A systematic review of physicians' survival predictions in terminally ill cancer patients

Paul Glare, Kiran Virik, Mark Jones, Malcolm Hudson, Steffen Eychmuller, John Simes, Nicholas Christakis

ELPS

This is an abridged version; the full version is on bmj.com

Abstract

Objective To systematically review the accuracy of physicians' clinical predictions of survival in terminally ill cancer patients.

Data sources Cochrane Library, Medline (1996-2000), Embase, Current Contents, and Cancerlit databases as well as hand searching.

Study selection Studies were included if a physician's temporal clinical prediction of survival (CPS) and the actual survival (AS) for terminally ill cancer patients were available for statistical analysis. Study quality was assessed by using a critical appraisal tool produced by the local health authority.

Data synthesis Raw data were pooled and analysed with regression and other multivariate techniques.

Results 17 published studies were identified; 12 met the inclusion criteria, and 8 were evaluable, providing 1563 individual prediction-survival dyads. CPS was generally overoptimistic (median CPS 42 days, median AS 29 days); it was correct to within one week in 25% of cases and overestimated survival by at least four weeks in 27%. The longer the CPS the greater the variability in AS. Although agreement between CPS and AS was poor (weighted κ 0.36), the two were highly significantly associated after log transformation (Spearman rank correlation 0.60, $P < 0.001$). Consideration of performance status, symptoms, and use of steroids improved the accuracy of the CPS, although the additional value was small. Heterogeneity of the studies' results precluded a comprehensive meta-analysis.

Conclusions Although clinicians consistently overestimate survival, their predictions are highly correlated with actual survival; the predictions have discriminatory ability even if they are miscalibrated. Clinicians caring for patients with terminal cancer need to be aware of their tendency to overestimate survival, as it may affect patients' prospects for achieving a good death. Accurate prognostication models incorporating clinical prediction of survival are needed.

Introduction

"How long do I have, doctor?" is a central question for patients with far advanced, incurable illnesses.[1] Several studies, however, have suggested that doctors are inaccurate and overly optimistic when predicting the survival of patients with terminal cancer.[2–4]

Do doctors overestimate or underestimate the survival of terminally ill cancer patients on average? How reliable are doctors in estimating survival? Do doctors' estimates of survival provide information above and beyond prognostic or risk factor models for outcome? We obtained individual patient data from studies identified by a systematic search strategy and did a meta-analysis to answer these questions.

Methods

We searched Ovid Premedline (Jan 2001) and Medline (1966-2000), Embase, Current Contents, Cochrane Library, and Cancerlit databases on 19 January 2001. See bmj.com for details of the search strategy.

We obtained potential papers to see if they met the following preset selection criteria: *(a)* the study involved patients with far advanced cancer; *(b)* the results section included a temporal survival prediction, given in days or weeks, made prospectively for each patient by a doctor; *(c)* the results section provided the patients' individual survival durations; and *(d)* the methods section provided an explanation of how the date of death was determined. If the raw data for clinical prediction of survival (CPS) and actual survival (AS) were not retrievable from the publication directly we contacted the authors to obtain them. We excluded papers if these data were neither retrievable from the publication nor obtainable from the authors.

Three of us then re-read studies selected for inclusion in the review and independently evaluated them for their quality by using the *Method for Evaluating Research and Guideline Evidence* (MERGE) guide for critical appraisal.[5] MERGE incorporates a four point coding system for appraising the quality of a study and scoring the risk of bias from "A" (low) through "B1" and "B2" to "C" (high).

The individual patient data for CPS and AS could be abstracted directly from papers if they were presented as a table or scatter plot. When CPS and AS were presented in summarised form we sought the individual patient data from the authors.

Results

Study characteristics

The electronic search produced 22 citations, yielding six papers of apparent relevance. The hand search identified 11 other studies. After reading these 17 papers, we excluded five that did not meet the inclusion criteria.

Data on 1594 patients (80.3% of total) were available for analysis from eight of the studies, and we entered these into the meta-analysis.[2 3 6 7–10 11] Because some individual CPS or AS data were missing, 1563 complete CPS-AS dyads were available for analysis. We extracted individual patient data from tables or figures of four studies (n=296),[2 9 10 11] and the authors generously provided us with their original data for the other four (n=1280). Four studies were from the United Kingdom,[2 9 10 11] three were from Italy,[6–8] and one was from the United States (table 1).[3] Two studies involved referring doctors,[7 10] and the rest involved

Department of Palliative Care, Royal Prince Alfred Hospital, Camperdown, NSW 2050, Australia
Paul Glare *head of department*
Kiran Virik *research fellow*

NHMRC Clinical Trials Centre, University of Sydney, Sydney, Australia
Mark Jones *biostatistician*
John Simes *director*

Department of Statistics, Macquarie University, Sydney
Malcolm Hudson *professor*

Department of Palliative Care, Kantonsspital, St Gallen, Switzerland
Steffen Eychmuller *medical director*

Department of Health Care Policy, Harvard Medical School, Boston, MA, USA
Nicholas Christakis *professor*

Correspondence to:
P Glare
paul@email.cs.nsw.gov.au

BMJ 2003;327:195–8

Table 1 Summary of the eight studies included in the systematic review

Study	Quality rating*	No of sites/doctors	Individual patient data	Median (IQR) CPS (days)	Median (IQR) AS (days)	Rank correlation	Weighted κ
(1) Parkes, 1972[3]	C	1/?	71	28 (24-56)	21 (9-34)	0.49	0.31
(2) Evans, 1985[15]	C	1/6	42	81 (28-182)	120 (43-180)	0.69	0.40
(3) Heyse-Moore,1987[16]	C	1/?	50	56 (33-84)	14 (7-28)	0.26	0.06
(4) Maltoni, 1994[13]	B1	1/4	100	42 (28-56)	32 (13-63)	0.60	0.34
(5) Maltoni, 1995[8]	B1	22/?	530	42 (28-70)	32 (13-62)	0.70	0.44
(6) Oxenham, 1998[21]	C	1/5	21	21 (14-35)	15 (9-25)	0.73	0.52
(7) Maltoni, 1999[14]	B1	14/?	451†	42 (21-70)	33 (14-62)	0.70	0.44
(8) Christakis, 2000[4]	B1	5/343	326	77 (28-133)	24 (12-58)	0.50	0.25
Overall	—	—	1591	42 (28-84)	29 (13-62)	0.60	0.36

AS=actual survival; CPS=clinical prediction of survival; IQR=interquartile range; ?=number of clinicians making predictions either not published or not known.
*According to MERGE document: B1=low-moderate risk of bias; C=high risk of bias.[6]
†Included 36 patients who were censored (that is, still alive).

"receiving" doctors (palliative care specialists). Three studies involved patients in hospital,[2 10 11] and the rest involved patients being cared for at home. All studies involved patient populations that were heterogeneous for the primary cancer site.

Quantitative data synthesis

Validity assessment

All eight studies were assessed as being biased (selection biases and misclassification biases), with half being at a high risk. Selection biases included a narrow spectrum of patients and failure to use an inception cohort. Misclassification biases included the timing of the prediction in relation to recruitment, variations in clinical experience of the doctor making the prediction, access to other clinical information when making the prediction, and involvement of the predicting doctor in providing ongoing care to the patient.

Simple summary results

When all 1563 evaluable CPS-AS dyads were pooled, the median CPS was 42 days and the median AS was 29 days (table 1), a difference of 13 days. Overall, CPS was correct to within one week in 25% of cases, correct to within two weeks in 43%, and correct to within four weeks in 61%. CPS overestimated AS by at least four weeks in 27% of cases and underestimated it by at least four weeks in 12% of cases. Although the level of agreement between CPS and AS was only fair (weighted κ 0.36), the log transformation of CPS was significantly correlated with the log transformation of AS (Spearman rank correlation 0.60, $t_{1540}=32.3$, $P<0.001$).

Statistical aggregation

The patients in study 2 survived much longer than the patients in the other seven studies, and studies 3 and 6 had the shortest survivals. With the exception of study 2, CPS consistently overestimated AS. A lack of uniformity in the median difference between AS and CPS is apparent (figure). Because of the strong indication of heterogeneity, combining the data of the eight studies for extensive statistical analysis was not appropriate, which limited the aim of doing a comprehensive meta-analysis.

Modelling CPS and other prognostic factors

In the subset of 981 patients with data for multiple prognostic variables, log(CPS) was statistically significantly correlated with log(AS). The R square value of 0.51 indicates that greater than 50% of the variation in log(AS) was explained by log(CPS). Next, we generated a model based on 15 patient based prognostic factors. Using backwards elimination we found that palliative steroid use, anorexia, dyspnoea, and log(KPS) all contributed additional value to log(CPS) when predicting log(AS), but the additional value was small (R square 0.54).

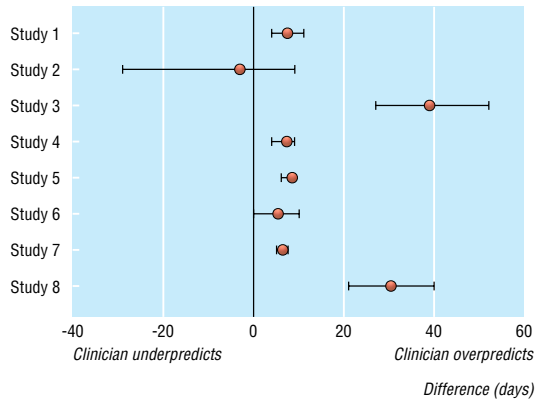
Prediction of AS according to health status

We repeated the models described above with the patients divided into three subgroups based on Karnofsky performance status scores: <40 (n=330), 40-50 (n=457), and ≥60 (n=194). For each model, log(CPS) explains more of the variation in log(AS) as the patient becomes sicker (table 2). The additional value provided by the other prognostic factors (anorexia, dyspnoea, steroid use) changes little, irrespective of how poor the patient's performance is.

Discussion

Statement of principal findings

Doctors' predictions for terminally ill cancer patients (a population very close to death with a median survival of approximately four weeks) were inaccurate. Doctors consistently overestimated the duration of survival in seven of the eight studies. Despite being inaccurate, clinical predictions are clinically useful; clinical prediction of survival (CPS) and actual survival (AS) were



Difference between actual survival and clinical prediction of survival for terminally ill cancer patients (median and 95% confidence interval)

strongly correlated. Furthermore, our independent modelling of supplementary data from two large Italian studies included in the review indicated that CPS seems to be better than conventional prognostic variables factors used in this population, such as performance status and symptoms, although CPS was more accurate in patients with worse performance status. These factors may help to refine the clinician's prediction to a limited extent.

Table 2 R square values obtained for three multiple linear regression models in 981 patients for whom data on multiple prognostic variables were available

Model	KPS <40	KPS 40-50	KPS ≥60
CPS alone	0.46	0.35	0.24
Other prognostic factors alone	0.25	0.15	0.08
CPS and other prognostic factors	0.50	0.38	0.27

CPS=clinical prediction of survival; KPS=Karnofsky performance status score.

Strengths and weaknesses: comparison with previous studies

One previous qualitative systematic review on this topic also concluded that CPS is one of the best predictors of survival and is correlated with AS.[4] Our review extends those conclusions by focusing on several questions relating to the characteristics of the CPS and providing numerical answers to better understand its clinical usefulness as well as its limitations.

Our electronic search strategy lacked sensitivity. Only one in three relevant studies was located electronically.

For appraising the quality of the studies two over-riding issues arose. The first was deciding what criteria to use, and the second was deciding how to apply them. Although predicting survival has to do with prognosis, studies to compare the accuracy of CPS with AS are closer in concept to the evaluation of a diagnostic test than to studies of prognosis. However, unlike other test evaluations, no reference standard exists with which CPS can be compared, other than the outcome itself. This makes blinding and verification bias irrelevant, but it reduces the usefulness of applying quality criteria when appraising studies of CPS. As a form of a diagnostic test, CPS predicts for a future health state and so is similar to screening in its evaluation. Therefore, the study population needs to be a well defined inception cohort, and spectrum bias and loss to follow up are important validity concerns. Information about the experience, specialty, and training of the clinician making the predictions may also be relevant and needs to be available. As associated decisions about the application or withholding of life sustaining treatments such as fluids or antibiotics will also affect survival, the physician or investigator making the prediction should not be responsible for the patient's clinical care. These are the types of problems with the quality of the studies in the review, and, in the absence of established criteria, our quality ratings may not be valid.

The heterogeneity of the studies prevented us from doing a comprehensive meta-analysis. However, some pooling of the data was still possible and we believe our principal findings are valid.

Implications

The key issue with CPS is not so much whether or how to improve physicians' discriminatory ability; rather it is how to supplement or support them in their formulation of prognosis and, in particular, how to enhance their calibration. Doctors need to be aware of their tendency to overestimate prognosis in cancer patients who are approaching death. This optimism may have serious implications for the patient in terms of inappropriate application of disease controlling treatment and delays in referral to a hospice or palliative care. The results of the meta-analysis suggest that survival of patients is typically 30% shorter than predicted, but arbitrarily assigning a "correction factor" of 0.7 to their CPS cannot be recommended.

Unanswered questions and future research

Because CPS seems to be related to AS, further studies that merely look at the accuracy of predictions or document the miscalibration are not warranted. Further research is needed on whether the demographics, training, or experience of the doctor makes a difference; whether the nature of the doctor-patient relationship is important; whether predictions made at follow up are superior to initial ones; and ways to enhance the CPS. On the basis of our findings, CPS could now be used as the reference standard for evaluating other methods for predicting survival, and it has been used for this purpose.[12] Understanding how doctors formulate their predictions, and interventions that train inexperienced doctors to make better predictions are also worthy of consideration.

If doctors are better able to anticipate death, they will be likely to be better able to make judicious use of medical treatments and optimise the use of palliative care, avoiding unnecessary treatments near the end of life. They will also help patients to achieve a good death if for no other reason than that they help to fulfil patients' own expectations about the kind of information they want. Although not all patients want all the prognostic information all of the time, most patients want most of the information most of the time.[3] Doctors face two challenges in prognosticating near the end of life: formulating accurate predictions

What is already known on this topic

Accurate prediction of the timing of death is important for good clinical decision making in the care of patients with a terminal illness

Doctors' survival predictions are not very accurate and often overestimate survival

Though inaccurate, doctors' predictions correlate with survival

What this study adds

Doctors' survival predictions become more accurate closer to the date of death

Though inaccurate, predictions of up to six months in length are nevertheless reliable, as they are highly correlated with actual survival

Traditional prognostic indicators such as performance status, anorexia, and breathlessness add little information to that contained in the physician's prediction

and communicating them. The former act, which has been the subject of this review, is a predicate for the latter, but we believe that both are necessary for patients to achieve a good death.

Contributors: See bmj.com

Funding: None.

Competing interests: None declared.

Ethical approval: Not needed.

1 Steinhauser KE, Christakis NA, Clipp EC, McNeilly M, McIntyre L, Tulsky JA. Factors considered important at the end of life by patients, family, physicians and other care providers. *JAMA* 2000;284:2476-82.
2 Parkes CM. Accuracy of predictions of survival in later stages of cancer. *BMJ* 1972;ii:29-31.
3 Christakis N, Lamont E. Extent and determinants of error in doctors' prognoses. *BMJ* 2000;320:469-73.
4 Vigano A, Dorgan M, Buckingham J, Bruera E, Suarez-Alzamor ME. Survival prediction in terminal cancer patients: a systematic review of the medical literature. *Palliat Med* 2000;14:363-74.
5 Liddel J, Williamson M, Irwig L. *Method for evaluating research and guideline evidence.* Sydney: NSW Health Department, 1996. Available at internal.health.nsw.gov.au/public-health/crcp/publications/general/mergetot.pdf
6 Maltoni M, Pirovano M, Scarpi E, Marinari M, Indelli M, Arnolodi E, et al. Prediction of survival of patients terminally ill with cancer. *Cancer* 1995;75:2613-22.
7 Maltoni M, Nanni O, Derni S, Innocenti MP, Fabbri L, Riva N, et al. Clinical prediction of survival is more accurate than the KPS in estimating life span of terminally ill cancer patients. *Eur J Cancer* 1994;30A:764-6.
8 Maltoni M, Nanni O, Pirovano M, Scarpi E, Indelli M, Martini C, et al. Successful validation of the palliative prognostic score in terminally ill cancer patients. *J Pain Symptom Manage* 1999;17:240-7.
9 Evans C, McCarthy M. Prognostic uncertainty in terminal care: can the Karnofsky index help? *Lancet* 1985;i:1204-6.
10 Heyse-Moore L, Johnson-Bell VE. Can doctors accurately predict the life expectancy of patient with terminal cancer? *Palliat Med* 1987;1:165-6.
11 Oxenham D, Cornbleet MA. Accuracy of prediction of survival by different professional groups in a hospice. *Palliat Med* 1998;12:117-82.
12 Morita T, Tsunoda J, Inoue S, Chihara S. Improved accuracy of physicians' survival predictions for terminally ill cancer patients using the palliative prognostic index. *Palliat Med* 2001;15:419-24.

(Accepted 12 June 2003)

ELPS

This is an abridged version; the full version is on bmj.com

Novel consent process for research in dying patients unable to give consent

Elizabeth Rees, Janet Hardy

Department of Palliative Medicine, Royal Marsden Hospital, Sutton, Surrey SM2 5PT
Elizabeth Rees *research nurse in palliative care*
Janet Hardy *consultant in palliative care*

Correspondence to: J Hardy

BMJ 2003;327:198–200

Abstract

Objectives To develop a process of advance consent to enable research to be undertaken in patients in the terminal phase.
Design Feasibility study of an advance consent process to support a randomised controlled trial of two antimuscarinic drugs (hyoscine hydrobromide and glycopyrronium bromide) in the management of noisy respirations associated with retained secretions ("death rattle").
Setting Palliative care wards in a major cancer centre.
Participants Patients admitted to a palliative care ward who may develop "death rattle" and thus be eligible for randomisation.
Main outcome measures Patient accrual; acceptability of the consent process.
Results Of the 107 patients approached to date, 58 patients gave advance consent to participate in the study. Of these, 15 patients developed death rattle and were randomised to receive either hyoscine or glycopyrronium; 16 patients died elsewhere; 15 patients died on the palliative care wards but were not randomised; 12 patients are still alive.
Conclusions Initial assessment suggests that this is a workable consent process allowing research to be undertaken in patients who are unable to give consent at the time of randomisation. Patient accrual rates to date are lower than needed to recruit adequate numbers in the time allotted to answer the research question.

P+

The patient information sheet is on bmj.com

Introduction

In order to participate in a clinical trial patients must receive, comprehend, and retain all the information necessary to allow them to give fully informed consent for that trial.[1] Only fully informed consent can protect patients' autonomy.[2] Obtaining such consent is often very difficult in some disciplines, such as emergency medicine, elderly care, and palliative care.[3–5]

Dying patients are often unable to clear secretions from their large airways, resulting in noisy breathing usually described as "death rattle." This can be distressing to relatives and people caring for dying patients. Two antimuscarinic drugs are commonly used for the control of this condition. Hyoscine hydrobromide, a tertiary amine that can cross the blood-brain barrier causing central nervous system side effects, has historically been the drug of choice. Glycopyrronium bromide is a quaternary amine that does not cross the blood-brain barrier.

Our aim was to undertake a study to assess the relative efficacy of hyoscine and glycopyrronium in the control of death rattle within the context of a randomised controlled trial. To do this, we needed a means of obtaining consent from patients who would be unable to give consent at the time of randomisation. In the United Kingdom no established legal means exist to obtain consent in such situations.

After consultation with ethicists and lawyers, our local ethics committee advised us that the development of an advance consent process was the only possible means of obtaining consent in this situation. This paper details a method of obtaining advance consent and the interim results of the recruitment process.

Methods

All patients admitted to the palliative care wards in the Royal Marsden Hospital are given an information sheet explaining that they might be approached about research studies during their admission. The "trial suitability" of patients is determined at pre-round