Clinical and Translational Radiation Oncology 37 (2022) 137–144

Contents lists available at ScienceDirect

Clinical and Translational Radiation Oncology

journal homepage: www.sciencedirect.com/journal/clinical-and-translational-radiation-oncology

Original Research Article

Replacing performance status with a simple patient-reported outcome in palliative radiotherapy prognostic modelling

Daniel Howdon [a,*], Wilbert van den Hout [b], Yvette van der Linden [c], Katie Spencer [a,d]

[a] Academic Unit of Health Economics, Leeds Institute of Health Sciences, Clarendon Way, Woodhouse, Leeds LS2 9LU, UK
[b] Biomedical Data Sciences, Leiden University Medical Center, the Netherlands
[c] Dept of Radiotherapy/Centre of Expertise in Palliative Care, Leiden University Medical Centre, the Netherlands
[d] Leeds Cancer Centre, Leeds Teaching Hospitals NHS Trust, UK

A B S T R A C T

Background and purpose: Prognostication is key to determining care in advanced incurable cancer. Although performance status (PS) has been shown to be a strong prognostic predictor, inter-rater reliability is limited, restricting models to specialist settings. This study assessed the extent to which a simple patient-reported outcome measure (PROM), the EQ-5D, may replace PS for prognosis of patients with bone metastases.
Materials and methods: Data from 1,011 patients in the Dutch Bone Metastasis Study were used. Cox proportional hazards models were developed to investigate the prognostic value of models incorporating PS alone, the EQ-5D SC dimension alone, all EQ-5D dimensions and EQ–VAS, and finally all dimensions and PS. Three prognostic groups were identified and performance assessed using the Harrell's C-index and Altman-Royston index of separation.
Results: Replacing performance status (PS) with the self-care (SC) dimension of the EQ-5D provides similar model performance. In our SC-based model, three groups are identified with median survival of 86 days (95 % CI 76–101), 174 days (95 % CI 145–213), and 483 days (95 % CI 431–539). Whilst not statistically significantly different, the C-index was 0.706 for the PS-only model, 0.718 for SC-only and 0.717 in our full model, suggesting patient-report outcome models perform as well as that based on PS.
Conclusion: Prognostic performance was similar across all models. The SC model provides prognostic value similar to that of PS, particularly where a prognosis of<6 months is considered. Larger, more contemporaneous studies are needed to assess the extent to which PROMs may be of prognostic value, particularly where specialist assessment is less feasible.

Introduction

Estimates of expected prognosis play a key role in clinical decision-making in advanced incurable cancer. It is, however, well documented that oncologists' predictions of prognosis are often inaccurate and specifically tend to be overly optimistic. [1–2] It has been suggested that optimistic prognostic estimates may play a role in overly aggressive care near the end of life. [3–5] As such, a number of prognostic models have been developed in order to support improved prognostication and thus better informed treatment decisions. [6–9] Multiple models have been developed in palliative radiotherapy where the decision to treat and to fractionate may vary with the expected prognosis of the patient.

Specialised radiotherapy treatment centres are, however, usually located in large cities. As a consequence, many patients will first be cared for by local non-specialist healthcare providers, often without specific experience in the assessment of prognosis in advanced incurable cancer. Travel to a treating centre, and the disruption associated with this, may therefore be undertaken before such an assessment takes place. Alternative prognostic models might help to avoid this in patients very close to the end of life. Whilst previously developed models vary, they broadly rely upon a number of key variables with primary cancer diagnosis and performance status being particularly notable for their presence in a wide range of models. [9].

Performance status is a measure of a patient's ability to carry out their activities of daily living. It has been demonstrated to be a strong predictor of prognosis, however, inter-rater reliability of performance status has been found to be very mixed. [10] Notably, concordance appears to be particularly reduced in comparisons between specialisations e.g. general or palliative care physician versus radiation oncologist. [10] This may in part reflect the routine use of performance status by oncologists distinct from its limited use in other medical and surgical specialities. It is notable that models incorporating performance status in the assessment of prognosis following palliative radiotherapy have relied upon performance status as determined by the treating oncologist. As such, within these models, the challenges of inter-rater reliability are likely to be reduced. [11] The impact of incorporating performance

* Corresponding author.
E-mail address: d.howdon@leeds.ac.uk (D. Howdon).

https://doi.org/10.1016/j.ctro.2022.09.008
Received 21 July 2022; Received in revised form 21 September 2022; Accepted 25 September 2022
Available online 3 October 2022

Table 1
Variables incorporated into alternative model specifications.

Model	Primary site of diagnosis	Non-bone metastases	Karnovsky performance score	EQ-5D self-care only	EQ-5D all dimensions	EQ-VAS
ps_only	✓	✓	✓			
sc_only	✓	✓		✓		
sc_ps	✓	✓	✓	✓		
alleq5d_vas	✓	✓			✓	✓
vas_only	✓	✓				✓
vas_ps	✓	✓	✓			✓
full	✓	✓	✓		✓	✓

status, as assessed by a wider clinical team, upon prognostic model performance is not clear. If an increase in measurement error (due to inter-rater reliability) is seen, model discrimination may deteriorate, reducing validity outside of the oncology clinic. [12] As such, developing prognostic models which are not reliant upon performance status will be valuable. Particularly in a setting where patients with very limited prognosis might experience significant disruption to undergo assessment, and potentially a subsequent intervention with limited benefit, very close to the end of life.

Patient reported outcomes have been demonstrated previously to have prognostic value in a range of diagnoses, including advanced incurable cancer. [13–16] There is also good reason to think that these outcomes might align with performance status; the measurement of performance status relies upon an assessment of an individual's ability to carry out daily activities, mobilise and perform self-care tasks. [17] These dimensions align closely with a range of patient reported outcome questionnaires and specifically are key to the EuroQoL group's EQ-5D questionnaire. [18].

Given the potential limitations to the use of performance status where prognostic models are used beyond the oncology clinic this study uses data collected within the Dutch Bone Metastasis Study to assess the extent to which patient-reported EQ-5D questionnaire results may be able to replace clinician assessed performance status in a patient-assessed prognostic model. [19–20].

Materials and methods

The DBMS dataset contains information on 1,157 patients treated within a randomised study assessing dose fractionation of palliative radiotherapy for uncomplicated painful bone metastases between 1996 and 1998. 146 patients with missing data within the baseline EQ-5D questionnaire were excluded, leaving 1,011 patients (complete case analysis). No significant differences were observed (at the 10 % level) in the survival time or baseline characteristics of included and excluded individuals. 264, or 26.1 % of, the patients were censored for survival (median follow-up time of censored patients 425 days). The median survival in the full sample was 188 days, compared to 190 days after exclusion of observations exhibiting missingness on the EQ-5D.

The EQ-5D provides a simple measure including five dimension questions assessing an individual's mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension of the EQ-5D in our dataset includes three levels of response (no, some, or severe problems). In addition, a 0–100 visual analogue scale (EQ-VAS) allows patients to report their self-rated overall health.

Multivariable Cox regression was used to estimate models incorporating Karnofsky performances status, the three-level EQ-5D (based on mailed patient-completed questionnaires [20–21], cancer diagnosis and disease extent. The palliative radiotherapy regimen used for bone metastases has been shown not to influence survival and was therefore not included within these models. [8,19] Models based on combinations of these variables are used to predict median survival times, with individuals grouped into terciles according to these survival times, resulting in three prognostic groups for each model [6].

A baseline model incorporates only the variables used by Chow et al [6]: a dichotomised indicator of Karnovsky performance score (<=60 vs > 60), the site of primary cancer diagnosis (breast, prostate, lung, other), and the presence of visceral metastases.

Further models incorporate these latter two variables as well as varying combinations of performance status, the EQ-5D dimensions and EQ-VAS [18].

The first alternative model aimed to use the EQ-5D dimensions to develop a maximally parsimonious model [6]. In doing so, we initially estimated the relationship between all dimensions, retaining a single dimension with the greatest significance in multivariable Cox proportional hazards regression. This resulted in a model which used only the

Table 2
Baseline characteristics of the study population. N = 1,011.

Variable	Mean / *n*	sd / %
Age (years)	64.78	11.27
Primary site of diagnosis		
Breast	*391*	*38.67 %*
Prostate	*239*	*23.64 %*
Lung	*246*	*24.33 %*
Other	*135*	*13.35 %*
Metastatic sites		
Bone metastases only	*737*	*72.90 %*
Non-bone metastases	*274*	*27.10 %*
Performance status		
KPS ≤ 60	*294*	*29.08 %*
KPS > 60	*717*	*70.92 %*
EQ-5D mobility		
1	*538*	*53.21 %*
2	*363*	*35.91 %*
3	*110*	*10.88 %*
EQ-5D self-care		
1	*512*	*50.64 %*
2	*396*	*39.17 %*
3	*103*	*10.19 %*
EQ-5D usual activities		
1	*135*	*13.35 %*
2	*429*	*42.43 %*
3	*447*	*44.21 %*
EQ-5D pain and discomfort		
1	*23*	*2.27 %*
2	*634*	*62.71 %*
3	*354*	*35.01 %*
EQ-5D anxiety and depression		
1	*450*	*44.51 %*
2	*479*	*47.38 %*
3	*82*	*8.11 %*
EQ-VAS		
<10	*45*	*4.45 %*
10–19	*94*	*9.30 %*
20–29	*155*	*15.33 %*
30–39	*122*	*12.07 %*
40–49	*172*	*17.01 %*
50–59	*141*	*13.95 %*
60–69	*117*	*11.57 %*
70–79	*104*	*10.29 %*
80+	*61*	*6.03 %*

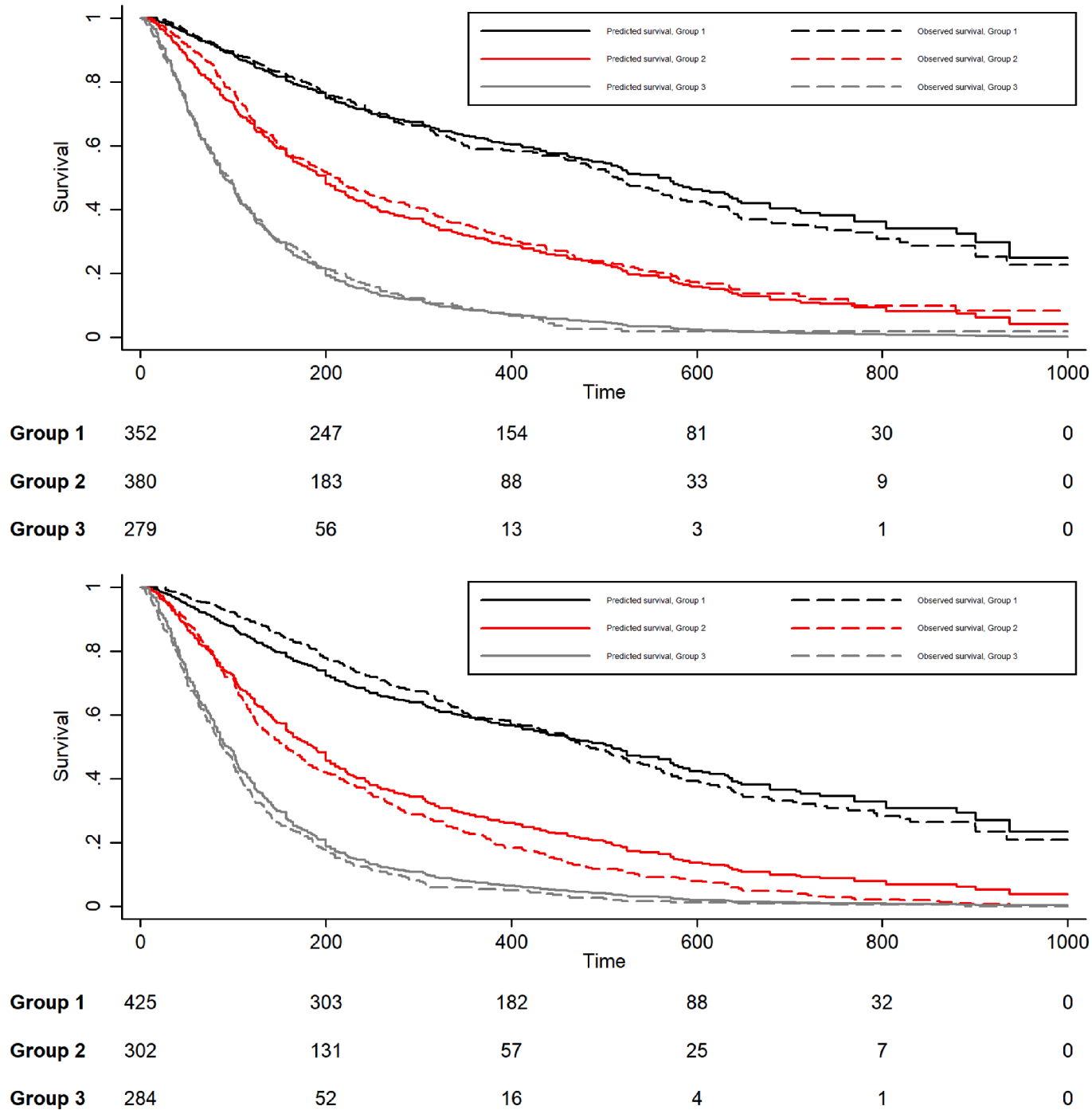

Fig. 1. Kaplan-Meier survival curves in observed data and model predicted survival in a) PS-only and b) SC-only models.

self-care EQ-5D dimension.

Further alternative EQ-5D based models incorporated only the EQ-VAS and separately all EQ-5D dimensions (as categorical variables) and the EQ-VAS in order to examine the additional benefits available from these more complex models. The incremental benefit of including performance status over and above the EQ-5D questionnaire information was assessed for each EQ-5D based model, and a final model includes all information included in any of our models. Table 1 summarises variable inclusion.

We present two diagnostic measures of model performance for all models: the Harrell's C-index and the Altman-Royston index of separation. [22–23] The Harrell's C-index quantifies the extent to which the model accurately allocates pairs of individuals to appropriate groups (discrimination). The Altman-Royston index of separation gives an indication of the meaningfulness of the three groups produced by each model, defined by the mean survival probability to a particular point in time in the group with the highest survival probability minus that of the group with the lowest survival probability. We calculate this value at intervals of 50 days in order to assess whether the most preferred model may vary depending on the period of prognosis required.

Rather than present a single split of the sample into training and validation, we carry out bootstrap resampling with replacement of the observed data. This involves creating 10,000 datasets of 1,011 sampled observations. We then split each bootstrap sample into training and validation (assigning as close to 50 % to each as is possible given the odd number of observations in our sample), and present diagnostic statistics

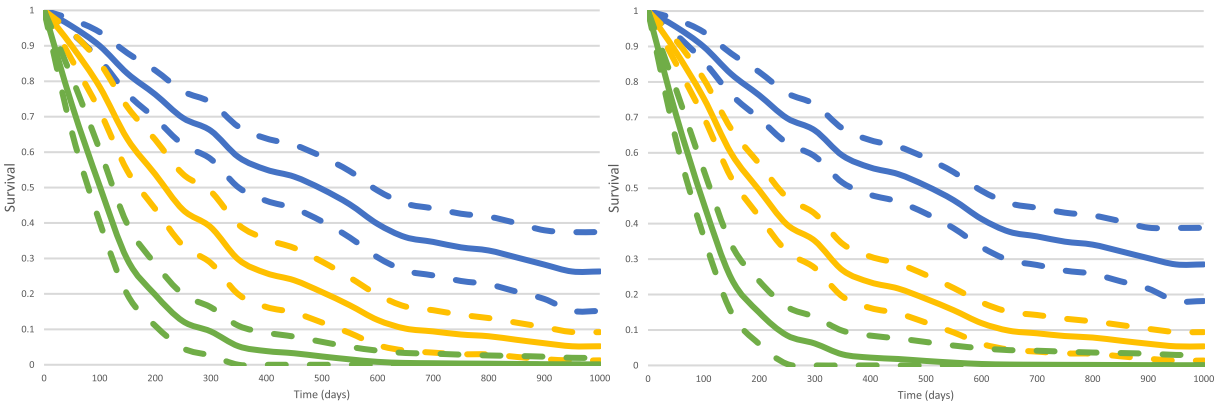

Fig. 2. Estimated out of sample survival for a) PS-only and b) SC-only models (solid lines, point estimates; dashed lines, 95% confidence intervals).

on these in order to characterise uncertainty. This guards against overfitting, and allows internal validation of our results [24].

Results

Table 2 presents baseline characteristics of the 1,011 included patients. Median survival time for these 1,011 individuals was 190 days (95 % CI 176–209). Based on the baseline performance status model, the observed survival times for the three groups in this population were 95 days (95 % CI 79–106), 210 days (95 % CI 181–245) and 515 days (95 % CI 458–572) for groups 1, 2 and 3 respectively (model C-statistic 0.706).

When replacing PS in the multi-variable Cox model, the EQ-5D self-care dimensions is found to be a significant predictor of survival probability ($p < 0.001$). Based on this model the median survival time for group 1 was 86 days (95 % CI 76–101), for group 2 174 days (95 % CI 145–213), and for group 3 483 days (95 % CI 431–539) (C-statistic 0.718). For the baseline (PS only) and SC-only models we present predicted and observed survival curves by group in Fig. 1. Bootstrapped uncertainty around our survival curves is presented in Fig. 2. Compared to those reporting a level of 1 in the EQ-5D self-care dimension, in our SC-only model, individuals reporting levels 2 and 3 have an elevation in their hazard of death of around 47 % and 95 % respectively ($p < 0.001$).

Table 3 presents the model estimates for all considered models and Harrel's C-statistic of each. For all variables in our EQ-5D + EQ-VAS model with significance of $p < 0.05$, we observe relationships with the hazard of death in expected directions – i.e. with worse self-reported health implying a higher hazard of survival. In this model, only one level of the mobility dimension is found to be significant at the 5 % level, with self-care continuing to offer the strongest such relevance.

Whilst no statistical difference is observed between models, the estimated discrimination of the EQ-5D dimensions models (self-care only (C-index 0.718) and EQ-5D dimensions plus EQ-VAS (0.717)) appears superior to that of the baseline performance status model (0.706), whilst the EQ-VAS alone performs worse (0.702). Conversely, the incorporation of performance status into these patient-assessed models provides only a marginal improvement in the estimated model performance.

Notably, the Altman-Royston separation index demonstrates limited separation at early time points: the average difference in the probability of dying between groups is inevitably small at this point due to a low number of deaths. With longer follow-up, differences between indices emerge (Fig. 3, also Table A1), although no statistically significant difference was observed. In the first six months, dimension-based models (including the maximally parsimonious self-care only model) appear to perform best whilst in longer follow-up those incorporating the EQ-VAS deliver improved separation with performance status providing a further marginal improvement in some cases.

Discussion

We demonstrate that as a quick and easy to complete patient questionnaire the EQ-5D offers valuable prognostic information which may be able to replace the role of clinician-assessed performance status in a simple prognostic model. Whilst we do confirm the prognostic value of clinician-assessed performance status by itself, we find that its additional value is reduced in the presence of the EQ-5D, its inclusion providing a non-significant marginal benefit in terms of model performance.

The model including performance status alone performs poorly in this sample, both in terms of discrimination and separation. In all EQ-5D based models, the addition of performance status appears to deliver a non-significant marginal improvement in discrimination. Similarly, model separation appears minimally impacted by the addition of performance status. Indeed, during the first 5 months of follow-up, the model incorporating the EQ-5D self-care dimension alone delivered the best separation. The simplicity of the EQ-5D questionnaire and independence of the predictors from clinician judgement offer a key advantage over existing models, potentially ensuring that such a model is accessible to non-specialist clinicians. The extent to which the extra discrimination delivered by performance status justifies its inclusion, as compared to a model which is independent of specialist clinician judgement, will be dependent upon the proposed use of the model.

By introducing the use of bootstrap resampling methods, we are able to demonstrate robust evidence of separation between groups. Our results suggest that depending upon the relevant time-period for prediction, the optimal variables for inclusion in the model may vary. For example, any benefit of including the EQ-VAS as a predictor variable only becomes apparent for predicted survival beyond 150 days. Conversely, compared to more complex models the most parsimonious model appears to provide similar separation between prognostic groups where very short prognosis is considered. It must be noted, however, that these differences are not found to be significant at conventional levels, and our results point to the need for further research in this area. Future studies should carefully define the question to be answered by the developed model and proposed environment for implementation. For example, defining the probability of survival at a very short interval may be valuable in considering the role of radiotherapy. In this case independence from a specialist clinician assessment may be desirable to avoid unnecessary travel for specialist review near the end of life. Conversely, determining if fractionation is appropriate would require a longer time horizon and thus incorporation of other variables. For these later end-points, relevant to the specialist clinician, inclusion of a clinician-assessed performance status may remain appropriate. Model usage and thus priorities should be identified a priori and variables incorporated to deliver these.

Whilst performance status has inherent limitations due to inter-rater

Table 3
Cox proportional hazards model estimated coefficients from all 1,011 observations in our estimation sample.

	ps_only	sc_only	sc_ps	vas_only	vas_ps	alleq5d_vas	full
	HR / 95 % CI / p	HR / 95 % CI / p	HR / 95 % CI / p	HR / 95 % CI / p	HR / 95 % CI / p	HR / 95 % CI / p	HR / 95 % CI / p
Primary diagnosis (breast omitted)							
Prostate	2.192	2.105	2.228	2.069	2.235	2.195	2.28
	[1.779,2.700]	[1.712,2.587]	[1.809,2.744]	[1.681,2.547]	[1.812,2.758]	[1.775,2.715]	[1.842,2.823]
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Lung	4.684	4.807	4.844	4.621	4.745	5.041	5.012
	[3.853,5.694]	[3.954,5.844]	[3.981,5.895]	[3.803,5.615]	[3.899,5.774]	[4.126,6.160]	[4.101,6.125]
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Other	3.533	3.867	3.833	3.356	3.499	4.044	3.92
	[2.805,4.449]	[3.062,4.884]	[3.035,4.841]	[2.662,4.232]	[2.776,4.412]	[3.182,5.140]	[3.085,4.980]
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Site of metastases (bone only omitted)							
Bone & non-bone	1.527	1.497	1.51	1.636	1.604	1.57	1.579
	[1.291,1.806]	[1.264,1.772]	[1.276,1.788]	[1.379,1.942]	[1.352,1.904]	[1.321,1.866]	[1.328,1.877]
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Performance status (>60 omitted)							
KPS <=60	2.176		1.768		2.02		1.673
	[1.864,2.542]		[1.481,2.111]		[1.718,2.376]		[1.392,2.011]
	(<0.001)		(<0.001)		(<0.001)		(<0.001)
EQ-5D self-care (1omitted)							
2		1.714	1.498			1.346	1.25
		[1.468,2.003]	[1.273,1.763]			[1.124,1.612]	[1.041,1.501]
		(<0.001)	(<0.001)			(0.001)	(0.017)
3		2.583	1.73			1.712	1.307
		[2.043,3.266]	[1.325,2.257]			[1.274,2.300]	[0.959,1.781]
		(<0.001)	(<0.001)			(<0.001)	(0.090)
EQ-5D mobility (1 omitted)							
2						1.301	1.196
						[1.093,1.548]	[1.002,1.426]
						−0.003	−0.047
3						1.319	1.287
						[0.994,1.749]	[0.972,1.703]
						(0.055)	(0.078)
EQ-5D usual activities (1 omitted)							
2						1.494	1.445
						[1.131,1.972]	[1.096,1.905]
						(0.005)	(0.009)
3						1.592	1.488
						[1.172,2.162]	[1.096,2.019]
						(0.003)	(0.011)
EQ-5D pain and discomfort (1 omitted)							
2						0.812	0.855
						[0.483,1.366]	[0.510,1.434]
						(0.433)	(0.553)
3						0.995	1.052
						[0.580,1.708]	[0.615,1.798]
						(0.987)	(0.854)
EQ-5D anxiety and depression (1 omitted)							
2						1.000	1.042
						[0.857,1.168]	[0.893,1.218]
						(0.996)	(0.599)
3						0.943	0.893
						[0.696,1.275]	[0.659,1.211]
						(0.701)	(0.468)
EQ-VAS (0–10 omitted)				(.)		(.)	(.)
10–19				1.101	1.148	1.19	1.162
				[0.742,1.633]	[0.774,1.703]	[0.799,1.771]	[0.780,1.731]
				(0.632)	(0.493)	(0.392)	(0.460)
20–29				0.893	0.995	1.027	1.035
				[0.614,1.296]	[0.684,1.446]	[0.701,1.503]	[0.706,1.517]
				(0.550)	(0.978)	(0.893)	(0.859)
30–39				0.692	0.815	0.914	0.924
				[0.469,1.021]	[0.551,1.206]	[0.604,1.383]	[0.611,1.398]
				(0.063)	(0.306)	(0.671)	(0.709)
40–49				0.710	0.845	1.040	1.035
				[0.489,1.030]	[0.581,1.229]	[0.697,1.552]	[0.694,1.543]
				(0.071)	(0.378)	(0.847)	(0.865)
50–59				0.732	0.968	1.039	1.129

(*continued on next page*)

Table 3 (*continued*)

	ps_only	sc_only	sc_ps	vas_only	vas_ps	alleq5d_vas	full
				[0.501,1.069]	[0.659,1.421]	[0.694,1.556]	[0.753,1.693]
				(0.106)	(0.867)	(0.851)	(0.556)
60–69				0.585	0.735	0.95	0.965
				[0.394,0.868]	[0.493,1.095]	[0.618,1.458]	[0.629,1.481]
				(0.008)	(0.130)	(0.813)	(0.871)
70–79				0.508	0.659	0.992	0.988
				[0.340,0.758]	[0.439,0.988]	[0.638,1.544]	[0.636,1.534]
				(0.001)	(0.044)	(0.972)	(0.956)
80+				0.428	0.564	0.846	0.861
				[0.265,0.690]	[0.348,0.914]	[0.507,1.413]	[0.515,1.437]
				(<0.001)	(0.020)	(0.523)	(0.566)
C-statistic	0.706	0.718	0.728	0.702	0.711	0.717	0.718

Exponentiated coefficients. 95% confidence intervals in square brackets; p-values in parentheses.

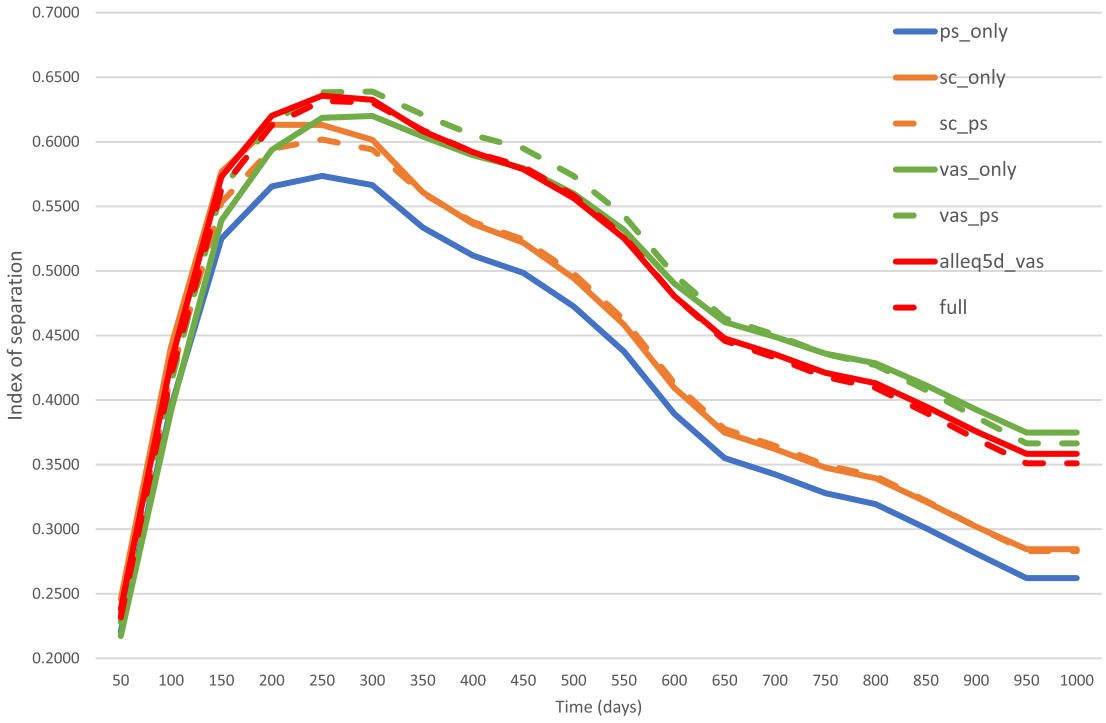

Fig. 3. Mean Altman-Royston separation indices for each model at varying time-points. Based on bootstrapped internal validation.

reliability it should be recognised that patient-reported outcomes in this context are also not without their challenges. The test–retest reliability of the EQ-5D dimensions has shown inconsistent results ranging from moderate to excellent. [25] It is generally accepted that the burden of questionnaires should be minimised for patients who are near the end of life. [26] Indeed, in clinical trials missing data is recognised to increase with proximity to death. [27] A number of studies have reported reduced questionnaire returns and lower completion rates with longer questionnaires, although this is not a consistent finding and has not been assessed specifically in a patient population near the end of life with advanced incurable cancer. [28].

These challenges raise an important potential limitation for the use of patient-reported outcomes in prognostic modelling, and highlight the probable existence of a trade-off between complexity of data collection and the prognostic value of what is obtained. We would argue, however, that the EQ-5D questionnaire used here offers the advantage of including only six questions even in its full form. Furthermore, the burden of questionnaire completion could be reduced, for instance, by the omission of the relatively complex EQ-VAS, leaving only five questions with a three-level response, and potentially making the questionnaire easier to complete. Longer questionnaires, such as the EORTC QLQ-C30 summary score have been shown to be prognostic: however, these may be considered too burdensome near the end of life. [29] The EQ-5D was designed to span the full range of quality of life, but in generic terms rather than, for instance, with particular regard to aspects of quality of life that may be specific to cancer. If it is deemed that greater data collection would be justified by additional prognostic value, additional items, such as fatigue, appetite and shortness of breath, or indeed patient-assessed performance status, may offer further value. [6] The incorporation of the remaining four EQ-5D dimensions (mobility, usual activities, pain/discomfort, and anxiety/depression) and EQ-VAS did not significantly improve model performance, although greater separation was observed at late time-points. Coefficient point-estimates for these dimensions were in expected directions (with worse health being associated with worse survival) and as such, we are unable to draw strong conclusions regarding their inclusion due to limited power.

Future studies, with larger patient cohorts should consider the inclusion of extra items or indeed, the five-level EQ-5D-5L whose additional levels may provide valuable prognostic information with limited impact upon questionnaire burden. The appropriateness of the inclusion of additional (or less) data collection is likely to be highly context-specific. This does not detract from the principal conclusion here,

however, that a simple patient-reported outcome measure (PROM) is able to replace performance status in predicting prognosis for patients with advanced incurable cancer. Indeed, the ability to replace performance status based only upon a short and generic questionnaire is a clear strength in a population near the end of life. There is a need for simplicity in prognostic models such that these are suitable for use in the clinical environment. The inclusion of greater numbers of PROM domains must be clearly justified by the model improvement delivered.

Conversely, the demonstration that the EQ-5D does not offer superior prognostication to that offered by specialist assessment of performance status, supports the continued use of the latter in models for use in specialist environments. Where models are developed for use outside this environment, the role of PROMs in place of clinician-assessed measures should be considered. The acceptability of this approach for patients and clinicians requires assessment: are the questions used acceptable?; are patients and clinicians willing to use the results to guide their decision-making?; is proxy PROM completion acceptable across relevant domains when necessary? [30–31]; when PROMs are used to inform wider routine care (e.g. through symptom control) are they robust to risks of well-meaning manipulation?

Beyond the limitations of power outlined above our study has other limitations worth considering. Our data are relatively old and given the transformational change that has been witnessed in, for example, NSCLC with the arrival of immunotherapies this may significantly impact upon the model, particularly given the extent to which primary diagnosis is a significant predictor of outcome. External validation within a separate, larger, more contemporaneous dataset is now required. Additionally, unlike Chow et al. our dataset is limited to a patient population with known painful bone metastases. Future studies should consider a broader population of patients with advanced incurable cancer where the domains not included here might well offer further valuable information. In addition, the potential for the EQ-5D to perform differently in the routine clinical setting, particularly if it were specifically stated to be used to support clinical decision-making, requires future consideration. It is notable that the results presented here reflect a comparison between patient-reported outcomes and specialist-assessed performance status, rather than non-specialist professional-assessed performance status. This latter comparison could not be made but would reasonably be expected to be detrimental to predictions based on specialist-assessed performance status and would, therefore, not be expected to change the conclusions of this study. Patient-reported performance status has, however, been shown to be a strong predictor of prognosis. How this performs as a prognostic factor in routine care should be considered. Finally, given that our study appears to be underpowered to detect any difference in performance between the different models under consideration, future studies should be designed in order to detect a difference deemed to meaningfully impact upon model use by both clinicians and patients.

Given the challenges of inter-rater reliability in performance status and the nature of the treatment referral pathway, including both specialist and non-specialist clinicians, flexibility is needed. These results support the development of prognostic models which incorporate patient-reported outcomes to circumvent these issues where necessary. In the specialist environment, performance status maybe the simplest predictor to incorporate, however, in non-specialist settings consideration should be given to alternative approaches incorporating PROMs. Larger, more contemporaneous datasets are now required to further optimise this approach and provide external validation.

Funding: This study was funded by the EuroQoL group. The funding organisation had no role in the design, conduct or reporting of this study.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

This work was undertaken on ARC3, part of the High Performance Computing facilities at the University of Leeds, UK.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ctro.2022.09.008.

References

[1] Glare P. A systematic review of physicians' survival predictions in terminally ill cancer patients. BMJ. 2003 Jul 24;327(7408):195–0.
[2] Chow E, Davis L, Panzarella T, Hayter C, Szumacher E, Loblaw A, et al. Accuracy of survival prediction by palliative radiation oncologists. Int J Radiat Oncol Biol Phys 2005;61(3):870–3.
[3] Gripp S, Mjartan S, Boelke E, Willers R. Palliative radiotherapy tailored to life expectancy in end-stage cancer patients: reality or myth? Cancer 2010 Jul 1;116 (13):3251–6.
[4] Tseng YD, Krishnan MS, Sullivan AJ, Jones JA, Chow E, Balboni TA. How radiation oncologists evaluate and incorporate life expectancy estimates into the treatment of palliative cancer patients: a survey-based study. Int J Radiat Oncol Biol Phys 2013 Nov 1;87(3):471–8.
[5] Weeks JC, Cook EF, O'Day SJ, Peterson LM, Wenger N, Reding D, et al. Relationship Between Cancer Patients' Predictions of Prognosis and Their Treatment Preferences. JAMA 1998 Jun 3;279(21):1709–14.
[6] Chow E, Abdolell M, Panzarella T, Harris K, Bezjak A, Warde P, et al. Predictive model for survival in patients with advanced cancer. J Clin Oncol 2008;26(36): 5863–9.
[7] Dosani M, Tyldesley S, Bakos B, Hamm J, Kong T, Lucas S, et al. The TEACHH model to predict life expectancy in patients presenting for palliative spine radiotherapy: external validation and comparison with alternate models. Support Care Cancer 2018;26(7):2217–27.
[8] Westhoff PG, de Graeff A, Monninkhof EM, Bollen L, Dijkstra SP, van der Steen-Banasik EM, et al. An Easy Tool to Predict Survival in Patients Receiving Radiation Therapy for Painful Bone Metastases. Int J Radiat Oncol 2014;90(4):739–47.
[9] Bollen L, van der Linden YM, Pondaag W, Fiocco M, Pattynama BPM, Marijnen CAM, et al. Prognostic factors associated with survival in patients with symptomatic spinal bone metastases: a retrospective cohort study of 1 043 patients. Neuro-Oncol 2014;16(7):991–8.
[10] Chow R, Chiu N, Bruera E, Krishnan M, Chiu L, Lam H, et al. Inter-rater reliability in performance status assessment among health care professionals: a systematic review. Ann. Palliat Med 2016;5(2):83–92.
[11] May CH, Lester JF, Lee S. PS score discordance and why it matters. Lung Cancer 2012 Jan;75:S18.
[12] Pajouheshnia R, van Smeden M, Peelen LM, Groenwold RHH. How variation in predictor measurement affects the discriminative ability and transportability of a prediction model. J Clin Epidemiol 2019 Jan;1(105):136–41.
[13] Park SM, Park MH, Won JH, Lee KO, Choe WS, Heo DS, et al. EuroQol and survival prediction in terminal cancer patients: a multicenter prospective study in hospice-palliative care units. Support Care Cancer 2006;14(4):329–33.
[14] Diouf M, Chibaudel B, Filleron T, Tournigand C, Hug de Larauze M, Garcia-Larnicol M-L, et al. Could baseline health-related quality of life (QoL) predict overall survival in metastatic colorectal cancer? The results of the GERCOR OPTIMOX 1 study. Health Qual Life Outcomes 2014;12(1):69.
[15] Clarke PM, Hayes AJ, Glasziou PG, Scott R, Simes J, Keech AC. Using the EQ-5D Index Score as a Predictor of Outcomes in Patients With Type 2 Diabetes. Med Care 2009 Jan;47(1):61–8.
[16] Mathews WC, May S. EuroQol (EQ-5D) measure of quality of life predicts mortality, emergency department utilization, and hospital discharge rates in HIV-infected adults under care. Health Qual Life Outcomes 2007 Jan 25;5(1):5.
[17] ECOG Performance Status [Internet]. ECOG-ACRIN. [cited 2020 Aug 10]. Available from: https://ecog-acrin.org/resources/ecog-performance-status.
[18] EuroQol Group. EQ-5D-3L [Internet]. [cited 2018 Oct 4]. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/.
[19] Steenland E, Leer J, van Houwelingen H, Post WJ, van den Hout WB, Kievit J, et al. The effect of a single fraction compared to multiple fractions on painful bone metastases: a global analysis of the Dutch Bone Metastasis Study. Radiother Oncol 1999;52(2):101–9.
[20] van den Hout WB, van der Linden YM, Steenland E, Wiggenraad RGJ, Kievit J, de Haes H, et al. Single- versus multiple-fraction radiotherapy in patients with painful bone metastases: cost-utility analysis based on a randomized trial. J Natl Cancer Inst 2003;95(3):222–9.
[21] van der Linden YM, Lok JJ, Steenland E, Martijn H, van Houwelingen H, Marijnen CAM, et al. Single fraction radiotherapy is efficacious: a further analysis of the Dutch Bone Metastasis Study controlling for the influence of retreatment. Int J Radiat Oncol 2004;59(2):528–37.
[22] Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Stat Med 2004 Mar 15;23(5):723–48.
[23] Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000 Feb 29;19(4):453–73.
[24] Riley RD, van der Windt D, Croft P, Prognosis MK. Research in Healthcare: Concepts, Methods, and Impact. Oxford University Press; 2019373 p..

[25] Buchholz I, Janssen MF, Kohlmann T, Feng YS. A Systematic Review of Studies Comparing the Measurement Properties of the Three-Level and Five-Level Versions of the EQ-5D. PharmacoEconomics 2018 Jun 1;36(6):645–61.
[26] Groenvold M, Petersen MA, Aaronson NK, Arraras JI, Blazeby JM, Bottomley A, et al. The development of the EORTC QLQ-C15-PAL: A shortened questionnaire for cancer patients in palliative care. Eur J Cancer 2006;42(1):55–64.
[27] Hussain JA, White IR, Langan D, Johnson MJ, Currow DC, Torgerson DJ, et al. Missing data in randomized controlled trials testing palliative interventions pose a significant risk of bias and loss of power: a systematic review and meta-analyses. J Clin Epidemiol 2016;74:57–65.
[28] Rolstad S, Adler J, Rydén A. Response Burden and Questionnaire Length: Is Shorter Better? A Review and Meta-analysis Value Health 2011 Dec 1;14(8):1101–8.
[29] Husson O, de Rooij BH, Kieffer JM, Oerlemans S, Mols F, Aaronson NK, et al. Independent prognostic value of the EORTC QLQ-C30 summary score on all-cause mortality: Results from the population-based PROFILES registry. J Clin Oncol. 2018 May 20;36(15_suppl):10070–10070.
[30] Hounsome N, Orrell M, Edwards RT. EQ-5D as a quality of life measure in people with dementia and their carers: evidence and key issues. Value Health J Int Soc Pharmacoeconomics Outcomes Res 2011 Apr;14(2):390–9.
[31] Pickard AS, Lin HW, Knight SJ, Sharifi R, Wu Z, Hung SY, et al. Proxy Assessment of Health-Related Quality of Life in African American and White Respondents With Prostate Cancer. Med Care 2009 Feb;47(2):176–83.