Review

Comparing the performance of the palliative prognostic (PaP) score with clinical predictions of survival: A systematic review

Patrick Stone [a,*], Nicola White [a], Linda J.M. Oostendorp [a], Henry Llewellyn [a], Victoria Vickerstaff [a,b]

[a] *Marie Curie Palliative Care Research Department, Division of Psychiatry, UCL, London, UK*
[b] *Primary Care and Population Health, Institute of Epidemiology and Health Care, University College London (UCL), London, UK*

Abstract **Background:** In patients with advanced cancer, prognosis is usually determined using clinicians' predictions of survival (CPS). The palliative prognostic (PaP) score is a prognostic algorithm that was developed to predict survival in patients with advanced cancer. The score categorises patients into three risk groups in accordance with their probability of surviving for 30 days. The relative accuracy of PaP and CPS is unclear.

Design: This was a systematic review of MEDLINE, Embase, AMED, CINAHL Plus and the Cochrane Database of Systematic Reviews and Trials from inception up to June 2021. The inclusion criteria were studies in adults with advanced cancer reporting data on performance of both PaP and CPS. Data were extracted on accuracy of prognoses and where available on discrimination (area under the receiver operating characteristic curve or C-index) and/or diagnostic performance (sensitivity, specificity).

Results: Eleven studies were included. One study reported a direct comparison between PaP risk groups and equivalent risk groups defined by CPS and found that PaP was as accurate as CPS. Five studies reported discrimination of PaP as a continuous total score (rather than using the previously validated risk categories) and reported C-statistics that ranged from 0.64 (95% confidence interval [CI] 0.54, 0.74) up to 0.90 (95% CI 0.87, 0.92). Other studies compared PaP against CPS using non-equivalent metrics (e.g. comparing probability estimates against length of survival estimates).

Conclusions: PaP risk categories and CPS are equally able to discriminate between patients with different survival probabilities. Total PaP scores show good discrimination between

---

* *Corresponding author*. Marie Curie Palliative Care Research Department, Division of Psychiatry, University College London, 6th Floor, Maple House, 149 Tottenham Court Road, London, W1T 7NF, UK.
E-mail address: p.stone@ucl.ac.uk (P. Stone).

patients in accordance with their length of survival. The role of PaP in clinical practice still needs to be defined.

Trial registration: PROSPERO (CRD42021241074, 5th March 2021).

1. Background

Accurate prognoses are critical to good care, particularly for patients with serious illnesses, a view shared by patients, families, clinicians and policymakers [1−5]. For example, accurate prognoses play critical roles in guiding decisions about the likely benefit or futility of medical interventions [6,7]. In addition, decisions about different forms of care (such as the decision to continue with a particular treatment), enrolment in clinical trials and access to services are influenced by life expectancy [8,9]. More accurate prognoses might facilitate more equitable access to care by distinguishing between patients based on their proximity to death [10,11]. A realistic sense of life expectancy may also enable patients and families to make better-informed decisions about their lives and future plans [12,13].

Over the last 20 years, several prognostic scores have been developed for use in palliative care patients with advanced cancer [14,15]. One of the best validated and most widely used tools is the palliative prognostic (PaP) score [16−18]. The total PaP score is calculated by summing partial scores for six variables: dyspnoea, anorexia, Karnofsky performance status, total white blood count, lymphocyte percentage and the clinician estimated survival (in terms of two-week incremental categories ranging from 0-2 weeks up to 12 weeks+). Total PaP scores range between 0 and 17.5 points. The score categorises patients into one of three risk groups in accordance with their probability of 30-day survival: Group A (score 0−5.5; >70% probability), group B (score 6−11; 30−70% probability) and group C (score >11; <30% probability). Since its original publication in 1999, numerous studies have evaluated PaP in palliative cancer populations and have generally reported that it seems to perform well [16,18,19]. However, before the PaP should be recommended for widespread clinical use, it ought to also demonstrate at least similar, if not superior, accuracy to current usual practice. In the context of advanced cancer, the default method of prognostication is to use clinicians' predictions of survival (CPS). Although CPS has been reported to be inaccurate and over-optimistic [20,21], it correlates reasonably well with actual survival [21], and as a minimum, any new method of prognostication should perform at least as well as this benchmark.

To evaluate the relative performance of PaP against usual practice in prognostication, we undertook a systematic review of the literature for studies which compared the accuracy of PaP against the accuracy of CPS.

2. Methods

This review was prospectively registered via PROSPERO (CRD42021241074, 5th March 2021).

2.1. Data sources and searches

Search terms for 'advanced cancer', 'PaP score', 'clinical prediction of survival' and 'prognostic studies' were developed based on previous literature and guidance in prognostic research [14,20,22−24] (see Table 1 for search terms used on the OVID platform). The following databases were searched from inception up until 8th June 2021: MEDLINE, Embase, AMED, CINAHL Plus and the Cochrane Database of Systematic Reviews and Trials. Grey Literature Report (www.greylit.org) and OpenGrey (www.opengrey.eu) were also searched to identify further potentially eligible studies. Forward (via Web of Science) and backward citation searches were completed for all included studies, as well as the two previous systematic reviews of prognostic tools [14,24]. If a relevant abstract was identified, authors were contacted to check whether a full-text article was available.

2.2. Study selection

2.2.1. Inclusion criteria

Studies were included if all the following criteria were satisfied:

- Study included original data.
- At least 50% of the patient population had advanced cancer, or patients with cancer were described as 'not curative', 'palliative', having a 'terminal illness' or other synonyms.
- Patients over 18 years old.
- Data were reported on the CPS and PaP score.

2.2.2. Exclusion criteria

Studies in abstract form were excluded. Retrospective studies (such as case note reviews), studies that were not in English and studies providing no quantitative data were excluded.

Table 1
Study and clinician characteristics.

Study characteristics				Clinician prediction of survival	
First author/year	Country	*N*	Setting	*N*	Types of clinicians
Ermacora 2019 [33]	Italy	334	Hospital	2	Oncologists (1 senior physician, 1 trainee)
			Hospice	1	Nurse
Glare 2004 [37]	Australia	100	Hospital	7	Oncologists
Hui 2016 [31]	USA	216	Hospital	18[a]	PC physicians
Hui 2020 [32]	USA	204	Hospital	nr	PC physicians
Mendis 2015 [38]	Australia	644	Hospital	1[b]	PC team
Maltoni 2012 [29]	Italy	549	Hospice	nr	Physicians
Numico 2011 [34]	Italy	208[c]	Hospital	nr	Physicians
Stiel 2010 [30]	Germany	83[d]	Hospital	nr	Physician
Stone 2021 [36]	UK	1833	Community	431[e]	Doctors
			Hospital	777	Nurses
			Hospice		
Tavares 2018 [35]	Portugal	38	Hospital	nr	Junior doctors
					PC physicians
Yoon 2020 [39]	South Korea	1534	Hospital	64	Internists
			PC unit	25	Surgeons
			Home	20	Anaesthesiologists
				58	PC physicians
				23	Other

nr, not reported; PC, palliative care.

[a] The article by Hui et al., 2016, did not report the number of physicians who provided estimates, but the article by Farinholt et al., 2017, describing the same study data, reported that 18 physicians were involved in patient evaluation.

[b] One MDT was involved in this study, consisting of palliative care specialist doctors, palliative care nurses, a social worker, a pharmacist and a pastoral care worker. Most input is provided by medical and nursing team members.

[c] 208 admissions for 166 patients.

[d] Out of the total sample size of 83, 79 patients had advanced cancer.

[e] The study by Stone et al., 2021, did not uniquely identify health professionals providing a CPS, and so the precise number of individuals providing a CPS was estimated.

2.3. Selection

After the removal of duplicates, records that were clearly ineligible based on information included in titles or abstracts were excluded. Full-text articles of remaining records were then screened against full eligibility criteria. At both screening stages, screening was conducted independently by two reviewers (N.W., L.O. or H.L.). Another reviewer was consulted for any arising ambiguities (P.S.).

2.4. Quality assessment

The 'Quality In Prognosis Studies' tool was used to assess the risk of bias [25], as recommended by the Cochrane Prognostic Methods Group. The domain of 'prognostic factor measurement' was scored both for the PaP score and CPS. Two reviewers (N.W. and L.O.) scored all studies independently, and disagreements were resolved by discussion. A third reviewer (P.S.) adjudicated as necessary. No studies were excluded based on these assessments, and the results are reported for transparency.

2.5. Extraction

The following data were extracted from each article:

- A description of study population (patients and clinicians).
- Performance of PaP and CPS including methods of analysis.

2.6. Data synthesis and analysis

Descriptive summary statistics were extracted from included studies. A narrative synthesis was undertaken to compare PaP with CPS. If published data were not sufficient for synthesis, authors were contacted for further information. Depending on the outcomes reported in included studies, data were summarised in one of the following ways:

2.6.1. Performance in predicting 30-day survival probabilities

In the first instance, a synthesis of evidence directly comparing the performance of PaP and CPS at predicting 30-day survival probability was attempted. Data were synthesised relating to the accuracy of PaP at categorising patients into three risk groups in accordance with 30-day survival probabilities and to compare this with the accuracy of clinicians at undertaking the same task using clinical judgement alone. Accuracy was judged by the proportion of patients in each risk group who survived for 30 days. Thus, if the proportion of
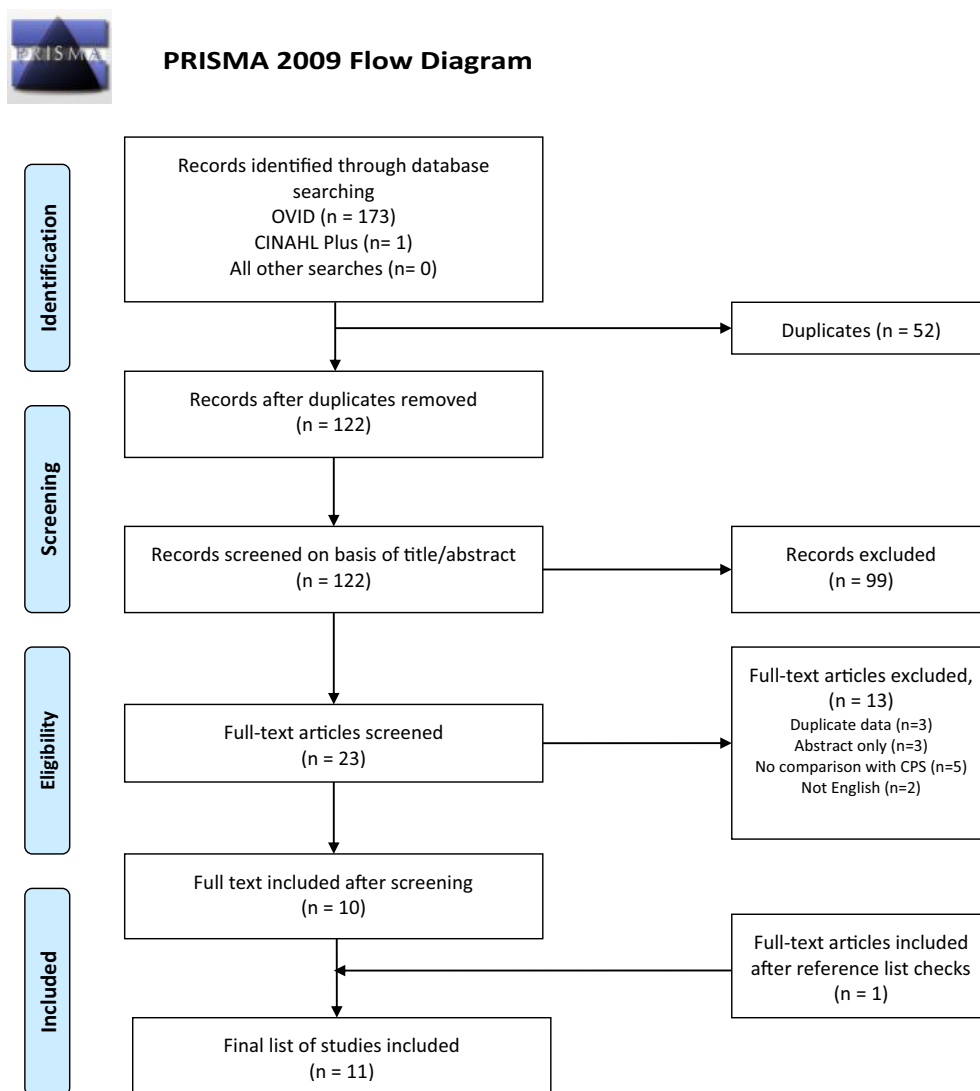
PRISMA 2009 Flow Diagram

Fig. 1. PRISMA flowchart.

patients surviving for 30 days in each of the three PaP risk groups was found to be in the correct range (group A: >70%, group B: 30−70% and group C: <30%), then PaP was judged to be accurate. The same metric was applied to clinicians' predictions about each risk group, and the two methods of categorising patients into three risk groups were directly compared.

2.6.2. Discriminatory performance

For studies that did not report a direct comparison of the accuracy of PaP and CPS at predicting 30-day survival in accordance with pre-established risk categories, we extracted and compared data on other performance measures as reported in the individual studies. Performance was usually measured in terms of discriminatory or diagnostic ability.

Discrimination in prognostic studies is usually reported in terms of area under the receiver operating characteristic curve (AUROC) or C-index. The AUROC is a measure of the ability of a tool to discriminate when there is a binary outcome (i.e. alive/dead at 30 days) [26]. The C-index is the probability that a randomly selected patient who dies is given a greater probability prediction of death than a randomly selected patient who does not die within the timeframe specified [27]. An AUROC or C-index of 0.5 means that the

prediction is no better than chance, whereas a value of 1.0 indicates perfect discrimination. Values between 0.7 and 0.8 are considered 'acceptable'; values between 0.8 and 0.9 are considered 'excellent', and values over 0.9 are considered 'outstanding' [28]. Sometimes discrimination of prognostic tools is depicted graphically using Kaplan-Meier Survival curves with accompanying log-rank test statistics.

The performance of diagnostic (and by extension prognostic) tools can be evaluated in terms of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). However, these concepts should usually only be applied to dichotomous outcomes (correct or incorrect predictions) rather than to the three risk groups generated by PaP.

3. Results

The initial search identified 174 studies. After deduplication, 122 were screened by the title and abstract. Twenty-three articles were screened at full text, of which 10 were included in this review. An additional article was identified through citation searches, making a final included total of 11 studies. See Fig. 1 for the PRISMA flowchart. In addition, two authors provided research data sets [29,30]. The reporting of outcome measurements and study confounders were the most common domains to achieve a high risk of bias, with 10 of 11 studies scoring at least moderate risk of bias (See Supplementary Material 1).

Table 1 provides a summary of the eleven included studies. Two studies were from North America [31,32], six were from Europe [29,30,33−36], two were from Oceania [37,38] and one from Asia [39]. Seven studies were hospital based, one study was hospice based and three studies had multiple settings. In seven studies, CPS was only provided by physicians [29−32,34,35,37], and in four studies, a multi-professional CPS was provided [33,36,38,39].

3.1. Performance of the PaP score

3.1.1. Direct comparisons of 30-day survival probabilities
3.1.1.1. PaP. Six articles [29,30,34,36−38] provided data about performance of PaP in the form in which it was originally presented and intended to be used. Mendis et al. [38] reported performance of PaP among inpatients and outpatients separately, resulting in 7 patient groups; however, they did not report survival data for the outpatient group because of small numbers. Across the remaining six patient groups [29,30,34,36−38], the proportion of patients surviving for 30 days among the three PaP risk groups was found to fall in the expected range (group A: >70%, group B: 30−70% and group C: <30%). Stiel et al. [30] reported that the proportion of patients who survived for 30 days in group A was lower than predicted (54% actual survival versus >70% predicted survival).

3.1.1.2. CPS. Only one study asked clinicians to undertake the same prognostic task (i.e. to estimate the 30-day survival probabilities of included participants) [36]. In this study, clinicians estimated 30-day survival probabilities, and patients' PaP risk categories were also calculated. The observed proportion of patients surviving for 30 days or more for each risk group was as follows: group A: 86.5% (687/794), group B: 46.8% (306/655) and group C: 15.4% (22/143). For comparison, patients were also divided into three groups in accordance with their probability of 30-day survival using CPS. The observed proportion of patients surviving for 30 days or more for each CPS risk group was as follows: group A: 85.3% (674/790),

Table 2
Discriminatory performance of PaP total scores and CPS.

First author	PaP		CPS	
	C-index (95% CI)	AUROC (30 days) (95% CI)	C-index (95% CI)	AUROC (30 days) (95% CI)
Ermacora [33]	nr	0.82 (0.77, 0.86)	nr	Nurse 0.78 (0.72−0.82) Doctor 1 0.77 (0.72−0.81) Doctor 2 0.76 (0.71−0.81)
Hui [31]	0.64 (0.54, 0.74)	0.73 (0.64, 0.82)	0.58 (0.47, 0.68)	0.58 (0.47, 0.68)
Hui [32]	0.70 (0.66, 0.74)	0.82 (0.73, 0.91)	0.75 (0.71, 0.79)	0.85 (0.78, 0.92)
Maltoni [29]	0.72 (0.70, 0.73)	nr	nr	nr
Yoon [39]				
PCT	0.80 (0.77, 0.82)	0.90 (0.87, 0.92)	0.78 (0.77, 0.82)	0.88 (0.86, 0.91)
PCU	0.75 (0.72, 0.78)	0.84 (0.80, 0.87)	0.71 (0.69, 0.74)	0.81 (0.77, 0.84)
HPC	0.73 (0.68, 0.78)	0.82 (0.76, 0.87)	0.73 (0.68, 0.77)	0.82 (0.76, 0.88)

PaP, palliative prognostic score; CPS, clinicians' predictions of survival; CI confidence interval; nr, not reported; PCT, inpatient palliative care consultation team; PCU, palliative care unit; HPC, home palliative care; AUROC, area under receiver operating characteristic curve; C-index, concordance index.

group B: 49.3% (292/593) and group C: 23.4% (49/209). Thus, this study showed that PaP risk categories were as good as (but no better than) CPS at discriminating which patients will or will not survive for 30 days.

3.1.2. Other comparisons of performance
3.1.2.1. PaP. Six studies [29,30,32,34,37,38] evaluated whether the three PaP categories were able to discriminate between patients with different survival risks using Kaplan-Meier survival curves and log-rank tests for trend [29,34,37,38] or chi-square tests for distribution of survival (Mantel−Cox) [30]. One study [32] generated Kaplan-Meier survival curves but did not report any tests for trend. In all patient samples, PaP categories showed statistically significant discrimination in terms of survival. Five studies used the AUROC or C-index as a measure of discrimination and reported that values ranged from 0.64 (95% confidence interval [CI] 0.54, 0.74) up to 0.90 (95% CI 0.87, 0.92), see Table 2.

3.1.2.2. CPS. None of these studies plotted equivalent survival curves for patients as per clinician predictions of 30-day survival probability. However, two studies plotted Kaplan-Meier survival curves as per clinician predictions about length of survival. Mendis et al. [38] plotted six survival curves (clinician prediction of survival in weeks: 1−2, 3−4, 5−6, 7−10, 10−12 and >12) for inpatient and outpatient samples separately. In both samples, the clinician predictions generated survival curves which showed good discrimination, and the log-rank tests were statistically significant for trend (p < 0.001). Hui et al. [39] plotted survival curves in accordance with whether clinician-predicted survival would be days (0−14 days), weeks (15−42 days) or months (≥43 days) but did not report log-rank tests.

Four studies reported the AUROC and/or C-indices for CPS and reported that values ranged from 0.58 (95% CI 0.47 to 0.68) up to 0.88 (95% CI 0.86 to 0.91).

3.1.3. Comparison of PaP and CPS
Yoon et al. [39] compared C-indices and the AUROC for an inpatient palliative care team, palliative care unit and home palliative care with CPS. They reported no statistically significant difference between them. Hui et al. [31] reported that the AUROC and C-index for PaP were significantly (p < 0.001) higher than equivalent figures for CPS (transformed into a score ranging between 0 and 8.5 rather than being treated as a continuous variable). In contrast, a 2020 study by the same author reported that the C-index and AUROC were higher for CPS than those for PaP (although differences were not statistically significant). Ermacora et al. [33] calculated separate AUROC values for three different clinicians and for PaP and reported that the AUROC was slightly higher for PaP (0.82 vs. 0.76−0.78), but differences were not statistically significant.

3.1.4. Diagnostic performance
Maltoni et al. [29] reported the sensitivity, specificity, PPV and NPV of PaP at predicting 30-day survival. To calculate these performance measures with a dichotomous outcome, the authors calculated the best cut-off for PaP at discriminating whether patients would be dead/ alive at 30 days. The optimum cut-off in their patient sample was a score of 5 (this does not correspond to the pre-established cut-offs for generating the three PaP risk categories). Using this cut-off, the authors reported that PaP had a sensitivity of 91.5% (95% CI 88.5−94.5), specificity of 57.7% (95% CI 51.2−64.3), PPV of 76.4% (95% CI 71.4−81.4), NPV of 81.9% (95% CI 75.9−88.0) and accuracy of 88.0% (95% CI 84.9−91.1). For comparison, they reported that the accuracy of CPS was 75.6% (no confidence intervals provided). No statistical test was reported for comparison with CPS.

Stiel et al. [30] also calculated the PPV and NPV for PaP. For their calculation, they used the originally presented cut-off scores for distinguishing PaP groups A, B and C. However, because the PPV and NPV can only be calculated for dichotomous variables, they merged groups B and C when comparing against group A and they merged groups A and B when comparing against group C. Moreover, they assumed that a patient in group A (>70% probability of surviving for 30 days) was definitively predicted to survive for 30 days and a patient in group C (<30% probability of surviving for 30 days) was definitively predicted to die within this timeframe. Using these definitions, they reported that group A predictions had a sensitivity of 78% and a specificity of 65%. Group C predictions had a sensitivity of 67% and a specificity of 100%. No directly equivalent analysis was undertaken for CPS.

4. Discussion

After a systematic search of the literature, we identified eleven studies that purported to compare the accuracy of PaP against accuracy of CPS. Only one of the identified studies made a direct comparison between PaP, in the format in which it was originally presented for use, and clinician predictions about probability of surviving for 30 days. This study found that performance of PaP was as good as, but not superior to, clinical prediction alone.

Most commonly, the overall performance of PaP was summarised using the AUROC or C-index. These summary statistics represent the ability of a score to discriminate between patients with different survival prospects. Prognostic scores which have poor discrimination are of no clinical use, whereas higher scores on these indexes are generally a sign that one prognostic tool is better than another. However, neither the AUROC nor C-index values are easy to interpret in the case of a prognostic tool such as PaP. The AUROC and

C-index represent the probability that PaP scores of two randomly selected patients (with different outcomes) would attribute the greater risk to the patient with the higher score. Thus, both the AUROC and C-index are calculated on the assumption that a patient with a total PaP score of 5 is predicted to have a greater risk of death than a patient with a PaP score of 2. However, as originally constructed (and as recommended for use), PaP predicts that both patients will have the same probability of surviving for 30 days. Because both patients would be categorised as being in risk group A, their probability of 30-day survival would both be >70%. Therefore, the assumptions underlying the calculation of the AUROC and C-index do not hold. The AUROC and C-index calculation assumes that the PaP score is a continuous variable with higher scores representing worse prognosis, whereas in fact the only meaning that can be attributed to a total PaP score is dependent on the risk group into which it places the patient. A few studies also calculated AUROC/C-indices for CPS. However, no studies based these calculations on clinician predictions of probability of surviving for 30 days, but rather they used clinician predictions of length of survival. Thus, we found that studies that purported to compare the discrimination of PaP with the discrimination of CPS were comparing the discrimination of the total PaP score (using an unvalidated scoring method) against the discrimination of clinicians undertaking a different prognostic task.

Sensitivity, specificity, PPV and NPV are useful summary statistics for evaluating the performance of diagnostic or prognostic tools. However, for these statistics to be calculated, a predictive tool needs to produce a definitive dichotomous outcome (either the test is positive or negative; either the patient is predicted to live or die). In the case of PaP, the prediction is neither definitive nor dichotomous. It is not definitive because a patient in PaP risk category A (>70% probability of 30-day survival) is not definitively predicted to survive for 30 days. The status of the survival prediction for patients in risk category B (30–70% probability of 30-day survival) is indeterminate—the prediction is equally compatible with either outcome (dead or alive at 30 days). Furthermore, PaP outcomes are not dichotomous because there are three PaP risk categories. Therefore, sensitivity, specificity, PPV and NPV can only be calculated by comparing one risk category against the other two combined.

Despite the limitations of previous studies with regard to comparing PaP against CPS, there is still ample evidence that PaP is a reliable and (within its own terms) accurate prognostic tool. The use of PaP has been reported in numerous studies spanning palliative care [40], oncology [18] and non-malignant disease [41]. Multiple studies [16,18,19] have demonstrated that PaP risk categories can discriminate between patients with different survival prospects, and the observed probabilities of survival at 30 days in each of the risk categories broadly correspond to the predicted risks (<30%, 30–70% and >70%). In this sense, it is fair to say that PaP is the most validated prognostic score in palliative care patients.

Even if PaP is only as good as (but not better than) CPS, it might have other features to recommend it for routine clinical use. For instance, PaP scores are more likely to be objective than unalloyed CPS and so there is likely to be less interobserver disagreement about which risk category patients belong to. PaP scores may also be more suitable for use by less experienced clinicians or may have a role acting as a second opinion. Despite these potential additional benefits, there are also features of PaP which suggest that it may be less useful than clinician predictions of survival. The range of the PaP risk categories is quite broad and may not be specific enough to inform clinical judgement. For example, there are presumably cases when clinicians' intuition is that patients have next to no chance of surviving for 30 days, and they can thus be quite confident in their decision-making. Compare this with being informed that a patient is in risk group C, with a <30% chance of surviving for 30 days. Is this level of certainty enough to base clinical decisions on? The situation is even more acute when one considers the clinical interpretation of being told that someone is in risk category B (30–70% probability of surviving for 30 days). What does such a prognosis mean and how should such information be communicated to patients?

Future research needs to focus on establishing the role of PaP in clinical practice. This will require so-called impact studies [42] in which routine use of PaP (and/or other prognostic tools) is compared against 'usual practice' (clinical predictions of survival). To establish its place in clinical practice, PaP will ultimately need to be able to demonstrate improvements in measurable, relevant clinical outcomes.

Funding

Conflict of interest statement

P.S., N.W., L.J.M.O., H.L. and V.V. declare no competing conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2021.08.049.

References

[1] Adams E, Boulton M, Watson E. The information needs of partners and family members of cancer patients: a systematic literature review. Patient Educ Counsel 2009;77(2):179—86. https://doi.org/10.1016/j.pec.2009.03.027 [published Online First: 2009/05/02].

[2] 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(7452): 1343. https://doi.org/10.1136/bmj.38103.423576.55 [published Online First: 2004/05/21].

[3] Kutner JS, Steiner JF, Corbett KK, et al. Information needs in terminal illness. Soc Sci Med 1999;48(10):1341—52. https://doi.org/10.1016/s0277-9536(98)00453-5 [published Online First: 1999/06/16].

[4] Steinhauser K, Christakis N, Clipp E, et al. Preparing for the end of life: preferences of patients, families, physicians, and other care providers. J Pain Symptom Manag 2001;22(3):727—37.

[5] Steinhauser KE, Christakis NA, Clipp EC, et al. Factors considered important at the end of life by patients, family, physicians, and other care providers. J Am Med Assoc 2000;284(19): 2476—82. https://doi.org/10.1001/jama.284.19.2476 [published Online First: 2000/11/14].

[6] Gill TM. The central role of prognosis in clinical decision making. J Am Med Assoc 2012;307(2):199—200. https://doi.org/10.1001/jama.2011.1992 [published Online First: 2012/01/12].

[7] Gospodarowicz M, Mackillop W, O'Sullivan B, et al. Prognostic factors in clinical decision making: the future. Cancer 2001;91(8 Suppl):1688—95. https://doi.org/10.1002/1097-0142(20010415)91: 8+<1688::aid-cncr1184>3.0.co;2-7 [published Online First: 2001/04/20].

[8] Prognostic indicator guidance (PIG). © the gold standards framework centre. In: Thomas K, et al., editors. End of life care CIC. 4th ed. Oct 2011.

[9] NICE. End of life care for adults (QS13). 2019. Published date: November 2011 Last updated: March 2017.

[10] Grande GE, Addington-Hall JM, Todd CJ. Place of death and access to home care services: are certain patient groups at a disadvantage? Soc Sci Med 1998;47(5):565—79. https://doi.org/10.1016/s0277-9536(98)00115-4 [published Online First: 1998/08/05].

[11] Lewis JM, DiGiacomo M, Currow DC, et al. Dying in the margins: understanding palliative care and socioeconomic deprivation in the developed world. J Pain Symptom Manag 2011;42(1): 105—18. https://doi.org/10.1016/j.jpainsymman.2010.10.265 [published Online First: 2011/03/16].

[12] Franks PJ, Salisbury C, Bosanquet N, et al. The level of need for palliative care: a systematic review of the literature. Palliat Med 2000;14(2):93—104. https://doi.org/10.1191/026921600669997774 [published Online First: 2000/06/01].

[13] Glare PA, Sinclair CT. Palliative medicine review: prognostication. J Palliat Med 2008;11(1):84—103. https://doi.org/10.1089/jpm.2008.9992 [published Online First: 2008/03/29].

[14] Simmons CPL, McMillan DC, McWilliams K, et al. Prognostic tools in patients with advanced cancer: a systematic review. e10 J Pain Symptom Manag 2017;53(5):962—70. https://doi.org/10.1016/j.jpainsymman.2016.12.330 [published Online First: 2017/01/08].

[15] Stone PC, Lund S. Predicting prognosis in patients with advanced cancer. Ann Oncol 2007;18(6):971—6. https://doi.org/10.1093/annonc/mdl343.

[16] Maltoni M, Nanni O, Pirovano M, et al. Successful validation of the palliative prognostic score in terminally ill cancer patients. J Pain Symptom Manag 1999;17(4):240—7. https://doi.org/10.1016/S0885-3924(98)00146-8.

[17] Pirovano M, Maltoni M, Nanni O, et al. A new palliative prognostic score: a first step for the staging of terminally ill cancer patients. J Pain Symptom Manag 1999;17(4):231—9. https://doi.org/10.1016/S0885-3924(98)00145-6.

[18] Baba M, Maeda I, Morita T, 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. https://doi.org/10.1016/j.ejca.2015.04.025.

[19] 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 Manag 2001; 22(5):891—8.

[20] White N, Reid F, Harris A, et al. A systematic review of predictions of survival in palliative care: how accurate are clinicians and who are the experts? PloS One 2016;11(8):e0161407. https://doi.org/10.1371/journal.pone.0161407 [published Online First: 2016/08/26].

[21] Glare P, Virik K, Jones M, et al. A systematic review of physicians' survival predictions in terminally ill cancer patients. BMJ 2003;327(7408):195—8.

[22] Cochrane Methods Prognosis [Available from: http://methods.cochrane.org/prognosis.

[23] Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001;323(7306):224—8.

[24] Lau F, Cloutier-Fisher D, Kuziemsky C, et al. A systematic review of prognostic tools for estimating survival time in palliative care. J Palliat Care 2007;23(2):93—112 [published Online First: 2007/09/15].

[25] Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158(4): 280—6.

[26] Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology 2010;21(1):128.

[27] Harrell Jr FE, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. J Am Med Assoc 1982;247(18):2543—6. https://doi.org/10.1001/jama.1982.03320430047030.

[28] Hosmer D, Lemeshow S. Applied logistic regression. 2nd ed. New York, NY: John Wiley & Sons; 2000. p. 160—4.

[29] Maltoni M, Scarpi E, Pittureri C, et al. Prospective comparison of prognostic scores in palliative care cancer populations. Oncol 2012;17(3):446—54. https://doi.org/10.1634/theoncologist.2011-0397 [published Online First: 2012/03/02].

[30] Stiel S, Bertram L, Neuhaus S, et al. Evaluation and comparison of two prognostic scores and the physicians' estimate of survival in terminally ill patients. Support Care Canc 2010;18(1): 43—9.

[31] Hui D, Park M, Liu D, et al. Clinician prediction of survival versus the Palliative Prognostic Score: which approach is more accurate? Eur J Cancer 2016;64:89—95.

[32] Hui D, Ross J, Park M, et al. Predicting survival in patients with advanced cancer in the last weeks of life: how accurate are prognostic models compared to clinicians' estimates? Palliat Med 2020. https://doi.org/10.1177/0269216319873261.

[33] Ermacora P, Mazzer M, Isola M, et al. Prognostic evaluation in palliative care: final results from a prospective cohort study. Support Care Cancer : Off J Multinat Assoc Supp Care Cancer 2019;27(6):2095—102. https://doi.org/10.1007/s00520-018-4463-z.

[34] Numico G, Occelli M, Russi EG, et al. Survival prediction and frequency of anticancer treatment in cancer patients hospitalized

due to acute conditions. Role of clinical parameters and PaP score. Support Care Canc 2011;19(11):1823—30.

[35] Tavares T, Oliveira M, Goncalves J, et al. Predicting prognosis in patients with advanced cancer: a prospective study. Palliat Med 2018;32(2):413—6. https://doi.org/10.1177/0269216317705788 [published Online First: 2017/05/11].

[36] Stone P, Vickerstaff V, Kalpakidou A, et al. Prognostic tools or clinical predictions: which are better in palliative care? PloS One 2021;16(4):e0249763.

[37] Glare PA, Eychmueller S, McMahon P. Diagnostic accuracy of the palliative prognostic score in hospitalized patients with advanced cancer.[Erratum appears in J Clin Oncol. 2005 Jan 1; 23(1):248]. J Clin Oncol 2004;22(23):4823—8.

[38] Mendis R, Soo WK, Zannino D, et al. Multidisciplinary prognostication using the palliative prognostic score in an Australian cancer center. Palliat Care 2015;9:7—14. https://doi.org/10.4137/PCRT.S24411 [published Online First: 2015/08/27].

[39] Yoon S-J, Suh S-Y, Hui D, et al. Accuracy of the palliative prognostic score with or without clinicians' prediction of survival in patients with far advanced cancer. J Pain Symptom Manag 2020. https://doi.org/10.1016/j.jpainsymman.2020.10.019.

[40] Tarumi Y, Watanabe SM, Lau F, et al. Evaluation of the palliative prognostic score (PaP) and routinely collected clinical data in prognostication of survival for patients referred to a palliative care consultation service in an acute care hospital. J Pain Symptom Manag 2011;42(3):419—31. https://doi.org/10.1016/j.jpainsymman.2010.12.013.

[41] Glare P, Eychmueller S, Virik K. The use of the palliative prognostic score in patients with diagnoses other than cancer. J Pain Symptom Manag 2003;26(4):883—5.

[42] Steyerberg EW, Moons KGM, van der Windt DA, et al. Prognosis research strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013;10(2):e1001381. https://doi.org/10.1371/journal.pmed.1001381.