9%	10	4.2%	
Residual lifetime (days) (mean)	115	101.2	238	85.0	0.276
Marital status					
Married	45	39.8%	42	47.7%	0.210
Never married	22	19.5%	11	12.5%	
Divorced	21	18.6%	22	25.0%	
Widow	25	22.1%	13	14.8%	
Live with a partner	55	48.7%	45	51.1%	0.777
Without a partner	58	51.3%	43	58.9%	
Have children	87	77.0%	70	79.6%	0.732
No children	26	23.0%	18	20.4%	
Social group (1–5) (mean)	112	3.2	87	3.2	0.861
KPS (mean)	112	48.5	161	47.3	0.736
MMSE (mean)	77	26.1	39	25.5	0.568
Physician ratings (mean)					
PF, physical functioning	114	21.0	159	18.0	0.135
EF, emotional functioning	113	71.2	156	78.9	**0.008**
CF, cognitive functioning	112	73.2	156	65.6	0.071
SF, social functioning	113	30.1	157	34.1	0.368
FA, fatigue	113	73.5	158	69.7	0.252
NV, nausea/vomiting	113	26.0	158	16.5	**0.008**
PA, pain	112	47.5	158	50.1	0.502
Item 8 'Dyspnoea'	114	26.9	159	22.9	0.475
Item 11 'Insomnia'	113	27.7	159	26.6	0.975
Item 13 'Appetite loss'	113	58.4	159	50.3	0.074
Item 16 'Constipation'	112	31.3	158	30.4	0.805
Item 17 'Diarrhoea'	112	12.8	158	8.0	0.296

a Percentages are of those with information on a particular variable.
b P-values <0.05 are in bold.

symptoms or better functioning compared to the patients' own assessments. The absolute mean differences ranged from 14 to 36. The smallest absolute differences, indicating best reliability, were observed for physical functioning, nausea/vomiting, and 'constipation', while the largest absolute differences were observed for 'insomnia' and the two 'psychosocial scales' emotional and social functioning. With exact agreement between physicians and patients for 21–26% of the patients the three 'psychosocial scales', cognitive, emotional, and social functioning, had the lowest proportions of exact agreement. The highest proportions of exact agreement were observed for the single items measuring dyspnoea, constipation, and diarrhoea with exact agreement ranging from 55% to 62%. Similarly, when looking at 'close to exact' agreements, i.e. agreement within one score level, the highest proportions were observed for the single item measures, while the poorest agreement was observed for emotional and social functioning. The correlations between the physician and patient assessments ranged from 0.23 (social functioning) to 0.76 (nausea/vomiting and dyspnoea) with an average of 0.54. The weighted kappa coefficients were between 0.15 (emotional and social functioning) and 0.60 (constipation). Using the widely applied classification of the kappa coefficient,[27] the agreements were 'slight' for emotional and social functioning, 'fair' for insomnia, appetite loss, diarrhoea, fatigue, and cognitive functioning, and 'moderate' for the remaining five domains.

Fig. 1 shows plots of the mean physician scores for each level of the patient reported scores. The lines for the cogni-

Table 3 – Results of the comparison of patient and physician assessments at the patients' first contact with the department

Variable	N	Mean[a]	Abs. mean[b]	Agree[c] (%)	±1 Level[d] (%)	Corr.[e]	Kappa[f]
PF, physical fct	108	−9.5***	14.1	25.9	74.1	0.74	0.47
EF, emotional fct	102	13.4***	27.1	25.5	52.0	0.28	0.15
CF, cognitive fct	104	14.4***	24.0	21.2	61.5	0.46	0.24
SF, social fct	99	−18.4***	35.9	21.2	44.5	0.23	0.15
FA, fatigue	107	−7.0**	18.2	39.3	58.9	0.54	0.37
NV, nausea/vomit	108	−9.3***	16.0	43.5	73.1	0.76	0.54
PA, pain	102	−9.6***	20.4	35.3	60.8	0.67	0.47
Dyspnoea	111	−15.3***	20.7	55.0	84.7	0.76	0.51
Insomnia	110	−13.0***	28.8	38.2	77.3	0.38	0.23
Appetite loss	109	−9.8**	24.5	43.1	86.2	0.57	0.40
Constipation	111	−8.7***	16.5	61.3	90.1	0.74	0.60
Diarrhoea	108	−14.5***	21.3	62.0	82.4	0.40	0.33

a Mean difference between patient and physician assessments. For functional scales a positive difference reflects that the physicians estimate better functioning than the patients do while for the symptom measures a positive difference reflects that the physicians estimate more severe symptoms. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$, for a Wilcoxon signed rank test.
b Absolute mean difference between patient and physician assessments.
c Percent exact agreements between patients and physicians.
d Patient–physician assessments differing at most one score level. One score level is 33.3 points for the single items measures and defined as 16.7 points for the scales corresponding to one level for a 2-item scale.
e Correlation between patient and physician assessments.
f Weighted kappa measure of agreement between patients and physicians.

tive and emotional functioning were close to horizontal indicating that physician ratings were almost independent of patient ratings. The physicians generally overestimated the cognitive and emotional functioning of the patients most pronounced for patients with very poor functioning. For the physical and social functioning scales, the physicians generally underestimated the patients' functioning. The better the physical and social functioning reported by the patients, the poorer did the physicians agree with the patients. The agreements across the range of patient reported scores were generally better for the three symptoms scales, FA, PA, and NV, than for the functional scales. The general trend was that the physician scores increased with increasing patient scores. Nevertheless, all three scales showed large deviations (>20) for some levels of symptomatology. The same trend was seen for the five single-item symptom measures although less distinct. The lines were more horizontal and below the diagonal reflecting moderate to severe underestimation of the symptoms, especially for patients reporting 'very much'.

For the evaluation of agreement *after* T0 (1–13 weeks after admission to the department), a total of 263 patient–physician pairs were available from 107 patients. The mean differences and absolute differences observed at T1–T13 deviated less than two points from those observed at T0. The percentages of exact and close to exact agreements at T1–T13 deviated less than 4% from those at T0, and on average the kappas at T1–T13 were 0.1 smaller than the kappas at T0. Thus, overall the level of agreement observed at T1–T13 was similar to or slightly poorer than the agreement at T0.

4. Discussion

Patients may be cognitively impaired, deny symptoms and problems, or exaggerate the level of symptoms. Despite this, it is generally agreed that the patients are the most valid source of information about their quality of life.[4] Therefore, we viewed the patient reported scores as the 'gold standard' to which the physician assessments should be compared. That is, deviations between the physician and patient assessments would indicate that using physician assessments might result in erroneous conclusions (this point is, however, debated later in this section).

In this study, the agreement between patients and physicians was far from optimal. Significant bias was observed for all domains assessed. The mean difference between patient and physician scores ranged from −15 to 18 with an average of 12 points across the domains (results from the first assessment). Further, plots indicated that the bias from using physician ratings would be severe (>20 points) for some levels of functioning/symptomatology for all domains. A difference of 10 points has been suggested as a clinically relevant difference.[28] These findings indicate that using physician ratings may result in assessments that are both statistically and clinically different from the palliative care patients' own assessments of their health-related quality of life.

The agreements were, however, not equally poor across the domains. The poorest agreements were observed for the 'psychosocial scales' social and emotional functioning while the best agreements were observed for nausea/vomiting and constipation. This is in line with previous findings.[17] The correlations for nausea/vomiting and constipation were 0.76 and 0.74, respectively, indicating relatively good agreement. But plots revealed that for patients reporting severe symptoms (score = 100), the physicians underestimated the level of symptoms with more than 25 points on average. That is, even for the domains with the best agreement there were pronounced deviations for important subgroups of patients. This is a good example of how the evaluation methods often supplemented each other; a lack of agreement overlooked by one method was elucidated by another. The agreements found at the patients' first contact with the

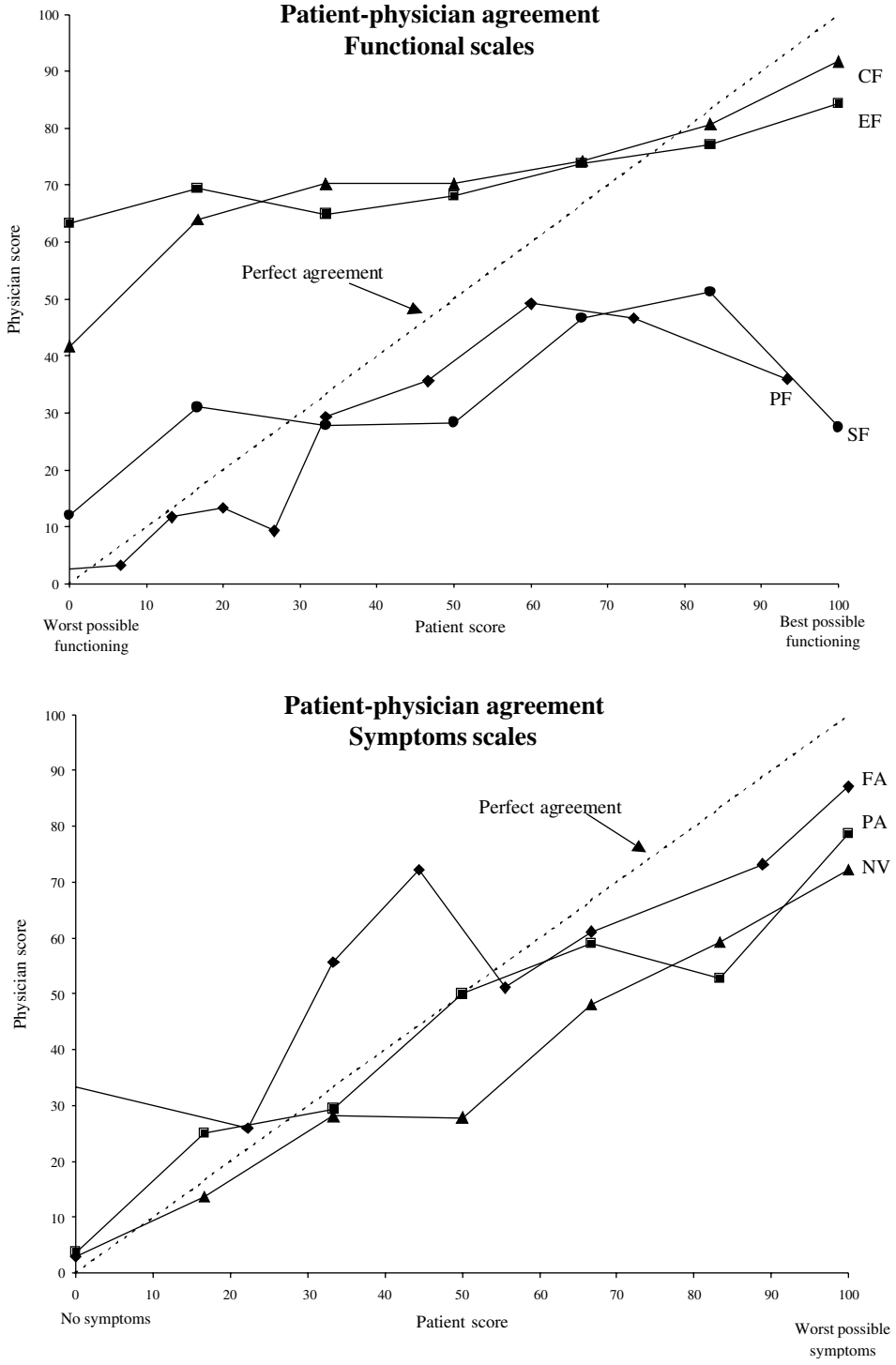

Fig. 1 – The average physician scores for each level of the patient scores at the patients' first contact with the department. *Note.* The broken lines represent perfect average agreement between physicians and patients. PF, physical functioning; EF, emotional functioning; CF, cognitive functioning; SF, social functioning; FA, fatigue; NV, nausea/vomiting; PA, pain.

department were confirmed by the analyses of data from week 1 to week 13.

It should be recognised that although the patients are considered the most valid source of information about their quality of life, patient assessments are subjective. For example, what one patient regards as a severe problem may be considered only a little problem by another patient. This is why it is preferable that patients assess their problems themselves. However, physicians may provide a broader, more objective evaluation of the patients' problems and symptoms, which may provide important, additional information. Therefore, even though our findings suggest that physician assessments cannot replace patient assessments, physician assessments may still have a value as a supplement to patient

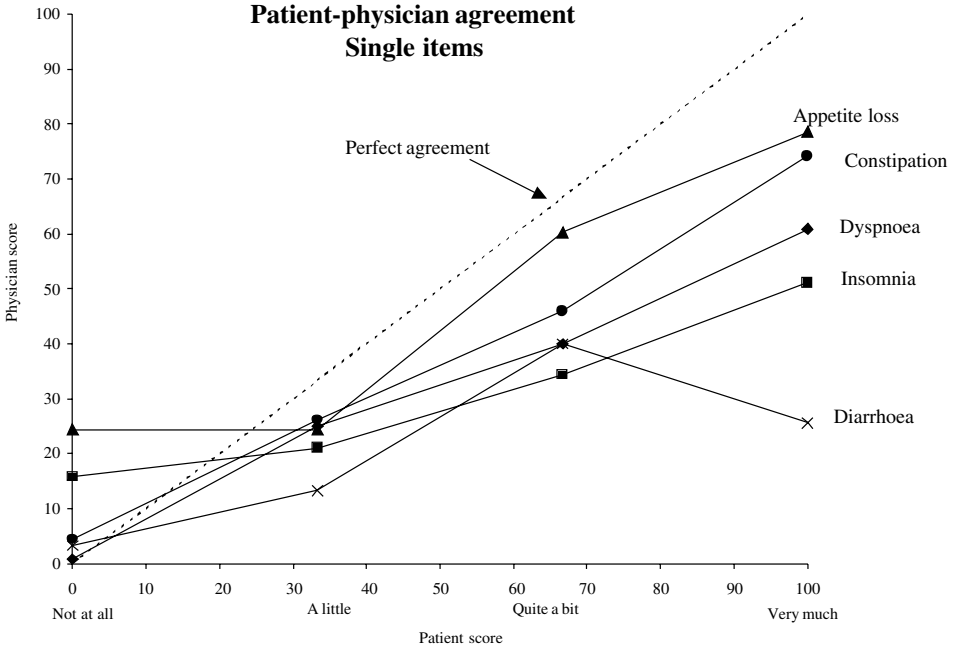

Fig. 1 – *continued*

assessments. Furthermore, physician assessments may still replace patient assessments in palliative care if the physician assessments can be adjusted to remove the bias. Adjusting scores seems particularly promising if the deviations are systematic. Whether agreement can be improved by adjusting the physician assessments is a relevant topic for future research.

Analyses showed that the main difference between participants and non-participants was the patient's affiliation to the department; inpatients more often participated than other patients. We did not find any significant differences in the agreement across the affiliations. This indicates that our findings of agreement also apply to the non-participants. That is, there is no reason to believe that better agreement would have been found had all patients participated.

Previous studies have found poor agreement between patients and physicians in palliative care.[7,16] Our study indicated even poorer agreement than that found in those studies. One can speculate whether the physicians in our study have not been sufficiently motivated or whether the time available for the assessments may have been too limited. However, we have no reason to believe that the situation in our department is different from other departments of palliative medicine.

In another cohort of patients, we have previously investigated whether the symptoms and problems reported by patients in self-assessment questionnaires were covered by the medical records[29] or the nursing records.[30] To our surprise we found considerable discrepancies: most symptoms or problems reported by patients were only mentioned by physicians or nurses in less than half of the cases. Many of these were potentially treatable symptoms/problems. In line with this, we found in the present study that the physicians underestimated the symptoms/problems for 10 of the 12 domains. This may reflect that physicians tend to overlook problems/symptoms that are not obvious or not mentioned explicitly by the patients. More systematic screening for symptoms may improve the physicians' insight into the patients' symptoms. To our knowledge, it has never been investigated whether there is any association between patient–physician agreement and the adequacy of the care. However, it has been found that physician access to patient ratings improved communication and physician awareness of the patients' problems.[31,32]

In summary, our results showed that even though physicians are specialised in palliative care and see palliative care patients on a daily basis, their perception of the patients' health-related quality of life was quite different from the patients' own view. Based on these results and previous findings[7,16] physician assessments cannot be recommended as a substitute for patient self-assessment in palliative care.

Conflict of interest statement

None declared.

Acknowledgement

The study was supported by the Danish Cancer Society, Grant PP 03 013.

REFERENCES

1. Johnston G, Abraham C. The WHO objectives for palliative care: to what extent are we achieving them? *Palliat Med* 1995;**9**(2):123–37.
2. McWhinney IR, Bass MJ, Donner A. Evaluation of a palliative care service: problems and pitfalls. *BMJ* 1994;**309**(6965):1340–2.

3. Rinck GC, van den Bos GA, Kleijnen J, de Haes HJ, Schade E, Veenhof CH. Methodologic issues in effectiveness research on palliative cancer care: a systematic review. *J Clin Oncol* 1997;**15**(4):1697–707.
4. Higginson IJ. Can professionals improve their assessments? *J Pain Symptom Manage* 1998;**15**(3):149–50.
5. 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**(7):743–60.
6. Sneeuw KC, Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *J Clin Epidemiol* 2002;**55**(11):1130–43.
7. Brunelli C, Costantini M, Di Giulio P, Gallucci M, Fusco F, Miccinesi G, et al. Quality-of-life evaluation: when do terminal cancer patients and health-care providers agree. *J Pain Symptom Manage* 1998;**15**(3):151–8.
8. Curtis AE, Fernsler JI. Quality of life of oncology hospice patients: a comparison of patient and primary caregiver reports. *Oncol Nurs Forum* 1989;**16**(1):49–53.
9. Field D, Douglas C, Jagger C, Dand P. Terminal illness: views of patients and their lay carers. *Palliat Med* 1995;**9**(1):45–54.
10. Higginson I, Wade A, McCarthy M. Palliative care: views of patients and their families. *BMJ* 1990;**301**(6746):277–81.
11. Higginson IJ, McCarthy M. Validity of the support team assessment schedule: do staffs' ratings reflect those made by patients or their families? *Palliat Med* 1993;**7**(3):219–28.
12. Kristjanson LJ, Nikoletti S, Porock D, Smith M, Lobchuk M, Pedler P. Congruence between patients' and family caregivers' perceptions of symptom distress in patients with terminal cancer. *J Palliat Care* 1998;**14**(3):24–32.
13. Lobchuk MM, Kristjanson L, Degner L, Blood P, Sloan JA. Perceptions of symptom distress in lung cancer patients: I. Congruence between patients and primary family caregivers. *J Pain Symptom Manage* 1997;**14**(3):136–46.
14. McCusker J, Stoddard AM. Use of a surrogate for the Sickness Impact Profile. *Med Care* 1984;**22**(9):789–95.
15. Spiller JA, Alexander DA. Domiciliary care: a comparison of the views of terminally ill patients and their family caregivers. *Palliat Med* 1993;**7**(2):109–15.
16. 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.
17. McPherson CJ, Addington-Hall JM. Judging the quality of care at the end of life: can proxies provide reliable information? *Soc Sci Med* 2003;**56**(1):95–109.
18. Stromgren AS, Goldschmidt D, Groenvold M, Petersen MAa, Jensen PT, Pedersen L, et al. Self-assessment in cancer patients referred to palliative care: a study of feasibility and symptom epidemiology. *Cancer* 2002;**94**(2):512–20.
19. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, 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**(5):365–76.
20. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 2002;**324**(7351): 1417–1419.
21. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. *The EORTC QLQ-C30 Scoring Manual*. Brussels: European Organisation for Research and Treatment of Cancer; 2001.
22. Bjorner JB, Petersen MAa, Groenvold M, Aaronson N, Ahlner-Elmqvist M, Arraras JI, et al. Use of item response theory to develop a shortened version of the EORTC QLQ-C30 emotional function scale. *Qual Life Res* 2004;**13**(10):1683–97.
23. Karnofsky DT, Abelman WH, Craver LF, Burchenal JH. The use of the nitrogen mustards in the palliative treatment of carcinoma – with particular reference to bronchogenic carcinoma. *Cancer* 1948;**1**(4):634–56.
24. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;**12**(3):189–98.
25. Hansen EJ. Inequality in the welfare state. In: Erikson R, Hansen EJ, Ringen S, Uusitalo H, editors. *The Scandinavian Model: Welfare States and Welfare Research*. New York: M.E. Sharpe; 1987.
26. Stephens RJ, Hopwood P, Girling DJ, Machin D. Randomized trials with quality of life endpoints: are doctors' ratings of patients' physical symptoms interchangeable with patients' self-ratings? *Qual Life Res* 1997;**6**(3):225–36.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**(1):159–74.
28. Osoba D. A taxonomy of the uses of health-related quality-of-life instruments in cancer care and the clinical meaningfulness of the results. *Med Care* 2002;**40**(Suppl. 6):III31–=0?>III38.
29. Stromgren AS, Groenvold M, Pedersen L, Olsen AK, Spile M, Sjogren P. does the medical record cover the symptoms experienced by cancer patients receiving palliative care? A comparison of the record and patient self-rating. *J Pain Symptom Manage* 2001;**21**(3):189–96.
30. Stromgren AS, Groenvold M, Sorensen A, Andersen L. Symptom recognition in advanced cancer. A comparison of nursing records against patient self-rating. *Acta Anaesthesiol Scand* 2001;**45**(9):1080–5.
31. Detmar SB, Muller MJ, Schornagel JH, Wever LD, Aaronson NK. Health-related quality-of-life assessments and patient–physician communication: a randomized controlled trial. *JAMA* 2002;**288**(23):3027–34.
32. Velikova G, Brown JM, Smith AB, Selby PJ. Computer-based quality of life questionnaires may contribute to doctor–patient interactions in oncology. *Br J Cancer* 2002;**86**(1): 51–9.