



HHS Public Access

Author manuscript

Eur J Cancer. Author manuscript; available in PMC 2017 September 01.

Published in final edited form as:

Eur J Cancer. 2016 September ; 64: 89–95. doi:10.1016/j.ejca.2016.05.009.

Clinician Prediction of Survival versus The Palliative Prognostic Score: Which Approach is More Accurate?

David Hui^a, Minjeong Park^b, Diane Liu^b, Carlos Eduardo Paiva^c, Sang-Yeon Suh^d, Tatsuya Morita^e, and Eduardo Bruera^a

^aDepartment of Palliative Care, Rehabilitation and Integrative Medicine, MD Anderson Cancer Center, Houston, USA, 77030

^bDepartment of Biostatistics, MD Anderson Cancer Center, Houston, USA, 77030

^cDepartment of Clinical Oncology, Barretos Cancer Hospital, Barretos, São Paulo, Brazil

^dDepartment of Medicine, Dongguk University, Seoul, South Korea

^eDepartment of Palliative and Supportive Care, Palliative Care Team and Seirei Hospice, Seirei Mikatahara General Hospital, Shizuoka, Japan

Abstract

Background—Clinician prediction of survival (CPS) has low accuracy in the advanced cancer setting, raising the need for prediction models such as the palliative prognostic (PaP) score that includes a transformed CPS (PaP-CPS) and 5 clinical/laboratory variables (PaP-without CPS). However, it is unclear if the PaP score is more accurate than PaP-CPS, and whether PaP-CPS helps to improve the accuracy of PaP score. We compared the accuracy among PaP-CPS, PaP-without CPS and PaP-total score in patients with advanced cancer.

Patients and Methods—In this prospective study, PaP score was documented in hospitalization patients seen by palliative care. We compared the discrimination of PaP-CPS versus PaP-total and PaP-without CPS versus PaP-total using 4 indices: concordance statistics, area under the receiver-operating characteristics curve (AUC), net reclassification index and integrated discrimination improvement for 30-day survival and 100-day survival.

Results—216 patients were enrolled with a median survival of 109 days (95% confidence interval [CI] 71–133 days). The AUC for 30-day survival was 0.57 (95%CI 0.47–0.67) for PaP-CPS, 0.78 (95%CI 0.7–0.87) for PaP-without CPS, and 0.73 (95%CI 0.64–0.82) for PaP-total score. PaP-total was significantly more accurate than PaP-CPS according to all 4 indices for both

Corresponding author: David Hui, MD, MSc, Department of Palliative Care, Rehabilitation and Integrative Medicine, Unit 1414, University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX, USA 77030, Phone: 713-606-3376, Fax: 713-792-6092, dhui@mdanderson.org.

Conflicts of interest

None

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

30-day and 100-day survival ($P < 0.001$). PaP-without CPS was significantly more accurate than PaP-total for 30-day survival ($P < 0.05$).

Conclusion—We found that PaP-score was more accurate than CPS, and the addition of CPS to the prognostic model reduced its accuracy. This study highlights the limitations of clinical gestalt and the need to use objective prognostic factors and models for survival prediction.

Keywords

Clinical prediction rule; forecasting; prognosis; neoplasms; statistical data analysis; survival

Introduction

Accurate prediction of survival is essential for clinical decision making [1, 2]. This becomes particularly important as cancer patients approach the end-of-life, in which the recommendations for chemotherapy and palliative procedures (e.g. gastrostomy) could differ substantially among patients with months, weeks or days of life expectancy [3]. When asked to estimate survival, clinicians often rely on their clinical experience and intuition. Although clinician prediction of survival (CPS) is quick and simple, the literature shows that clinicians consistently over-estimate survival and that the accuracy of CPS (i.e. expected survival within 33% of actual survival) varies between 20% and 30% [4–6].

To improve our accuracy, multiple groups have proposed actuarial prediction of survival that uses prognostic factors and mathematical models [7]. The Palliative Prognostic Score (PaP score) represents one of the most validated prognostic models in the advanced cancer setting [8–12]. It consists of CPS and 5 clinical and laboratory variables (i.e. performance status, dyspnea, anorexia, leukocyte count and lymphocyte percentage). This score has been well studied in multiple clinical settings, and its accuracy as measured by the concordance statistic (C-index) has been reported to be between 72% and 89% [11, 12].

One fundamental question is whether the PaP score is more accurate than clinical gestalt. Few studies have made a direct comparison between the two approaches and applied the same metrics. Indeed, the only group that has attempted to investigate this issue reported that their correlation with actual survival to be similar [13]. Given that the formulation of CPS often incorporates many established prognostic factors in the PaP score, it is unclear if PaP-total score is superior to CPS alone. Conversely, PaP-total score consists of both subjective (i.e. CPS) and objective (i.e. 5 clinical and laboratory variables, PaP-without CPS) components, and many investigators have questioned whether CPS increases or decreases the accuracy of PaP-total score. A better understanding of these fundamental issues would help clinicians to decide how to better prognosticate. The objective of this study is to compare the accuracy between CPS and PaP-total score and between PaP-without CPS and PaP-total score in patients with advanced cancer. We hypothesized that PaP-total score is more accurate than CPS, and that CPS improves the accuracy of PaP-without CPS.

Patients and Methods

Study Setting and Criteria

This is planned analysis of a prospective study examining novel prognostic markers in patients with advanced cancer [14]. Briefly, we enrolled patients with a diagnosis of advanced cancer who were ≥ 18 years of age, hospitalized at MD Anderson Cancer Center, seen by the palliative care mobile team for consultation, and received parenteral hydration. Patients with delirium, contraindications to bioelectric impedance analysis or inability to use the hand dynamometer were excluded. The study protocol was reviewed and approved by the Institutional Review Board at MD Anderson Cancer Center. All participants provided written informed consent and were enrolled between 9/22/2011 and 1/26/2013.

Data Collection

We prospectively collected baseline patient demographics at the time of study enrollment, such as age, sex, race, cancer diagnosis. The PaP score, a validated prognostic model for patients with advanced cancer, consists of 6 variables: CPS, dyspnea (absence/presence), anorexia (absence/presence), Karnofsky Performance status, total leukocyte count and lymphocyte percentage [8].

The raw data for CPS was obtained by asking the palliative care specialists most responsible for the patient to provide an estimate of survival. This was then transformed into PaP score categories following the original scoring system: >12 weeks = 0 point; 11–12 weeks = 2 points; 7–10 weeks = 2.5 points; 5–6 weeks = 4.5 points; 3–4 weeks = 6 points; and ≤ 2 weeks = 8.5 points) [8, 9].

According to the PaP score, patients also received 1 point for the presence of dyspnea, 1.5 points for the presence of anorexia, and 2.5 points for KPS $\geq 40\%$. This prognostic model also includes two laboratory variables, white blood cell count ($8501\text{--}11000\text{ cell/mm}^3 = 0.5$ point; $>11000\text{ cell/mm}^3 = 1.5$ point) and lymphocyte percentage ($12\%\text{--}19.9\% = 1$ point; $0\text{--}11.9\% = 2.5$ points). The total PaP score ranges between 0 and 17.5 points [8]. A higher total score indicates worse prognosis.

Survival from time of study enrollment was collected from institutional databases and electronic health records.

Statistical Analysis

The sample size justification was reported previously and was based on having at least 10 events (i.e. deaths) for each prognostic variable in a multivariable Cox Proportional Hazards regression model [14]. All patients with PaP score were included in this study.

We summarized the baseline demographics using descriptive statistics, including means, medians, proportions, standard deviations (SD), interquartile ranges (IQR), and 95% confidence intervals (95% CI).

We conducted analyses to examine the accuracy of the following 4 prognostic approaches:

1. Raw-CPS = CPS in days obtained from clinicians before transformation to PaP score units;
2. PaP-CPS = CPS that has been transformed into PaP score units, range 0–8.5 points;
3. PaP-without CPS = the combined score of the 5 clinical and laboratory variables in the PaP score, which is equivalent to the PaP-total score minus PaP-CPS, range 0–9 points; and
4. PaP-total score = PaP-CPS + PaP-without CPS, range 0–17.5 points

For each prognostic approach, we assessed discrimination ability using the following metrics: the C-index, the area under the receiver operating characteristics curve (AUC), net reclassification index (NRI) and integrated discrimination improvement (IDI).

Discrimination reflects how well a prognostic tool differentiates between patients who died and remained alive by a specific time frame. With the exception of C-index, we conducted the analyses for 30-day survival (alive, dead) and 100-day survival (alive, dead). 30-day survival was chosen as a cutoff because it has practical implications for assessing the quality of end-of-life care and has been used in other studies examining the accuracy of PaP score [15]. 100-day survival approximates the median overall survival for this patient cohort, and many palliative procedures are contraindicated if a patient's life expectancy is less than this time frame.

The C-index is a measure of predictive discrimination, defined as the proportion of patient pairs in which the predicted and observed survival outcomes are concordant [16]. We employed the bootstrap validation method to estimate the bias-corrected or over fitting-corrected predictive accuracy of the model. C-index ranges from 0.5 to 1, indicating no discrimination and a perfect discrimination ability to predict survival time, respectively.

Receiver operating characteristic (ROC) curve examines the performance of a binary outcome (e.g. alive or dead at 30 days) with various discrimination thresholds by plotting sensitivity (y-axis) against 1-specificity (x-axis). In contrast to the C-index which examines a continuous outcome, AUC quantifies the overall performance for discriminating a binary outcome. The AUC ranges between 0.5 and 1, with a higher value suggesting better discrimination.

Because C-index and AUC are not always sensitive to the addition of a novel prognostic marker to an existing model, we conducted reclassification statistical analyses with NRI and IDI [17, 18]. Specifically, we examined if the addition of PaP-without CPS to PaP-CPS improved its accuracy, and also whether the addition of PaP-CPS to PaP-without CPS improved its accuracy. NRI is calculated based on two components, event and non-event group. For the event group, we assign +1 if an individual is reclassified into a higher risk category in a new model compared to an existing model, -1 if an individual is reclassified into a lower risk category, and 0 if the risk category does not change. For the non-event group, we assign +1 if an individual is reclassified into lower risk category in a new model compared to an existing model, -1 if an individual is reclassified into higher risk category and 0 if the risk category does not change. Individual scores are summed separately for event

and non-event group, referred to as score(event) and score(non-event). NRI is calculated as follows: $\text{score}(\text{event})/(\text{total number of events}) + \text{score}(\text{non-event})/(\text{total number of non-events})$. IDI is computed based on integrated sensitivity and specificity which can be defined as a difference in discrimination slope between two models, an existing model and a new model [19]. Discrimination slope is calculated by subtracting an average predicted probability of non-event from an average predicted probability of event.

Statistical analyses are carried out in Statistical Analysis System (SAS version 9.2, SAS Institute, Cary, North Carolina) and R version 3.1.3. A P-value of <0.05 is considered statistically significant.

Results

Patient characteristics

216 of 222 (97%) patients enrolled onto this study had full PaP score data. We were unable to obtain CPS from the clinicians for the remaining 6 patients and thus their PaP scores could not be computed. Table 1 shows the patient characteristics and distribution of PaP score. 136/216 (63%) patients have died at the time of analysis, with a median follow-up of 239 days (IQR 186–261 days). The median overall survival was 109 days (95% CI 71–133 days).

Accuracy of CPS, PaP-without CPS and PaP-total score

The C-index for raw-CPS, PaP-CPS, PaP-without CPS and PaP-total score was 0.58, 0.56, 0.65 and 0.64, respectively (Table 2). The lower limit of the 95% confidence intervals of raw-CPS and PaP-CPS were below 0.5, suggesting that they had limited prognostic utility. The area under the ROC curve analysis showed similar findings for both 30-day survival and 100-day survival, with PaP-without CPS having the highest accuracy, followed by PaP-total score, raw-CPS, and PaP-CPS (Table 2 and Figure 1).

Accuracy of PaP-CPS versus PaP-total

As shown in Table 3, PaP-total score was significantly more accurate than PaP-CPS for both 30-day and 100-day survival. Specifically, our analyses with C-index ($P<0.0001$), AUC (30 day: <0.0001 ; 100 day: 0.0004), NRI (30 day: <0.0001 ; 100 day: 0.0006) and IDI (30 day: <0.0001 ; 100 day: <0.0001) supported that the addition of PaP-without CPS to PaP-CPS improved its accuracy.

The net reclassification table shows that when the probability of survival from PaP-total score was compared directly to PaP-CPS, PaP-total score was better in 59–75% of the patients (Table 4).

Accuracy of PaP-without CPS versus PaP-total

A comparison of PaP-total score to PaP-without CPS revealed that PaP-without CPS was significantly more accurate than PaP-total score with AUC ($P=0.03$), NRI ($P=0.0003$) and IDI ($P=0.0004$) for 30 day survival but not 100-day survival (Table 3). The net

reclassification table shows that PaP-without CPS was better at predicting the survival outcomes than PaP-total score in 53–71% of patients (Table 4).

Discussion

Using multiple statistical techniques, we found that PaP-total score was consistently more accurate than PaP-CPS alone, and that PaP-without CPS was more accurate than PaP-total score. Our findings support that clinicians should consider using prognostic scores in their routine clinical practice and research design. Furthermore, this study highlights the need to develop prognostic tools with greater accuracy based on objective prognostic markers.

In this study, experienced palliative care specialists were asked to estimate survival for hospitalized patients with advanced cancer. Consistent with the literature, the accuracy of Raw-CPS was suboptimal: only 21% (n=45) of the predictions fell within 33% of actual survival, the AUC was 0.58–0.62, and the C-index was 0.58. We found no significant difference between Raw-CPS and PaP-CPS. To our knowledge, this is the first study to examine the accuracy of CPS using multiple statistical approaches, which allows us to compare against other methods of prognostication. Our findings point to the clear need to identify better ways to predict survival more accurately. Some investigators proposed that framing the questions differently such as with the use of the surprise question and probabilistic question may improve accuracy [5, 20, 21]. Others have focused on identifying novel objective prognostic makers and developing more sophisticated prognostic models [2].

Prognostic scores are potentially attractive because they are objective, independent of clinical experience, and transferable across settings. At the same time, they are cumbersome to calculate and often difficult to interpret. Importantly, it has not been clarified if they are more accurate than clinical judgement. Thus, clinicians continue to rely on CPS to formulate prognosis. Stiel et al. examined the performance of the Palliative Prognostic Index (PPI), PaP score and PaP-CPS in 84 cancer patients. PPI had the highest correlation coefficient with actual survival (0.68), followed by PaP score (0.58) and CPS (0.56) [13]. In contrast, our robust analyses revealed that both the C-index and AUC for PaP-total score were superior to Raw-CPS and PaP-CPS. The use of reclassification statistics yielded the same conclusions. The discrepancy between these two studies may be related to patient population and statistical methods.

Why is a prognostic model more accurate than CPS, particularly when the formulation of CPS often already incorporates many known prognostic factors, such as performance status, anorexia and cachexia? Potential explanations may include (1) clinicians do not always include all relevant prognostic variables, (2) clinicians may assign different weights to prognostic variables, and (3) CPS include some other variables that may decrease its accuracy (e.g. pain, emotional connection with the patient). Further research is needed to examine these possibilities. Because PaP-total already includes PaP-CPS as part of its score, our study supports that inclusion of clinical/laboratory variables to CPS could enhance its accuracy.

Interestingly, the reverse may not be true. We were somewhat surprised to find that PaP-without CPS was more accurate than PaP-total score, suggesting that addition of PaP-CPS to PaP-without CPS may actually reduce its accuracy. In fact, PaP-CPS was given a heavy weight in PaP-total score, contributing up to 8.5 of 17.5 points. This raises the question of whether clinicians should consider using PaP-total or PaP-without CPS. Since no studies have validated or examined the interpretation of PaP-without CPS, PaP-total may be the best option at this time.

Although PaP-total score and PaP-without CPS were significantly more accurate than CPS alone, the accuracy of these scores were less than ideal. Thus, there is a need to identify novel objective prognostics markers, such as phase angle [14]. Several groups have also developed prognostic models based on objective prognostic factors only, although further validation is needed [22].

Our study has several limitations. First, patients were recruited from a tertiary care cancer center and the findings may not be generalizable to other settings. Second, our findings are dependent on multiple variables that could impact the accuracy of CPS, such as who provided the survival estimates, where was the study setting, how far along the disease trajectory were the patients, and what question was used to elicit CPS. Prognostic scores may not be significantly better if CPS had a higher accuracy with different study conditions. Thus, further research is needed to examine this question in other settings and patient populations. Importantly, the statistical approaches employed in this study may be applicable to research addressing similar questions.

In this prospective study, we found that PaP-without CPS was most accurate, followed by PaP-total score and PaP-CPS. Our findings support the use of actuarial prognostication over CPS in clinical practice to augment clinical decision making and the need to identify novel objective prognostic factors and to develop better prognostic models.

Acknowledgments

Funding

This work was supported by the National Institutes of Health Cancer Center Support Grant (CA016672 to M.P. and D.L.). D.H. is supported in part by a National Institutes of Health grant (R21CA186000-01A1), an American Cancer Society Mentored Research Scholar Grant in Applied and Clinical Research (MRS-14-1418-01-CCE), and an institutional startup grant (#18075582).

References

1. Hui D, Con A, Christie G, Hawley PH. Goals of care and end-of-life decision making for hospitalized patients at a canadian tertiary care cancer center. *J Pain Symptom Manage*. 2009; 38(6): 871–81. [PubMed: 19811887]
2. Hui D. Prognostication of Survival in Patients With Advanced Cancer: Predicting the Unpredictable? *Cancer Control*. 2015; 22(4):489–97. [PubMed: 26678976]
3. Hui D, Bansal S, Reddy A, Park M, Cortes J, Fossella F, et al. Differences in Attitudes and Beliefs toward End-of-Life Care Between Hematologic and Solid Tumor Oncology Specialists. *Ann Oncol*. 2015; 26(7):1440–6. [PubMed: 26041765]
4. Christakis NA, Lamont EB. Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ*. 2000; 320(7233):469–72. [PubMed: 10678857]

5. Hui D, Kilgore K, Nguyen L, Hall S, Fajardo J, Cox-Miller TP, et al. The accuracy of probabilistic versus temporal clinician prediction of survival for patients with advanced cancer: a preliminary report. *The Oncologist*. 2011; 16(11):1642–8. [PubMed: 21976316]
6. Perez-Cruz PE, Dos Santos R, Silva TB, Crovador CS, Nascimento MS, Hall S, et al. Longitudinal Temporal and Probabilistic Prediction of Survival in a Cohort of Patients With Advanced Cancer. *J Pain Symptom Manage*. 2014
7. Glare PA, Sinclair CT. Palliative Medicine review: prognostication. *J Palliat Med*. 2008; 11(1):84–103. [PubMed: 18370898]
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. Italian Multicenter Study Group on Palliative Care. *J Pain Symptom Manage*. 1999; 17(4):240–47. [PubMed: 10203876]
9. 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. Italian Multicenter and Study Group on Palliative Care. *J Pain Symptom Manage*. 1999; 17(4):231–39. [PubMed: 10203875]
10. Glare PA, Eychmueller S, McMahon P. Diagnostic accuracy of the palliative prognostic score in hospitalized patients with advanced cancer. *J Clin Oncol*. 2004; 22(23):4823–28. [PubMed: 15570085]
11. Maltoni M, Scarpi E, Pittureri C, Martini F, Montanari L, Amaducci E, et al. Prospective comparison of prognostic scores in palliative care cancer populations. *Oncologist*. 2012; 17(3):446–54. [PubMed: 22379068]
12. Baba M, Maeda I, Morita T, Inoue S, Ikenaga M, Matsumoto Y, et al. Survival prediction for advanced cancer patients in the real world: A comparison of the Palliative Prognostic Score, Delirium-Palliative Prognostic Score, Palliative Prognostic Index and modified Prognosis in Palliative Care Study predictor model. *Eur J Cancer*. 2015; 51(12):1618–29. [PubMed: 26074396]
13. Stiel S, Bertram L, Neuhaus S, Nauck F, Ostgathe C, Elsner F, et al. Evaluation and comparison of two prognostic scores and the physicians' estimate of survival in terminally ill patients. *Support Care Cancer*. 2010; 18(1):43–9. [PubMed: 19381693]
14. Hui D, Bansal S, Morgado M, Dev R, Chisholm G, Bruera E. Phase angle for prognostication of survival in patients with advanced cancer: preliminary findings. *Cancer*. 2014; 120(14):2207–14. [PubMed: 24899148]
15. Hui D, Kim SH, Roquemore J, Dev R, Chisholm G, Bruera E. Impact of timing and setting of palliative care referral on quality of end-of-life care in cancer patients. *Cancer*. 2014; 120(11):1743–9. [PubMed: 24967463]
16. 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(4):361–87. [PubMed: 8668867]
17. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014; 35(29):1925–31. [PubMed: 24898551]
18. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010; 21(1):128–38. [PubMed: 20010215]
19. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008; 27(2):157–72. discussion 207–12. [PubMed: 17569110]
20. Moss AH, Lunney JR, Culp S, Auber M, Kurian S, Rogers J, et al. Prognostic significance of the “surprise” question in cancer patients. *J Palliat Med*. 2010; 13(7):837–40. [PubMed: 20636154]
21. Hamano J, Morita T, Inoue S, Ikenaga M, Matsumoto Y, Sekine R, et al. Surprise Questions for Survival Prediction in Patients With Advanced Cancer: A Multicenter Prospective Cohort Study. *Oncologist*. 2015; 20(7):839–44. [PubMed: 26054631]
22. Suh SY, Choi YS, Shim JY, Kim YS, Yeom CH, Kim D, et al. Construction of a new, objective prognostic score for terminally ill cancer patients: a multicenter study. *Support Care Cancer*. 2010; 18(2):151–7. [PubMed: 19381691]

Highlights

- Clinician prediction of survival has low accuracy in the advanced cancer setting.
- For the first time, we directly compared the Palliative Prognostic Score to clinician prediction of survival using multiple statistical indices, and found that the prognostic score was more accurate.
- Palliative Prognostic Score without clinician prediction of survival was even more accurate than when clinician prediction of survival was included as part of the score.
- This study highlights the need to use objective prognostic factors/ models for prognostication instead of relying on clinical gestalt.

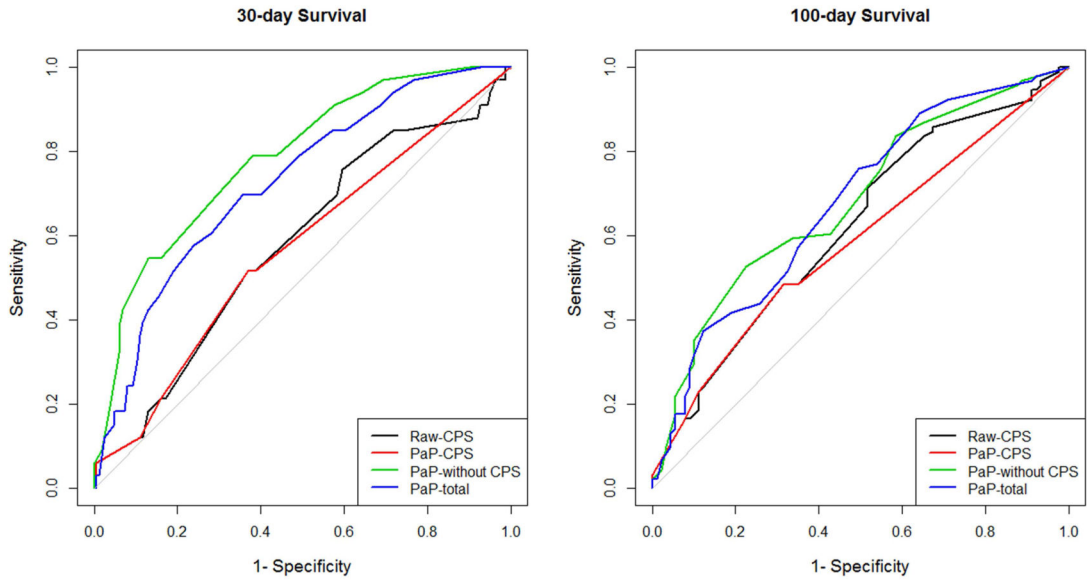


Figure 1. Discrimination of Raw-CPS, PaP-CPS, PaP-without CPS and PaP-total score
 These receiver-operating characteristics curves plot sensitivity vs. 1-specificity for (A) 30-day survival and (B) 100-day survival. PaP-without CPS (green) has the largest area under the curve and thus the best performance compared to the other variables.

Table 1

Patient characteristics (N=216)

| Characteristics | N (%) [*] |
|--|--------------------|
| Age, average (range) | 54.9 (22 – 79) |
| Sex | |
| Female | 126 (58) |
| Male | 90 (42) |
| Ethnicity | |
| White | 143 (66) |
| Black | 43 (20) |
| Hispanic | 28 (13) |
| Others | 2 (1) |
| Education | |
| High school or lower | 115 (53) |
| College | 72 (33) |
| Advanced | 29 (14) |
| Cancer | |
| Breast | 27 (13) |
| Gastrointestinal | 71 (33) |
| Genitourinary | 19 (9) |
| Gynecological | 23 (11) |
| Head and neck | 10 (5) |
| Hematological | 12 (6) |
| Others | 18 (8) |
| Respiratory | 36 (17) |
| Overall survival in days, median (95% confidence interval) | 109 (71–133) |
| Clinician prediction of survival (Raw-CPS) in days, median (interquartile range) | 90 (60–155) |
| Palliative prognostic score (PaP-total score) | |
| 0–5.5 points | 152 (70) |
| 5.6–11 points | 16 (8) |
| 11.1–17.5 points | 48 (22) |
| Clinician prediction of survival transformed into PaP score categories (PaP-CPS) | |
| >12 weeks (0 point) | 128 (59) |
| 11–12 weeks (2 points) | 3 (1) |
| 7–10 weeks (2.5 points) | 48 (22) |
| 5–6 weeks (4.5 points) | 11 (5) |
| 3–4 weeks (6 points) | 23 (11) |
| ≤2 weeks (8.5 points) | 3 (1) |
| Dyspnea in PaP score categories | |
| Absent (0 point) | 125 (58) |
| Present (1 point) | 91 (42) |
| Anorexia in PaP score categories | |

| Characteristics | N (%) [*] |
|--|--------------------|
| Absent (0 point) | 187 (87) |
| Present (1.5 points) | 29 (13) |
| Karnofsky performance status in PaP score categories | |
| 50% (0 point) | 182 (84) |
| 40% (2.5 points) | 34 (16) |
| Leukocytosis in PaP score categories | |
| 8500 cell/mm ³ (0 point) | 130 (60) |
| 8501–11000 cell/mm ³ (0.5 point) | 42 (20) |
| >11000 cell/mm ³ (1.5 points) | 44 (20) |
| Lymphocytopenia in PaP score categories | |
| 20% (0 point) | 42 (20) |
| 12%-19.9% (1 point) | 46 (21) |
| 0–11.9% (2.5 points) | 128 (59) |

Abbreviations: SD, standard deviation

* unless otherwise specified

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Discriminatory ability of raw-CPS, PaP-CPS, PaP-without CPS, PaP-total

| Model | Concordance index (95% CI) | Area under the ROC curve (95% CI) for 30-day survival | Area under the ROC curve (95% CI) for 100-day survival |
|-----------------|----------------------------|---|--|
| Raw-CPS | 0.58 (0.47, 0.68) | 0.58 (0.47, 0.68) | 0.62 (0.54, 0.70) |
| PaP-CPS | 0.56 (0.46, 0.66) | 0.57 (0.47, 0.67) | 0.59 (0.51, 0.66) |
| PaP-without CPS | 0.65 (0.55, 0.77) | 0.78 (0.70, 0.87) | 0.68 (0.61, 0.76) |
| PaP-total score | 0.64 (0.54, 0.74) | 0.73 (0.64, 0.82) | 0.68 (0.60, 0.76) |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Relative Performance of PaP-total score, PaP-CPS and PaP-without CPS

| | PaP-total score vs. PaP-CPS (i.e. can we improve the accuracy of PaP-CPS by adding PaP-without CPS?) | | PaP-total score vs. PaP-without CPS (i.e. can we improve the accuracy of PaP-without CPS by adding PaP-CPS?) | |
|---|--|---------|--|---------|
| | Difference (95% CI) | P-value | Difference (95% CI) | P-value |
| Concordance index [*] | 0.08 (0.05, 0.11) | <0.0001 | -0.01 (-0.04, 0.02) | 0.53 |
| Area under the ROC curve for 30-day survival [†] | 0.16 (0.09, 0.22) | <0.0001 | -0.06 (-0.11, -0.005) | 0.03 |
| Area under the ROC curve for 100-day survival [†] | 0.09 (0.04, 0.14) | 0.0004 | -0.005 (-0.06, 0.05) | 0.86 |
| Net reclassification improvement for 30-day survival [‡] | 77% (42%, 112%) | <0.0001 | -70% (-105%, -34%) | 0.0003 |
| Net reclassification improvement for 100-day survival [‡] | 51% (23%, 79%) | 0.0006 | -29% (-58%, -0.3%) | 0.05 |
| Integrated discrimination improvement for 30-day survival [§] | 0.08 (0.04, 0.11) | <0.0001 | -0.09 (-0.14, -0.04) | 0.0004 |
| Integrated discrimination improvement for 100-day survival [§] | 0.06 (0.03, 0.08) | <0.0001 | -0.01 (-0.04, 0.02) | 0.41 |

Abbreviations: CI, confidence interval; ROC, receiver operating characteristics

^{*} A positive value indicates better discrimination for concordance index[†] A positive value indicates better discrimination for area under the ROC curve[‡] The percentage of improvement by adding the new variable (PaP-is shown based on the net reclassification improvement index. A positive value indicates better discrimination[§] The relative change in slope is shown based on the integrated discrimination improvement. A positive value indicates better discrimination

Table 4

Net Reclassification Table

| NRI raw data | PaP-total score vs. PaP-CPS | | PaP-total score vs. PaP-without CPS | |
|----------------------|--|--|--|--|
| | PaP-total score better than PaP-CPS N (%) [*] | PaP-CPS better than PaP-total score N (%) [*] | PaP-total score better than PaP-without CPS N (%) [*] | PaP-without CPS better than PaP-total score N (%) [*] |
| Died within 30 days | 21 (64) | 12 (36) | 12 (36) | 21 (64) |
| Alive after 30 days | 122 (75) | 41 (25) | 47 (29) | 116 (71) |
| Died within 100 days | 54 (59) | 37 (41) | 43 (47) | 48 (53) |
| Alive after 100 days | 59 (66) | 30 (34) | 34 (38) | 55 (62) |

* We applied logistic regression modeling to compute the probability of an outcome of interest (e.g. death within 30 days) for each prognostication approach (i.e. PaP-CPS, PaP-without CPS or PaP-total). We then compared the approaches in pairs for each outcome. Each cell shows the number of patients in which the probability of having the outcome based on one prognostication approach is closer to predicting the outcome than the other approach, along with the row percentage in parenthesis.