

Evaluation and comparison of two prognostic scores and the physicians' estimate of survival in terminally ill patients

S. Stiel · L. Bertram · S. Neuhaus · F. Nauck ·
C. Ostgathe · F. Elsner · L. Radbruch

Received: 16 December 2008 / Accepted: 23 March 2009 / Published online: 21 April 2009
© Springer-Verlag 2009

Abstract

Background Most terminally ill patients request information about their remaining life span. Professionals are not generally willing to provide prognosis on survival, even though they are expected to be able to do so from their clinical experience. This study aims to find out whether the standardized instruments Palliative Prognostic Index (PPI) and the Palliative Prognostic Score (PaP-S) are appropriate, specific, and sensitive to estimate survival time in patients receiving inpatient palliative care in Germany.

Method PPI and PaP-S were assessed in addition to the core documentation data set of the Hospice and Palliative Care Evaluation for patients admitted to the palliative care units in Aachen, Bonn, and Cologne. Time of survival was assessed with repeated phone calls to the family and was defined as the difference between the day of completion of the instruments (excluded) and the day of death (included).

Results Survival time was compared with physicians' estimations and prognostic scores in 83 patients. Whereas the estimates of the PPI and the PaP-S correlate highly, even higher correlations are found for the physicians' prognosis and the scores. Correlations between survival time and the prognostic scores or physicians' prognosis were lower. Physicians' estimations overestimated survival time on average fourfold. Estimations were more often correct for very good and very bad prognosis.

Discussion The prognostic scores are not able to produce a precise reliable prognosis for the individual patient. Nevertheless, they can be used for ethical decision making and team discussions. Estimating survival time from clinical experience seems to be easier for very bad or very good prognosis for physicians.

Keywords Palliative care · Prognostication · Prognostic scores · Terminally ill · Survival time · Physicians' estimation · Palliative Prognostic Index (PPI) · Palliative Prognostic Score (PaP-S)

F. Elsner and L. Radbruch contributed equally to the study and share responsibility for the workgroup.

S. Stiel (✉) · L. Bertram · S. Neuhaus · F. Elsner · L. Radbruch
Department of Palliative Medicine, RWTH Aachen University,
Pauwelsstraße 30,
52074 Aachen, Germany
e-mail: sstiel@ukaachen.de

F. Nauck
Department of Palliative Medicine,
University Hospital Göttingen,
Göttingen, Germany

F. Nauck
Palliative Care Center, Malteser Hospital Bonn,
Bonn, Germany

C. Ostgathe
Department of Palliative Medicine, University Hospital Cologne,
Cologne, Germany

Introduction

Prognostication, along with diagnosis and treatment, is one of the most central issues in palliative care [1] and is a key factor in decision making for caregivers, patients, and their families in all medical settings. "One of the key roles for palliative care services is to initiate discussions on prognosis and goals of care" [1]. On the other hand, prognostication in palliative care may be challenged as irrelevant for patients with restricted life expectancy or even as unethical, as communication of a definite and limited time frame may burden the patient.

However, indications of specific treatments, choice of medication, and decisions on care and discharge planning are usually

based on a balance of potential burden and benefit of available options [2]. This balance may shift considerably in relation to the estimated survival time. Estimation of survival time and prognostic uncertainty were named as major factors in decision making on antibiotic treatment in palliative care [3].

The rationale of the estimations by health care professionals is rarely defined, even though it forms the basis of these important decisions. In a study by Fallowfield et al., 88% of terminally ill patients ask for information about their remaining life span while less than 8% are willing to receive only good or no detailed information [4]. The question “How long do I have?” addresses patients’ needs and wishes for the residual life span, such as setting up a proxy for decision making, putting financial and social affairs in order, or organizing funeral details [5].

Communicating prognostic information may raise ethical problems and cause tense dilemmas. It is challenging for physicians to balance between realistic information without increasing anxiety and reducing trust, on one hand, and allowing hope and optimism without creating unrealistic expectations, on the other hand [6].

Professionals are not generally willing to provide prognostic information and do not feel adequately prepared for this task, even though they are expected to know from their clinical experience. More than half feel poorly trained and report difficulty in formulating and communicating prognosis to patients [1].

Some studies on prognostication report mean ratios of predicted to observed survival time with a factor of 3–5 [7], showing overoptimistic estimations of the physicians. Sex, years of medical practice, board certification, cultural background, or self-rated optimism have been investigated as possible bias toward optimism, but no homogenous results could be found [4, 7].

To improve assessment of prognosis and alleviate communication with terminally ill patients, several workgroups have developed prognostic scoring systems on survival time in palliative care [1, 8]. To our knowledge, these scores have not been tested in a head-to-head comparison yet. There is also no information whether the prognostic scores can be used in German palliative patients.

This study compared the Palliative Prognostic Index (PPI) and Palliative Prognostic Score (PaP-S) with physicians’ estimation and survival time in palliative care patients in Germany and investigated how appropriate, specific, and sensitive the prognostic scores were in this setting.

Materials and methods

Study plan

The two prognostic scores PPI and PaP-S were assessed by experienced physicians as part of the Hospice and Palliative

Evaluation (HOPE) in 2005 [9, 10]. HOPE is a standardized documentation system with a core documentation form and a range of modules, among them is the prognostic module with the two prognostic scores. HOPE is used in a nationwide survey with a 3-month period of recruitment each year, as a quality assurance program for inpatient palliative care units in Germany [11]. Thirty patients recruited consecutively to each of the palliative care units in Aachen, Bonn, and Cologne in the evaluation period from the 15th of June to the 15th of October 2005 were included in this survey. Additional information was collected on survival times for these patients. The departments in Aachen (nine beds) and Bonn (nine beds) are run by physicians with a subspecialty in palliative medicine and a specialist degree in anesthesiology; the department in Cologne (11 beds) is run by a physician with a subspecialty in palliative medicine and a specialist degree in neurology.

Informed consent was not obtained from the patients, as the assessment was an additional part of the regular documentation and quality assurance program of the units. Therefore, an evaluation of the ethics committee was not necessary. Nevertheless, an information sheet for the patients and their families was provided to explain the background and method of the study. All data were entered anonymously in an electronic database.

Evaluation of the survival time of patients was essential for the research project. For patients treated as inpatients until the time of death, the survival time could be calculated from the clinical documentation. Patients discharged home to hospice or nursing home were contacted repeatedly, and time of death was elicited by phoning families, general practitioners, nursing home staff, and other relevant carers until June 2008. Phone calls were included if possible in the regular follow-up procedure for discharged patients. Survival time was defined as the difference between the day of completion of the instruments (excluded) and the day of death (included).

Prognostic scores

The PaP-S [12] allows identification of a subgroup of patients with advanced disease and very bad prognosis. After validating the PaP-S in a multicenter study with Italian palliative care patients, the rating system was modified in relation to the empirical results [13]. The PaP-S contains six items: dyspnea, anorexia, the Karnofsky performance status, clinician prediction of survival, total white blood count, and lymphocyte percentage. Although the clinicians’ prognosis is imprecise, it is included as an important parameter in the final assessment tool. The sum score of all items in a range from 0 to 17.5 points is separated into three categories: category 3 (0–5.5 points) predicts a probability higher than 70% of at least 30 days

survival time, category 2 (5.6–11 points) predicts a probability between 30% and 70%, and finally, category 1 (11.1–17.5 points) predicts a chance lower than 30% of the 4-week survival time. The PaP-S was validated for Italian hospice–home care patients in a multicenter study and successfully used for hospitalized patients with terminal cancer, advanced cancer, and various end-stage noncancer illnesses.

The PPI has been constructed in extension of the Palliative Performance Scale and contains five variables: the performance status, oral intake, edema, dyspnea at rest, and delirium [8]. The performance status is a modification of the Karnofsky index. A performance status score lower than 50% predicts a chance of survival of 10% for 6 months [14]. The sum score of the PPI ranges from 0 to 15 points. Sum scores above six points (category 1) are associated with a 20% chance of survival for 3 weeks; a sum score above four points (category 2) is equivalent to a chance of survival of 20% for 6 weeks. Validated in a Japanese palliative care inpatient unit, the PPI predicted survival times with a sensitivity of 91% and a specificity of 67% [8]. For this study, sum scores lower than four points were used as a third category to test whether the survival time of this group of patients is significantly longer than in the other two categories.

The *physicians' estimation of survival time* was documented in the PaP-S, as this instrument includes an item requiring physicians to estimate survival time on a six-step categorical scale (1–2, 3–4, 5–6, 7–8, 9–10, 11–12, or more than 12 weeks).

Sensitivity (right positive/(right positive + false negative)) and specificity (right negative/(right negative + false positive)) was calculated for each PPI category as well as for PaP-S categories 1 and 3 separately.

The positive predictive value as the amount of patients with a poor prognosis (category 1) who died within the given timeframe and the negative predictive value as the amount of patients with a good prognosis (category 3) who survived the given timeframe are given for both prognostic instruments.

Empirical analyses were carried out using SPSS for Windows, version 14.0.

Results

Study population

A total of 90 patients were recruited. Data on 83 patients were used for analysis; seven data sets were incomplete. The palliative care unit in Aachen included 30%, Bonn 35%, and Cologne 35% of the analyzed study population.

Mean patient age was 63 years (standard deviation (SD) 13 years, range 37 to 87 years), with 53% female and 47% male patients. Patients had a mean survival time of 50 days (SD = 98, range 0 to 737 days). Main diagnoses are shown in Table 1.

The PaP-S assigned 19% of all patients to category 1 (poor prognosis), 27% to category 2 (intermediate prognosis), and 47% to category 3 (good prognosis); 7% were missing values. No patient in category 1 (standard less than 30% survivors), 32% in category 2 (standard 30–70%), and

Table 1 First diagnosis (ICD) of the study population

Main diagnosis	Number of patients	Percentage
Malignant neoplasms		
Digestive organs	19	23
Respiratory and intrathoracic organs	16	19
Breast	9	11
Lymphoid, hematopoietic, and related tissue	7	9
Female genital organs	7	9
Ill-defined, secondary, and unspecified sites	4	5
Male genital organs	4	5
Urinary tract	3	4
Skin	3	4
Thyroid and other endocrine glands	2	2
Mesothelial and soft tissue	2	2
Lip, oral cavity, and pharynx	2	2
Eye, brain, and other parts of central nervous system	1	1
Others		
Diseases of the circulatory system	2	2
Diseases of the respiratory system	2	2
Total	83	100

54% in category 3 (standard more than 70%) survived 30 days.

The PPI assigned 29% of all patients to category 1 (poor prognosis), 19% to category 2 (intermediate prognosis), and 46% to category 3 (good prognosis); 6% were missing values. Eight percent of the patients in category 1 survived for more than 3 weeks (standard 20%), 12% of those in category 2 (standard 20%), and 50% in category 3 (standard >20%) survived for more than 6 weeks. Data on sensitivity/positive rate and specificity/negative rate can be seen in Table 2.

The positive predictive value as the amount of patients with a poor prognosis (category 1) who died within the given timeframe of (a) 30 days of the PaP-S reaches 100% (16 out of 16) and (b) 3 weeks of the PPI reaches 92% (22 out of 24). The negative predictive value as the amount of patients with a good prognosis (category 3) who survived the given timeframe of (a) 30 days of the PaP-S reaches 54% (21 out of 39) and (b) 6 weeks of the PPI reaches 50% (19 out of 38).

The survival functions (Kaplan–Meier) as the decreasing cumulative survival of patients in dependence on lifetime is pictured separately for the scoring categories 1: bad, 2: moderate, and 3: good prognosis of (a) the PaP-S and (b) the PPI in Fig. 1a, b. Chi-square for the distribution of survival (Mantel–Cox) for the curves of the different score categories was highly significant ($p < 0.01$) for both prognostic instruments.

The prognosis scores PaP-S and PPI correlated highly ($r_{(\text{PaP-S}; \text{PPI})} = 0.704$), as well as physicians' estimation with the instruments scores: $r_{(\text{PPI}; \text{physician})} = 0.726$ and $r_{(\text{PaP-S}; \text{physician})} = 0.857$. However, only moderate correlations were found between survival time and both scores as well as between survival time and physicians' estimation. For data on single correlations, please see Table 3. Number of data sets is varying because missing values in single items of one prognostic instrument prevent from calculating sum scores and the analysis of their correlations.

Physicians' estimation of survival time

Comparing the six categories of the PaP-S variable on the physician's prediction of survival with survival times 31% of the physicians' ratings were in the correct category, 5%

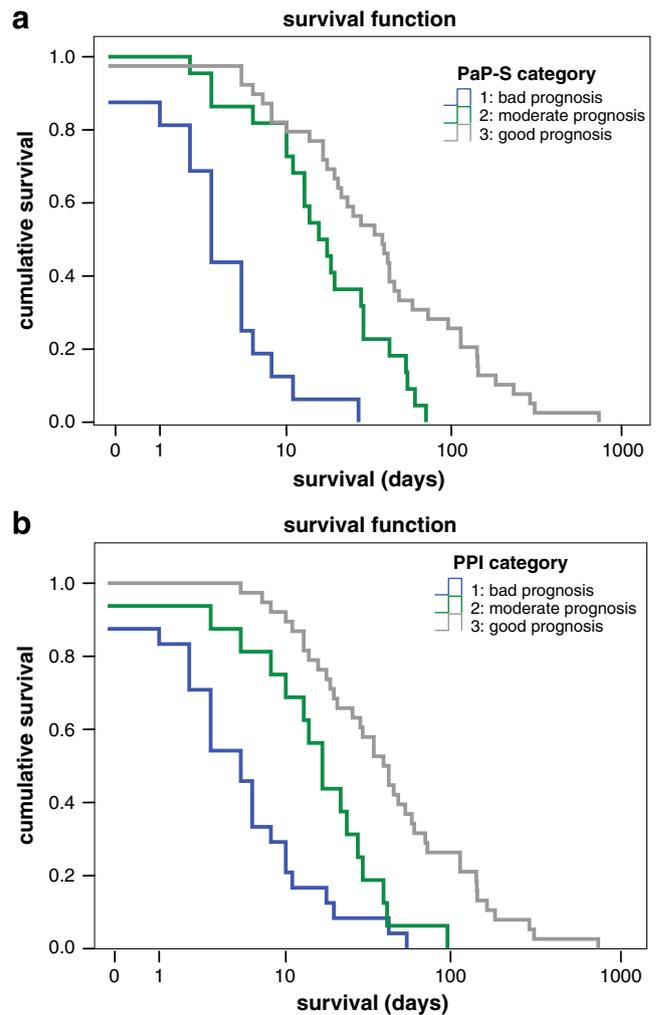


Fig. 1 **a** Survival function as the decreasing cumulative survival of patients in dependence on time (days) separately for the scoring categories 1: bad, 2: moderate, and 3: good prognosis of the PaP-S. **b** Survival function as the decreasing cumulative survival of patients in dependence on time (days) separately for the scoring categories 1: bad, 2: moderate, and 3: good prognosis of the PPI

were overpessimistic, and 63% were too optimistic; 1% were missing values. Physicians overestimated survival times with a mean factor of 4.2 (SD = 6.5; range 0.3–43). Overestimation was subject to a centre effect, as the mean factor was 5.2 in Aachen (52% overoptimistic and 36%

Table 2 Sensitivity and specificity of PaP-S and PPI categories 1, 2, and 3

	Category	Sensitivity	Specificity
PaP-S	1 (<30% chance for 30 days survival)	67%	100%
	2 (30–70% chance for 30 days survival)	–	–
	3 (>70% chance for 30 days survival)	78%	65%
PPI	1 (20% chance for 3-week survival)	51%	94%
	2 (20% chance for 6-week survival)	46%	84%
	3 (>20% chance for 6-week survival)	83%	66%

Table 3 Spearman rank correlations of prognostic scores PaP-S and PPI, survival of palliative patients, and physicians' estimate of survival

	PaP-S	PPI	Survival of palliative patients	Physicians estimate of survival (PaP-S)
PaP-S		0.704* (N=72)	-0.575* (N=77)	0.857* (N=77)
PPI	0.704* (N=72)		-0.678* (N=78)	0.726* (N=74)
Survival of palliative patients	-0.575* (N=77)	-0.678* (N=78)		-0.556* (N=79)

* $p < 0.01$

accurate hits), 3.7 in Bonn (62% overoptimistic und 35% accurate hits), and 3.9 in Cologne (72% overoptimistic and 24% accurate hits). Accurate hits were more frequent in categories 1 and 3 of the PaP-S (46% accurate hits each) than in category 2 (8% accurate hits). A similar distribution was found for the PPI with 56% accurate hits in category 1, 0% in category 2, and 44% in category 3.

Discussion

Discussion of survival time is one of the most sensitive and ethically challenging issues in palliative care. Patients may be torn between anxiety and fear on one side and hope for survival or a chance to manage their remaining life span despite the worst prognosis on the other side.

Relatives and carers may need accurate information on prognosis, as important decisions may depend on it. As an example from our experience, the son of a patient was asking for the expected time of survival, as he wanted to use his accumulated holiday leave for the time his father died, and was afraid that if he took his leave too soon, he would have to work at the time of his father's death. Other examples include a visiting relative from Canada asking whether she should reschedule her return flight in 3 weeks time, or a patient asking whether he should cancel the holiday trip planned in 2 months.

Information on prognosis influences the patients' own conceptions of their future and the ideas on maintaining quality of life for the remaining lifetime as well as their choices in medical technologies and treatments. Patients prefer an "individualized and realistic disclosure from a confident and supportive cancer specialist" [6] which is understood as hope giving. Physicians, on the other hand, are reluctant to communicate prognosis unless asked and try to be unspecific, but optimistic [5]. These ideas are slightly contradictory. Additionally, there is a discrepancy between the formulated prognosis (as discussed in the team) and what is communicated to patient and family. However, there can be no doubt that it is the patient's right to make plans for the residual life span. In consequence, he can also expect a prognosis as precise as possible.

Currently, the estimation of survival at the palliative care units participating in this study depends mostly on "gut

feeling" as results of clinical observation and experience of the teams. Patients and their families get information about the prognosis as far as they want to know. However, as shown by the results of this study, estimations of survival times often were incorrect, and most were often too optimistic. This confirms other trials showing overoptimistic prognosis by clinicians. Christakis et al. reported that only 20% of predictions were accurate and 63% overoptimistic by a factor of 5.3 in the physicians' estimation of survival time [7]. Similar results are found by Gripp et al. who found 71–96% of estimates of survival on patients who died within 1 month to be overoptimistic [15].

Christakis et al. observed the effect that a stronger physician–patient relationship is associated with lower prognostic accuracy and that more experienced physicians were more accurate. Considering the conflict between experience and relationship, a trend toward second opinion among colleagues seems meaningful as a more neutral source of opinion without personal investigation into the outcome for the patient [7].

Prognostic scores would offer additional accuracy for prognostication and thus could facilitate ethical decision making and communication with patient and family. However, even if prognosis scores provide meaningful results for groups of patients with satisfactory specificity and sensitivity, results should be considered carefully for the individual patient, as considerable variation from default is possible.

To our knowledge, available prognostic scores have not been tested in a head-to-head comparison yet. Results from this study show no advantage between PaP-S and PPI. As the PaP-S and the PPI correlate highly in their prognostication, any one of them should be adequate for clinical use in addition to the physicians' estimation. As both only have a moderate correlation with survival time, clinicians should not rely only on them as the only or main instrument of prognosis. The PPI correlated better with survival time than the PaP-S; however, sensitivity for bad and moderate prognosis was lower than for the PPI.

Physicians' estimation of survival time reached the lowest correlation with the actual survival. This may be due to the form of this specific question, which was included as an item in the PaP-S. Physicians had to estimate life time in two weekly steps, which seemed to be difficult.

The single item on physician estimation correlated closely with both prognostic instruments. This leads to the conclusion that this single item may replace the test. However, the degree of overestimation in the single item and the low correlation with survival time call for additional input for prognostication.

Detailed analysis showed that clinicians were able to score poor or good prognosis more accurately than intermediate prognosis. In line with this finding, the PPI and the PaP-S were as well suited to predict which patients had the poorest survival, but they were less adequate to distinguish between good and intermediate prognosis. This points toward a huge uncertainty with intermediate prognosis. Prognostic assessment tools seem to be most needed for intermediate prognosis. However, in our study, both instruments were not suited for intermediate prognosis.

Regarding the limitations of scoring systems, Henderson et al. tried to establish a best statistical model for prediction which surprisingly included the clinicians' prediction as an additional risk factor as well, although not accurate [16].

However, the inclusion of the physicians' estimate of survival as on the items of the PaP-S seems to produce a dilemma. A prognostic index would be most useful for those occasions where the physician is uncertain of the patient's survival and relying on a score which give large weighting to the clinicians estimate may be a major drawback [5]. Following the results of our study, scoring systems on prognosis of survival time seem to be least useful for those patients [5] with neither a good nor a poor prognosis because here both physicians' estimates as well as the existing prognostic instruments lack precision.

Physicians' estimations may have been biased by the format of the item, requiring them to differentiate steps of 2 weeks survival time which demands a very specific estimate. Using broader categories might have resulted in less overestimation. However, clinical practice frequently requires prognostication with this accuracy, for example, if family leave, travel, or holiday schedules or care plans have to be realized.

Limitations of this study are a rather small sample size, and correlations may have been masked by outliers in the restricted sample. However, methodological considerations restricted the number of participating centers and the sample size should have been large enough to demonstrate the usefulness of the scores in clinical practice.

The parallel consideration of the physicians' estimate as an item of one of the scores (PaP-S), on one hand, and as a comparator for clinical prognostication, on the other hand, may have confounded the evaluation. Using another scale for the physicians' estimate of survival and not the small-step scoring system of the PaP-S item on the physicians' estimate could have given more insight into the derivation and determination of their clinical prognostication.

Our results demonstrate the need for more robust and better fitting prognostic tools that can be used in clinical palliative care practice to supplement the clinical estimation of survival. Whereas poor scores in both prognostic scales adequately identified patients with poor prognosis, patients with good scores often did not have longer survival times. Both scores are constructed to differentiate prognostic chances for several weeks. In clinical practice, it would be valuable to have a set of different scores with different time frames. More specifically, a sensitive prognostic score with a shorter time frame, allowing identification of patients in the final days of life, would be helpful for decision making in the terminal phase.

Prognostication with adequate sensitivity and specificity is only the first step in the care of patients with short life expectancy. Not only using prognosis carefully in ethical decision making but also communication strategies with patients and families on prognosis and on the best course of therapy have to supplement prognostication.

Using the available prognostic scores in clinical practice, with due consideration of the shortcomings, will train clinicians in these skills and will facilitate discussion and communication on prognosis.

References

1. Glare PA, Sinclair CT (2008) Palliative medicine review: prognostication. *J Palliat Med* 11:84–103. doi:10.1089/jpm.2008.9992
2. Liao S, Arnold RM (2008) Prognosticating: the end of a series. Introduction. *J Palliat Med* 11:82–83. doi:10.1089/jpm.2008.9993
3. Pestinger MOC, Bausewein C et al (2004) Antibiotica in der Palliativmedizin: Ergebnisse einer Fokusgruppe. *Z Palliativmedizin* 5:68–74. doi:10.1055/s-2004-828408
4. Fallowfield LJ, Jenkins VA, Beveridge HA (2002) Truth may hurt but deceit hurts more: communication in palliative care. *Palliat Med* 16:297–303. doi:10.1191/0269216302pm575oa
5. Stone PC, Lund S (2007) Predicting prognosis in patients with advanced cancer. *Ann Oncol* 18:971–976. doi:10.1093/annonc/mdl343
6. Hagerty RG, Butow PN, Ellis PM et al (2005) Communicating with realism and hope: incurable cancer patients' views on the disclosure of prognosis. *J Clin Oncol* 23:1278–1288. doi:10.1200/JCO.2005.11.138
7. Christakis NA, Lamont EB (2000) Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ* 320:469–472. doi:10.1136/bmj.320.7233.469
8. Morita T, Tsunoda J, Inoue S, Chihara S (1999) Survival prediction of terminally ill cancer patients by clinical symptoms: development of a simple indicator. *Jpn J Clin Oncol* 29:156–159. doi:10.1093/jjco/29.3.156
9. <http://www.hope-clara.de/download/HOPE2007Basisbogen.pdf>. Accessed September 2008
10. <http://www.hope-clara.de/download/HOPE2007Expose.pdf>. Accessed September 2008
11. Lindena G, Nauck F, Bausewein C et al (2005) Quality assurance in palliative medicine—results of the core documentation of 1999–2002. *Z Arztl Fortbild Qualitatssich* 99:555–565

12. Glare PA, Eychmueller S, McMahon P (2004) Diagnostic accuracy of the palliative prognostic score in hospitalized patients with advanced cancer. *J Clin Oncol* 22:4823–4828. doi:10.1200/JCO.2004.12.056
13. Pirovano M, Maltoni M, Nanni O et al (1999) A new palliative prognostic score: a first step for the staging of terminally ill cancer patients. Italian Multicenter and Study Group on Palliative Care. *J Pain Symptom Manage* 17:231–239. doi:10.1016/S0885-3924(98)00145-6
14. Anderson F, Downing GM, Hill J, Casorso L, Lerch N (1996) Palliative performance scale (PPS): a new tool. *J Palliat Care* 12:5–11
15. Gripp S, Moeller S, Bolke E et al (2007) Survival prediction in terminally ill cancer patients by clinical estimates, laboratory tests, and self-rated anxiety and depression. *J Clin Oncol* 25:3313–3320. doi:10.1200/JCO.2006.10.5411
16. Henderson R, Keiding N (2005) Individual survival time prediction using statistical models. *J Med Ethics* 31:703–706. doi:10.1136/jme.2005.012427