

# Development of Machine Learning Algorithms Incorporating Electronic Health Record Data, Patient-Reported Outcomes, or Both to Predict Mortality for Outpatients With Cancer

Ravi B. Parikh, MD, MPP<sup>1,2,3,4,5</sup>; Jill S. Hasler, MS<sup>6</sup>; Yichen Zhang, PhD<sup>4</sup>; Manqing Liu, MHS<sup>7</sup>; Corey Chivers, PhD<sup>8</sup>; William Ferrell, MPH<sup>1</sup>; Peter E. Gabriel, MD<sup>3</sup>; Caryn Lerman, PhD<sup>9</sup>; Justin E. Bekelman, MD<sup>1,2,3,4</sup>; and Jinbo Chen, PhD<sup>6</sup>

**PURPOSE** Machine learning (ML) algorithms that incorporate routinely collected patient-reported outcomes (PROs) alongside electronic health record (EHR) variables may improve prediction of short-term mortality and facilitate earlier supportive and palliative care for patients with cancer.

**METHODS** We trained and validated two-phase ML algorithms that incorporated standard PRO assessments alongside approximately 200 routinely collected EHR variables, among patients with medical oncology encounters at a tertiary academic oncology and a community oncology practice.

**RESULTS** Among 12,350 patients, 5,870 (47.5%) completed PRO assessments. Compared with EHR- and PRO-only algorithms, the EHR + PRO model improved predictive performance in both tertiary oncology (EHR + PRO *v* EHR *v* PRO: area under the curve [AUC] 0.86 [0.85-0.87] *v* 0.82 [0.81-0.83] *v* 0.74 [0.74-0.74]) and community oncology (area under the curve 0.89 [0.88-0.90] *v* 0.86 [0.85-0.88] *v* 0.77 [0.76-0.79]) practices.

**CONCLUSION** Routinely collected PROs contain added prognostic information not captured by an EHR-based ML mortality risk algorithm. Augmenting an EHR-based algorithm with PROs resulted in a more accurate and clinically relevant model, which can facilitate earlier and targeted supportive care for patients with cancer.

JCO Clin Cancer Inform 6:e2200073. © 2022 by American Society of Clinical Oncology

Licensed under the Creative Commons Attribution 4.0 License 

## INTRODUCTION

For patients with cancer, early supportive care interventions, including serious illness conversations and palliative care, are evidence-based practices that improve quality of life and goal-concordant care.<sup>1-4</sup> However, timely identification of patients who may benefit from early supportive care is challenging: Oncology clinicians are often unable to identify patients at risk of 6-month mortality and overestimate life expectancy for up to 70% of their patients.<sup>5-8</sup> Interventions on the basis of electronic health record (EHR)-based machine learning (ML) prognostic algorithms increase serious illness conversations and palliative care referral and could lead to more goal-concordant cancer care for patients with cancer.<sup>9-19</sup> However, such ML algorithms usually rely on structured EHR data, including laboratories, demographics, and diagnosis codes, which provide limited insight into patient symptoms or functional status.<sup>20</sup>

Patient-reported outcomes (PROs), which are independently associated with mortality,<sup>21</sup> may augment such ML algorithms. Routine PRO assessment is now more common and may allow oncology clinicians to better identify patients with high symptom burden or

declining functional status.<sup>22-26</sup> However, the role of PROs in risk stratification remains unexplored. Incorporating PROs may improve traditional risk stratification tools used for supportive and end-of-life care planning.

In this study, we trained and compared algorithms on the basis of EHR data alone, PRO data alone, and EHR plus PRO data, to estimate 6-month risk of mortality among patients seen in either a large tertiary academic practice, or a community-based general oncology clinic. We hypothesized that adverse PROs would be independently associated with 6-month mortality, and that integrating routinely collected PROs into EHR-based ML algorithms would improve predictive performance compared with ML algorithms on the basis of EHR or PRO data alone in both oncology settings.

## METHODS

### Data Source

We derived our cohort from patients receiving care at the University of Pennsylvania Health System (UPHS) who were listed in Clarity, an EPIC reporting database, which contains individual electronic medical records for patients containing demographic, comorbidity, and laboratory data. Our study followed the transparent reporting of a multivariable prediction model for

## ASSOCIATED CONTENT

### Appendix

### Data Supplement

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on October 21, 2022 and published at [ascopubs.org/journal/cci](https://ascopubs.org/journal/cci) on December 8, 2022: DOI <https://doi.org/10.1200/CCI.22.00073>

## CONTEXT

### Key Objectives

To train and compare algorithms on the basis of electronic health record (EHR) data alone, patient-reported outcome (PRO) data alone, and EHR plus PRO data, to estimate 6-month risk of mortality among patients seen in routine oncology practice.

### Knowledge Generated

Compared with EHR- and PRO-only algorithms, the EHR + PRO model improved predictive performance in both tertiary oncology (EHR + PRO  $\nu$  EHR  $\nu$  PRO: area under the curve 0.86 [0.85-0.87]  $\nu$  0.82 [0.81-0.83]  $\nu$  0.74 [0.74-0.74]) and community oncology (area under the curve 0.89 [0.88-0.90]  $\nu$  0.86 [0.85-0.88]  $\nu$  0.77 [0.76-0.79]) settings. Performance was superior across patient- and clinician-specific risk threshold preferences and did not result in increased false-positives.

### Relevance

Incorporating routinely collected PROs into automated EHR-based mortality prediction algorithms significantly improves performance and may aid in targeting supportive care interventions in oncology.

individual prognosis or diagnosis (TRIPOD; Data Supplement) checklist for prediction model development and validation.<sup>27</sup> We obtained approval and waiver of informed consent from the University of Pennsylvania institutional review board before conducting this study.

### Study Population

Our cohort consisted of patients age 18 years or older who had outpatient medical oncology encounters at the Perelman School of Advanced Medicine (PCAM), a large tertiary practice with disease-specific oncology clinics, and Pennsylvania Hospital (PAH), a community oncology practice, between July 1, 2019, and January 1, 2020. We chose patients in these medical oncology clinics because (1) there has been routine collection of PROs in these clinics since mid-2019, (2) an EHR-based ML algorithm has been prospectively validated and implemented in these clinics as part of an initiative to increase serious illness conversations,<sup>28,29</sup> and (3) a tertiary academic practice and a community oncology site are representative of the majority of oncology care settings. Details of PRO collection can be found in the Data Supplement. Patients were not required to have received cancer-directed treatment to be included in this study. We excluded patients who had benign hematology or genetics encounters, < 2 encounters during the study period, or no laboratory or comorbidity data within 6 months of the encounter. Our final cohort consisted of 12,350 patients (8,555 at PCAM and 3,795 at PAH); Appendix Figure A1 represents our cohort selection. In all statistical analysis and modeling, we used the first hematology/oncology encounter in the study period for each patient as the index encounter for statistical modeling. We chose not to incorporate PRO data from subsequent encounters because we found that trends in PROs were not meaningfully associated with mortality (Appendix Fig A2).

### Features

Features included EHR and PRO data. Our EHR data set included three broad classes of features: (1) demographic

variables, (2) comorbidities<sup>30</sup>; and (3) laboratory data. Our final feature set consisted of approximately 200 variables from the EHR (Appendix Table A1). PRO features were derived from the PRO version of The Common Terminology Criteria for Adverse Events (PRO-CTCAE)<sup>31</sup> and the Patient-Reported Outcomes Measurement Information System (PROMIS) Global v.1.2<sup>32</sup> scales (Appendix Table A2). Further details on features are available in the Data Supplement.

### Outcome

The primary outcome was death within 180 days of the index encounter at an oncology practice. We chose 180-day mortality because it is a common indicator of short-term mortality and is often used as a criterion for hospice referral.<sup>16</sup> Date of death was derived from the first date of death recorded in either the EHR (Clarity database) or the Social Security Administration Death Master File, matched to UPHS patients by social security number and date of birth.<sup>33</sup>

### Training and Validation Set Split

In the PCAM cohort, the study population was randomly split into a training cohort (70%), in which the mortality risk algorithms were derived, and a validation cohort (30%), in which the algorithms were applied and tested. Patients could not appear in both the training and validation sets. In the PAH cohort, splitting the data set into a training and testing set was not feasible because of the much lower number of cases.

### Algorithm Development

To develop an algorithm on the basis of EHR variables alone (EHR algorithm), we fitted a logistic regression model with the adaptive LASSO algorithm to ensure consistent variable selection.<sup>34</sup> To develop an algorithm on the basis of PROs alone (PRO algorithm), we fit a logistic regression model where all of the PROs are included as covariates, with observed 180-day mortality as the outcome. To develop an algorithm that includes both EHR and PRO variables (EHR + PRO algorithm), we applied a two-phase method to

**TABLE 1.** Baseline Demographic Characteristics of Patients

Practice Site	Tertiary Oncology Practice					Community Oncology Practice				
	Overall (n = 8,555)	PRO Data Present (n = 4,677)	PRO Data Absent (n = 3,878)	Standardized Mean Difference	P	Overall (n = 3,795)	PRO Data Present (n = 1,193)	PRO Data Absent (n = 2,602)	Standardized Mean Difference	P
No. of encounters, median (IQR)	3.00 (2.00-6.00)	3.00 (2.00-7.00)	3.00 (2.00-6.00)	0.165	< .001	5.00 (3.00-8.00)	6.00 (3.00-10.00)	4.00 (3.00-7.00)	0.391	< .001
Age, years, median (IQR)	64.41 (55.09-72.16)	63.45 (53.93-71.05)	65.66 (56.34-73.38)	0.174	< .001	61.24 (44.71-71.83)	61.34 (47.27-70.95)	61.14 (43.38-72.36)	0.044	.532
Age, years, No. (%)				0.193	< .001				0.124	.005
< 50	1,461 (17.1)	876 (18.7)	585 (15.1)			1,190 (31.4)	343 (28.8)	847 (32.6)		
50-60	1,676 (19.6)	986 (21.1)	690 (17.8)			612 (16.1)	216 (18.1)	396 (15.2)		
60-70	2,682 (31.4)	1,501 (32.1)	1,181 (30.5)			865 (22.8)	297 (24.9)	568 (21.8)		
> 70	2,736 (32.0)	1,314 (28.1)	1,422 (36.7)			1,128 (29.7)	337 (28.2)	791 (30.4)		
Sex, No. (%)				0.013	.57				0.076	.032
Male	4,140 (48.4)	2,277 (48.7)	1863 (48.0)			1,221 (32.2)	413 (34.6)	808 (31.1)		
Female	4,415 (51.6)	2,400 (51.3)	2015 (52.0)			2,574 (67.8)	780 (65.4)	1794 (68.9)		
Comorbidity count, median (IQR)	2.00 (2.00, 4.00)	3.00 (2.00, 4.00)	2.00 (2.00, 4.00)	0.219	< .001	2.00 (0.00, 2.00)	2.00 (1.00, 3.00)	2.00 (0.00, 2.00)	0.457	< .001
ECOG performance status, No. (%)				0.125	< .001				0.282	< .001
0	1,018 (11.9)	610 (13.0)	408 (10.5)			628 (16.5)	241 (20.2)	387 (14.9)		
1	1,021 (11.9)	613 (13.1)	408 (10.5)			529 (13.9)	216 (18.1)	313 (12.0)		
≥ 2	402 (4.7)	198 (4.2)	204 (5.3)			103 (2.7)	47 (3.9)	56 (2.2)		
Missing	6,114 (71.5)	3,256 (69.6)	2,858 (73.7)			2,535 (66.8)	689 (57.8)	1846 (70.9)		
Self-reported race/ethnicity, No. (%) <sup>b</sup>				0.179	< .001				0.228	< .001
Non-Hispanic White	6,304 (73.7)	3,586 (76.7)	2,718 (70.1)			1,946 (51.3)	703 (58.9)	1,243 (47.8)		
Non-Hispanic Black	1,480 (17.3)	669 (14.3)	811 (20.9)			1,465 (38.6)	380 (31.9)	1,085 (41.7)		
Hispanic	143 (1.7)	70 (1.5)	73 (1.9)			153 (4.0)	43 (3.6)	110 (4.2)		
Other	628 (7.3)	352 (7.5)	276 (7.1)			231 (6.1)	67 (5.6)	164 (6.3)		
Insurance, No. (%)				0.153	< .001				0.232	< .001
Medicare	4,399 (51.4)	2,277 (48.7)	2,122 (54.7)			1,808 (47.6)	560 (46.9)	1,248 (48.0)		
Medicaid	486 (5.7)	239 (5.1)	247 (6.4)			553 (14.6)	120 (10.1)	433 (16.6)		
Managed care	2,671 (31.2)	1,582 (33.8)	1,089 (28.1)			1,102 (29.0)	380 (31.9)	722 (27.7)		
Commercial insurance	861 (10.1)	495 (10.6)	366 (9.4)			264 (7.0)	110 (9.2)	154 (5.9)		
Other	21 (0.2)	15 (0.3)	6 (0.2)			0 (0.0)	0 (0.0)	0 (0.0)		
Missing	117 (1.4)	69 (1.5)	48 (1.2)			68 (1.8)	23 (1.9)	45 (1.7)		

(continued on following page)

**TABLE 1.** Baseline Demographic Characteristics of Patients (Continued)

Practice Site	Tertiary Oncology Practice				Community Oncology Practice					
	Overall (n = 8,555)	PRO Data Present (n = 4,677)	PRO Data Absent (n = 3,878)	Standardized Mean Difference	P	Overall (n = 3,795)	PRO Data Present (n = 1,193)	PRO Data Absent (n = 2,602)	Standardized Mean Difference	P
Practice site, No. (%) <sup>a</sup>				0.16	< .001					
Breast	1,035 (12.1)	616 (13.2)	419 (10.8)							
GI	1,123 (13.1)	578 (12.4)	545 (14.1)							
Thoracic	928 (10.8)	474 (10.1)	454 (11.7)							
Genitourinary	801 (9.4)	429 (9.2)	372 (9.6)							
Gynecologic	567 (6.6)	256 (5.5)	311 (8.0)							
Leukemia	935 (10.9)	531 (11.4)	404 (10.4)							
Lymphoma	1,260 (14.7)	707 (15.1)	553 (14.3)							
Melanoma	252 (2.9)	153 (3.3)	99 (2.6)							
Myeloma	1,067 (12.5)	595 (12.7)	472 (12.2)							
Neuro-oncology	349 (4.1)	213 (4.6)	136 (3.5)							
Thyroid	202 (2.4)	108 (2.3)	94 (2.4)							
Missing	36 (0.4)	17 (0.4)	19 (0.5)							

Abbreviations: ECOG, Eastern Cooperative Oncology Group; IQR, interquartile range; PRO, patient-reported outcome.

<sup>a</sup>Data on practice site were only available in the tertiary oncology practice.

<sup>b</sup>Hispanic includes self-reported Hispanic Black and White individuals. Other includes the following race/ethnicity groups: Asian, East Indian, Pacific Islander, American Indian, and unknown.

fit the prediction algorithm and estimate the area under the receiver operating characteristic curve (AUC) and area under the precision-recall curve (AUPRC) that makes full use of all available EHR (N = 8,555 PCAM; N = 3,795 PAH) and PRO (N = 4,677 PCAM; N = 1,193 PAH) data.<sup>35-37</sup> Rationale and approaches for deriving the EHR, PRO, and EHR + PROs model are provided in the Data Supplement.

### Statistical Analysis

We used descriptive statistics to compare the characteristics of the study population, stratified by whether PROs were collected. All algorithm analyses were performed separately for the PCAM and PAH cohorts using Rstudio software. We first explored correlations among individual PRO features using the aggregated PCAM and PAH data. We then fit logistic regression models with 180-day mortality as the outcome and each PRO as the only covariate. We also fit two-variable logistic regression models that assessed the association between each PRO and mortality, adjusted for the continuous 180-day mortality risk from the EHR algorithm. These exploratory analyses informed independent associations between PROs and mortality and the potential of PROs to augment ML performance.

Then the performance of the three different algorithms (EHR, PRO, and EHR + PRO) was assessed by calculating AUC and AUPRC, our primary performance metrics. True-positive rate (TPR) and false-positive rate at a previously specified 10% risk threshold<sup>29</sup> were secondary performance metrics. The 95% CIs for each performance metric were derived using bootstrapping, where each of the two cohorts was repeatedly sampled with replacement to generate 1,000 data sets of the same size. Performance for the EHR model was evaluated for all individuals in the testing set for PCAM (n = 2,566) and in the entire cohort for PAH (n = 3,795). Performance of the PRO and EHR + PRO models were evaluated only for those who had complete PRO data (n = 1,387 for PCAM and n = 1,193 for PAH). Because estimation of the performance metrics for the EHR + PRO algorithm corrected for the potential non-representativeness of the subset of individuals with complete PRO data, the EHR + PRO results are therefore representative of the complete test set. As a sensitivity analysis, we obtained predictive accuracy metrics for all models from the test set for PCAM using only the subset of individuals with PRO data available (n = 1,387). Finally, we conducted a decision curve analysis (see the Data Supplement for methodology) to assess the clinical impact of using each model to identify high-risk patients for the purpose of directing earlier supportive care.<sup>38-40</sup>

## RESULTS

### Cohort Description

The study cohort consisted of 8,555 patients who had 50,590 encounters from the tertiary oncology practice and 3,795 patients who had 32,805 encounters from the community oncology practice (median encounters per patient 4,

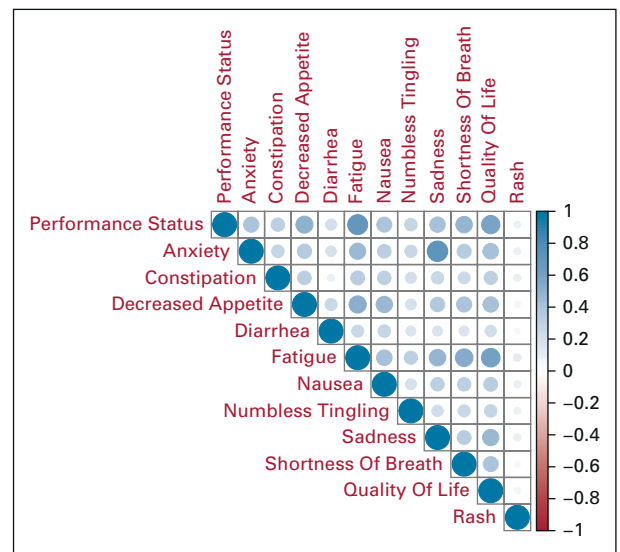
interquartile range 2-7 during study period). The median age was 63.6 years (interquartile range 52.6-72.1 years), 6,989 (56.6%) were female, 8,250 (66.8%) were non-Hispanic White, 2,945 (23.8%) were non-Hispanic Black, 296 (2.4%) were Hispanic, and 6,207 (50.3%) had Medicare insurance. Among patients in the tertiary and community oncology practice cohort, 485 (5.7%) and 122 (3.2%), respectively, died during the 180-day follow-up period. 4,677 (54.7%) patients at the tertiary oncology practice and 1,193 (31.4%) patients at the community oncology practice had completed PRO assessments. Compared with patients who did not complete PRO assessments, patients who completed PRO assessments were more likely to be White (4,289 [73.1%] v 3,961 [61.1%];  $P < .001$ ) and have managed care insurance (1,962 [33.4%] v 1,811 [27.9%]; Table 1).

### Correlation Among PRO Variables

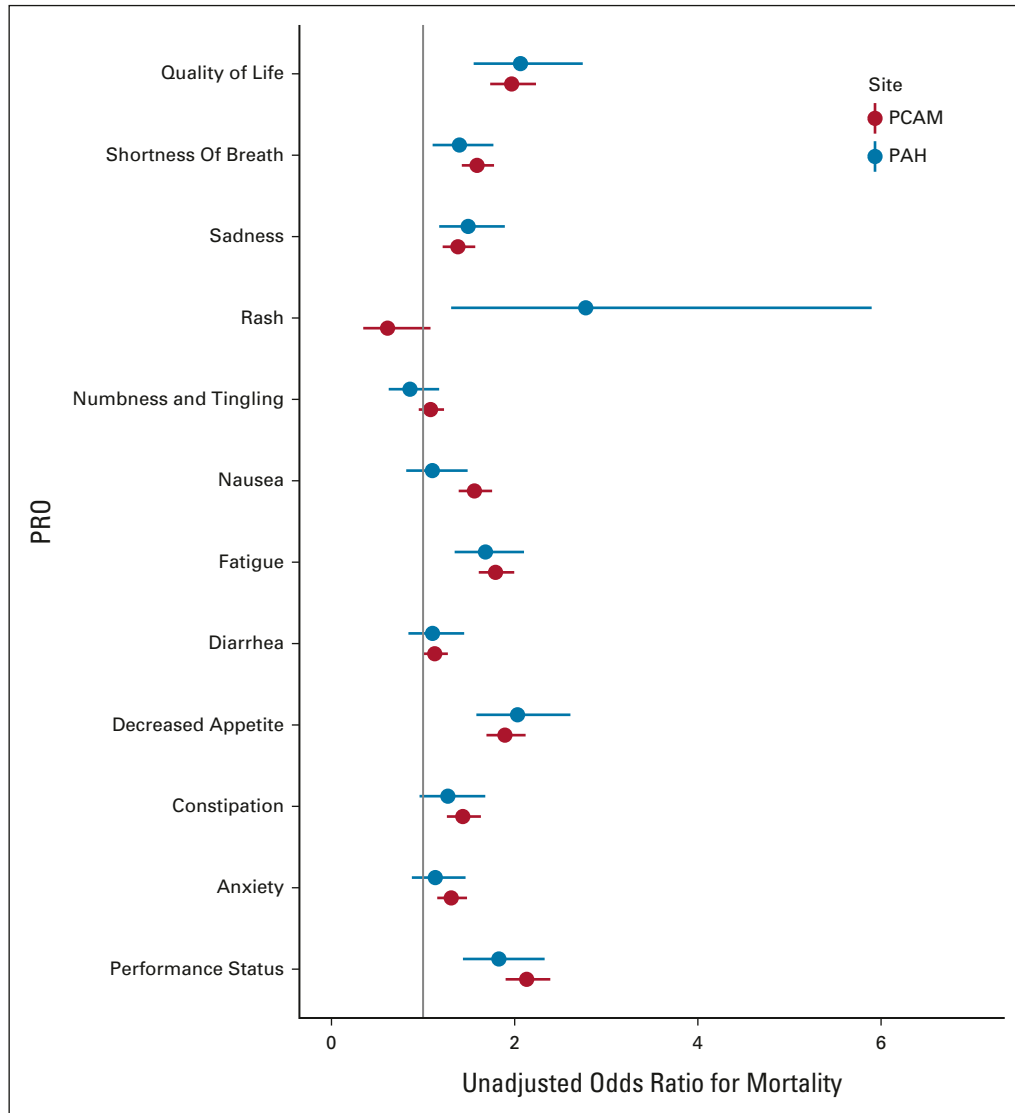
In the combined tertiary and community oncology practice cohorts, decreased performance status was strongly correlated with fatigue ( $r = 0.69$ ), decreased appetite ( $r = 0.5$ ), and poorer quality of life ( $r = 0.58$ ); fatigue was also strongly correlated with poorer quality of life ( $r = 0.6$ ) and decreased appetite ( $r = 0.51$ ; Fig 1). Increased anxiety was strongly correlated with increased sadness ( $r = 0.72$ ). The correlation for all other PRO variable pairs was weak or moderate ( $r < 0.5$ ). These results were consistent in practice-specific subset analyses (Appendix Figs A3 and A4).

### PRO Associations With Observed Mortality

In unadjusted analyses, worse patient-reported performance status (odds ratio [OR], 2.13; 95% CI, 1.90 to 2.39), quality of life (OR, 1.97; 95% CI, 1.72 to 2.23), decreased appetite



**FIG 1.** The correlations between the PROs in the data set. Darker and larger dots indicate stronger correlations. For example, the correlation between decreased performance status and fatigue was 0.69, while the correlation between anxiety and sadness was 0.72. PRO, patient-reported outcome.



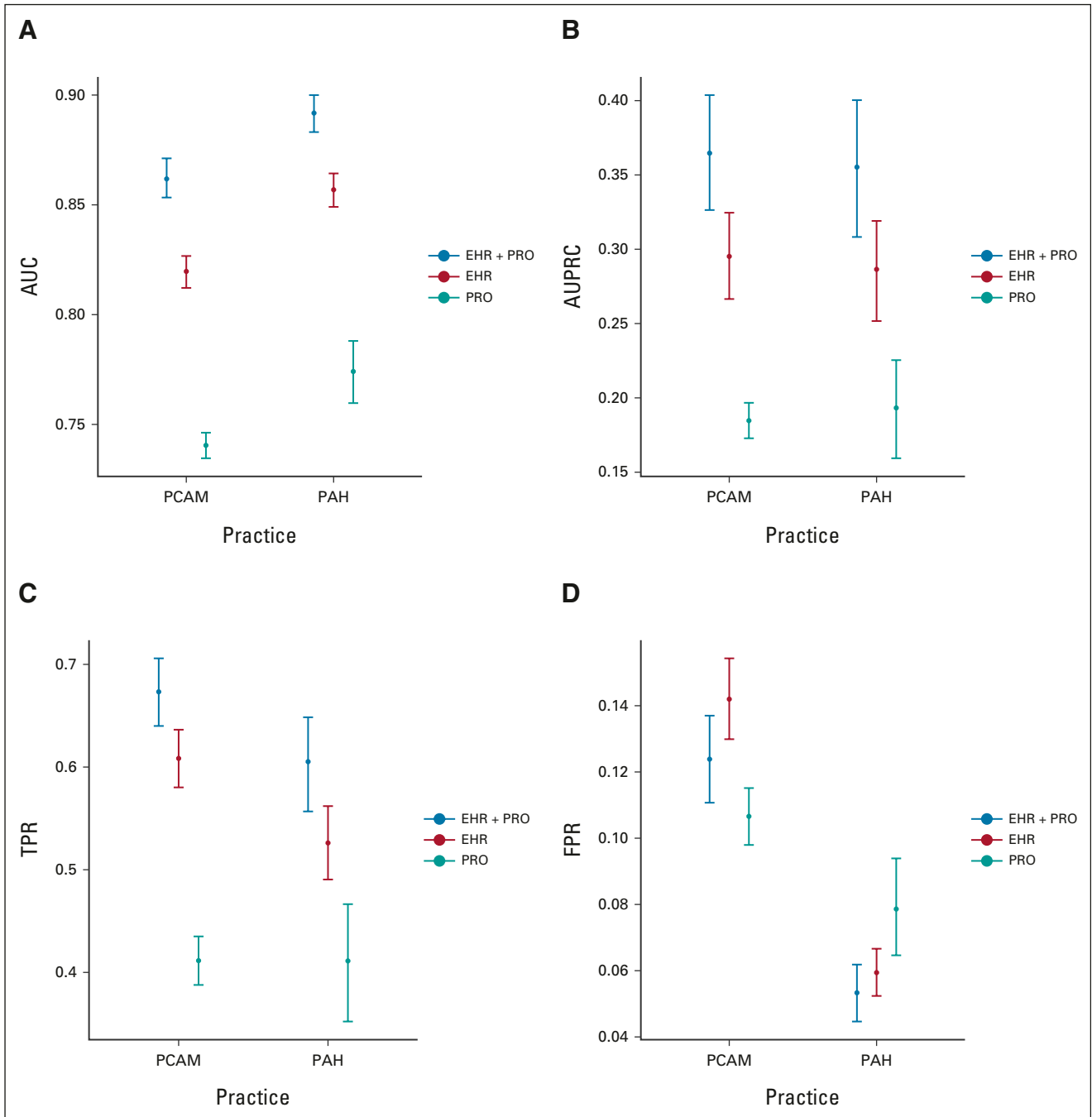
**FIG 2.** Univariable associations between PROs and 180-day mortality. 180-day mortality was defined as a binary indicator variable. PROs were coded on a 1-5 Likert scale, with greater values indicating more severe symptoms, with the exception of rash, which was coded on a 0-1 scale (present/absent). PAH, Pennsylvania Hospital; PCAM, Perelman School of Advanced Medicine; PRO, patient-reported outcome.

(OR, 1.89; 95% CI, 1.69 to 2.12), and fatigue (OR, 1.79; 95% CI, 1.61 to 2.00) had the strongest associations with observed mortality (Fig 2). After adjusting for EHR mortality risk, associations between adverse PROs and observed mortality remained significant for performance status, quality of life, fatigue, shortness of breath, anxiety, sadness, constipation, decreased appetite, and nausea (range of adjusted ORs 1.18-1.53; Appendix Table A3). We observed a similar pattern with community oncology practice data, although fewer associations were statistically significant (Appendix Table A4). Adverse PROs were also associated with higher EHR mortality risk for all PROs except rash.

### Algorithm Performance

The final EHR + PRO model included the logit of the predicted probabilities from the EHR model, performance

status, quality of life, numbness and tingling, and nausea (Appendix Fig A5). For the tertiary oncology practice data, the AUC of the EHR + PRO algorithm (0.86; 95% CI, 0.85 to 0.87) was significantly higher than that of the EHR (0.82; 95% CI, 0.81 to 0.83) and PRO (0.74; 95% CI, 0.73 to 0.75) algorithms (Fig 3A). The AUPRC of the EHR + PRO algorithm (0.36; 95% CI, 0.33 to 0.40) was significantly higher than that of the EHR (0.30; 95% CI, 0.27 to 0.32) and PRO (0.18; 95% CI, 0.17 to 0.20) algorithms (Fig 3B). The TPR of the EHR + PRO algorithm (0.67; 95% CI, 0.64 to 0.71) was significantly higher than that of the EHR (0.61; 95% CI, 0.58 to 0.64) and PRO (0.41; 95% CI, 0.39 to 0.44) algorithms (Fig 3C). There was no difference in false-positive rates among the EHR + PRO (0.12; 95% CI, 0.11 to 0.14), EHR (0.14; 95% CI, 0.13 to 0.15) and PRO (0.11;

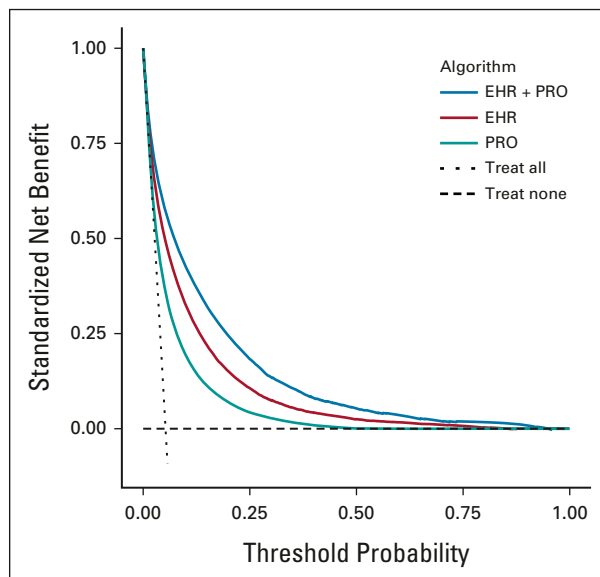


**FIG 3.** Comparison of model performance metrics between the EHR + PRO, EHR, and PRO algorithms at tertiary oncology (PCAM) and community oncology (PAH) practices. Model performance metrics include (A) AUC, (B) AUPRC, (C) TPR, and (D) FPR. TPR and FPR were calculated using a 10% mortality risk threshold, which corresponds to the risk threshold currently used in clinical practice. AUC, area under the curve; AUPRC, area under the precision-recall curve; EHR, electronic health record; FPR, false-positive rate; PAH, Pennsylvania Hospital; PCAM, Perelman School of Advanced Medicine; PRO, patient-reported outcome; TPR, true-positive rate.

95% CI, 0.10 to 0.12) algorithms (Fig 3D). The results were similar in the community oncology practice cohort (Figs 3A-D). In the sensitivity analysis among patients with only PRO data, the EHR + PRO model had consistently higher performance than the PRO model (Appendix Table A5).

### Decision Curve Analysis

In both the tertiary and community oncology practice data sets, the decision curve for the EHR + PRO algorithm dominated the decision curves for the EHR and PRO algorithms, indicating that the EHR + PRO algorithm achieves



**FIG 4.** Decision curve analysis showing standardized net benefit of EHR + PRO, EHR, and PRO algorithms across several model risk thresholds at the tertiary oncology practice (PCAM). A standardized decision curve plots the net benefit to the population of using a risk model against a range of risk threshold values for identifying high-risk patients. A risk threshold chosen to assess the net benefit reflects the user's perspective on the relative cost of a false-positive and false-negative prediction of patients' high-risk status. The standardized net benefit is defined as the difference between the true-positive and the weighted false-positive rates, where the weight is calculated as the odds of the risk threshold multiplied by the inverse odds of the outcome prevalence. It has a maximum of one and can be interpreted as the fraction of maximum utility achieved by the model at the given risk threshold where the maximum utility is achieved when TPR = 1 and FPR = 0. EHR, electronic health record; FPR, false-positive rate; PCAM, Perelman School of Advanced Medicine; PRO, patient-reported outcome; TPR, true-positive rate.

greater clinical utility than the EHR and PRO algorithms regardless of risk preferences (Fig 4). At the clinically relevant mortality risk threshold of 10%, the standardized net benefit was higher for the EHR + PRO algorithm compared with the EHR or PRO algorithms in the tertiary oncology (0.42 [0.37 to 0.47] for EHR + PRO v 0.32 [0.27 to 0.38] for EHR v 0.19 [0.14 to 0.24] for PRO) and community oncology (0.43 [0.37 to 0.48] for EHR + PRO v 0.33 [0.27 to 0.37] for EHR v 0.22 [0.14 to 0.29] for PRO) cohorts.

## DISCUSSION

Among patients with cancer treated at both a tertiary cancer center and community oncology practice, ML algorithms on the basis of combined structured EHR and PRO data outperformed algorithms on the basis of EHR or PRO data alone in predicting short-term mortality. Adverse PROs had strong associations with 180-day mortality. Decision curve analysis revealed that the EHR + PRO algorithm was consistently superior when considering patient-specific

threshold preferences. Collectively, these findings suggest that routinely collected patient-reported symptoms, quality of life, and performance status have considerable independent prognostic value over and above structured EHR data and augment ML models on the basis of EHR data alone.

Accurate identification of patients at risk of short-term mortality is important in oncology, given guidelines around early palliative care, advance care planning, and serious illness communication for high-risk patients with cancer.<sup>41,42</sup> EHR-based ML algorithms linked to automated clinician alerts increase rates of serious illness conversations and palliative care consultation among patients with cancer, with good acceptability among oncologists and no impact on conversation quality.<sup>29,43-45</sup> Underidentification of high-risk patients is a barrier, as TPRs are generally below 50% in EHR algorithms. Incorporating PROs could improve accuracy, aid oncologist clinicians' risk assessments, and prompt clinician discussions about goals and end-of-life preferences. On the basis of our results, in a hypothetical population of 1,000 patients, integrating routinely collected PROs with EHR data would correctly identify an additional 60-80 patients at high risk for short-term mortality, compared with using EHR data alone. Routinely collected PROs add value to existing supportive care triggers in outpatient oncology.

Although routine PRO collection is recommended by consensus guidelines for clinical symptom management and toxicity monitoring during clinical trials,<sup>24,46</sup> use of routinely collected PROs as part of risk stratification, including prognostic risk stratification, is rare in practice. Prior retrospective studies have found that adverse quality of life and symptoms such as depression, fatigue, and pain are independently associated with poorer survival.<sup>21,47,48</sup> However, few studies have demonstrated the independent prognostic value of PROs in contemporary machine learning algorithms. Our study suggests that PROs are only modestly correlated with EHR-predicted mortality risk, and there is likely additional independent prognostic value of PROs that would be of benefit in ML algorithms. Although natural language processing for clinician notes is another potential option to elicit symptoms, there is significant discordance between actual patient-reported symptoms and clinicians' documentation in the EHR.<sup>49,50</sup> Relying on routinely collected PROs is likely a better way to capture symptoms to maximally improve performance of predictive algorithms.

A strength of our two-phase methodology is its flexible approach, using PRO data when available and EHR data for all patients. This differs from traditional complete-case analyses, which may not use representative populations, or imputation-based approaches, which would perform poorly in a setting with a high missingness of PROs. Other advantages of this two-phase methodology are detailed in the Data Supplement.



There are several potential limitations to this study. First, although we trained EHR + PRO algorithms across academic and community oncology cohorts, validation across other institutions, including those with greater Hispanic representation, would be valuable. However, the EHR features used in our models are all commonly available in structured data fields in all health system EHRs, and the PRO features we used were based on standard instruments using Likert scale values. We did not derive any features from unstructured data, and thus, we would not expect semantic differences in coding between different systems. Nevertheless, our approach should be externally validated as other issues of data quality, including completeness of features and heterogeneity in coding practices between institutions, are well known and should be accounted for.<sup>51,52</sup> Second, we did not validate a specific model in an external institutional cohort, but rather used a two-phase approach to test similar algorithms in two unique practices. This approach is justified because the purpose of our study was not to validate a specific algorithm, but rather to validate the conclusion that integrating PROs into routine prognostic algorithms improves risk prediction. Third, there is no gold

standard reference for mortality prediction, and it is unclear how our EHR + PRO model compares with other published mortality prediction tools. However, we used the same features used in a validated EHR algorithm that is in routine use in medical oncology practices within our cancer center to prompt serious illness conversations.<sup>28</sup> Fourth, although we expect that our institutional registry and Social Security death data captured most deaths, we were unable to use more robust death data including National Death Index and obituary data. Fifth, although algorithm performance in our PCAM sample is reported on a typical holdout test set, we were unable to use a train/test split using the PAH data because of the much smaller number of cases in that data set.

In conclusion, among 12,350 patients with cancer seen in tertiary and community oncology practices, ML algorithms to predict short-term mortality that integrated routinely collected patient-reported outcomes with electronic health record features significantly improved predictive performance, compared with algorithms on the basis of EHR or PRO data alone. Our findings suggest that PROs can significantly improve performance of predictive algorithms in oncology.

## AFFILIATIONS

<sup>1</sup>Department of Medical Ethics and Health Policy, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

<sup>2</sup>Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

<sup>3</sup>Penn Center for Cancer Care Innovation, Abramson Cancer Center, University of Pennsylvania, Philadelphia, PA

<sup>4</sup>Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA

<sup>5</sup>Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA

<sup>6</sup>Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, Philadelphia, PA

<sup>7</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

<sup>8</sup>Penn Medicine, University of Pennsylvania, Philadelphia, PA

<sup>9</sup>USC Norris Comprehensive Cancer Center, Los Angeles, CA

## CORRESPONDING AUTHOR

Ravi B. Parikh, MD, MPP, Department of Medical Ethics and Health Policy, Perelman School of Medicine at the University of Pennsylvania, 423 Guardian Dr, Philadelphia, PA 19104; e-mail: Ravi.Parikh@pennmedicine.upenn.edu.

## DISCLAIMER

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## EQUAL CONTRIBUTION

R.B.P. and J.S.H. contributed equally to this work as co-first authors. C.L., J.E.B., and J.C. contributed equally to this work as co-senior authors.

## SUPPORT

Supported by the National Cancer Institute (R35197461) C.L., National Cancer Institute (K08CA263541) R.B.P., National Cancer Institute (P30-014089) C.L., NHLBI (R01-HL138306) J.C., National Cancer Institute (R01-CA236468) Jinbo Chen, NIGMS (R01-GM140463) J.C., and National Center for Advancing Translational Sciences (UL1-TR001878) J.C.

## DATA SHARING STATEMENT

The data that support the findings of this study are available on request. All statistical analysis was performed in R version 3.6.0. The validation for the EHR algorithm has been previously published (<https://jamanetwork.com/journals/jamaoncology/article-abstract/2770698>), and source code is available at <https://github.com/pennsignals/eol-nc>.

## AUTHOR CONTRIBUTIONS

**Conception and design:** Ravi B. Parikh, Peter E. Gabriel, Caryn Lerman, Justin E. Bekelman, Jinbo Chen

**Financial support:** Ravi B. Parikh, Caryn Lerman, Justin E. Bekelman, Jinbo Chen

**Administrative support:** Ravi B. Parikh, William Ferrell

**Provision of study materials or patients:** Ravi B. Parikh, Jinbo Chen

**Collection and assembly of data:** Ravi B. Parikh, Yichen Zhang, Corey Chivers, William Ferrell, Peter E. Gabriel, Jinbo Chen

**Data analysis and interpretation:** Ravi B. Parikh, Jill S. Hasler, Yichen Zhang, Manqing Liu, Corey Chivers, Caryn Lerman, Justin E. Bekelman, Jinbo Chen

**Manuscript writing:** All authors

**Final approval of manuscript:** All authors

**Accountable for all aspects of the work:** All authors

## AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the

subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to [www.asco.org/rwc](http://www.asco.org/rwc) or [ascopubs.org/cci/author-center](http://ascopubs.org/cci/author-center).

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](http://OpenPayments.gov)).

#### Ravi B. Parikh

**Stock and Other Ownership Interests:** Merck, Google, GNS Healthcare, Onc.AI  
**Consulting or Advisory Role:** GNS Healthcare, Cancer Study Group, Onc.AI, Thyme Care, Humana, NanOlogy, Merck

**Research Funding:** Humana

**Patents, Royalties, Other Intellectual Property:** Technology to integrate patient-reported outcomes into electronic health record algorithms

**Travel, Accommodations, Expenses:** The Oncology Institute of Hope and Innovation

#### Jill S. Hasler

**Patents, Royalties, Other Intellectual Property:** Patent currently pending for machine learning systems using electronic health record data and patient-reported outcomes

#### William Ferrell

**Research Funding:** Humana

#### Peter E. Gabriel

**Travel, Accommodations, Expenses:** Varian Medical Systems

#### Justin E. Bekelman

**Stock and Other Ownership Interests:** Reimagine Care

**Honoraria:** National Comprehensive Cancer Network

**Consulting or Advisory Role:** UnitedHealthcare, Reimagine Care

No other potential conflicts of interest were reported.

## ACKNOWLEDGMENT

The authors thank Jay Fein for assistance with manuscript preparation.

## REFERENCES

- Bernacki R, Paladino J, Neville BA, et al: Effect of the serious illness care program in outpatient oncology: A cluster randomized clinical trial. *JAMA Intern Med* 179:751-759, 2019
- Paladino J, Bernacki R, Neville BA, et al: Evaluating an intervention to improve communication between oncology clinicians and patients with life-limiting cancer: A cluster randomized clinical trial of the serious illness care program. *JAMA Oncol* 5:801-809, 2019
- Temel JS, Greer JA, Muzikansky A, et al: Early palliative care for patients with metastatic non-small-cell lung cancer. *N Engl J Med* 363:733-742, 2010
- Wright AA: Associations between end-of-life discussions, patient mental health, medical care near death, and caregiver bereavement adjustment. *JAMA* 300:1665-1673, 2008
- Krishnan M, Temel J, Wright A, et al: Predicting life expectancy in patients with advanced incurable cancer: A review. *J Support Oncol* 11:68-74, 2013
- Chow E, Harth T, Hruby G, et al: How accurate are physicians' clinical predictions of survival and the available prognostic tools in estimating survival times in terminally ill cancer patients? A systematic review. *Clin Oncol* 13:209-218, 2001
- 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* 11:e0161407, 2016
- Christakis NA: Extent and determinants of error in doctors' prognoses in terminally ill patients: Prospective cohort study commentary: Why do doctors overestimate? Commentary: Prognoses should be based on proved indices not intuition. *BMJ* 320:469-473, 2000
- Tang ST, Chen CH, Wen FH, et al: Accurate prognostic awareness facilitates, whereas better quality of life and more anxiety symptoms hinder end-of-life care discussions: A longitudinal survey study in terminally ill cancer patients' last six months of life. *J Pain Symptom Manage* 55:1068-1076, 2018
- Nipp RD, Greer JA, El-Jawahri A, et al: Coping and prognostic awareness in patients with advanced cancer. *J Clin Oncol* 35:2551-2557, 2017
- Lundquist G, Rasmussen BH, Axelsson B: Information of imminent death or not: Does it make a difference? *J Clin Oncol* 29:3927-3931, 2011
- Zhang B, Wright AA, Huskamp HA, et al: Health care costs in the last week of life: Associations with end-of-life conversations. *Arch Intern Med* 169:480-488, 2009
- El-Jawahri A, Traeger L, Park ER, et al: Associations among prognostic understanding, quality of life, and mood in patients with advanced cancer: Prognosis, QoL, and Mood in Advanced CA. *Cancer* 120:278-285, 2014
- Finlay E, Casarett D: Making difficult discussions easier: Using prognosis to facilitate transitions to hospice. *CA Cancer J Clin* 59:250-263, 2009
- Parikh RB, Manz C, Chivers C, et al: Machine learning approaches to predict 6-month mortality among patients with cancer. *JAMA Netw Open* 2:e1915997, 2019
- Elfiky AA, Pany MJ, Parikh RB, Obermeyer Z: Development and application of a machine learning approach to assess short-term mortality risk among patients with cancer starting chemotherapy. *JAMA Netw Open* 1:e180926, 2018
- Bertsimas D, Dunn J, Pawlowski C, et al: Applied informatics decision support tool for mortality predictions in patients with cancer. *JCO Clin Cancer Inform* 2:1-11, 2018
- Gensheimer MF, Henry AS, Wood DJ, et al: Automated survival prediction in metastatic cancer patients using high-dimensional electronic medical record data. *J Natl Cancer Inst* 111:568-574, 2019
- Gensheimer MF, Aggarwal S, Benson KR, et al: Automated model versus treating physician for predicting survival time of patients with metastatic cancer. *J Am Med Inform Assoc* 28:1108-1116, 2021
- Baum LVM, Friedman D: The uncertain science of predicting death. *JAMA Netw Open* 3:e201736, 2020
- Stukenborg GJ, Blackhall LJ, Harrison JH, et al: Longitudinal patterns of cancer patient reported outcomes in end of life care predict survival. *Support Care Cancer* 24:2217-2224, 2016
- Bubis LD, Davis L, Mahar A, et al: Symptom burden in the first year after cancer diagnosis: An analysis of patient-reported outcomes. *J Clin Oncol* 36:1103-1111, 2018
- Basch E, Deal AM, Kris MG, et al: Symptom monitoring with patient-reported outcomes during routine cancer treatment: A randomized controlled trial. *J Clin Oncol* 34:557-565, 2016
- Basch E, Barbera L, Kerrigan CL, Velikova G: Implementation of patient-reported outcomes in routine medical care. *Am Soc Clin Oncol Ed Book* 38:122-134, 2018

25. Yang LY, Manhas DS, Howard AF, Olson RA: Patient-reported outcome use in oncology: A systematic review of the impact on patient-clinician communication. *Support Care Cancer* 26:41-60, 2018
26. Basch E, Deal AM, Dueck AC, et al: Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. *JAMA* 318:197-198, 2017
27. Collins GS, Reitsma JB, Altman DG, et al: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *BMJ* 350:g7594, 2015
28. Manz CR, Chen J, Liu M, et al: Validation of a machine learning algorithm to predict 180-day mortality for outpatients with cancer. *JAMA Oncol* 6:1723-1730, 2020
29. Manz CR, Parikh RB, Small DS, et al: Effect of integrating machine learning mortality estimates with behavioral nudges to clinicians on serious illness conversations among patients with cancer: A stepped-wedge cluster randomized clinical trial. *JAMA Oncol* 6:e204759, 2020
30. Elixhauser A, Steiner C, Harris DR, Coffey RM: Comorbidity measures for use with administrative data. *Med Care* 36:8-27, 1998
31. Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). <https://healthcaresdelivery.cancer.gov/pro-ctcae/>
32. Hays RD, Bjorner JB, Revicki DA, et al: Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Qual Life Res* 18:873-880, 2009
33. NTIS: Limited Access Death Master File Download. <https://dmf.ntis.gov/>
34. Friedman J, Hastie T, Tibshirani R, et al: glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models, 2021. <https://dx.doi.org/10.18637/jss.v033.i01>
35. Breslow NE, Cain KC: Logistic regression for two-stage case-control data. *Biometrika* 75:11-20, 1988
36. Cain KC, Breslow NE: Logistic regression analysis and efficient design for two-stage studies. *Am J Epidemiol* 128:1198-1206, 1988
37. Scott AJ, Wild CJ: Fitting logistic models under case-control or choice based sampling. *J R Stat Soc Series B Stat Methodol* 48:170-182, 1986
38. Vickers AJ, Elkin EB: Decision curve analysis: A novel method for evaluating prediction models. *Med Decis Making* 26:565-574, 2006
39. Fitzgerald M, Saville BR, Lewis RJ: Decision curve analysis. *JAMA* 313:409-410, 2015
40. Kerr KF, Brown MD, Zhu K, Janes H: Assessing the clinical impact of risk prediction models with decision curves: Guidance for correct interpretation and appropriate use. *J Clin Oncol* 34:2534-2540, 2016
41. Ferrell BR, Temel JS, Temin S, et al: Integration of palliative care into standard oncology care: American Society of Clinical Oncology clinical practice guideline update. *J Clin Oncol* 35:96-112, 2017
42. Temel JS, Greer JA, El-Jawahri A, et al: Effects of early integrated palliative care in patients with lung and GI cancer: A randomized clinical trial. *J Clin Oncol* 35:834-841, 2017
43. Courtright KR, Chivers C, Becker M, et al: Electronic health record mortality prediction model for targeted palliative care among hospitalized medical patients: A pilot quasi-experimental study. *J Gen Intern Med* 34:1841-1847, 2019
44. Parikh RB, Manz CR, Nelson MN, et al: Clinician perspectives on machine learning prognostic algorithms in the routine care of patients with cancer: A qualitative study. *Support Care Cancer* 30:4363-4372, 2022
45. Li E, Manz C, Liu M, et al: Oncologist phenotypes and associations with response to a machine learning-based intervention to increase advance care planning: Secondary analysis of a randomized clinical trial. *PLoS One* 17:e0267012, 2022
46. Basch E, Snyder C, McNiff K, et al: Patient-reported outcome performance measures in oncology. *JCO Oncol Pract* 10:209-211, 2014
47. Saint-Maurice PF, Troiano RP, Bassett DR, et al: Association of daily step count and step intensity with mortality among US adults. *JAMA* 323:1151-1160, 2020
48. Seventer EE, Fish MG, Fosbenner K, et al: Associations of baseline patient-reported outcomes with treatment outcomes in advanced gastrointestinal cancer. *Cancer* 127:619-627, 2021
49. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL: Agreement between patient-reported symptoms and their documentation in the medical record. *Am J Manag Care* 14:530-539, 2008
50. Yeung AR, Pugh SL, Klopp AH, et al: Improvement in patient-reported outcomes with intensity-modulated radiotherapy (RT) compared with standard RT: A report from the NRG Oncology RTOG 1203 study. *J Clin Oncol* 38:1685-1692, 2020
51. Fu S, Wen A, Schaeferle GM, et al: Assessment of data quality variability across two EHR systems through a case study of post-surgical complications. *AMIA Annu Symp Proc AMIA Symp* 2022:196-205, 2022
52. Horsky J, Drucker EA, Ramelson HZ: Accuracy and completeness of clinical coding using ICD-10 for ambulatory visits. *AMIA Annu Symp Proc AMIA Symp* 2017:912-920, 2017
53. Chivers, C: pennsignals/eol-onc. <https://github.com/pennsignals/eol-onc>



**APPENDIX**

**TABLE A1.** Electronic Health Record Feature Set

Laboratories	Units	Features Included
Age	Years	
Sex	Male/female	
Elixhauser comorbidities	Count	Count of No. of times comorbidity was ever coded (total) Count of No. of times comorbidity was coded in the past 180 days (recent)
Metastatic cancer		
Diabetes, uncomplicated		
Alcohol abuse		
Rheumatoid arthritis/collagen vascular diseases		
Weight loss		
Iron deficiency anemia		
Lymphoma		
Cardiac arrhythmias		
Pulmonary circulation disorder		
Liver disease		
Obesity		
Chronic pulmonary disease		
Hypertension, uncomplicated		
Fluid and electrolyte disorders		
Renal failure		
Other neurologic disorders		
Pulmonary valvular disorder		
Peptic ulcer disease excluding bleeding		
Hypothyroidism		
Valvular disease		
Congestive heart failure		
Depression		
Paralysis		
Drug abuse		
Diabetes, complicated		
Coagulopathy		
Blood loss anemia		
Solid tumor without metastasis		
AIDS/HIV		
Hypertension, complicated		
Psychoses		
EKG values		Count
T axis	Degrees	First value checked Last value checked
R axis	Degrees	Maximum value Minimum value
P axis	Degrees	Proportion of values ordered as STAT
Q-T interval	ms	
QRS duration	ms	
P-R interval	ms	
QTC calculation (Bazett)	ms	
Atrial rate	BPM	
Ventricular rate	BPM	
Laboratories		Count
Basophils manual	%	First value checked Last value checked
Eosinophils manual	%	Maximum value Minimum value
Reticulocyte count	%	

(continued in next column)

**TABLE A1.** Electronic Health Record Feature Set (Continued)

Laboratories	Units	Features Included
Band neutrophils	%	Proportion of values ordered as STAT
Lymphocytes manual	%	
Eosinophils	%	
Basophils	%	
Lymphocyte variants	%	
Segmented neutrophils manual	%	
Monocytes	%	
Lymphocytes	%	
Myelocytes	%	
Monocytes manual	%	
Alkaline phosphatase	IU/L	
WBC-corrected	THO/ $\mu$ L	
Abs retic	THO/ $\mu$ L	
absolute eosinophils manual	THO/ $\mu$ L	
absolute segmented neutrophils manual	THO/ $\mu$ L	
absolute basophils manual	THO/ $\mu$ L	
absolute monocytes manual	THO/ $\mu$ L	
absolute atypical lymphocytes	THO/ $\mu$ L	
absolute lymphocytes manual	THO/ $\mu$ L	
AST	U/L	
Lactate dehydrogenase	U/L	
ALT	U/L	
UA pH	pH	
absolute band neutrophils	cells/ $\mu$ L	
MCV	fL	
RDW	fL	
Gamma-globulin fraction	g/dL	
MCHC	g/dL	
Beta-globulin fraction	g/dL	
Albumin fraction	g/dL	
Alpha2-globulin fraction	g/dL	
Alpha1-globulin fraction	g/dL	
% nucleated RBCs	k/cumm	
% neutrophils	k/ $\mu$ L	
eGFR, AA	mL/min/ 1.73 m <sup>2</sup>	
eGFR, non-AA	mL/min/ 1.73 m <sup>2</sup>	
Magnesium	mg/dL	
Immunoglobulin G	mg/dL	
Immunoglobulin M	mg/dL	
Fibrinogen	mg/dL	
BUN	mg/dL	
Urea nitrogen	mg/dL	
POC glucose (fingerstick)	mg/dL	
Glucose	mg/dL	
Calcium	mg/dL	
Bilirubin, indirect	mg/dL	
Uric acid	mg/dL	
Bilirubin, direct	mg/dL	
Phosphate	mg/dL	

(continued on following page)

**TABLE A1.** Electronic Health Record Feature Set (Continued)

Laboratories	Units	Features Included
Creatinine	mg/dL	
Hemoglobin	mg/dL	
Triglycerides	mm/L	
Total cholesterol	mm/L	
pCO <sub>2</sub> art	mmHg	
Carbon dioxide	mmol/L	
Anion gap	mmol/L	
Lactic acid	mmol/L	
Chloride	mmol/L	
Sodium	mmol/L	
Potassium	mmol/L	
Carcinoembryonic antigen	ng/mL	
Ferritin	ng/mL	
MCH	pg	
INR	ratio	
PTT	second(s)	
PT	seconds	
TSH	$\mu$ U/mL	
Tacrolimus level	$\mu$ g/L	
absolute basophils	$\times 10^3/\mu$ L	
absolute neutrophils	$\times 10^3/\mu$ L	
absolute monocytes	$\times 10^3/\mu$ L	
absolute eosinophils	$\times/\mu$ L	
WBCs	$\times 10^3/\mu$ L	
Platelets	$\times 10^3/\mu$ L	
absolute lymphocytes	$\times 10^3/\mu$ L	
RBCs	$\times 10^6/\mu$ L	
Albumin	g/dL	
Bilirubin, total	mg/dL	
Immunoglobulins A	mg/dL	
Protein, total	g/dL	

NOTE. Although all variables in the comorbidity and laboratory sections were included in some form, not all features were included after feature selection. For full list of variables included after feature selection, see GitHub, Inc.<sup>53</sup>

Abbreviations: eGFR, estimated glomerular filtration rate; EKG, electrocardiogram; INR, international normalized ratio; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; POC, point-of-care; PT, prothrombin time; PTT, partial thromboplastin time; RDW, red cell distribution width; TSH, thyroid-stimulated hormone; UA, urinalysis.

**TABLE A2.** PRO Features

PRO	Scale	Label
Performance status <sup>a</sup>	1-5	Higher scores indicate worse symptoms
Quality of life <sup>b</sup>		
Fatigue <sup>a</sup>		
Shortness of breath <sup>a</sup>		
Anxiety <sup>a</sup>		
Sadness <sup>a</sup>		
Constipation <sup>a</sup>		
Decreased appetite <sup>a</sup>		
Diarrhea <sup>a</sup>		
Nausea <sup>a</sup>		
Numbness and tingling <sup>a</sup>		
Rash <sup>a</sup>	0-1	1: yes, 0: no

Abbreviation: PRO, patient-reported outcome.

<sup>a</sup>PRO version of The Common Terminology Criteria for Adverse Events (PRO-CTCAE).<sup>31</sup>

<sup>b</sup>Patient-Reported Outcomes Measurement Information System (PROMIS) Global v.1.2.<sup>32</sup>

**TABLE A4.** Univariable and Adjusted Associations Between PROs and Mortality (Pennsylvania Hospital)

PRO	Unadjusted		Adjusted for Baseline Mortality Risk <sup>a</sup>	
	OR	95% CI	OR	95% CI
Performance status	1.83	1.44 to 2.33	1.53	1.16 to 2.02
Quality of life	2.06	1.55 to 2.74	1.47	1.07 to 2.02
Fatigue	1.68	1.34 to 2.10	1.62	1.24 to 2.11
Shortness of breath	1.40	1.10 to 1.77	1.32	1.00 to 1.74
Anxiety	1.13	0.88 to 1.46	1.26	0.95 to 1.66
Sadness	1.49	1.18 to 1.89	1.47	1.12 to 1.93
Constipation	1.27	0.96 to 1.68	1.09	0.79 to 1.51
Decreased appetite	2.03	1.58 to 2.61	1.66	1.26 to 2.19
Diarrhea	1.10	0.84 to 1.45	0.91	0.66 to 1.25
Nausea	1.10	0.82 to 1.49	1.01	0.72 to 1.42
Numbness and tingling	0.86	0.62 to 1.18	0.74	0.52 to 1.04
Rash	2.78	1.31 to 5.90	2.85	1.19 to 6.85

NOTE. Regressions estimate the association between a one-point increase in each PRO scale and 180-day mortality.

Abbreviations: EHR, electronic health record; OR, odds ratio; PRO, patient-reported outcome.

<sup>a</sup>Adjusted for continuous mortality risk as predicted from the EHR-only model.

**TABLE A3.** Univariable and Adjusted Associations Between PROs and Mortality (Perelman School of Advanced Medicine)

PRO	Unadjusted		Adjusted for Baseline Mortality Risk <sup>a</sup>	
	OR	95% CI	OR	95% CI
Performance status	2.13	1.90 to 2.39	1.53	1.34 to 1.75
Quality of life	1.97	1.72 to 2.23	1.43	1.25 to 1.65
Fatigue	1.79	1.61 to 2.00	1.37	1.21 to 1.55
Shortness of breath	1.59	1.42 to 1.78	1.22	1.07 to 1.38
Anxiety	1.31	1.15 to 1.48	1.25	1.09 to 1.44
Sadness	1.38	1.21 to 1.57	1.25	1.08 to 1.45
Constipation	1.43	1.26 to 1.63	1.18	1.02 to 1.37
Decreased appetite	1.89	1.69 to 2.12	1.36	1.19 to 1.55
Diarrhea	1.13	1.00 to 1.27	1.00	0.88 to 1.15
Nausea	1.56	1.39 to 1.75	1.30	1.14 to 1.48
Numbness and tingling	1.08	0.95 to 1.23	0.90	0.78 to 1.04
Rash	0.61	0.35 to 1.08	0.74	0.41 to 1.33

NOTE. Regressions estimate the association between a one-point increase in each PRO scale and 180-day mortality.

Abbreviations: EHR, electronic health record; OR, odds ratio; PRO, patient-reported outcome.

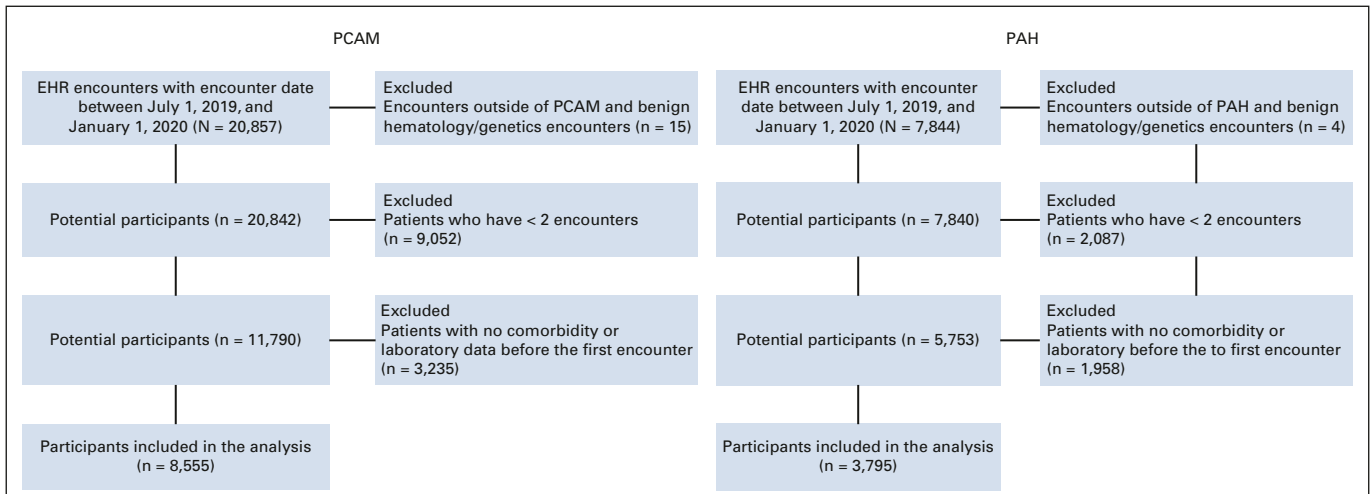
<sup>a</sup>Adjusted for continuous mortality risk as predicted from the EHR-only model.

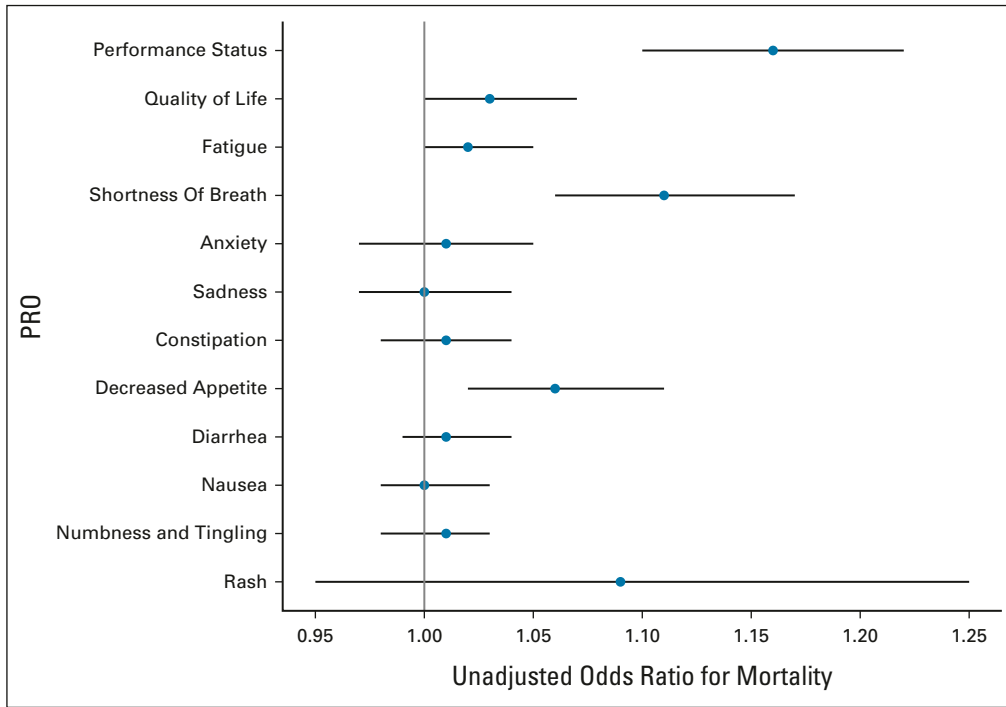
**TABLE A5.** Distribution of PROs Within Electronic Health Record Mortality Risk Qualities

PRO	Tertiary Oncology Practice					Community Oncology Practice				
	Q1 (N = 2,108)	Q2 (N = 2,117)	Q3 (N = 2,209)	Q4 (N = 2,121)	P <sup>a</sup>	Q1 (N = 964)	Q2 (N = 994)	Q3 (N = 898)	Q4 (N = 939)	P <sup>a</sup>
Mortality risk	< 0.011	0.011-0.023	0.023-0.058	> 0.058		< 0.005	0.005-0.011	0.011-0.027	> 0.027	
Performance status, mean (SD)	2.49 (0.75)	2.59 (0.80)	2.74 (0.86)	3.16 (1.01)	.056	1.65 (0.86)	1.77 (0.95)	1.71 (0.85)	2.06 (1.05)	.166
Quality of life, mean (SD)	2.24 (0.97)	2.41 (0.99)	2.51 (0.95)	2.88 (1.01)	.038	3.61 (0.99)	3.56 (0.96)	3.53 (0.98)	3.14 (1.05)	.138
Fatigue, mean (SD)	3.00 (1.02)	3.10 (1.06)	3.27 (1.08)	3.66 (1.14)	.044	2.29 (1.18)	2.29 (1.16)	2.26 (1.17)	2.58 (1.18)	.283
Shortness of breath, mean (SD)	1.32 (0.70)	1.42 (0.81)	1.52 (0.85)	1.81 (1.06)	.037	5.48 (0.86)	5.48 (0.93)	5.44 (0.94)	5.32 (1.01)	.110
Anxiety, mean (SD)	2.53 (0.85)	2.54 (0.86)	2.57 (0.88)	2.66 (0.92)	.083	1.66 (0.97)	1.69 (0.99)	1.63 (0.98)	1.67 (1.01)	.973
Sadness, mean (SD)	2.44 (0.78)	2.45 (0.81)	2.48 (0.77)	2.60 (0.89)	.095	1.51 (0.86)	1.55 (0.89)	1.53 (0.91)	1.69 (0.97)	.160
Constipation, mean (SD)	2.35 (0.70)	2.39 (0.76)	2.44 (0.78)	2.61 (0.89)	.065	1.53 (0.89)	1.38 (0.75)	1.47 (0.82)	1.63 (0.92)	.503
Decreased appetite, mean (SD)	2.26 (0.62)	2.35 (0.69)	2.52 (0.82)	2.84 (1.04)	.035	1.40 (0.75)	1.46 (0.79)	1.52 (0.82)	1.71 (0.99)	.058
Diarrhea, mean (SD)	2.60 (0.93)	2.64 (0.97)	2.72 (1.02)	2.77 (1.05)	.006	1.50 (0.87)	1.63 (0.95)	1.60 (0.93)	1.73 (1.04)	.095
Nausea, mean (SD)	2.37 (0.74)	2.36 (0.73)	2.51 (0.87)	2.70 (1.00)	.075	1.52 (0.87)	1.45 (0.80)	1.45 (0.86)	1.57 (0.90)	.663
Numbness and tingling, mean (SD)	2.59 (0.88)	2.69 (0.94)	2.72 (0.94)	2.86 (1.01)	.022	1.61 (0.92)	1.65 (0.92)	1.71 (0.98)	1.80 (1.04)	.015
Rash, mean (SD)	2.07 (0.25)	2.08 (0.28)	2.08 (0.27)	2.08 (0.27)	.488	1.06 (0.23)	1.07 (0.26)	1.06 (0.24)	1.09 (0.29)	.202

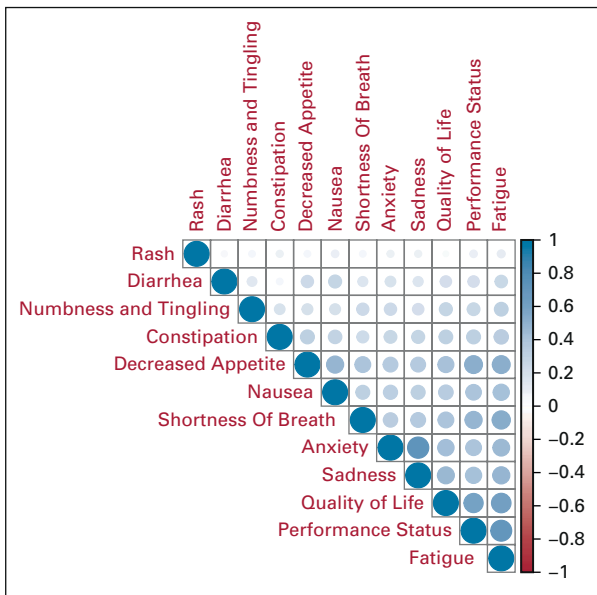
Abbreviations: PRO, patient-reported outcome; Q, quarter; SD, standard deviation.

<sup>a</sup>Linear model, where outcome is the categories of distribution of mortality risk (quarter 1-quarter 4), and predictor is the mean of PRO in each quarter.

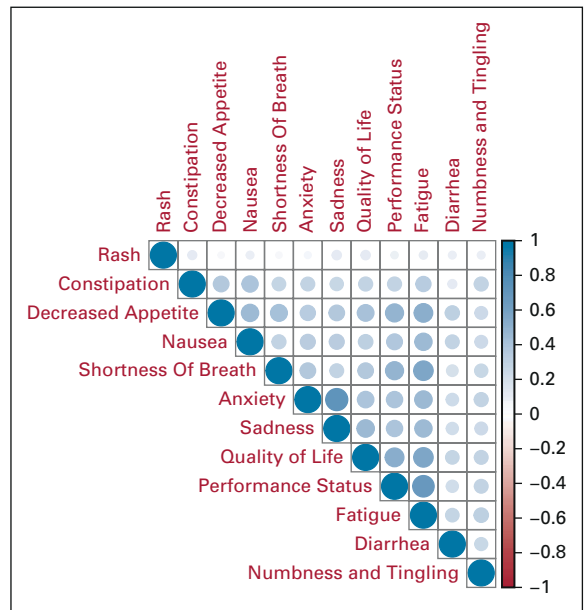
**FIG A1.** Cohort selection criteria. PAH, Pennsylvania Hospital; PCAM, Perelman School of Advanced Medicine.



**FIG A2.** Association between PRO trends and mortality. PRO, patient-reported outcome.

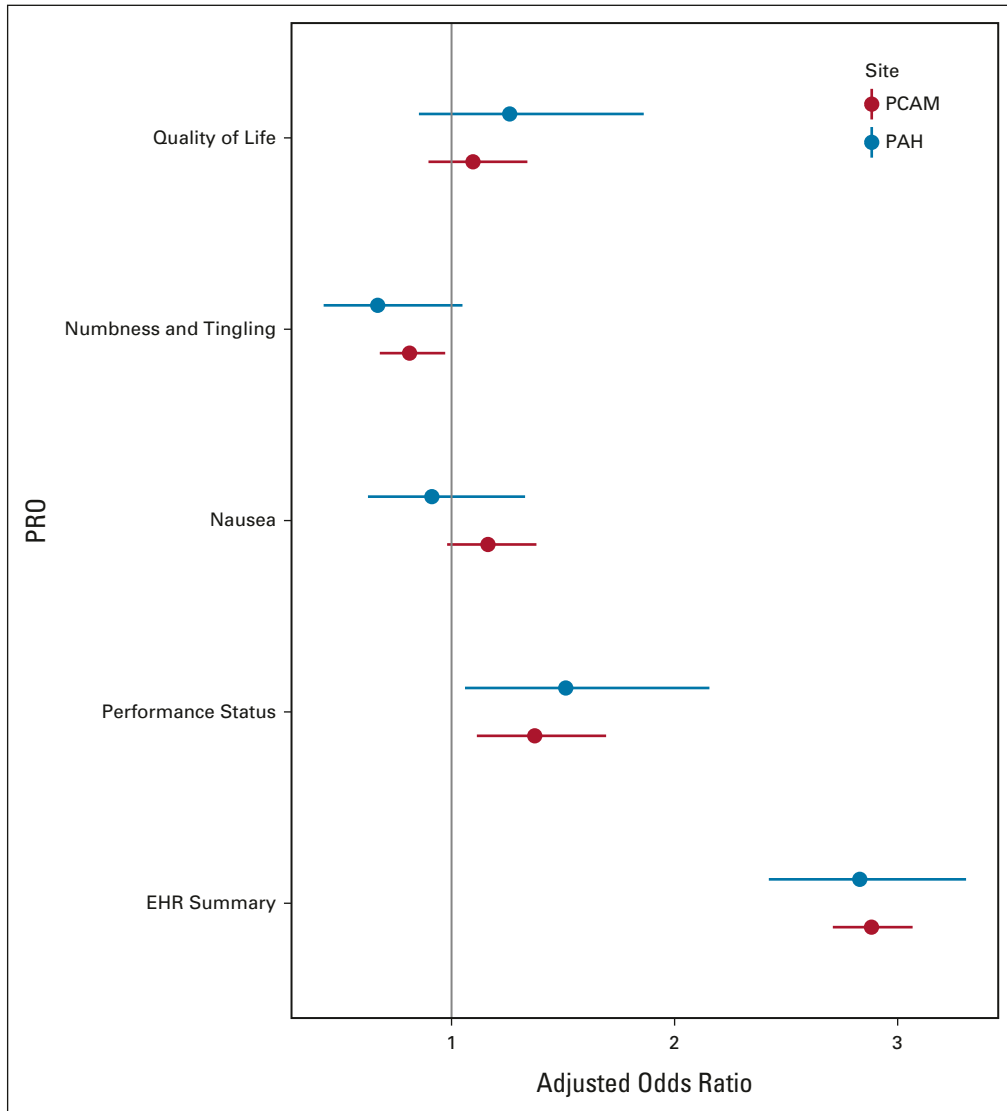


**FIG A3.** Associations between PROs at tertiary oncology practice. PRO, patient-reported outcome.



**FIG A4.** Associations between PROs at community oncology practice. PRO, patient-reported outcome.





**FIG A5.** PROs included in final model and associations with mortality. EHR, electronic health record; PAH, Pennsylvania Hospital; PCAM, Perelman School of Advanced Medicine; PRO, patient-reported outcome.