

Development of a Machine Learning Model Using Limited Features to Predict 6-Month Mortality at Treatment Decision Points for Patients With Advanced Solid Tumors

George Chalkidis, PhD¹; Jordan McPherson, PharmD, MS²; Anna Beck, MD²; Michael Newman, MS³; Shuntaro Yui, PhD¹; and Catherine Staes, PhD, MPH⁴

PURPOSE Patients with advanced solid tumors may receive intensive treatments near the end of life. This study aimed to create a machine learning (ML) model using limited features to predict 6-month mortality at treatment decision points (TDPs).

METHODS We identified a cohort of adults with advanced solid tumors receiving care at a major cancer center from 2014 to 2020. We identified TDPs for new lines of therapy (LoTs) and confirmed mortality at 6 months after a TDP. Using extreme gradient boosting, ML models were developed, which used or derived features from a limited set of electronic health record data considering the literature, clinical relevance, variability, availability, and predictive importance using Shapley additive explanations scores. We predicted and observed 6-month mortality after a TDP and assessed a risk stratification strategy with different risk thresholds to support communication of chance of survival.

RESULTS Four thousand one hundred ninety-two patients were included. Patients had 7,056 TDPs, for which the 6-month mortality increased from 17.9% to 46.7% after starting first to sixth LoT, respectively. On the basis of internal validation, models using both 111 (Full) or 45 (Limited-45) features accurately predicted 6-month mortality (area under the curve ≥ 0.80). Using a 0.3 risk threshold in the Limited-45 model, the observed 6-month survival was 34% (95% CI, 28 to 40) versus 81% (95% CI, 81 to 82) among those classified with low or higher chance of survival, respectively. The positive predictive value of the Limited-45 model was 0.66 (95% CI, 0.60 to 0.72).

CONCLUSION We developed and validated a ML model using a limited set of 45 features readily derived from electronic health record data to predict 6-month prognosis in patients with advanced solid tumors. The model output may support shared decision making as patients consider the next LoT.

JCO Clin Cancer Inform 6:e2100163. © 2022 by American Society of Clinical Oncology

Creative Commons Attribution Non-Commercial No Derivatives 4.0 License 

INTRODUCTION

Despite new therapeutic options, many patients with advanced cancer transition to a phase of terminal illness in which intensive treatments might have limited ability to extend life or restore health.^{1,2} This final 6 months preceding death is generally referred to as the end of life (EOL).³ At EOL, patients with advanced cancer should have access to high-value, compassionate, and evidence-based care consistent with their wishes.^{4,5} Despite the desire for a good death and eligibility for hospice when the life expectancy is less than 6 months,⁶ patients may receive intensive treatments at EOL.^{1,2,7-9} When clinicians accurately estimate prognosis, they refer patients earlier for palliative care and hospice care.¹⁰ Similarly, when patients have an accurate perception of prognosis, they are less likely to choose aggressive care at EOL^{11,12} and more likely to receive care consistent with their preferences.¹³

Delivering high-value care for patients with advanced cancer requires that clinicians accurately communicate prognosis during treatment discussions, and patients express their values and goals of care.¹⁴ Despite the existence of prognostic tools,^{9,15} clinicians often overestimate survival,¹⁶ which may affect decision making about treatment options.^{5,17} In addition, prognostic tools developed before 2018 focused on patients who were hospitalized, receiving hospice care, or within a few weeks or months of death, potentially suboptimal time frames for treatment decision making.¹⁸⁻²¹

Advances in computational capacity and electronic health records (EHRs) provide opportunities for machine learning (ML) to deliver data-driven mortality predictions linked to cancer care decisions. Recent publications describe accurate ML-based approaches for predicting 6-month mortality for patients with

ASSOCIATED CONTENT

Data Supplement

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

Accepted on March 3, 2022 and published at ascopubs.org/journal/cci on April 25, 2022; DOI <https://doi.org/10.1200/CCI.21.00163>

CONTEXT

Key Objective

Is it possible to use limited electronic health record data to develop a risk model for use at treatment decision points that predicts whether patients with advanced solid tumors will be deceased within 6 months after starting a new line of therapy?

Knowledge Generated

We developed and evaluated three models using a limited set of widely available electronic health record data features. The models accurately identified patients with a low chance of survival, who were then found to have had suboptimal referrals for palliative care or hospice. Cancer-specific features were not strongly predictive, suggesting a common pathway at end of life among patients with advanced solid tumors.

Relevance

By flagging at-risk patients when new lines of therapy are considered, we hope to improve shared decision making between patients and clinicians. Mutual understanding of prognosis and the predicted outcome of treatment decisions may improve alignment of care with patient's expectations at end of life.

cancer.²²⁻²⁶ However, their relevance, scalability, and interpretability may be limited by inclusion of patients before diagnosis with advanced cancer, the hundreds of features used, or positive predictive values (PPVs) not reported or < 54%.²²⁻²⁶ One model has been implemented as a stand-alone application,²⁷ but requires manual data entry and is based on pre-2015 data.²² These tools appear to be designed to influence clinician behavior and are not patient-facing. To optimize advance planning and palliative care, prognostic tools should be available during treatment discussions.

At the Huntsman Cancer Institute (HCI), a team with expertise in oncology (A.B. and J.M.), informatics (C.S. and M.N.), and data science (G.C., M.N., and S.Y.) developed a Use Case to describe goals, stakeholders, data, and tasks for integrating a prognostic tool into the advanced cancer care workflow.²⁸ From these discussions, the team sought to develop an automated, EHR-integrated tool using a limited feature set to accurately predict 6-month prognosis and support shared decision making as clinicians and patients with any advanced solid tumor consider a new line of therapy (LoT). Our objective was to (1) develop and validate a ML model that predicts 6-month mortality for patients with an advanced solid tumor considering a new LoT, using only EHR data available before starting therapy, and (2) risk stratify patients to support usability of the output for clinicians and patients.

METHODS

Data Source

Data were obtained in February 2021 from the University of Utah Health (UHealth) enterprise data warehouse, which includes data from the (1) Epic EHR implemented in all inpatient or outpatient UHealth and HCI care settings, (2) HCI hospital-based cancer registry (HCI-CR),²⁹ and (3) Utah Vital Records. To identify the study population, we used HCI-CR²⁹ and EHR data. To assign a deceased

outcome, we queried EHR or vital records data for a death date; if not available, we queried HCI-CR. To assign an alive outcome, we queried EHR data for patient measurements. All predictive features were based on EHR data, including International Classification of Diseases (ICD) codes used to assign metastases.

Study Population and Treatment Decision Points

Our study cohort concerned patients with advanced solid tumors who received care within UHealth enterprise. Advanced solid tumor was defined as malignant brain or nervous system cancer or any other solid tumor with metastases (Data Supplement). The index date was defined as the first date when an advanced solid tumor diagnosis was recorded in EHR data. Eligible patients had (1) no history of hematologic malignancy or bone marrow transplant; (2) two or more visits with a UHealth medical oncologist from 6 months before the index date or later; (3) age 18 years or older on the index date; and (4) an index date between June 1, 2014, and June 1, 2020.

To determine the final study population with at least one LoT with a treatment decision point (TDP), additional preprocessing and inclusion criteria were applied. To identify a LoT and treatment start date, we extracted data from 12 months before the patient's index date through November 30, 2020, allowing for a minimum follow-up of 6 months. Any anticancer therapy (chemotherapy, biologics, targeted therapy, immunotherapy, and hormonal therapy) entered into EHR treatment plans was defined as an eligible LoT, consistent with a recently published framework.³⁰ Injectable and selected oral hormonal therapies for prostate and breast cancer were not included. Finally, to define a TDP, we identified a visit within 30 days before the LoT start date with values for required features. A description of the preprocessing strategy and a flow diagram illustrating the study population are available (Data Supplement).

Features

Initially, the team reviewed tools and publications concerning cancer prognostication, enumerated potential prognostic features mentioned, and identified additional potentially clinically relevant features (Data Supplement). Next, clinical experts (A.B. and J.M.) selected a subset deemed clinically important and available in EHR data and described logic to operationalize features. To assess usefulness, we investigated feature availability and variability; then, evaluated feature logic and importance during iterative model development. On the basis of this analysis, we identified 111 features to be operationalized for each TDP (Table 1). Treated cancer type was determined using ICD codes reported in the EHR treatment plan data and classified into 15 categories (Data Supplement). The presence of metastasis (seven categories) was determined using ICD codes in encounter data. Most features were directly used, whereas others were derived using logic. For example, LoTs were counted after the index date and included as a counter variable. Time to next treatment (TTNT), the difference between current and previous treatment start dates, is a surrogate end point for real-world progression-free survival³¹ and only calculated for a second or later LoT. Features related to history of encounters (eg, emergency visits, inpatient visits, and inpatient length of stay), radiation therapy, and blood transfusions were summarized for defined months before TDP.^{23,32} Binary variables were used to create features for palliative care history (eg, flag if an order was recorded in the EHR ever before or within three months before the TDP). Order dates, encounter types, and logic were used to derive features concerning hospice and advance directives. Laboratory values, body mass index (BMI), and weight were augmented with their 3-month percent change if data within 2.5-3.5 months before the TDP were available.³³ We arrived at 111 features to include in the full prediction model (Table 1). No imputation of missing data was performed. Data after the TDP and the primary outcome were not included.

Outcome

The primary outcome was 6-month mortality after a TDP. An outcome of deceased was assigned if a death date was documented within 6 months after a TDP. For patients with no recorded death date, an outcome status of alive was assigned if at least one of the following data points was available during a visit 6 months or longer after the TDP: BMI, diagnoses, laboratory measurement, patient performance score, or length of stay.

Sample Size

Patients might have had multiple TDPs, and each TDP was an independent observation. Use of multiple observations per patient is appropriate for prognostic model development.^{34,35} When performing analyses by cancer type and LoT, we required at least two observations for each outcome (deceased or alive) to ensure that the algorithm could be trained and evaluated for higher LoTs when data were scarce.

Machine Learning Algorithm

We used the extreme gradient boosting algorithm (xgb)³⁶ for three reasons. First, a previous relevant publication compared six ML strategies and demonstrated highest performance using xgb, even with missing values.²² Second, xgb was used in four recent publications for a similar context.^{23-25,37} Third, the Shapley additive explanations (SHAP)³⁸ can be used to identify the importance of features when aggregating individual predictions across the population, thus allowing for selection of features for a limited model to improve explainability for users of a future tool.

We used repeated Monte Carlo record-wise cross-validation simulations^{34,35,39} to train and evaluate the classifier. TDPs were randomly split accounting for cancer type, LoT, and 6-month mortality: 70% of observations were used to train and 30% were used to evaluate the model. Optimized hyperparameters were determined as described in the Data Supplement.³⁹

Limited Models

To develop a model with a limited set of features, we used an iterative strategy. While operationalizing features for the full model, we ranked features by SHAP score and requested clinical experts (A.B. and J.M.) to select 25-30 features on the basis of clinical importance, predictive impact, and expected availability across health care settings. Generally, features with scores above 0.1 were selected. Next, we created Limited models with and without features for cancer type and metastases ($n = 45$ v 23 features, respectively). Although these features were not highly ranked, they may be expected by users to be included.

Assessment of Performance, Risk Thresholds, and Risk Stratification

We used repeated Monte Carlo cross-validations to assess the mean and 95% CI for each performance measure.^{39,40} Since PPV, sensitivity, negative predictive value (NPV), and specificity depend on a risk threshold (ie, decision boundary) for assigning patients to either low or higher survival chance group at six months, we used clinical input and a data-driven approach to determine a clinically relevant risk threshold, rather than the default of 0.50. We calculated model performance and observed survival using incremental risk thresholds from 0.05 to 0.95.

Using the Limited-45 model and defined risk threshold of 0.30, patients were classified by chance of survival (low or higher). We then calculated mean observed 6-month survival, overall and stratified by LoT, cancer type, predicted chance of survival, and quality metrics such as referrals for palliative care or hospice and rates of hospitalization.

Statistical Analysis

We used descriptive analyses to summarize population characteristics and quality metrics. Delong's method was used to compare area under the curve (AUC) of different

TABLE 1. Summary of Features Included in the Full Predictive Model (111 features) and the Subset Used for Limited Models With 45 or 23 Features

Category/Feature	Description	Importance in the Full Model		Included in the Limited Model		Reason for Not Including in Limited Model(s) Other Than Low Rank
		SHAP ^a	Rank	-45	-23	
Population descriptors						
Age	Age at the time of TDP	0.24	8	Yes	Yes	
Sex	Gender indicator	0.04	57	No	No	
Race	Flag if White	0.01	79	No	No	
Ethnicity	Flag if Hispanic or Latino	0.01	83	No	No	
Primary cancer type						
Pancreas	Flag for the treated type of cancer determined by ICD codes associated with treatment plan data	0.10	29	Yes	No	
Lung		0.09	33	Yes	No	
Breast		0.08	39	Yes	No	
Brain/nervous system		0.05	47	Yes	No	
Prostate		0.03	60	Yes	No	
Colon, rectum, and small bowel		0.03	61	Yes	No	
Lip, oral cavity, and pharynx		0.03	63	Yes	No	
Others		0.02	69	Yes	No	
Ovary		0.01	74	Yes	No	
Melanoma/skin		0.01	84	Yes	No	
Kidney		0.01	89	Yes	No	
Urinary bladder and other urinary		0.00	97	Yes	No	
Stomach		0.00	101	Yes	No	
Corpus uteri		0.00	102	Yes	No	
Soft tissue	0.00	107	Yes	No		
Sites of metastasis						
Liver	Flag for the site of metastasis determined by ICD codes in encounter data	0.10	30	Yes	No	
Bone		0.09	32	Yes	No	
Lung		0.07	40	Yes	No	
Retroperitoneal		0.06	44	Yes	No	
Brain		0.05	48	Yes	No	
Others		0.04	56	Yes	No	
Mediastinum		0.01	82	Yes	No	
Timing of cancer diagnosis and progression						
Time since the index date	TDP minus index date	0.18	14	Yes	Yes	
Time since initial diagnosis	TDP minus first cancer diagnosis	0.20	11	No	No	Time since the index date is included
Time between initial diagnosis and index date	Index date minus first cancer diagnosis date	0.14	20	No	No	
Encounter information: Inpatient length of stay						
Previous 12 months before TDP	Calculated as the sum of inpatient length of stay in the previous months before TDP	0.11	27	No	No	May not be readily available at oncology centers
Previous 9 months before TDP		0.09	35	No	No	
Previous 6 months before TDP		0.03	59	No	No	
Previous 3 months before TDP		0.04	51	No	No	
Previous 1 month before TDP		0.02	68	No	No	
Encounter information: No. of inpatient visits						

(Continued on following page)

TABLE 1. Summary of Features Included in the Full Predictive Model (111 features) and the Subset Used for Limited Models With 45 or 23 Features (Continued)

Category/Feature	Description	Importance in the Full Model		Included in the Limited Model		Reason for Not Including in Limited Model(s) Other Than Low Rank
		SHAP ^a	Rank	-45	-23	
Previous 12 months before TDP	No. of inpatient visits at specified intervals before TDP	0.01	73	No	No	May not be readily available at oncology centers
Previous 9 months before TDP		0.01	92	No	No	
Previous 6 months before TDP		0.01	85	No	No	
Previous 3 months before TDP		0.01	70	No	No	
Previous 1 month before TDP		0.00	100	No	No	
Encounter information: No. of emergency department visits						
Previous 12 months before TDP	No. of emergency department visits at specified intervals before TDP	0.01	71	No	No	May not be readily available at oncology centers
Previous 9 months before TDP		0.01	87	No	No	
Previous 6 months before TDP		0.01	93	No	No	
Previous 3 months before TDP		0.01	90	No	No	
Previous 1 month before TDP		0.00	96	No	No	
Risk behaviors (tobacco history)						
Ever smoked tobacco	Yes/no	0.04	58	No	No	
Never or only passively smoked tobacco	Yes/no	0.02	66	No	No	
Ever smokeless tobacco use	Yes/no	0.01	91	No	No	
Never or passively only used smokeless tobacco	Yes/no	0.00	108	No	No	
BMI and other nutritional features						
BMI measurement		0.19	13	Yes	Yes	
Three-month weight loss	See the note at the end of the table ^b	0.09	36	Yes	Yes	
Weight, kg		0.12	22	No	No	BMI included
Three-month BMI percent change	See the note at the end of the table ^b	0.11	26	No	No	3-month weight loss included
Clinical metrics						
Pain score	Recorded pain scores	0.31	2	Yes	Yes	
MEWS score	Surrogate for abnormal vital signs	0.07	42	No	No	
CCI		0.10	28	No	No	SHAP initially < 0.1 during iterative model development. Model performance unchanged when included
Blood transfusion history						
Ever before TDP	Flag if received blood transfusions at specified intervals before TDP	0.01	76	No	No	Data may not be broadly available
Previous 12 months before TDP		0.01	86	No	No	
Previous 9 months before TDP		0.01	72	No	No	
Previous 6 months before TDP		0.00	98	No	No	
Previous 3 months before TDP		0.00	99	No	No	
Previous 1 month before TDP		0.00	104	No	No	
Cancer treatment: past						
TTNT, used as surrogate for progression-free survival	Planned treatment start date minus previous treatment start date; Note: Only calculated for a second or later LoT	0.24	7	Yes	Yes	
Chemotherapy within 6 months before the index date	Flag if the patient received chemotherapy within 6 months before the index date according to treatment plan data	0.04	53	No	No	

(Continued on following page)

TABLE 1. Summary of Features Included in the Full Predictive Model (111 features) and the Subset Used for Limited Models With 45 or 23 Features (Continued)

Category/Feature	Description	Importance in the Full Model		Included in the Limited Model		Reason for Not Including in Limited Model(s) Other Than Low Rank		
		SHAP ^a	Rank	-45	-23			
Cancer treatment: current								
Current LoT (ie, 1, 2, 3, etc)	Line of therapy recorded in EHR after the index date, derived from start dates in treatment plans	0.16	17	Yes	Yes			
Radiation therapy								
Ever before TDP	Flag if the patient received radiation therapy at specified intervals before TDP	0.01	88	No	No	Data may not be broadly available		
Within 12 months before TDP		0.05	49	No	No			
Within 9 months before TDP		0.01	81	No	No			
Within 6 months before TDP		0.01	80	No	No			
Within 3 months before TDP		0.00	94	No	No			
Within 1 month before TDP		0.00	95	No	No			
Laboratory data								
WBC count		0.12	21	Yes	Yes			
HgB		0.26	6	Yes	Yes			
Lymphocyte (%)		0.28	5	Yes	Yes			
Platelet count		0.16	18	Yes	Yes			
Monocyte (%)		0.11	25	Yes	Yes			
Absolute lymphocyte count		0.21	10	No	No	Lymphocyte % included		
Monocyte count		0.16	16	No	No	Monocyte % included		
Neutrophil count		0.01	75	No	No			
Albumin		0.81	1	Yes	Yes			
Creatinine		0.11	24	Yes	Yes			
LDH concentration		0.23	9	Yes	Yes			
ALP		0.29	4	Yes	Yes			
Sodium		0.19	12	Yes	Yes			
Calcium		0.17	15	Yes	Yes			
Bilirubin		0.15	19	Yes	Yes			
Three-month percent change in laboratory data								
Platelet count	See the note at the end of the table ^b	0.11	23	Yes	Yes	SHAP score initially ≥ 0.1 during iterative model development		
Hemoglobin		0.10	31	Yes	Yes			
Alkaline phosphatase		0.09	34	Yes	Yes			
Albumin		0.06	43	Yes	Yes			
WBC count		0.04	55	No	No			
Absolute lymphocyte count		0.04	54	No	No			
Neutrophil count		0.00	106	No	No			
Lymphocyte, %		0.08	38	No	No			
Monocyte count		0.03	62	No	No			
Monocyte, %		0.07	41	No	No			
Creatinine		0.06	45	No	No			
LDH		0.01	77	No	No			
Sodium		0.08	37	No	No			
Calcium		0.05	46	No	No			
Bilirubin		0.04	52	No	No			
Performance status								

(Continued on following page)

TABLE 1. Summary of Features Included in the Full Predictive Model (111 features) and the Subset Used for Limited Models With 45 or 23 Features (Continued)

Category/Feature	Description	Importance in the Full Model		Included in the Limited Model		Reason for Not Including in Limited Model(s) Other Than Low Rank
		SHAP ^a	Rank	-45	-23	
ECOG		0.01	78	No	No	Not routinely recorded
KPS		0.02	65	No	No	Not routinely recorded
Patient-reported outcomes—symptoms						
CMSA score		0.00	103	No	No	Not broadly available
CMSA score—physical subscore		0.02	67	No	No	Not broadly available
Palliative care, advance directives, and hospice						
Palliative care consultation before TDP	Flag for palliative care consultation(s) at specified intervals before TDP	0.30	3	No	No	Feature is dynamic and may vary by cancer center. For our Use Case, we want tool to prompt palliative care referral (outcome v input)
Palliative care consultation within 3 months before TDP		0.03	64	No	No	
Advance directive recorded	Flag if documented before TDP	0.00	105	No	No	
Time since advance directive documented	TDP minus advance directive date	0.05	50	No	No	
Hospice referral before TDP	Flag if referred before TDP	0.00	109	No	No	
Meets hospice criteria	Flag if most recent medical oncologist visit occurred within 14 days after last hospice referral	0.00	110	No	No	
Time since first hospice referral	TDP minus earliest available hospice referral date	0.00	111	No	No	

Abbreviations: ALP, alkaline phosphatase; BMI, body mass index; CCI, Charlson Comorbidity Index; CMSA, Condensed Memorial Symptom Assessment; ECOG, Eastern Cooperative Oncology Group; EHR, electronic health record; HgB, hemoglobin; ICD, International Classification of Diseases; KPS, Karnofsky Performance Score; LDH, lactate dehydrogenase; LoT, line of therapy; SHAP, Shapley additive explanations; TDP, treatment decision point; TTNT, time to next treatment.

^aThe SHAP scores presented in this table are the final scores derived using 100 simulations and the full 111-feature model after all features were defined and selected during model development. They vary slightly from the scores derived for the Limited model feature selection process.

^bThree-month percent change calculation, using weight loss as an example: $(\text{Current weight in kg} - \text{Previous weight in kg}) / \text{previous weight in kg} / 100$. For every 3-month percent change feature, the previous value is the measurement within 2.5-3.5 months before TDP that is closest to the 3-month point in time.

prediction models.⁴¹ We adjusted *P* values using Bonferroni correction.

This project was approved by the University of Utah and Hitachi Institutional Review Boards and followed Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis guidelines.⁴²

RESULTS

Among the 7,690 potentially eligible patients identified in the EHR, 5,167 (67%) had treatment data and 4,192 (54.5%) met the inclusion criteria for the final study cohort (Data Supplement). These 4,192 patients had a total of 7,056 TDPs distributed across the first to sixth LoTs (Table 2). The most common cancer types were breast (13.1%), lung (11.4%), melanoma (11.2%), colon, rectum and small bowel (10.1%), and prostate (8.9%). The observed 6-month mortality after a TDP increased incrementally from 17.9% after the first line to 46% after the sixth LoT. Additional

detailed patient characteristics and observed 6-month survival stratified by cancer type at each LoT are available (Data Supplement).

Among the 111 features included in the full model, albumin was most predictive (SHAP score 0.81), followed by selected laboratory values, pain score, palliative care history, time since index date, TTNT, and age. Cancer type, metastasis, MEWS score, Charlson comorbidity index, encounter history, sex, and race had lower importance (Table 1).

The subset of predictive and clinically relevant features selected by clinical experts resulted in two models: Limited-45 (with 45 features) and Limited-23 (with 23 features excluding types of cancers and metastases). The top 20 features for both models in descending order were as follows: albumin, lymphocyte %, pain score, hemoglobin, alkaline phosphatase, TTNT, time since index date, sodium, lactate dehydrogenase, BMI, age, 3-month % weight change, calcium, platelets, monocyte %, line of

TABLE 2. Patient Characteristics, Overall and by Line of Therapy Available in Electronic Health Record Data (n = 4,192 patients)

Characteristic	All Patients ^a	First Line	Second Line	Third Line	Fourth Line	Fifth Line	Sixth Line
TDPs, No. (% of total)	7,056 (100)	4,129 ^b (58.5)	1,750 (24.8)	762 (10.8)	302 (4.3)	98 (1.4)	15 (0.2)
Deceased in 6 months, No. (%)		740 (17.9)	444 (25.4)	267 (35.0)	109 (36.1)	41 (41.8)	7 (46.7)
Demographic							
Age, mean (95% CI), years	60.4 (60.0 to 60.8)	60.7 (60.3 to 61.1)	59.9 (59.2 to 60.5)	59.1 (58.1 to 60.1)	57.9 (56.4 to 59.4)	57.3 (54.5 to 60.2)	62.0 (54.7 to 69.3)
Female, No. (%)	2,076 (49.5)	2,054 (49.7)	930 (53.1)	428 (56.2)	181 (59.9)	74 (75.5)	15 (100.0)
Race, White, No. (%)	3,764 (89.8)	3,707 (89.8)	1,560 (89.1)	688 (90.3)	274 (90.7)	90 (91.8)	14 (93.3)
Cancer type, No. (%)							
Breast	550 (13.1)	540 (13.1)	287 (16.4)	144 (18.9)	74 (24.5)	33 (33.7)	15 (100.0)
Lung	477 (11.4)	469 (11.4)	206 (11.8)	76 (10.0)	20 (6.6)	0 (0.0)	0 (0.0)
Colon, rectum, and small bowel	425 (10.1)	420 (10.2)	178 (10.2)	87 (11.4)	35 (11.6)	11 (11.2)	0 (0.0)
Prostate	371 (8.9)	359 (8.7)	132 (7.5)	69 (9.1)	25 (8.3)	10 (10.2)	0 (0.0)
Melanoma	470 (11.2)	463 (11.2)	176 (10.1)	76 (10.0)	37 (12.3)	19 (19.4)	0 (0.0)
Brain and nervous system	241 (5.7)	231 (5.6)	95 (5.4)	37 (4.9)	12 (4.0)	0 (0.0)	0 (0.0)
Pancreas	230 (5.5)	230 (5.6)	89 (5.1)	28 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)
Kidney	175 (4.2)	171 (4.1)	86 (4.9)	51 (6.7)	20 (6.6)	0 (0.0)	0 (0.0)
Bladder and urinary	142 (3.4)	141 (3.4)	61 (3.5)	22 (2.9)	0 (0.0)	0 (0.0)	0 (0.0)
Ovary	166 (4.0)	165 (4.0)	79 (4.5)	40 (5.2)	26 (8.6)	15 (15.3)	0 (0.0)
Corpus uteri	106 (2.5)	105 (2.5)	42 (2.4)	17 (2.2)	12 (4.0)	0 (0.0)	0 (0.0)
Lip, oral cavity, and pharynx	111 (2.6)	111 (2.7)	25 (1.4)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Soft tissue	110 (2.6)	110 (2.7)	53 (3.0)	25 (3.3)	12 (4.0)	0 (0.0)	0 (0.0)
Stomach	73 (1.7)	72 (1.7)	39 (2.2)	13 (1.7)	0 (0.0)	0 (0.0)	0 (0.0)
Others	545 (13.0)	542 (13.1)	202 (11.5)	77 (10.1)	29 (9.6)	10 (10.2)	0 (0.0)

Abbreviation: TDP, treatment decision point.

^aAt the index date.

^bAmong the 4,192 patients, 4,129 (98.5%) had data available for their first line and the remaining 63 patients had data for a second or further line of therapy only.

TABLE 3. Performance Results for Three Machine Learning Models on Aggregate and Stratified by Line of Therapy^a

Model	AUC	Precision/PPV	Recall/Sensitivity	NPV	Specificity
	Mean (95% CI)				
Full (111 features)					
Aggregate	0.81 (0.79 to 0.82)	0.68 (0.62 to 0.73)	0.25 (0.22 to 0.29)	0.81 (0.81 to 0.82)	0.96 (0.95 to 0.97)
LoT 1	0.81 (0.78 to 0.83)	0.64 (0.56 to 0.75)	0.19 (0.15 to 0.25)	0.85 (0.84 to 0.85)	0.98 (0.97 to 0.99)
LoT 2	0.80 (0.76 to 0.83)	0.67 (0.57 to 0.76)	0.28 (0.22 to 0.34)	0.79 (0.78 to 0.81)	0.95 (0.93 to 0.97)
LoT 3	0.78 (0.73 to 0.83)	0.74 (0.61 to 0.88)	0.33 (0.24 to 0.42)	0.72 (0.69 to 0.75)	0.94 (0.89 to 0.97)
LoT 4	0.77 (0.70 to 0.84)	0.70 (0.53 to 0.88)	0.36 (0.19 to 0.49)	0.72 (0.68 to 0.76)	0.91 (0.86 to 0.97)
LoT 5	0.79 (0.63 to 0.91)	0.72 (0.43 to 1.00)	0.42 (0.17 to 0.70)	0.70 (0.61 to 0.82)	0.89 (0.76 to 1.00)
LoT 6	0.48 (0.00 to 1.00)	0.79 (0.00 to 1.00)	0.10 (0.00 to 0.50)	0.49 (0.33 to 0.67)	0.89 (0.50 to 1.00)
Limited-45 (45 features)					
Aggregate	0.80 (0.78 to 0.81)	0.66 (0.60 to 0.72)	0.24 (0.21 to 0.27)	0.81 (0.81 to 0.82)	0.96 (0.95 to 0.97)
LoT 1	0.80 (0.77 to 0.82)	0.62 (0.53 to 0.71)	0.17 (0.13 to 0.22)	0.84 (0.84 to 0.85)	0.98 (0.97 to 0.98)
LoT 2	0.79 (0.75 to 0.82)	0.66 (0.55 to 0.74)	0.27 (0.20 to 0.33)	0.79 (0.78 to 0.80)	0.95 (0.93 to 0.97)
LoT 3	0.78 (0.72 to 0.82)	0.74 (0.61 to 0.86)	0.34 (0.26 to 0.42)	0.72 (0.70 to 0.75)	0.93 (0.89 to 0.97)
LoT 4	0.76 (0.69 to 0.83)	0.67 (0.47 to 0.86)	0.35 (0.19 to 0.52)	0.72 (0.67 to 0.77)	0.90 (0.84 to 0.96)
LoT 5	0.78 (0.62 to 0.91)	0.72 (0.46 to 1.00)	0.46 (0.21 to 0.73)	0.71 (0.62 to 0.82)	0.88 (0.71 to 1.00)
LoT 6	0.46 (0.00 to 1.00)	0.85 (0.00 to 1.00)	0.10 (0.00 to 0.50)	0.51 (0.33 to 0.67)	0.93 (0.50 to 1.00)
Limited-23 (23 features)					
Aggregate	0.78 (0.76 to 0.80)	0.64 (0.58 to 0.70)	0.23 (0.20 to 0.26)	0.81 (0.80 to 0.81)	0.96 (0.95 to 0.97)
LoT 1	0.77 (0.74 to 0.79)	0.59 (0.48 to 0.69)	0.16 (0.12 to 0.20)	0.84 (0.83 to 0.85)	0.98 (0.97 to 0.98)
LoT 2	0.77 (0.74 to 0.80)	0.63 (0.54 to 0.73)	0.25 (0.19 to 0.32)	0.79 (0.78 to 0.80)	0.95 (0.93 to 0.97)
LoT 3	0.75 (0.69 to 0.79)	0.72 (0.57 to 0.82)	0.32 (0.25 to 0.40)	0.72 (0.69 to 0.74)	0.93 (0.88 to 0.97)
LoT 4	0.76 (0.68 to 0.84)	0.66 (0.49 to 0.83)	0.34 (0.19 to 0.48)	0.71 (0.67 to 0.76)	0.90 (0.83 to 0.96)
LoT 5	0.78 (0.63 to 0.93)	0.72 (0.50 to 1.00)	0.46 (0.18 to 0.73)	0.71 (0.60 to 0.84)	0.88 (0.71 to 1.00)
LoT 6	0.42 (0.00 to 1.00)	0.86 (0.00 to 1.00)	0.15 (0.00 to 0.50)	0.53 (0.33 to 0.67)	0.93 (0.50 to 1.00)

Abbreviations: AUC, area under the curve; LoT, line of therapy; NPV, negative predictive value; PPV, positive predictive value.

^aUsing a 30% threshold to stratify risk, AUC of the Limited-23 model was significantly different when compared with that of the full and Limited-45 model with Delong's test ($P < .05$).

therapy (LoT), bilirubin, 3-month % platelets change, 3-month % hemoglobin change, white blood count (Data Supplement).

Model Performance

For predicting 6-month mortality at any TDP aggregating across LoTs using internal validation, the full model had an AUC of 0.81 (95% CI, 0.79 to 0.82), which was not significantly different from the Limited-45 model (AUC: 0.80 [95% CI, 0.78 to 0.81]; Table 3). The Limited-23 model was accurate (AUC: 0.78 [95% CI, 0.76 to 0.80]) despite using no cancer features, but the AUC differed significantly from the other two models ($P < .05$).

To define a risk threshold, the Limited-45 model was used. PPV, sensitivity, NPV, and specificity varied across risk thresholds, as shown in Figure 1 and the Data Supplement. Similarly, the observed 6-month survival increased by 3% with each 10% change in risk threshold (eg, from 27% to 39% when the risk threshold was adjusted from < 0.10 to $<$

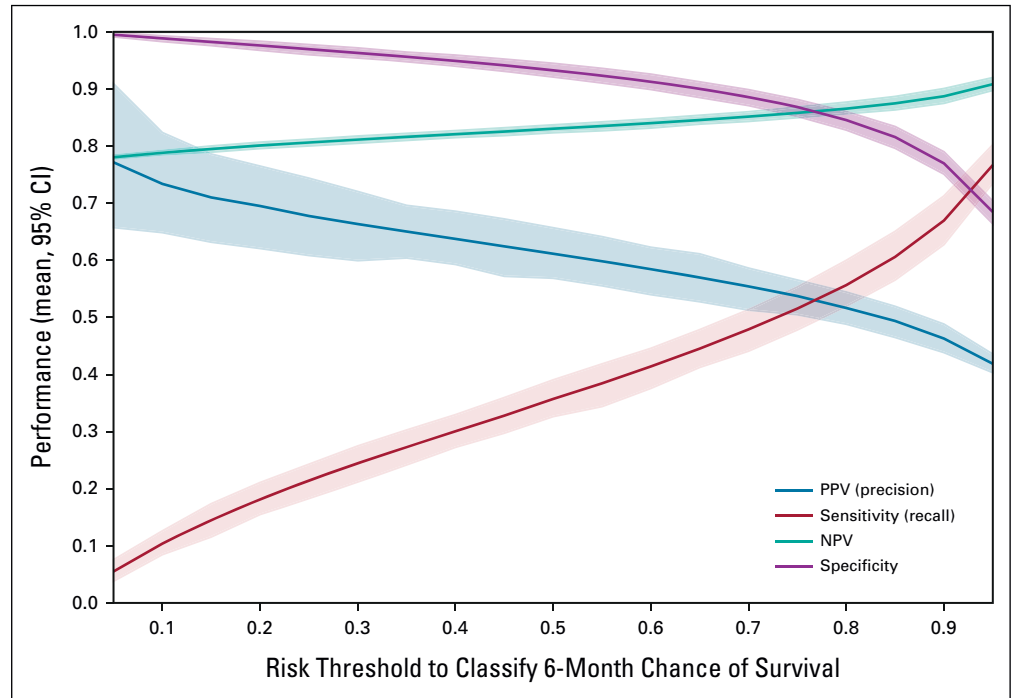
0.50; Fig 2A). Using a risk threshold of 0.30, only one in 3 (33%) patients who were classified as having low chance of survival had in fact survived 6 months, matching the mental model of clinical experts on the team. In addition, PPV was higher when using a threshold of 0.30 compared with the standard threshold of 0.50. Thus, given the priority for accurate prognosis over sensitivity, the risk threshold of 0.30 was used for all model performance evaluations.

Using the selected risk threshold of 0.30, predictive performance metrics varied by LoTs and were similar for all models (Table 3). PPV and sensitivity were lowest when prognosticating for first LoT but increased with each subsequent line. NPV and specificity were high for all models and subgroups.

Assessment of Risk Stratification Compared With Observed Outcomes

Using the Limited-45 model (and 0.30 risk threshold), on aggregate, the observed 6-month survival among

FIG 1. Performance of the Limited-45 model using incremental risk thresholds from 0.05 to 0.95. When the risk threshold is set at 0.30 (30%), only one of three patients with a low predicted chance of survival is, in fact, alive after 6 months (ie, PPV = 66%). NPV, negative predictive value; PPV, positive predictive value.



patients predicted with low chance of survival (33.6%) was less than half of that observed among patients predicted with higher chance of survival (81.1%; Fig 2B and

Table 4). The proportion of patients classified with a low chance of survival increased with each successive LoT (from 5% for the first LoT to 10%, 16%, 19% to 25% for the

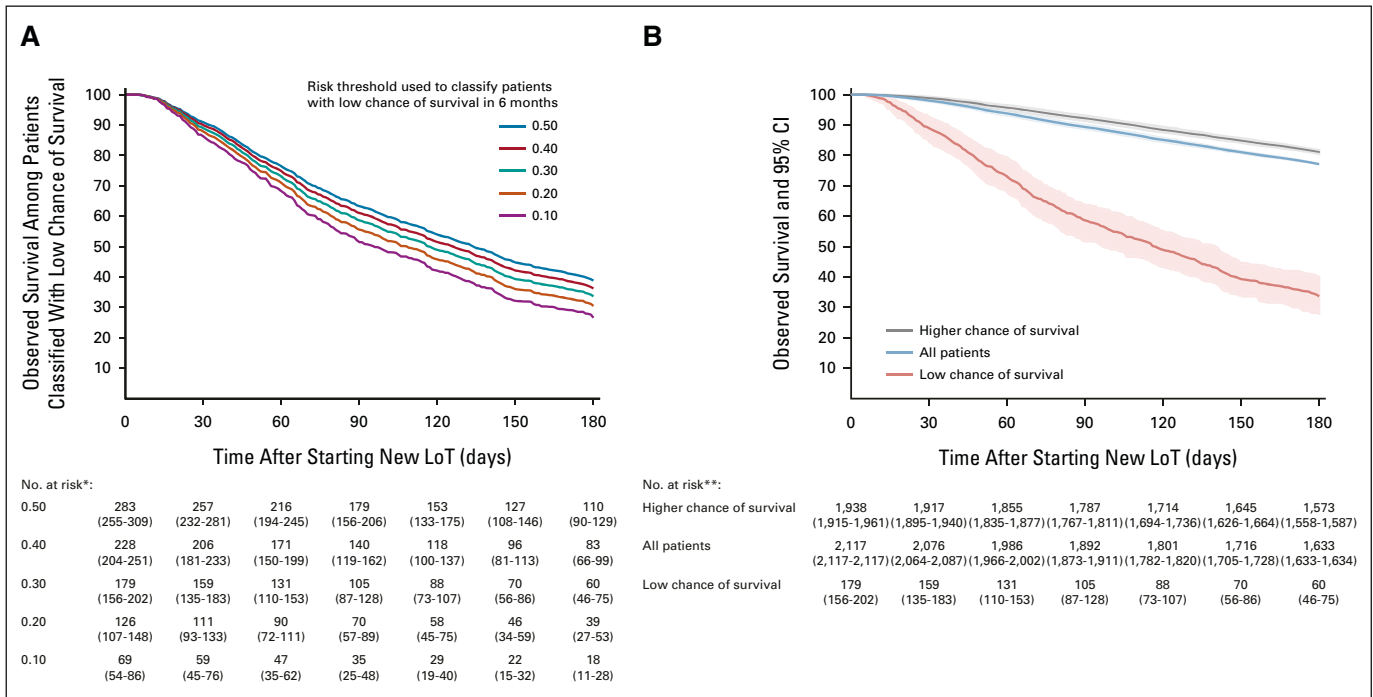


FIG 2. Observed survival during six months after starting a new line of therapy stratified by (A) risk thresholds used in the Limited-45 model to assign low chance of survival and (B) chance of survival (low or higher) on the basis of the use of a risk threshold of 0.30 for the Limited-45 model. *No. at risk at TDPs when classified as low chance of survival, mean No. (95% CI). Based on 100 Monte Carlo simulations of repeated 70%-30% train-test splits. The evaluation results in the table are based on the TDPs in the test split that were classified as low chance of survival. **No. at risk at TDPs, mean No. (95% CI). Based on 100 Monte Carlo simulations of repeated 70%-30% train-test splits. The test split upon which the evaluation results in the table are based uses 30% of our total of 7056 TDPs (ie, 2,117 TDPs on aggregate for all patients when rounded to the nearest integer at time 0). LoT, line of therapy; TDP, treatment decision point.

TABLE 4. Observed 180-Day Survival and Patient Group Size Stratified by 6-Month Mortality Risk and Line of Therapy Using the Limited-45 Model^a

Line of Therapy	Survival Chance	Population		Observed 6-Month Survival Mean % (95% CI)
		Mean No. (95% CI)	Mean % (95% CI)	
Aggregate	Higher	1,938 (1,915 to 1,961)	91.5 (90.5 to 92.6)	81.1 (80.5 to 81.8)
	Low	179 (156 to 202)	8.5 (7.4 to 9.5)	33.6 (28.0 to 40.0)
First	Higher	1,179 (1,163 to 1,194)	95.0 (93.7 to 96.1)	84.2 (83.6 to 84.9)
	Low	63 (48 to 79)	5.0 (3.9 to 6.3)	37.7 (28.5 to 47.0)
Second	Higher	471 (459 to 484)	89.6 (87.3 to 92.0)	79.2 (77.8 to 80.5)
	Low	55 (42 to 67)	10.4 (8.0 to 12.7)	34.4 (25.5 to 45.4)
Third	Higher	192 (183 to 201)	83.7 (79.9 to 87.6)	72.4 (69.6 to 74.9)
	Low	37 (28 to 46)	16.3 (12.4 to 20.1)	26.5 (14.5 to 39.4)
Fourth	Higher	71 (65 to 78)	81.1 (73.9 to 89.2)	71.5 (67.1 to 77.1)
	Low	16 (9 to 23)	18.9 (10.8 to 26.1)	32.6 (14.3 to 53.1)
Fifth	Higher	21 (16 to 25)	74.6 (56.1 to 86.0)	71.4 (61.9 to 82.4)
	Low	7 (4 to 13)	25.4 (14.0 to 43.9)	27.6 (0.0 to 53.7)

^aOn the basis of 100 Monte Carlo simulations of repeated 70%-30% train-test splits and using a 0.30 risk threshold: low is < 0.30; higher is ≥ 0.30.

fifth LoT); however, the observed 6-month survival was similar regardless of LoT (range: 26% to 37%; [Table 4](#)).

Among 7,056 TDPs experienced by our patient population, an average of 8.5% were flagged with low chance of survival. Among those with low chance of survival, on average, 66% were deceased within 6 months (PPV = 66%). Of those patients predicted to have low chance of survival who died within 6 months of starting a new LoT, estimated quality metrics were suboptimal: an average of 15% were referred for hospice care, 51% were referred for palliative care, and 60% were hospitalized between the TDP and death.

DISCUSSION

The observed 6-month mortality rate for our cohort of patients with advanced solid tumors increased incrementally from 18% after starting the first LoT to 47% after starting the sixth line, demonstrating the need for improved prediction of mortality when considering new LoTs. We developed an accurate prognosis model that relied on a limited set of features from EHR data available at a patient's visit to discuss treatment options. Our unique approach for feature selection using SHAP resulted in a set of most influential features, including albumin, consistent with those identified by others.²²⁻²⁴ However, we identified additional predictive features that could be operationalized using EHR data, including treatment response time (using TTNT as a proxy), palliative care consultation, and pain scores over time. Most importantly, the predictive performance of the model was similar with or without cancer-related features and continued to have a PPV of 60% or higher regardless of the model or LoT. This suggests that patients with advanced solid tumors may converge to a common pathway at EOL regardless of the cancer type, at which point patient-specific factors unrelated to cancer are most important.

We acknowledge that previously reported ML models have similar predictive capabilities; however, these models required hundreds of data points, potentially limiting scalability and interpretability.^{23,24} Our Limited-45 model achieved equivalent and accurate performance with only 45 readily derivable features. The 45 features may be implementable using the FHIR standard,⁴³ maximizing portability, and displayed in a user interface to support interpretability. Transparent display of input values may allow users to manually update values to recalculate prognosis, if needed. These factors may improve presentation of objective data used by providers to discuss the difficult subject of prognosis with their patients.

We selected a decision threshold consistent with how clinicians explain low or higher chance of survival. Regardless of LoT, observed 6-month survival approximated one third for every subset assigned to low chance of survival, illustrating model consistency discrimination. In addition, these patients had suboptimal EOL quality metrics.⁴⁴ Automating identification of at-risk patients could improve care.

Our risk stratification was designed to be patient- and provider-facing to support shared decision making. We are currently developing a user interface for clinicians and patients to visualize model inputs and observed survival among patients with similar risk. A clinical trial assessing the effect on shared decision making and EOL quality metrics is planned.

Our study has limitations. The model was trained on patients who underwent therapy because TDPs were defined using treatment data. Therefore, the model can only make predictions on the basis of the assumption that anticancer therapy will be continued. The model was trained on data from one institution and may not reflect care patterns in other settings or for minority populations.⁴⁵ Furthermore, LoTs were only calculated on the basis of treatment data

available in the EHR. Our limited sample size, particularly at higher LoTs, rendered predictions with higher uncertainty, but reporting 95% CIs improved transparency. Our model was only validated internally; external validation will be necessary. Finally, although pain score was a highly predictive feature, these data are subjective and inconsistently recorded. Despite limitations, our approach is expected to be scalable, interpretable, and clinically relevant. A

prospective evaluation and deployment at other health systems is needed to validate these assumptions.

In conclusion, we developed and validated a ML model using a limited set of 45 features readily derived from EHR data to predict 6-month prognosis. Risk stratification using this model may be used to support shared decision making as patients with advanced cancer consider the next LoT.

AFFILIATIONS

¹Hitachi, Ltd, Tokyo, Japan

²Huntsman Cancer Institute, Salt Lake City, UT

³University of Utah Health, Salt Lake City, UT

⁴College of Nursing, University of Utah, Salt Lake City, UT

CORRESPONDING AUTHOR

Catherine Staes, PhD, MPH, University of Utah, College of Nursing, 10 South 2000 East, Salt Lake City, UT 84112; e-mail: Catherine.Staes@hsc.utah.edu.

SUPPORT

Supported by Hitachi, Ltd.

AUTHOR CONTRIBUTIONS

Conception and design: George Chalkidis, Jordan McPherson, Anna Beck, Shuntaro Yui, Catherine Staes

Collection and assembly of data: George Chalkidis, Jordan McPherson, Anna Beck, Michael Newman, Catherine Staes

Data analysis and interpretation: George Chalkidis, Jordan McPherson, Anna Beck, Catherine Staes

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 or 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](#)).

George Chalkidis

Employment: Hitachi, Ltd

Stock and Other Ownership Interests: Pfizer, GlaxoSmithKline, Medtronic, 3M, Johnson & Johnson/Janssen

Research Funding: Hitachi, Ltd

Patents, Royalties, Other Intellectual Property: Patents pending: US20190032126A1, US20210271924A1, US20200311612A1, US20210310951A1 (Inst)

Jordan McPherson

Research Funding: Hitachi, Ltd (Inst)

Shuntaro Yui

Employment: Hitachi, Ltd

Research Funding: Hitachi, Ltd

Patents, Royalties, Other Intellectual Property: Patents pending: US20190032126A1, US20210271924A1, US20200311612A1 (Inst)

Catherine Staes

Stock and Other Ownership Interests: Pfizer (I), Quest Diagnostics (I), Merck, CVS Health (I)

Research Funding: Hitachi, Ltd (Inst)

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

We acknowledge support for this work from Kensaku Kawamoto, MD, PhD, MHS, Associate Chief Medical Information Officer, University of Utah Health, and Andrew Post, MD, PhD, Research Informatics Director, Huntsman Cancer Institute.

REFERENCES

- Kim H, Keating NL, Perloff JN, et al: Aggressive care near the end of life for cancer patients in Medicare accountable care organizations. *J Am Geriatr Soc* 67:961-968, 2019
- Falchhook AD, Dusetzina SB, Tian F, et al: Aggressive end-of-life care for metastatic cancer patients younger than age 65 years. *J Natl Cancer Inst* 109:djx028, 2017
- Hui D, Nooruddin Z, Didwaniya N, et al: Concepts and definitions for "actively dying," "end of life," "terminally ill," "terminal care," and "transition of care": A systematic review. *J Pain Symptom Manage* 47:77-89, 2014
- Institute of Medicine: *Dying in America: Improving Quality and Honoring Individual Preferences Near the End of Life*. Washington, DC, The National Academies Press, 2015
- Schnipper LE, Smith TJ, Raghavan D, et al: American Society of Clinical Oncology identifies five key opportunities to improve care and reduce costs: The top five list for oncology. *J Clin Oncol* 30:1715-1724, 2012
- Choices for Care When Treatment May Not Be an Option. National Cancer Institute, 2018. *Advanced Cancer*. <https://www.cancer.gov/about-cancer/advanced-cancer/care-choices>
- Wasp GT, Alam SS, Brooks GA, et al: End-of-life quality metrics among medicare decedents at minority-serving cancer centers: A retrospective study. *Cancer Med* 9:1911-1921, 2020

8. Morden NE, Chang CH, Jacobson JO, et al: End-of-life care for Medicare beneficiaries with cancer is highly intensive overall and varies widely. *Health Aff (Millwood)* 31:786-796, 2012
9. Bestvina CM, Wroblewski KE, Daly B, et al: A rules-based algorithm to prioritize poor prognosis cancer patients in need of advance care planning. *J Palliat Med* 21:846-849, 2018
10. Lamont EB, Christakis NA: Physician factors in the timing of cancer patient referral to hospice palliative care. *Cancer* 94:2733-2737, 2002
11. Temel JS, Greer JA, Admane S, et al: Longitudinal perceptions of prognosis and goals of therapy in patients with metastatic non-small-cell lung cancer: Results of a randomized study of early palliative care. *J Clin Oncol* 29:2319-2326, 2011
12. Weeks JC, Cook EF, O'Day SJ, et al: Relationship between cancer patients' predictions of prognosis and their treatment preferences. *JAMA* 279:1709-1714, 1998
13. Mack JW, Weeks JC, Wright AA, et al: End-of-Life discussions, goal attainment, and distress at the end of life: Predictors and outcomes of receipt of care consistent with preferences. *J Clin Oncol* 28:1203-1208, 2010
14. Serious Illness Care. Ariadne Labs. <https://www.ariadnelabs.org/areas-of-work/serious-illness-care/>
15. Hui D: Prognostication of survival in patients with advanced cancer: Predicting the unpredictable? *Cancer Control* 22:489-497, 2015
16. Glare P, Virik K, Jones M, et al: A systematic review of physicians' survival predictions in terminally ill cancer patients. *BMJ* 327:195-198, 2003
17. Prigerson HG, Bao Y, Shah MA, et al: Chemotherapy use, performance status, and quality of life at the end of life. *JAMA Oncol* 1:778-784, 2015
18. Adelson K, Lee DKK, Velji S, et al: Development of imminent mortality predictor for advanced cancer (IMPAC), a tool to predict short-term mortality in hospitalized patients with advanced cancer. *J Oncol Pract* 14:e168-e175, 2018
19. Lage DE, Nipp RD, D'Arpino SM, et al: Predictors of posthospital transitions of care in patients with advanced cancer. *J Clin Oncol* 36:76-82, 2018
20. Ramchandran KJ, Shega JW, Von Roenn J, et al: A predictive model to identify hospitalized cancer patients at risk for 30-day mortality based on admission criteria via the electronic medical record: Assessment of mortality risk using the EMR. *Cancer* 119:2074-2080, 2013
21. Maltoni M, Scarpi E, Pittureri C, et al: Prospective comparison of prognostic scores in palliative care cancer populations. *Oncologist* 17:446-454, 2012
22. 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
23. Elfiky AA, Pany MJ, Parikh RB, et al: 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
24. 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
25. 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
26. 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
27. OncoMortality. alexandria Health. <https://apps.apple.com/us/app/oncomortality/id1445192538?platform=ipad>
28. Use Cases. usability.gov. <https://www.usability.gov/how-to-and-tools/methods/use-cases.html>
29. Utah Cancer Registry. <https://uofuhealth.utah.edu/utah-cancer-registry/>
30. Saini KS, Twelves C: Determining lines of therapy in patients with solid cancers: A proposed new systematic and comprehensive framework. *Br J Cancer* 125:155-163, 2021
31. Walker B, Boyd M, Aguilar K, et al: Comparisons of real-world time-to-event end points in oncology research. *JCO Clin Cancer Inform* 5:45-46, 2021
32. Makar M, Ghassemi M, Cutler DM, et al: Short-term mortality prediction for elderly patients using medicare claims data. *Int J Mach Learn Comput* 5:192-197, 2015
33. Martin L, Senesse P, Gioulbasanis I, et al: Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol* 33:90-99, 2015
34. Saeb S, Lonini L, Jayaraman A, et al: The need to approximate the use-case in clinical machine learning. *Gigascience* 6:1-9, 2017
35. Little MA, Varoquaux G, Saeb S, et al: Using and understanding cross-validation strategies. Perspectives on Saeb et al. *Gigascience* 6:1-6, 2017
36. Chen T, Guestrin C: XGBoost: A scalable tree boosting system, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD '16. New York, NY, ACM Press, 2016, pp 785-794
37. Subbiah IM, Xu C, Lu SC, et al: Development of patient reported outcomes-based machine learning algorithm for the six-month mortality prediction in patients with advanced cancer. *J Clin Oncol* 39, 2021 (28_suppl; abstr 273)
38. Lundberg SM, Lee SI: A unified approach to interpreting model predictions, in Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, Red Hook, NY, Curran Associates, Inc, pp 4768-4777
39. Kuhn M, Johnson K: Applied Predictive Modeling. New York, NY, Springer New York, 2013
40. Nicholls A: Confidence limits, error bars and method comparison in molecular modeling. Part 1: The calculation of confidence intervals. *J Comput Aided Mol Des* 28:887-918, 2014
41. DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 44:837-845, 1998
42. Collins GS, Reitsma JB, Altman DG, et al: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med* 162:55-63, 2015
43. HL7.org: HL7 FHIR release 4. <https://www.hl7.org/fhir/overview.html>
44. 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
45. Gianfrancesco MA, Tamang S, Yazdany J, et al: Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med* 178:1544-1547, 2018

