View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by St George's Online Research Archive

BMJ

BMJ 2011;343:d4920 doi: 10.1136/bmj.d4920

RESEARCH

Development of Prognosis in Palliative care Study (PiPS) predictor models to improve prognostication in advanced cancer: prospective cohort study

OPEN ACCESS

Bridget Gwilliam *clinical research fellow*[1], Vaughan Keeley *consultant in palliative medicine*[2], Chris Todd *professor of primary care and community health and director of research*[3], Matthew Gittins *statistician*[4], Chris Roberts *reader in biostatistics*[4], Laura Kelly *Macmillan consultant in palliative medicine*[5], Stephen Barclay *Macmillan post-doctoral research fellow*[6], Patrick C Stone *reader in palliative medicine*[1]

[1]Division of Population, Health Sciences and Education, St George's University of London, London SW17 0RE, UK; [2]Royal Derby Hospital, Derby, UK; [3]School of Nursing, Midwifery and Social Work, University of Manchester, Manchester, UK; [4]Health Sciences, School of Community Based Medicine, University of Manchester; [5]Macmillan Consultant in Palliative Care Team, East Surrey Hospital, Surrey and Sussex Healthcare NHS Trust, Redhill, Surrey, UK; [6]General Practice and Primary Care Research Unit, Department of Public Health and Primary Care, Institute of Public Health, Cambridge, UK

Abstract

Objective To develop a novel prognostic indicator for use in patients with advanced cancer that is significantly better than clinicians' estimates of survival.

Design Prospective multicentre observational cohort study.

Setting 18 palliative care services in the UK (including hospices, hospital support teams, and community teams).

Participants 1018 patients with locally advanced or metastatic cancer, no longer being treated for cancer, and recently referred to palliative care services.

Main outcome measures Performance of a composite model to predict whether patients were likely to survive for "days" (0-13 days), "weeks" (14-55 days), or "months+" (>55 days), compared with actual survival and clinicians' predictions.

Results On multivariate analysis, 11 core variables (pulse rate, general health status, mental test score, performance status, presence of anorexia, presence of any site of metastatic disease, presence of liver metastases, C reactive protein, white blood count, platelet count, and urea) independently predicted both two week and two month survival. Four variables had prognostic significance only for two week survival (dyspnoea, dysphagia, bone metastases, and alanine transaminase), and eight variables had prognostic significance only for two month survival (primary breast cancer, male genital cancer, tiredness, loss of weight, lymphocyte count, neutrophil count, alkaline phosphatase, and albumin). Separate prognostic models were created for patients without (PiPS-A) or with (PiPS-B) blood results. The area under the curve for all models varied between 0.79 and 0.86. Absolute agreement between actual survival and PiPS predictions was 57.3% (after correction for over-optimism). The median survival across the PiPS-A categories was 5, 33, and 92 days and survival across PiPS-B categories was 7, 32, and 100.5 days. All models performed as well as, or better than, clinicians' estimates of survival.

Conclusions In patients with advanced cancer no longer being treated, a combination of clinical and laboratory variables can reliably predict two week and two month survival.

Introduction

Patients with advanced cancer and their carers often wish to know how long they have left to live.[1 2] Accurate prognostic information can allow patients adequate time to prepare for their impending death.[3] Qualitative studies show that patients in palliative care want to be given honest and accurate prognostic information but that this information needs to be shared sensitively and in a way that respects patients' desire to maintain hope.[4 5]

Prognostic information is also important for clinicians. Realistic survival estimates can inform decisions about the appropriateness of medical interventions and the timing of referral to specialist palliative care services or admission to a hospice. Clinicians' predictions are routinely used to prioritise patients who are suitable for inclusion in programmes such as the Gold Standards Framework,[6] to determine which patients

Correspondence to: P C Stone pstone@sgul.ac.uk

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

are suitable for "fast-tracking" arrangements for referral to community care, and to determine eligibility for clinical trials.

Clinicians' predictions of survival are inaccurate and over-optimistic.[7] Moreover, their estimates are affected by factors such as training, experience, seniority, and level of acquaintance with the patient.[8] For these reasons, a clinician's prediction is not a very reliable or robust method of predicting survival. None the less, clinicians' predictions are widely used, and any attempt to produce a more standardised answer to the question "How long have I got?" should be able to show that it is at least as reliable as this approach.

Previous studies have identified several clinical and laboratory variables that predict survival in patients with advanced cancer.[9 10] Prognostic tools have been created by using scoring systems derived from combinations of these variables.[11] Some of these tools include clinicians' subjective estimates.[12 13] Some rely on observers' ratings of clinical variables but do not consider the added value of including laboratory data.[14 15] Some tools include laboratory data but consequently can be applied only when such data are available.[13 16 17] Some tools were developed by using only competent patients,[13 16 17] whereas confused patients were included in the process of scale development of other tools.[14 18] One important criticism of existing tools is that their performance has not been "benchmarked" against clinicians' predictions of survival, making it difficult to judge the usefulness of these scales in clinical practice.

We sought to develop a prognostic tool that could be easily applied in clinical practice. Patients with advanced cancer are often frail, vulnerable, or confused. Many patients are unwilling to undergo further procedures (even blood tests). To produce a practical prognostic tool, we aimed to create a composite scoring system that could be used in both competent and incompetent patients and regardless of whether laboratory data were available. We wanted to develop a scoring system that was applicable across a range of palliative care settings (hospice, hospital, and community). Moreover, we wanted to produce a scale that did not rely on clinicians' estimates of survival but was at least as accurate as their best predictions.

Methods

Study settings

This was a multicentre study involving 18 palliative care services across England. Participating units included hospital support teams, hospice inpatient units, day centres, and community palliative care teams. We collected data between March 2006 and August 2009 and followed up all patients for a minimum of three months after recruitment.

Patient population

Patients were eligible to participate in this study if they had been newly referred to the relevant palliative care service with a diagnosis of advanced (locally extensive or metastatic), incurable cancer. Eligible patients were no longer receiving active treatment for cancer, and no further disease modifying treatment was planned. Patients with hormone sensitive tumours who were still receiving hormone therapy but who had developed hormone resistant disease were eligible to participate, as were patients receiving palliative radiotherapy that was not expected to prolong survival (for example, for bone metastases). Both competent and incompetent patients were eligible. We restricted study assessments in incompetent patients to observers' ratings of clinical status and extraction of information from case notes (questionnaires and blood tests were not done in incompetent patients).

Study assessments

Systematic reviews of studies involving patients with advanced cancer identified several variables with good a priori evidence of prognostic utility.[9 10] These variables formed the basis for our assessments in study participants.

Observer rated symptom checklist, performance status, and global health status

In consultation with the clinical team, a researcher completed a checklist of the following symptoms: pain, breathlessness at rest, loss of appetite, dry mouth, difficulty swallowing, and tiredness. We recorded these symptoms as being either present or absent. We recorded the patient's Eastern Co-operative Oncology Group (ECOG) performance status.[19] ECOG scores vary between 0 and 4 (0=normal functional abilities, 4=confined to a bed or chair and requires all care). We recorded global health status by using a study specific seven point scale (1=extremely poor health, 7=normal health).

Clinical observations

We measured weight and height when possible and asked patients or their carers whether weight loss had occurred over the previous month. We recorded pulse rate and the presence of clinically apparent dependent pitting oedema or ascites.

Abbreviated mental test score

We used the abbreviated mental test score to assess cognitive status.[20] Although not a comprehensive assessment of cognitive function, this 10 item scale is often used in clinical practice to screen patients for confusion. We chose it for its ease of administration and because it reflects the type of assessment commonly done in clinical practice. For ethical reasons, patients who were deemed to be incompetent were not interviewed for this study but were attributed a score of 0.

Clinician's prediction of survival

We asked clinicians to predict the survival of the patient by using one of four categories: "days" (that is, less than 14 days), "weeks" (from two weeks to less than eight weeks), "months" (from two months to less than 12 months), and "years" (12 months or more). We chose these categories as they seemed to have the greatest face validity among palliative care practitioners. Clinicians seldom make precise predictions, and asking clinicians to predict survival to the nearest day or week would have resulted in spuriously accurate prognostic predictions. Moreover, the Department of Health's Prognostic Indicator Guidance, issued as part of the Gold Standards Framework,[6] recommends these categories. Researchers obtained a clinician's estimate independently from both a doctor and a nurse. When these estimates agreed, we also took this as the "agreed multi-professional estimate." When a discrepancy existed between the clinicians, we asked them to discuss the case and arrive at an explicitly agreed estimate. Because only a small number of participants were estimated to be likely to survive for "years," we included these patients with those with an estimated survival of "months" in all analyses.

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Demographic, disease related, and treatment related variables

We recorded the age, sex, ethnicity, marital status, normal residence, diagnosis, and extent of disease (site of primary tumour and presence and sites of metastatic disease). We also recorded information on the extent of previous treatments, the time since the original diagnosis, and the time since the last treatment for cancer had been administered. We collected data on survival from entry to the study by flagging all patients for mortality with the NHS Information Centre.

Comorbidity

We recorded comorbidity by using the Charlson Co-morbidity Index.[21] This provides a measure of the overall extent of comorbid disease by using a four point scoring system (0=no comorbid diseases, 3=presence of multiple or severe comorbidity).

Laboratory variables (when available)

We asked all competent patients who agreed to participate in the study to provide a specimen of blood for analysis of haematological and biochemical parameters. We did not require incompetent patients to provide a blood specimen for analysis; however, where such results were available from clinical practice (within four days of study entry), we entered them into the study database and included them in the analysis.

Additional data

We asked competent patients to complete the symptom checklist, ECOG performance status, and global health status. We also asked competent patients whether they wished to estimate their own prognosis. We did not use these data in the generation of the prognostic scores reported here, and they will be reported elsewhere.

Procedures

We maintained a screening log of all clinical referrals to participating units. We identified eligible competent patients and, with the agreement of the clinical team, provided them with information and invited them to participate. We similarly informed the relatives or carers of eligible incompetent patients and asked them for assent.

We assessed study participants on two occasions one week apart: blood tests were not repeated specifically for the study, as this was judged to be too onerous for participants. We used only data from the baseline assessment to construct the prognostic tools described in this paper, and data from the second assessment are not presented. We calculated survival (in days) from the date of study entry.

Statistical methods

We constructed a database and checked data for accuracy and missing values. We included in the model building analysis those variables that we deemed a priori to be predictors of survival on the basis of clinical knowledge and for which only small amounts of data were missing (<6% of data missing for all included variables except C reactive protein, for which 13% of data were missing). We included participants with incomplete data in the analysis by using multiple imputation by chained equations,[22 23] with 20 imputations. No consensus exists about the best method for selecting variables for inclusion in a predictive tool, but backwards elimination is generally the preferred method, and we chose it for this analysis.[24] To reduce the computational burden, we used single imputation for the initial stages of backward selection with a cut-off P value of 0.3 before completing the selection of variables by using multiple imputation with a cut-off P value of 0.05. We did this analysis for outcomes at two weeks' (14 day) and two months' (56 day) survival in both the full dataset (PiPS model A) and the restricted dataset obtained from participants for whom blood results were available (PiPS model B), giving four models in all. An alternative modelling strategy would have used a single ordinal logistic model, but the data did not satisfy the proportional odds assumption, evident from the difference between the two week and two month models for both PiPS models A and B. We estimated the level of over-optimism in each model by using a bootstrap technique,[22 25-27] in which we drew bootstrap samples from the original dataset and then repeated the complete model selection procedure to develop a new model. We then estimated optimism as the difference in performance between the bootstrap sample and the original dataset. We used the average of the optimism over all 1000 bootstraps as a correction factor for the performance of the original model based on the full dataset.

We assessed the goodness of fit of the models by using the area under the receiver operating characteristics curve. We combined the models' predictions for the two week and two month cut-off points to produce a categorical prediction of survival ("days," "weeks," or "months/years") that could be directly compared with clinician's estimates. One limitation of fitting two logistic models is that some participants may be classified as having a survival of both less than two weeks and more than two months. This happened on only one occasion, and we classified this case as having a survival between two weeks and two months. We used linear weighted κ to compare the performance of the clinicians with that of the models (after adjustment for over-optimism), as this averages the κ coefficients for the two week and two month threshold.[28] We plotted Kaplan-Meier survival curves for each of the three risk groups ("days," "weeks," and "months+") and used Harrel's C statistic to measure concordance.[29]

Results

During the study period, 7017 consecutive eligible patients were identified across the 18 participating palliative care services. The research team was able to access only 2401 of the eligible patients (fig 1⇓). The most common reason for failure to access patients was discharge or death before being approached by the research team. The clinical team denied the researchers access to eligible patients on 1102 occasions, because they judged that involvement in a research study would prove to be too distressing for patients or their carers. Of the 2401 patients/carers who were approached by the research team, 1023 (43%) agreed to participate (780 competent, 243 incompetent patients). No significant differences in age, sex, or distribution of diagnoses existed between recruited patients and the eligible patients who were not recruited. Five patients were removed from the dataset: one was subsequently diagnosed as having a new primary tumour and no longer met the eligibility criteria, two chose to withdraw, and two were recruited twice in error. The final sample thus consisted of 1018 participants. Tables 1⇓ and 2⇓ show the characteristics of the study sample. The median survival of the group was 34 days.

Development of PiPS-A and PiPS-B models

We developed the PiPS-A models by using the core dataset that was collected from all participants (both competent and incompetent). We developed separate models to predict survival

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

of two weeks (14 days) or more (PiPS-A14) and to predict survival of two months (56 days) or more (PiPS-A56). We developed the PiPS-B models by using data from only those participants for whom blood results had also been obtained. Again, we developed separate models for predicting two week (PiPS-B14) and two month (PiPS-B56) survival. Tables 3⇓ and 4⇓ show the results of the backward stepwise logistic regression.

Performance of models and cross validation

The areas under the curve of the four models all exceeded 0.79 (tables 3⇓ and 4⇓), which shows good discrimination.[24] Logistic regression models provide estimates of survival at specified "cut-off points." However, clinicians are not usually interested in just whether a patient will survive for more than, for example, two weeks (the information provided by PiPS-A14). For this reason, the clinical question is often framed in terms of whether the patient is expected to survive for "days" (<14 days), "weeks" (14-56 days), or "months/years" (>56 days). To consider this question, we combined the PiPS-A14 and PiPS-A56 models (and the PiPS-B14 and PiPS-B56 models) to predict whether a patient was likely to survive for more than two weeks but less than two months. We then compared these predictions with the multi-professional estimates of survival (tables 5⇓ and 6⇓). Using this more demanding measure of clinical utility, the PiPS-A models performed at least as well as the clinicians (PiPS-A predictions correct on 59.6% of occasions and multi-professional predictions correct on 57.5%). The PiPS-B models performed significantly better than did either the doctors (61.5% *v* 52.6%; P=0.0135) or the nurses (61.5% *v* 52.3%; P=0.012) but were not significantly better than the multi-professional estimate (61.5% *v* 53.7%; P=0.188).

Using the PiPS-A models (fig 2⇓), the median survival of patients predicted to survive "days" was five days, that of those predicted to survive "weeks" was 33 days, and that of those predicted to survive "months+" was 92 days (Harrell's concordance (C) statistic=0.69). For the PiPS-B models (fig 3⇓), the equivalent median survival was seven days, 32 days, and 100.5 days (C=0.67).

Discussion

In this large, prospective, multicentre study, we have developed and validated four prognostic models for predicting survival in palliative care patients with advanced cancer. These models are able to identify reliably those patients with expected prognoses of "days," "weeks," or "months/years" and can be used in either competent or incompetent patients and in circumstances when blood results are available and when additional investigations would be inappropriate. When combined, the prognostic models were at least as good as a multi-professional clinical estimate of survival; when blood results are available, the models were significantly better than either a doctor's or a nurse's prediction (but not a multi-professional estimate).

Strengths of study

Our study had several strengths. The primary aim was to develop a prognostic scoring system that was significantly better than (and independent of) clinicians' predictions of survival. No previous studies have attempted to "benchmark" their performance against current best clinical practice in this way. As a result, we have been able to show that the PiPS scores are better than the best uni-professional specialist clinical predictions of survival. All of the variables assessed as part of the PiPS study were previously identified in systematic reviews as being of likely prognostic significance,[9 10] and the final variables included in our models have been found to be reliable predictors of survival in a large heterogeneous group of patients with advanced cancer in a variety of different settings (hospital, hospice, and community). This supports the generalisability of our results. Our decision to combine the prognostic models to produce a "categorical" estimate of survival (in terms of "days," "weeks," or "months+") can also be considered to be a strength of the study. The resulting PiPS algorithms produce an estimate of survival that is clinically meaningful and that can be directly compared with clinicians' own estimates.

Weaknesses of study

Our study also had several limitations. Although we tried to study all evaluable patients, some were not accessible because of "gatekeeping" by clinical staff. The phenomenon of "gatekeeping" is a common problem in palliative care studies,[30 31] reflecting a desire by clinical staff to protect patients from the perceived burden of involvement in clinical studies. Such selective recruitment may distort findings if it renders the research sample unrepresentative of the population of interest. Aware that this was a potential limitation to our study, we invested considerable effort in accurately documenting "flows" of patients through the research process. Most previous palliative prognostic studies have not needed to rely on patients' consent before recruitment or have reported only the results from "evaluable" patients.[12 13 18 32] In contrast, we recorded data on all referrals to participating units and attempted to document the reasons why we could not evaluate some patients.

The fact that many potentially eligible patients for our study were non-evaluable was disappointing, but this reflects the reality of dealing with patients who have advanced disease. We did not find any significant differences in age, sex, or diagnoses between evaluable and non-evaluable patients. The most common reasons why patients were not evaluated were that they were under the care of the palliative care team for only a very short time (n=1899), so insufficient time was available to obtain informed consent; that the healthcare team deemed that discussion of prognosis would be too distressing (n=1102); or that the patient declined the offer from the healthcare team to talk to a researcher about the study (n=578). All of these reasons for failure to evaluate patients with the PiPS score also pertain to "real world" situations. We could therefore argue that the sample of patients recruited to our study is representative of the sort of patients who would be offered a PiPS evaluation in clinical practice.

Another potential limitation of our study is that the results have not yet been evaluated in an independent population. Developing prognostic models and testing them in the same population is known to produce over-optimism. The usual approach to cross validation is to develop the model in one half of the dataset (the training set) and to validate it in the other half (the testing set). Bootstrapping is an alternative cross validation technique that maximises the study data. It is a re-sampling procedure involving taking repeated samples (with replacement) from the dataset. The bootstrap validation in this study allowed us to "correct" for over-optimism in the development of the models and still to show that the PiPS predictions are robust. However, external validation in a separate cohort is still needed to confirm the predictive accuracy of the proposed models.

Our study was limited to determining the statistical accuracy of the PiPS prognostic algorithms. We did not assess whether introduction of the PiPS scores into clinical practice results in any demonstrable improvements in patient care. Future studies

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

should include an assessment of the clinical utility as well as the statistical accuracy of the PiPS scores.

Relation to other studies

The PiPS models offer some definite advantages over existing methods of predicting survival in this population of patients. The most common method of prognostication is a clinician's prediction of survival. However, as this is a subjective assessment, predictions cannot be standardised across sites or over time. This lack of consistency means that relying on clinicians' predictions of survival is not an entirely suitable method of identifying patients for inclusion in clinical trials, for referral to specialist services, or for discussions with patients and families about the future. The PiPS-A scoring algorithms, although no more accurate than a multi-professional prognostic estimate, do provide an objective and repeatable output and would at the very least provide an excellent reproducible method for assessing eligibility for research studies. When blood results are available, the PiPS-B scores are significantly more accurate than either a doctor's or a nurse's estimate and thus could provide a robust rationale for making many decisions about treatment.

Other than clinicians' estimates, the most widely studied method of predicting survival in patients with advanced cancer is probably the palliative prognostic score.[12 13] This score is calculated from six variables, one of which is itself a clinician's prediction. Each variable is allotted a "partial score" that depends on the size of the regression coefficient. Total palliative prognostic scores can range between 0 and 17.5 and define three prognostic categories representing different probabilities of surviving 30 days (<30%, 30-70%, and >70%). The palliative prognostic score has been subsequently validated in several clinical studies.[33-37] One of its drawbacks is the heavy weighting that is attributed to the clinician's prediction of survival; the partial score for the clinician's prediction can represent 8.8/17.5 (49%) of the total score. Moreover, clinicians are required to provide an unrealistically accurate prognostic estimate in two week intervals. A patient who is expected to live for three to four weeks thus scores 1.5 points more than a patient expected to survive for five to six weeks and 3.5 points more than a patient expected for live for seven to 10 weeks. Most clinicians are reluctant to offer prognostic estimates with this degree of precision. Another limitation of the palliative prognostic score is the difficulty of converting the prognostic categories into meaningful clinical information. How should a clinician (or patient) interpret the information that they have a 30-70% chance of surviving 30 days? PiPS overcomes these problems.

Meaning of study

The PiPS-A score can be calculated for any patient with advanced cancer who is no longer receiving disease modifying treatment, and it is at least as good as, but not significantly better than, a clinician's estimate of survival. The PiPS-B prognostic estimate requires a blood test, but it is statistically significantly better than the best uni-professional estimate of survival. Although the absolute increase in prognostic accuracy is not large, PiPS-B shares with PiPS-A the advantages of being independent of the clinician's opinion and of being reproducible and comparable across settings. In a clinical context, we believe that PiPS estimates would usually be used to inform and augment clinicians' own subjective estimates (rather than to replace them). However, our study suggests that the PiPS-B estimate of survival could now act as the "benchmark" against which new prognostic tools are assessed. We hope that this will act as the first step in an incremental process of improving prognostic accuracy as novel indicators are tested against, and then incorporated into, the PiPS models.

Unanswered questions and further research

Although the PiPS prognostic scores have been developed in a robust fashion, further validation work is needed before the scales can be recommended for use in routine clinical practice. The clinical information needed to calculate the PiPS scores is easily obtained. However, although the scores can be calculated with minimal arithmetical effort, they cannot be directly or easily calculated at the bedside. We have therefore produced a simple computer based interface for the PiPS instruments.[38] We plan to develop this tool so that the algorithms will be available across a range of platforms and will allow users to rapidly convert clinical data into prognostic estimate categories.

As previously described, we have tested the PiPS models by using cross validation techniques, but their accuracy has not yet been assessed in independent datasets. The validation of prognostic tools emerged as the highest priority among clinicians and researchers in a recent consensus workshop on prognostication.[39] We are planning a large prospective multicentre external validation study to include an assessment of users' (clinicians, patients, and carers) views about the models and the best way for data to be presented to accurately reflect the degree of uncertainty inherent in the models.

We thank the following colleagues for their help with this study: Rehana Bakawala, Mike Bennett, Teresa Beynon, Cath Blinman, Patricia Brayden, Helen Brunskill, Kate Crossland, Alison Cubbitt, Rachel Glascott, Anita Griggs, Anne Harbison, Debra Hart, Philip Lomax, Caroline Lucas, Wendy Makin, Oliver Minton, Paul Perkins, Marek Plaskota, Dai Roberts, Katie Richies, Susan Salt, Ileana Samanidis, Margaret Saunders, Jennifer Todd, Catherine Waight, Nicola Wilderspin, Gail Wiley, and Julie Young. We also thank John Ellershaw for chairing the steering committee and Robert Godsill for providing a service user's perspective. Thanks go to Rosie Head for administrative support and data management. Thanks also go to the following hospices and palliative care units for their participation in the study: Arthur Rank House (Cambridge), Worcestershire Royal Hospital, St John's Hospice (Lancaster), Gloucestershire Hospitals NHS Foundation Trust, Pasque Hospice (Luton), Guy's and St Thomas' NHS Foundation Trust (London), Princess Alice Hospice (Esher), Bolton Hospice, St Catherine's Hospice (Crawley), St George's Hospital NHS Trust (London), Surrey and Sussex Healthcare NHS Trust, St Ann's Hospice (Manchester), Christie Hospital NHS Foundation Trust (Manchester), Nightingale Macmillan Unit (Derby), Trinity Hospice (London), and Trinity Hospice (Blackpool).

Contributions: PCS, BG, VK, CT, CR, LK, and SB contributed to the conception and design of the study. CR, MG, and BG contributed to the analysis of data. All authors contributed to the interpretation of data, the drafting or revising of the manuscript, and final approval for publication. PCS is the guarantor.

Funding: This study was funded by Cancer Research UK (grant number C11075/A6126). SB is funded by Macmillan Cancer Support and the NIHR CLAHRC (Collaborations for Leadership in Applied Health Research and Care) for Cambridgeshire and Peterborough.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: support from Cancer Research UK (CRUK) for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: This study was approved by the Wandsworth Multi-centre Research Ethics Committee. Site specific approval and

What is already known on this topic

Prognostic information is valued by patients, carers, and healthcare professionals

Clinicians' predictions of survival are the mainstay of current practice but are unreliable, over-optimistic, and subjective

What this study adds

Two prognostic scores have been created, both of which are able to predict whether patients will survive for "days," "weeks," or "months"

Both scores are independent of clinicians' subjective estimates of survival, and both are at least as accurate as a clinician's estimate

One of the prognostic scores (which requires a blood test) is significantly better than an individual doctor's or nurse's prediction, but neither scale is significantly more accurate than a multi-professional estimate of survival

research and development approval were obtained for participating units. Exemption from Section 60 of the Health and Social Care Act (2001) was obtained from the Patient Information Advisory Group (PIAG). This allows the records of incompetent patients to be accessed without explicit consent so that study variables can be recorded and patients "flagged" for mortality purposes with the NHS Information Centre. Consent was given by competent patients and assent by the relatives/carers of incompetent patients.

Data sharing: No additional data available.

1 Adams E, Boulton M, Watson E. The information needs of partners and family members of cancer patients: a systematic literature review. *Patient Educ Couns* 2009;77:179-86.
2 Degner LF, Kristjanson LJ, Bowman D, Sloan JA, Carriere KC, O'Neill J, et al. Information needs and decisional preferences in women with breast cancer. *JAMA* 1997;277:1485-92.
3 Steinhauser KE, Christakis NA, Clipp EC, McNeilly M, Grambow S, Parker J, et al. Preparing for the end of life: preferences of patients, families, physicians, and other care providers. *J Pain Symptom Manage* 2001;22:727-37.
4 Kirk P, Kirk I, Kristjanson LJ. What do patients receiving palliative care for cancer and their families want to be told? A Canadian and Australian qualitative study. *BMJ* 2004;328:1343.
5 Kutner JS, Steiner JF, Corbett KK, Jahnigen DW, Barton PL. Information needs in terminal illness. *Soc Sci Med* 1999;48:1341-52.
6 Thomas K, Free A. Gold Standards Framework prognostic indicator guidance. 2008. www.goldstandardsframework.nhs.uk/OneStopCMS/Core/CrawlerResourceServer.aspx?resource=B8424129-940E-4AFD-B97A-7A9D9311BC25&mode=link&guid=eae1921fd0694340841dbfc891be1047.
7 Glare P, Virik K, Jones M, Hudson M, Eychmuller S, Simes J, et al. A systematic review of physicians' survival predictions in terminally ill cancer patients. *BMJ* 2003;327:195-8.
8 Christakis NA, Lamont EB. Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ* 2000;320:469-72.
9 Vigano A, Dorgan M, Buckingham J, Bruera E, Suarez-Almazor ME. Survival prediction in terminal cancer patients: a systematic review of the medical literature. *Palliat Med* 2000;14:363-74.
10 Maltoni M, Caraceni A, Brunelli C, Broeckaert B, Christakis N, Eychmueller S, et al. Prognostic factors in advanced cancer patients: evidence-based clinical recommendations—a study by the Steering Committee of the European Association for Palliative Care. *J Clin Oncol* 2005;23:6240-8.
11 Stone PC, Lund S. Predicting prognosis in patients with advanced cancer. *Ann Oncol* 2007;18:971-6.
12 Pirovano M, Maltoni M, Nanni O, Marinari M, Indelli M, Zaninetta G, et al. A new palliative prognostic score: a first step for the staging of terminally ill cancer patients. *J Pain Symptom Manage* 1999;17:231-9.
13 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.
14 Morita T, Tsunoda J, Inoue S, Chihara S. The Palliative Prognostic Index: a scoring system for survival prediction of terminally ill cancer patients. *Support Care Cancer* 1999;7:128-33.
15 Chuang R-B, Hu W-Y, Chiu T-Y, Chen C-Y. Prediction of survival in terminal cancer patients in Taiwan: constructing a prognostic scale. *J Pain Symptom Manage* 2004;28:115-22.
16 Kelly L, White S, Stone PC. The B12/CRP index as a simple prognostic indicator in patients with advanced cancer: a confirmatory study. *Ann Oncol* 2007;18:1395-9.
17 Stone P, Kelly L, Head R, White S. Development and validation of a prognostic scale for use in patients with advanced cancer. *Palliat Med* 2008;22:711-7.
18 Anderson F, Downing GM, Hill J, Casorso L, Lerch N. Palliative performance scale (PPS): a new tool. *J Palliat Care* 1996;12:5-11.
19 Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982;5:649-55.
20 Hodkinson HM. Evaluation of a mental test score for assessment of mental impairment in the elderly. *Age Ageing* 1972;1:233-8.
21 Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373-83.
22 Steyerberg EW. Clinical prediction models: a practical approach to development, validation and updating. Springer, 2009.
23 Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
24 Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. *BMJ* 2009;338:b604.
25 Chen CH, George SL. The bootstrap and identification of prognostic factors via Cox's proportional hazards regression model. *Stat Med* 1985;4:39-46.
26 Schumacher M, Hollander N, Sauerbrei W. Resampling and cross-validation techniques: a tool to reduce bias caused by model building? *Stat Med* 1997;16:2813-27.
27 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
28 Roberts C, McNamee R. Assessing the reliability of ordered categorical scales using kappa-type statistics. *Stat Methods Med Res* 2005;14:493-514.
29 Harrel FE, Califf RM, Pryor DB, Lee KL. Evaluating the yield of a medical test. *JAMA* 1982;247:2543-6.
30 White C, Gilshenan K, Hardy J. A survey of the views of palliative care healthcare professionals towards referring cancer patients to participate in randomized controlled trials in palliative care. *Support Care Cancer* 2008;16:1397-405.
31 Ewing G, Rogers M, Barclay S, McCabe J, Martin A, Todd C. Recruiting patients into a primary care based study of palliative care: why is it so difficult? *Palliat Med* 2004;18:452-9.
32 Geissbuhler P, Mermillod B, Rapin CH. Elevated serum vitamin B12 levels associated with CRP as a predictive factor of mortality in palliative care cancer patients: a prospective study over five years. *J Pain Symptom Manage* 2000;20:93-103.
33 Glare P, Eychmueller S, Virik K. The use of the palliative prognostic score in patients with diagnoses other than cancer. *J Pain Symptom Manage* 2003;26:883-5.
34 Glare P, Virik K. Independent prospective validation of the PaP score in terminally ill patients referred to a hospital-based palliative medicine consultation service. *J Pain Symptom Manage* 2001;22:891-8.
35 Glare PA, Eychmueller S, McMahon P. Diagnostic accuracy of the palliative prognostic score in hospitalized patients with advanced cancer [published correction in: *J Clin Oncol* 2005;23:248]. *J Clin Oncol* 2004;22:4823-8.
36 Caraceni A, Nanni O, Maltoni M, Piva L, Indelli M, Arnoldi E, et al. Impact of delirium on the short term prognosis of advanced cancer patients. *Cancer* 2000;89:1145-9.
37 Naylor C, Cerqueira L, Costa-Paiva LHS, Costa JV, Conde DM, Pinto-Neto AM. Survival of women with cancer in palliative care: use of the palliative prognostic score in a population of Brazilian women. *J Pain Symptom Manage* 2010;39:69-75.
38 Prognosis in Palliative Care Study. The PiPS prognosticator. 2011. www.pips.sgul.ac.uk.
39 Stevinson C, Preston N, Todd C. Defining priorities in prognostication research: results of a consensus workshop. *Palliat Med* 2010;24:462-8.

Accepted: 4 July 2011

Cite this as: *BMJ* 2011;343:d4920

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: http://creativecommons.org/licenses/by-nc/2.0/ and http://creativecommons.org/licenses/by-nc/2.0/legalcode.

Tables

Table 1| **Characteristics of study sample. Values are numbers (percentages) unless stated otherwise**

Characteristic	Value (n=1018)
Patient competent	775 (76.1)
Male sex	557 (54.7)
Location:	
Home	123 (12.1)
Hospital	182 (17.9)
Hospice	701 (68.9)
Other	25 (2.5)
Marital status single	524 (51.5)
White British ethnicity	904 (88.8)
Site of primary cancer:	
Digestive organs	305 (30.0)
Respiratory/intrathoracic	238 (23.4)
Independent multiple sites/unknown primary/other	103 (10.1)
Breast	74 (7.3)
Male genital organs (including prostate)	71 (7.0)
Urinary tract	68 (6.7)
Female genital organs	52 (5.1)
Eye/brain/central nervous system	38 (3.7)
Lymphatic/haematological	36 (3.5)
Oropharynx/mouth	33 (3.2)
Site of metastatic disease:	
Anywhere	684 (67.2)
Liver	276 (27.1)
Bone	268 (26.3)
Lung	213 (20.9)
Brain	90 (8.8)
Other/not recorded	82 (8.1)
Ascites	68 (6.7)
Pleural effusion	50 (4.9)
Adrenal	25 (2.5)
Skin	20 (2.0)
Pericardial effusion	3 (0.3)
Regional lymph node spread	209 (20.5)
Receiving hormone therapy	101/1017 (9.9)
Receiving palliative radiotherapy	32/1017 (3.1)
Charlton comorbidity score:	(n=1011)
0	453 (44.8)
1	395 (39.1)
2	117 (11.6)
3	46 (4.5)
Drugs:	
Strong opioids	688/1017 (67.6)
Corticosteroids	449/1013 (44.3)
Bisphosphonates	99/1012 (9.8)
Antibiotics	173/1015 (17.0)
Weight loss in previous month	693/998 (69.4)

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Table 1 (continued)

Characteristic	Value (n=1018)
Clinically apparent ascites	137/1015 (13.5)
Clinically apparent pitting oedema	391/1015 (38.5)
Lack of appetite	663/1011 (65.6)
Breathlessness	280/1015 (27.6)
Dry mouth	626/1014 (61.7)
Difficulty swallowing	292/1014 (28.8)
Pain	879/1014 (86.7)
Tired	879/1015 (86.6)
Eastern Co-operative Oncology Group score:	(n=1015)
0	2 (0.2)
1	81 (8.0)
2	216 (21.3)
3	435 (42.9)
4	281 (27.7)
Global health status:	(n=1015)
1	125 (12.3)
2	154 (15.2)
3	336 (33.1)
4	290 (28.6)
5	94 (9.3)
6	16 (1.6)
7	0 (0)
Mental test score:	(n=1009)
≤3	233 (23.1)
>3	776 (76.9)
Mean (SD) pulse rate (beats/min)	86 (15.6)

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Table 2| **Mean (SD) laboratory results for study sample**

Variable	Result
Full blood count	
Haemoglobin (g/dL)	11.04 (2.0) (n=786)
Mean cell volume (fl)	88.75 (7.8) (n=779)
White blood count ($\times 10^9$/L)	11.96 (7.3) (n=786)
Neutrophil count ($\times 10^9$/L)	10.55 (10.3) (n=783)
Lymphocyte count ($\times 10^9$/L)	1.41 (2.0) (n=783)
Platelets ($\times 10^9$/L)	333.9 (162.3) (n=786)
Biochemistry	
Sodium (mmol/L)	135.75 (5.0) (n=797)
Potassium (mmol/L)	4.44 (0.8) (n=784)
Chloride (mmol/L)	98.2 (7.1) (n= 628)
Bicarbonate (mmol/L)	26.0 (4.6) (n=610)
Urea (mmol/L)	9.25 (7.0) (n=798)
Creatinine (µmol/L)	88.3 (66.5) (n=797)
Glucose (mmol/L)	6.79 (3.0) (n=611)
Bilirubin (µmol/L)	16.1 (38.9) (n=765)
Alanine transaminase (U/L)	36.0 (58.1) (n= 761)
Alkaline phosphatise (U/L)	263.9 (340.0) (n=778)
Albumin (g/L)	29.4 (6.8) (n=787)
Calcium (mmol/L)	2.18 (0.2) (n=749)
Magnesium (mmol/L)	0.84 (0.1) (n=627)
C reactive protein (mg/L)	82.0 (80.1) (n=696)

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Table 3| **PiPS-A backwards elimination models (n=1018)**

Variable	PiPS-A14*		PiPS-A56†	
	Odds ratio (95% CI)	P value	Odds ratio (95% CI)	P value
Mental test score >3	3.571 (2.301 to 5.541)	<0.001	2.341 (1.445 to 3.794)	0.001
Pulse rate	0.977 (0.965 to 0.989)	<0.001	0.978 (0.967 to 0.988)	<0.001
Presence of distant metastases	0.608 (0.374 to 0.987)	0.044	0.666 (0.467 to 0.949)	0.024
Site of metastases—liver	0.584 (0.374 to 0.912)	0.018	0.552 (0.374 to 0.815)	0.003
ECOG score	0.559 (0.430 to 0.754)	<0.001	0.803 (0.652 to 0.989)	0.039
Global heath score	1.567 (1.269 to 1.935)	<0.001	1.732 (1.449 to 2.071)	<0.001
Loss of appetite	0.463 (0.291 to 0.736)	0.001	0.657 (0.477 to 0.903)	0.01
Site of metastases—bone	1.681 (1.061 to 2.662)	0.027	–	–
Difficulty breathing	0.622 (0.419 to 0.922)	0.018	–	–
Difficulty swallowing	0.583 (0.389 to 0.873)	0.009	–	–
Primary breast cancer	–	–	1.853 (1.050 to 3.269)	0.033
Primary male genital cancer (including prostate)	–	–	4.379 (2.375 to 8.085)	<0.001
Weight loss	–	–	0.601 (0.433 to 0.834)	0.002

ECOG=Eastern Co-operative Oncology Group.

*Area under curve=0.866.

†Area under curve=0.793.

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Table 4| **PiPS-B backwards elimination models (n=796)**

Variable	PiPS-B14*		PiPS-B56†	
	Odds ratio (95% CI)	P value	Odds ratio (95% CI)	P value
Pulse rate	0.983 (0.968 to 0.999)	0.033	0.987 (0.974 to 0.999)	0.044
White blood count	0.941 (0.914 to 0.969)	<0.001	0.958 (0.923 to 0.995)	0.026
Platelets	1.003 (1.001 to 1.004)	0.001	1.001 (1.000 to 1.003)	0.017
Urea	0.944 (0.919 to 0.970)	<0.001	0.940 (0.908 to 0.974)	0.001
C reactive protein	0.994 (0.992 to 0.999)	<0.001	0.993 (0.990 to 0.996)	<0.001
Global heath score	1.482 (1.145 to 1.919)	0.003	1.606 (1.339 to 1.927)	<0.001
Alanine transaminase	0.996 (0.993 to 0.999)	0.007	–	–
Mental test score >3	2.591 (1.344 to 4.997)	0.004	–	–
Distant metastases	0.434 (0.247 to 0.761)	0.004	–	–
Site of metastases—bone	2.153 (1.240 to 3.739)	0.006	–	–
Lack of appetite	0.507 (0.295 to 0.874)	0.014	–	–
ECOG score	0.588 (0.407 to 0.849)	0.005	–	–
Neutrophils	–	–	0.969 (0.944 to 0.996)	0.025
Lymphocytes	–	–	1.177 (1.014 to 1.366)	0.032
Alkaline phosphatase	–	–	0.999 (0.998 to 0.999)	0.001
Albumin	–	–	1.041 (1.010 to 1.073)	0.009
Primary male genital cancer (including prostate)	–	–	4.757 (2.354 to 9.610)	<0.001
Tired	–	–	0.510 (0.301 to 0.863)	0.012

ECOG=Eastern Co-operative Oncology Group

*Area under curve=0.859.

†Area under curve=0.809.

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Table 5| **Comparison of PiPS-A models against clinicians' predictions**

	Actual survival				Absolute agreement (%)	Linear weighted κ	P value for comparison of model A with:		
Predictions	**Days**	**Weeks**	**Months+**	**Total**			**Doctors' predictions**	**Nurses' predictions**	**Multi-professional predictions**
Doctor									
Days	137	16	3	156					
Weeks	82	126	86	294					
Months+	36	139	204	379					
Total	255	281	293	829	56.3	0.442	NA	0.154	0.284
Nurse									
Days	135	19	7	161					
Weeks	84	116	79	279					
Months+	52	186	276	514					
Total	271	321	362	954	55.2	0.412	0.154	NA	0.007
Multi-professional									
Days	147	17	4	168					
Weeks	96	136	92	324					
Months+	35	176	284	495					
Total	278	329	380	987	57.5	0.457	0.284	0.007	NA
PiPS-A models									
Days	149	32	16	197					
Weeks	86	191	136	413					
Months+	22	93	229	344					
Total	257	316	381	954	59.6 (57.3*)	0.470 (0.441*)	0.958	0.249	0.515

NA=not applicable.

*Adjusted for over-optimism by using bootstrap.

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Table 6| **Comparison of PiPS-B models against clinicians' predictions**

	Actual survival				Absolute agreement (%)	Linear weighted κ	P value for comparison of model B with:		
Predictions	**Days**	**Weeks**	**Months+**	**Total**			**Doctors' predictions**	**Nurses' predictions**	**Multi-professional predictions**
Doctor									
Days	49	12	1	62					
Weeks	62	113	79	254					
Months+	33	124	183	340					
Total	144	249	263	656	52.6	0.322	NA	0.304	0.385
Nurse									
Days	50	13	6	69					
Weeks	57	102	69	228					
Months+	49	164	241	454					
Total	156	279	316	751	52.3	0.293	0.304	NA	0.029
Multi-professional									
Days	54	13	1	68					
Weeks	72	118	84	274					
Months+	35	154	244	433					
Total	161	285	329	775	53.7	0.336	0.385	0.029	NA
PiPS-B models									
Days	46	18	5	69					
Weeks	58	159	88	305					
Months+	11	69	192	272					
Total	115	246	285	646	61.5 (57.3*)	0.436 (0.373*)	0.0135	0.012	0.188

NA=not applicable.

*Adjusted for over-optimism by using bootstrap.

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

Figures

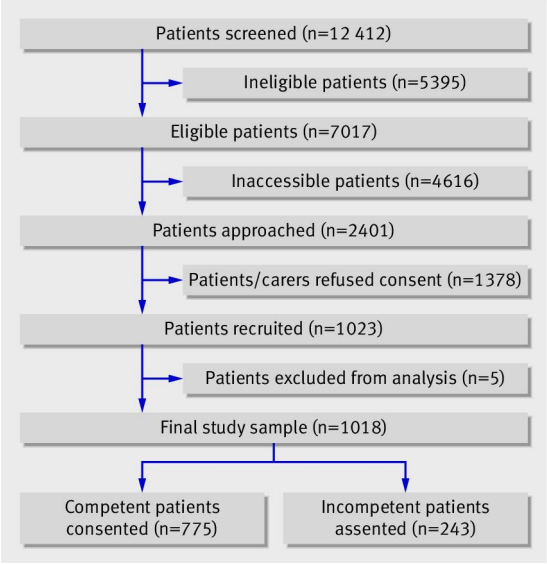

Fig 1 Selection of study sample

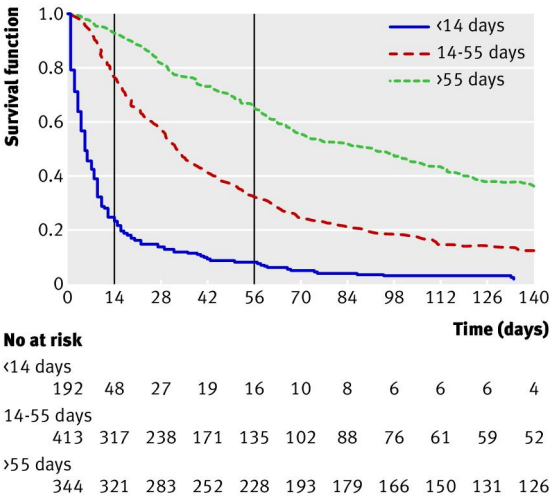

Fig 2 Kaplan-Meier survival curves for PiPS-A models. Graph shows survival curves for three prognostic groups identified by PiPS-A scores. Vertical lines indicate survival at specific "cut-off points" of 14 and 56 days. Harrell's C index is defined as proportion of pairs of participants in which predictions and outcomes are concordant, and C=0.6894 indicates that PiPS-A scores can correctly order survival times for pairs of participants 68.9% of time

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

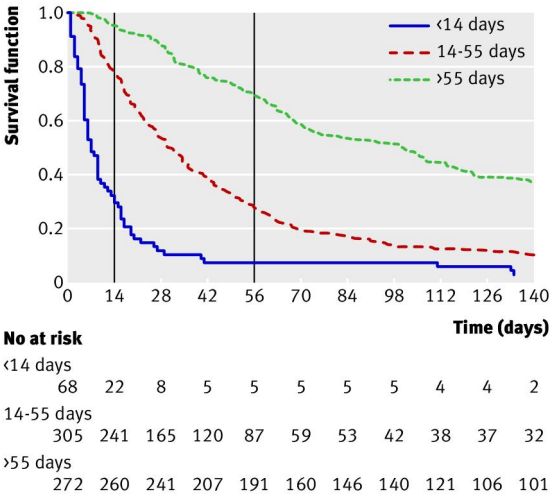

Fig 3 Kaplan-Meier survival curves for PiPS-B models. Graph shows survival curves for three prognostic groups identified by PiPS-B scores. Vertical lines indicate survival at specific "cut-off points" of 14 and 56 days. Harrell's C index is defined as proportion of pairs of participants in which predictions and outcomes are concordant, and C=0.6745 indicates that PiPS-B scores can correctly order survival times for pairs of participants 67.5% of time

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe