# Thirty-Day Unplanned Hospital Readmissions in Patients With Cancer and the Impact of Social Determinants of Health: A Machine Learning Approach

Nickolas Stabellini, BS[1,2,3,4]; Aziz Nazha, MD[5]; Nikita Agrawal, BA[6]; Merilys Huhn, BA[6]; John Shanahan, BA[7]; Nelson Hamerschlak, MD, PhD[8]; Kristin Waite, PhD[9]; Jill S. Barnholtz-Sloan, PhD[9,10]; and Alberto J. Montero, MD, MBA[2]

DOI https://doi.org/10.1200/CCI.22.00143

## ABSTRACT

**PURPOSE** Develop a cancer-specific machine learning (ML) model that accurately predicts 30-day unplanned readmissions in patients with solid tumors.

**METHODS** The initial cohort included patients 18 years or older diagnosed with a solid tumor. Two distinct cohorts were generated: one with and one without detailed social determinants of health (SDOHs) data. For each cohort, data were temporally partitioned in 70% (training), 20% (validation), and 10% (testing). Tree-based ML models were developed and validated on each cohort. The metrics used to evaluate the model's performance were receiver operating characteristic curve (ROC), area under the ROC curve, precision, recall (R), accuracy, and area under the precision-recall curve.

**RESULTS** We included 13,717 patients in this study in two cohorts (5,059 without SDOH data and 8,658 with SDOH data). Unplanned 30-day readmission occurred in 21.3% of the cases overall. The five main non-SDOH factors most highly associated with an unplanned 30-day readmission (R, 0.74; IQR, 0.58-0.76) were: number of previous unplanned readmissions; higher Charlson comorbidity score; nonelective index admission; discharge to anywhere other than home, hospice, or nursing facility; and higher anion gap during the admission. Neighborhood crime index, neighborhood median home values, annual income, neighborhood median household income, and wealth index were the main five SDOH factors important for predicting a high risk for an unplanned hospital readmission (R, 0.66; IQR, 0.56-0.72). The models were not directly comparable.

**CONCLUSION** Key drivers of unplanned readmissions in patients with cancer are complex and involve both clinical factors and SDOH. We developed a cancer-specific ML model that with reasonable accuracy identified patients with cancer at high risk for an unplanned hospital readmission.

ACCOMPANYING CONTENT

Data Supplement

Accepted May 24, 2023
Published July 18, 2023

JCO Clin Cancer Inform
7:e2200143




## INTRODUCTION

Hospital readmissions are an important measure of health care quality.[1,2] In 2005, the WHO defined reducing hospital readmission rates as a top strategic health care priority.[3] In the United States, the Affordable Care Act's Hospital Readmission and Reduction Program, which applies financial penalties for excess hospital readmissions of Medicare patients, made the reduction of unplanned hospital admissions a health care priority.[4,5] Unplanned hospital readmissions are costly and are estimated to cost the US health care system approximately 15-20 billion dollars annually.[6,7] Preventing avoidable readmissions therefore has the potential to profoundly both improve patient quality of life and help reduce hospital costs.[1,6]

Inpatient costs are one of the top drivers of cancer patient health care costs, given that patients with cancer have higher hospitalization and unplanned readmission rates than the general population.[8-12] Unplanned hospital readmission rates in patients with cancer have been reported as high as 30% and are known to be influenced by many key drivers such as underlying malignancy, presence of metastatic disease, cancer therapies received, underlying comorbidities, previous emergency department visits and/or hospitalizations, and socioeconomic conditions.[13-16] In addition to the financial burden on the health care system, high rates of

## CONTEXT

**Key Objective**

To develop a cancer-specific machine learning (ML) model to predict 30-day unplanned readmissions in patients with solid tumors.

**Knowledge Generated**

Independent of the cancer type and stage, the key clinical drivers of unplanned 30-day hospital readmissions in our model included: intensity of health care utilization, underlying comorbidities, hospitalization-related information, and specific laboratory values during the index admission.

**Relevance**

Unplanned hospital readmissions are costly and are estimated to cost the US health care system billion dollars annually. Patients with cancer have higher hospitalization and unplanned readmission rates than the general population. We developed a ML algorithm that was able to predict 30-day unplanned hospital readmissions in patients with solid tumor. The use of a cancer-specific ML model can assist providers in identifying patients at high risk for an unplanned hospital readmission to improve discharge planning and equitable allocation of additional health care resources to patients at higher risk of hospital readmission.



unplanned hospital readmissions represent an additional burden on patients with cancer and their families.[17]

Machine learning (ML)–based models have been explored as a tool to reduce hospital readmissions, given the transition to electronic medical records (EMRs) and the increasing amount of digital health care data.[18,19] Multiple ML models have already been developed and applied to predict unplanned readmissions in patients from the general population, with better predictive ability than the LACE index and the HOSPITAL score of widely used conventional models on the basis of a limited set of easily available data.[20-23] However, general hospital readmission ML models have little predictive ability in the setting of patients with cancer because of many factors related to the underlying disease and treatments that they receive that are not accounted for in these general readmission models. To date, few, if any, cancer-specific ML readmission models have been tested, with most being nonexplainable and nonreproducible as they are based on data that are not readily available across health systems.[20,24] Therefore, the primary objective of this work is to develop and test two distinct (and noncomparable) explainable cancer-specific ML models capable of predicting 30-day unplanned readmissions in patients with solid tumor: one model that takes into account only demographic and medical data without any social determinant of health (SDOH) and another that also includes detailed patient-level SDOH.

## METHODS

### Data Source

The study setting was the University Hospitals Seidman Cancer Center (UHSCC). All patient data were obtained from the UHSCC data repository on the basis of the CAISIS platform, which consists of an open-source, web-based cancer data management system that integrates disparate sources of cancer patient data (ie, Soarian, NGS Labs, Sunrise Clinical Manager, Tumor Registry, Via Oncology, OnCore, MosaiQ, PRO tools, and others).[25-27] All patient records were deidentified, and the study was approved by the University Hospitals (UH) of Cleveland Institutional Review Board. All the data were complemented with electronic health record (EHR) information captured via Electronic Medical Record Search Engine to obtain the most accurate and complete information per patient.[28]

The initial cohort included patients 18 years and older diagnosed with a solid tumor (Data Supplement [Table 1]) between January 1, 2005, and March 31, 2022, admitted at UHSCC. Patients were excluded from the analysis if they had a cancer diagnosis before 2015, an unknown diagnosis date, unknown age at diagnosis, or no sex recorded in the EHR.

### Outcome

The primary outcome was a 30-day unplanned hospital readmission, defined as a nonelective readmission (therefore, a hospitalization categorized as trauma, emergency, or of an urgent nature in the EHR) that occurred within 30 days from the last discharge.[29]

### Covariates

The information retrieved for each admission (Data Supplement [Table 1]) included: system admission details, patient's demographics, patient's medical history, and patient's individual SDOH. The admission details included admission type, admission reason, admission date, discharge date, discharge site, payer, vitals, laboratory values, and medications administered.

Patient's demographics included: age at initial cancer diagnosis, sex, race, ethnicity, cancer type, and diagnosis date. Patient's medical history included smoking status, Charlson comorbidity index, cancer treatment, treatment dates, TNM stage group, surgeries, surgery dates, emergency department visits, emergency department visit dates, medications prescribed, previous hospital admissions, previous 30-day unplanned readmissions, and percent of appointments attended.[30,31]

Patient's individual SDOH were obtained from LexisNexis, a provider of legal, government, business, and high-tech information sources, and included data on the following main domains: economic stability, education access and quality, neighborhood and built environment, and social and community context.[32,33] The LexisNexis data set includes a combination of all adult patients discharged from a UH facility over a 2.5-year timeframe and all adult patients who are members of the Accountable Care Organization.[34] This data set is composed of a combination of multiple public and private records that are updated in different frequencies. The data obtained referred to the most current records available.

## Cohorts

Two distinct cohorts were generated (Fig 1). The first (cohort without SDOH data) included patients who did not have information regarding SDOH, aiming to develop an algorithm replicable in other institutions given the difficulty of obtaining these data. The second (cohort with SDOH data) included patients who had information regarding SDOH to analyze its impact on the outcome of interest. The models developed on these distinct cohorts were not designed to be comparable in terms of performance change with the addition of SDOH.

## Population Description

The data were presented as absolute values and percentages for categorical variables and as median and quartiles for continuous variables and stratified according to the occurrence of 30-day unplanned readmission. The Pearson chi-square test was used to compare categorical variables. Data distribution assumptions for continuous variables were confirmed using histograms and the Kolmogorov-Smirnov test, followed by Student's t-tests for normally distributed factors and nonparametric Kruskal-Wallis tests for non-normal distributed factors.

## Preprocessing

### *Deduplication*

Patients with multiple records were selected by the most recent admission or the most recent one that generated an unplanned 30-day readmission. We selected the records by patient and not by admission to avoid the demographic characteristics of high utilizers (patients frequently hospitalized) to bias the models.[35] High utilizers were accounted



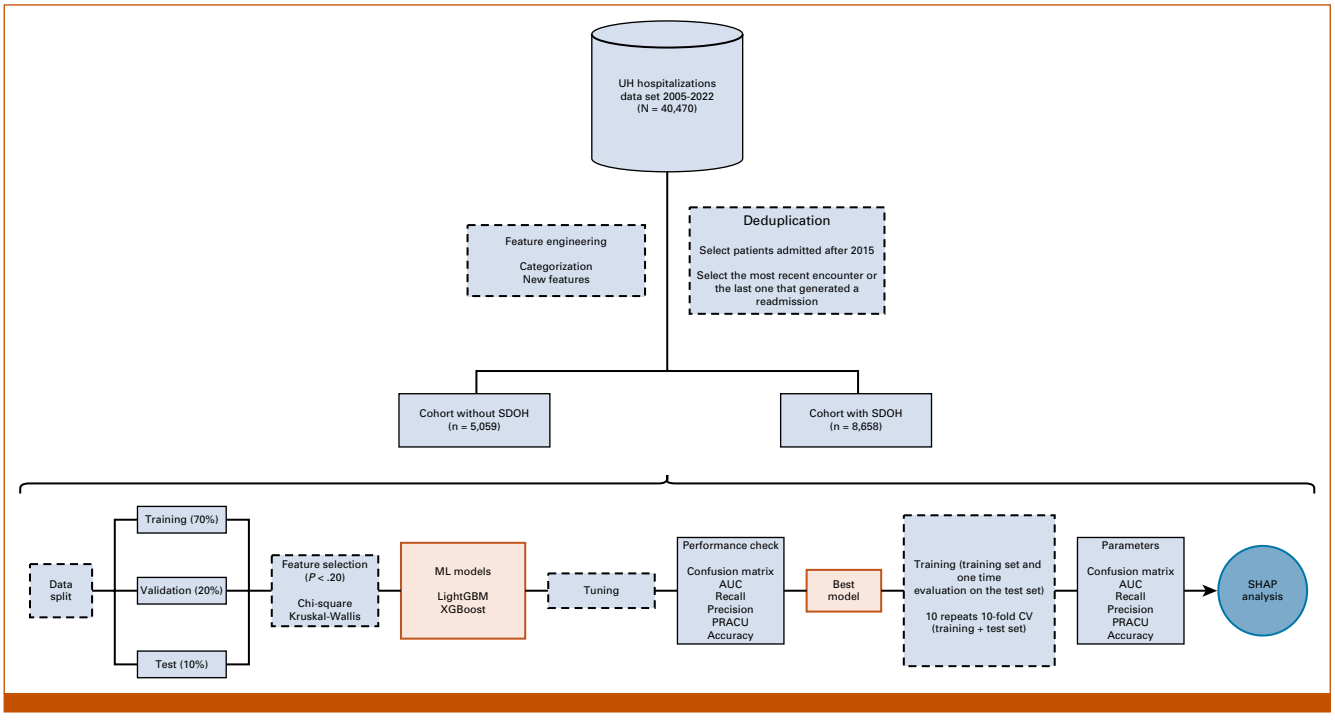


**FIG 1.** Machine learning framework applied in the study. Two distinct cohorts were created (with and without social determinants of health data), data were temporally partitioned within cohort, and the model was submitted to a 10-times 10-fold cross-validation. Recall was the key performance metric. SHAP provided explainability. AUC, area under the ROC curve; CV, cross-validation; ML, machine learning; PRAUC, area under the precision-recall curve; ROC, receiver operating characteristic curve; SDOH, social determinants of health; SHAP, SHapley Additive exPlanations; UH, University Hospital.

by the creation of variables for previous admissions and previous 30-day unplanned readmissions according to the admission index.

### *Data Split*

For each cohort, data were partitioned in 70% for training, 20% for validation (tuning), and 10% for testing. To replicate a real-world data collection, partitioning was performed temporally.[36] The readmissions rate was similar in all the partitions, and oversampling was applied to deal with the imbalanced outcome.

### *Feature Engineering*

Vitals were categorized as low, normal, or high according to current guidelines, and minimum and maximum values for each admission were obtained.[37,38] Surgeries, medications, and emergency department visits were accounted for the past 30, 60, 90, 180, and 360 days. Admission's International Classification of Diseases (ICDs) was converted to the ninth revision (ICD-9) and subsequently categorized according to the respective chapters. In addition, variables with a high number of missingness (≥20%) were excluded from the data set.

### *Feature Selection*

Features were selected according to the results of comparisons (from the Pearson chi-square test and Kruskal-Wallis test) between groups with and without an unplanned 30-day readmission. The variables selected were those that achieved $P < .20$. SDOH variables were included despite the $P$ value (Fig 1 and Data Supplement [Table 3]).

## ML Modeling

We tested tree-based algorithms LightGBM and XGBoost (Fig 1), which showed good performance results in previous predictive studies and are designed to handle missing data, an important characteristic for EHR-based data.[39,40] For each cohort, the models were trained on the training set and were further developed and tuned using the validation set. Hyperparameter tuning (model enhancement via parameters adjustments) with 10-fold cross-validation (CV) was optimized to the recall metric, aiming to develop an algorithm with the ability to capture as many true positives (TP) as possible. Subsequently, models were evaluated using 10 times 10-fold CV, on a combination of training and validation data sets. On the basis of these validation results, a final model was selected from the XGBoost and LightGBM models and its performance was evaluated on the testing data set. LASSO regressions were performed on each cohort as a reference model without the increased complexity to compare the improvement in the performance achieved by the use of ML. Correlation values, a supplemental model (without SDOH data) on the cohort with SDOH data available, and a model with only SDOH covariates were included to analyze whether the SDOH data are correlated with other clinical characteristics or if they are independently informative. The LACE index was calculated for each cohort to compare the improvement in the performance achieved by the use of a cancer-specific model in patients with cancer.

## Performance Metrics and Explainability

The metrics used to evaluate the model performance for each cohort were TP, true negatives, false positives (FP), false negatives, receiver operating characteristic curve (ROC), area under the ROC curve (AUC), precision, recall, accuracy, area under the precision-recall curve, and calibration plots (using Platt scaling) measured on the validation and test cohorts. The results from the 10 times 10-fold CV were represented as median and IQR. The variables associated with the outcome and its importance scores were calculated and represented via SHapley Additive exPlanations results from the training set.[41] Adversity cutoffs for continuous variables were defined according to partial dependence plots. All the models were optimized for recall and followed the TRIPOD development and validation checklist.[42]

## Software and Packages

All the analyses were performed using RStudio software (POSIT, Boston, MA).[43] The packages touch and icdcoder were used for ICD conversion and categorization.[44-46] The ML models were developed with the package mlr3.[47]

## Ethical Approval Information

Patient records were deidentified, and the study was approved by the UH of Cleveland Institutional Review Board.

# RESULTS

## Population

We included 13,717 patients. The median age for the entire cohort was 68 years (IQR, 59-76). Patients were predominantly White (73.4%), with slightly more women (52.7%) than men. The median Charlson comorbidity score was 6 (IQR, 3-9). The most common solid tumor diagnosis was breast cancer (21.8%); approximately one third (33.6%) of patients in the entire cohort had stage I disease, and 11% had stage IV disease. Radiotherapy was performed in 22.5% of the cohort, whereas 17.9% received chemotherapy, 4.6% received immunotherapy, and 11.7% received endocrine therapy. Admissions were mostly emergency (57.4%). Most patients were discharged home (67.9%), and the median length of stay was 3 days (IQR, 2-6). The unplanned 30-day readmission rate in the entire cohort was 21.3%.

## Cohort Without SDOH Data

The cohort without available SDOH data included 5,059 patients (Table 1) with an unplanned 30-day readmission

**TABLE 1.** Descriptive Statistics for the Two Study Cohorts: With and Without SDOH

| | Cohort Without SDOH Data (n = 5,059) | | | Cohort With SDOH Data (n = 8,658) | | |
|---|---|---|---|---|---|---|
| | Unplanned Readmission | No Unplanned Readmission | | Unplanned Readmission | No Unplanned Readmission | |
| Covariate | 1,057 (20.9%) | 4,002 (79.1%) | *P* | 1,869 (21.6%) | 6,789 (78.4%) | *P* |
| Age at diagnosis, median (IQR) | 68 (59-77) | 68 (59-77) | .63 | 68 (59-76) | 68 (59-76) | .87 |
| Sex, No. (%) | | | | | | |
| Male | 515 (48.7) | 1,898 (47.4) | .47 | 950 (50.8) | 3,129 (46.1) | <.001 |
| Race, No. (%) | | | | | | |
| White | 721 (68.2) | 2,976 (74.4) | <.001 | 1,314 (70.3) | 5,052 (74.4) | .001 |
| Black | 283 (26.8) | 851 (21.3) | | 468 (25) | 1,453 (21.4) | |
| Others | 53 (5) | 175 (4.4) | | 87 (4.7) | 284 (4.2) | |
| Ethnicity, No. (%) | | | | | | |
| Non-Hispanic | 1,034 (97.8) | 3,884 (97.1) | .04 | 1,843 (98.6) | 6,609 (97.3) | .004 |
| Smoking status, No. (%) | | | | | | |
| Yes | 141 (13.3) | 462 (11.5) | <.001 | 272 (14.6) | 801 (11.8) | <.001 |
| Former smoker | 371 (35.1) | 1,229 (30.7) | | 707 (37.8) | 2,067 (30.4) | |
| Never smoker | 339 (32.1) | 1,274 (31.8) | | 568 (30.4) | 2,245 (33.1) | |
| Unknown | 206 (19.5) | 1,037 (25.9) | | 322 (17.2) | 1,676 (24.7) | |
| Charlson comorbidity score, median (IQR) | 8 (6-10) | 6 (3-8) | <.001 | 8 (6-10) | 6 (3-8) | <.001 |
| Cancer site, No. (%) | | | | | | |
| Bladder | 86 (8.1) | 228 (5.6) | <.001 | 163 (8.7) | 432 (6.3) | <.001 |
| Lung | 146 (13.8) | 571 (14.2) | | 236 (12.6) | 959 (14.1) | |
| Brain/CNS | 137 (12.9) | 417 (10.4) | | 228 (12.2) | 525 (7.7) | |
| Stomach | 12 (1.1) | 49 (1.2) | | 21 (1.1) | 78 (1.1) | |
| Liver/intrahepatic bile duct | 78 (7.3) | 205 (5.1) | | 105 (5.6) | 268 (3.9) | |
| Colon | 94 (8.8) | 390 (9.7) | | 200 (10.7) | 715 (10.5) | |
| Esophagus | 35 (3.3) | 113 (2.8) | | 50 (2.6) | 118 (1.7) | |
| Kidney/other urinary organs | 60 (5.6) | 254 (6.3) | | 94 (5) | 409 (6) | |
| Pancreas | 37 (3.5) | 95 (2.3) | | 39 (2) | 133 (1.9) | |
| Breast | 198 (18.7) | 840 (13.5) | | 362 (19.3) | 1,594 (23) | |
| Prostate | 99 (9.3) | 541 (13.5) | | 39 (2) | 1,090 (16) | |
| Thyroid | 52 (4.9) | 189 (4.7) | | 100 (5) | 278 (4) | |
| Melanoma | 13 (1.2) | 51 (1.2) | | 28 (1.5) | 119 (1.7) | |
| Rectum/anus | 10 (0.9) | 59 (1.4) | | 18 (0.9) | 71 (1) | |
| Stage, No. (%) | | | | | | |
| 0 | 0 (4.4) | 89 (5.2) | .3 | 36 (5.8) | 201 (7.3) | .02 |
| I | 108 (28) | 559 (32.9) | | 216 (34.8) | 954 (34.6) | |
| II | 115 (29.8) | 481 (28.3) | | 179 (28.9) | 823 (29.9) | |
| III | 86 (22.3) | 345 (20.3) | | 111 (17.9) | 540 (19.6) | |
| IV | 60 (15.5) | 225 (13.2) | | 78 (12.6) | 238 (8.6) | |

(continued on following page)

**TABLE 1.** Descriptive Statistics for the Two Study Cohorts: With and Without SDOH (continued)

| | Cohort Without SDOH Data (n = 5,059) | | | Cohort With SDOH Data (n = 8,658) | | |
|---|---|---|---|---|---|---|
| | Unplanned Readmission | No Unplanned Readmission | | Unplanned Readmission | No Unplanned Readmission | |
| Covariate | 1,057 (20.9%) | 4,002 (79.1%) | *P* | 1,869 (21.6%) | 6,789 (78.4%) | *P* |
| Radiotherapy 30 days before admission, No. (%) | 56 (5.3) | 114 (2.8) | <.001 | 78 (4.2) | 138 (2) | <.001 |
| Chemotherapy 30 days before admission, No. (%) | 24 (2.3) | 46 (1.1) | .008 | 30 (1.6) | 55 (0.8) | .003 |
| Immunotherapy 30 days before admission, No. (%) | 11 (1) | 11 (0.3) | .001 | 8 (0.4) | 16 (0.2) | .24 |
| Endocrine therapy 30 days before admission, No. (%) | 2 (0.2) | 13 (0.3) | .68 | 5 (0.3) | 15 (0.2) | .6 |
| Surgeries in the year before admission, median (IQR) | 2 (1-3) | 2 (1-3) | .05 | 2 (1-3) | 2 (1-3) | .75 |
| ED visits in the year before admission, median (IQR) | 1 (1-2) | 1 (1-2) | <.001 | 1 (1-2) | 1 (1-2) | <.001 |
| % appointments attended, median (IQR) | 50 (0-75) | 62.5 (16-81) | <.001 | 62.5 (42-76) | 70 (50-84) | <.001 |

NOTE. Groups were compared according to the outcome 30-day unplanned readmission status.
Abbreviation: SDOH, social determinants of health.

rate of 20.9%. Patients with an unplanned hospital readmission when compared with patients without any unplanned readmissions were significantly more likely to be Black (26.8% *v* 21.3%), have a smoking history (13.3% *v* 11.5%; $P < .001$), and have a higher median Charlson comorbidity score (8 [IQR, 6-10] *v* 6 [IQR, 3-8]; $P < .001$). Higher rates of radiotherapy (5.3% *v* 2.8%; $P < .001$) and immunotherapy (1% *v* 0.3%; $P = .001$) 30 days before the admission and lower appointment completion rates (median of 50 [IQR, 0-75] *v* 62.5 [IQR, 16-81]; $P < .001$) were also noted among patients with a 30-day unplanned readmission.

Patients with unplanned readmissions (Table 2) were more likely to have emergency admissions (74.5% *v* 50.7%; $P < .001$), less likely to be discharged home (56.8% *v* 66.1%; $P < .001$), have a higher number of previous admissions (median of 1 [IQR, 0-2] *v* 0 [IQR, 0-1]; $P < .001$), have a higher median length of stay (4 [IQR, 2-7] *v* 3 days [IQR, 2-6]; $P < .001$), and have received a significantly higher median number of different medications during admission (46 [IQR, 23-85] *v* 37 [IQR, 17-66]; $P < .001$).

LightGBM was the ML method that performed best on this cohort (Data Supplement [Table 4]). The final model achieved an AUC of 0.80 (IQR, 0.75-0.84) and a recall of 0.74 (IQR, 0.58-0.76; Fig 2; Data Supplement [Table 5 and Fig 1]). The five clinical factors most highly correlated with an unplanned 30-day readmission included a higher number of previous 30-day unplanned readmissions; higher Charlson comorbidity score; nonelective admission; discharge disposition other than home, hospice, or nursing facility; and higher anion gap during the admission (Fig 3; Data Supplement [Table 6]). Adversity cutoffs were defined as >1 for previous 30-day unplanned readmissions, a Charlson comorbidity score of >5, a anion gap during admission of >15 mmol/L, and a hematocrit of <31% (Data Supplement [Fig 1]). The LASSO regression for this cohort achieved a recall of 0.51 (95% CI, 0.47 to 0.49).

### Cohort With SDOH Data

The cohort with SDOH data included 8,658 patients, with an unplanned readmission rate of 21.6%—similar to 20.9% previously described for the cohort without SDOH. Both SDOH and non-SDOH cohorts were similar with respect to patient characteristics (Tables 1 and 2).

In the SDOH cohort (Data Supplement [Table 2]), readmitted patients were significantly more likely to be unmarried (79% *v* 75.9%; $P = .004$), live in a household with fewer members ($P < .001$), rent (46.3% *v* 43%; $P = .004$), and have a lower household income (31.2% *v* 28.8%; $P = .05$).

LightGBM was the ML method that performed best on the SDOH cohort (Data Supplement [Table 4]). The final model achieved an AUC of 0.78 (IQR, 0.76-0.83) and a recall of 0.66 (IQR, 0.56-0.72; Fig 2; Data Supplement [Table 5 and Fig 1]). The five most highly predictive SDOH factors in this model for unplanned hospital readmissions were: neighborhood crime index, neighborhood median home values, annual income, neighborhood median household income, and wealth index (Fig 3; Data Supplement [Table 6]). The adversity cutoffs for the top three SDOH factors were >115, <$375,000 in US dollars (USD), and <$85,000 USD (Data Supplement [Fig 2]). The LASSO regression for this cohort achieved a recall of 0.44 (95% CI, 0.37 to 0.39). The supplemental model without SDOH data on this population achieved an AUC of 0.79 (IQR, 0.75-0.83) and a recall of 0.72 (IQR, 0.57-0.72; Data Supplement [Tables 4, 5 and Fig 3]). Among the five most highly predictive factors in this model for unplanned hospital readmission (Data Supplement [Table 6 and Fig 3]), four (higher number of previous 30-day unplanned readmissions, higher Charlson comorbidity score, nonelective admission, and discharge disposition other than home, hospice, or nursing facility) were similar to the top five highly predictive factors from the model with SDOH data. The model with only SDOH covariates achieved an AUC of 0.50 (IQR, 0.47-0.56) and a recall of 0.85 (IQR, 0.56-0.74; Data Supplement [Table 8]).

### LACE Index

Among patients with SDOH data, the LACE index achieved an AUC of 0.61 (IQR, 0.60-0.68) and a recall of 0.30 (IQR, 0.02-0.89; Data Supplement [Table 9]). In the cohort without SDOH data, the LACE index achieved an AUC of 0.65 (IQR, 0.60-0.70) and a recall of 0.91 (IQR, 0.79-0.91; Data Supplement [Table 8]).

## DISCUSSION

This study aimed to develop a cancer-specific ML algorithm for predicting 30-day unplanned hospital readmissions in patients with solid tumors and to evaluate the impact of SDOH using distinct and noncomparable models. Validated readmission prediction models are used in health care, such as the LACE index and HOSPITAL scores.[20-23] However, these models do not perform well in patients with cancer, as shown recently in a study by Jones et al,[48] where both LACE and HOSPITAL underperformed in patients with colon and lung cancers relative to the general population. This highlights the importance and the need for cancer-specific models.[48]

Independent of the cancer type and stage, the key clinical drivers of unplanned 30-day hospital readmissions in our model included: intensity of health care utilization (number of previous 30-day unplanned readmissions, number of previous admissions, and appointment completion rates), underlying comorbidities (Charlson score and BMI), hospitalization-related information (admission type, discharge type, and number of medications administered), and specific laboratory values during the index admission (anion gap, hematocrit, platelets, minimum corpuscular volume, sodium, creatinine, hemoglobin, and bicarbonate). With respect to SDOH, patients with unplanned readmissions had lower social and community

**TABLE 2.** Admission Variables' Descriptive Statistics for the Two Study Cohorts: With and Without SDOH Data

| | Cohort Without SDOH Data (n = 5,059) | | | Cohort With SDOH Data (n = 8,658) | | |
|---|---|---|---|---|---|---|
| | Unplanned Readmission | No Unplanned Readmission | | Unplanned Readmission | No Unplanned Readmission | |
| Covariate | 1,057 (20.9%) | 4,002 (79.1%) | *P* | 1,869 (21.6%) | 6,789 (78.4%) | *P* |
| Admission type, No. (%) | | | | | | |
| Emergency | 787 (74.5) | 2,029 (50.7) | <.001 | 1,371 (73.4) | 3,691 (54.4) | <.001 |
| Elective | 143 (13.5) | 1,586 (39.6) | | 299 (16) | 2,580 (38) | |
| Urgent | 124 (11.7) | 358 (8.9) | | 182 (9.7) | 429 (6.3) | |
| Trauma | 3 (0.3) | 29 (0.7) | | 17 (0.9) | 89 (1.3) | |
| Discharge site, No. (%) | | | | | | |
| Home | 600 (56.8) | 2,644 (66.1) | <.001 | 1,064 (56.9) | 5,012 (73.8) | <.001 |
| Nursing facility | 177 (16.7) | 579 (14.5) | | 286 (15.3) | 750 (11) | |
| Hospice | 35 (3.3) | 291 (7.3) | | 76 (4.1) | 376 (5.5) | |
| Others | 245 (23.2) | 488 (12.2) | | 443 (23.7) | 651 (9.6) | |
| Payer, No. (%) | | | | | | |
| Medicare | 430 (40.7) | 1,638 (40.9) | .01 | 708 (37.9) | 1,515 (37) | .28 |
| Medicaid | 31 (2.9) | 64 (1.6) | | 34 (1.8) | 94 (1.4) | |
| Other | 596 (56.4) | 2,300 (57.5) | | 1,127 (60.3) | 4,180 (61.6) | |
| No. of previous admissions, median (IQR) | 1 (0-2) | 0 (0-1) | <.001 | 1 (0-3) | 0 (0-1) | <.001 |
| No. of previous 30-day unplanned readmissions, median (IQR) | 0 (0-1) | 0 (0-0) | <.001 | 0 (0-1) | 0 (0-0) | <.001 |
| Length of stay in days, median (IQR) | 4 (2-7) | 3 (2-6) | <.001 | 4 (2-7) | 3 (1-5) | <.001 |
| Medications administered during admission, median (IQR) | 46 (23-85) | 37 (17-66) | <.001 | 46 (24-86) | 36 (18-63) | <.001 |

NOTE. Within-cohort comparisons are shown by 30-day unplanned readmission status.
Abbreviation: SDOH, social determinants of health.

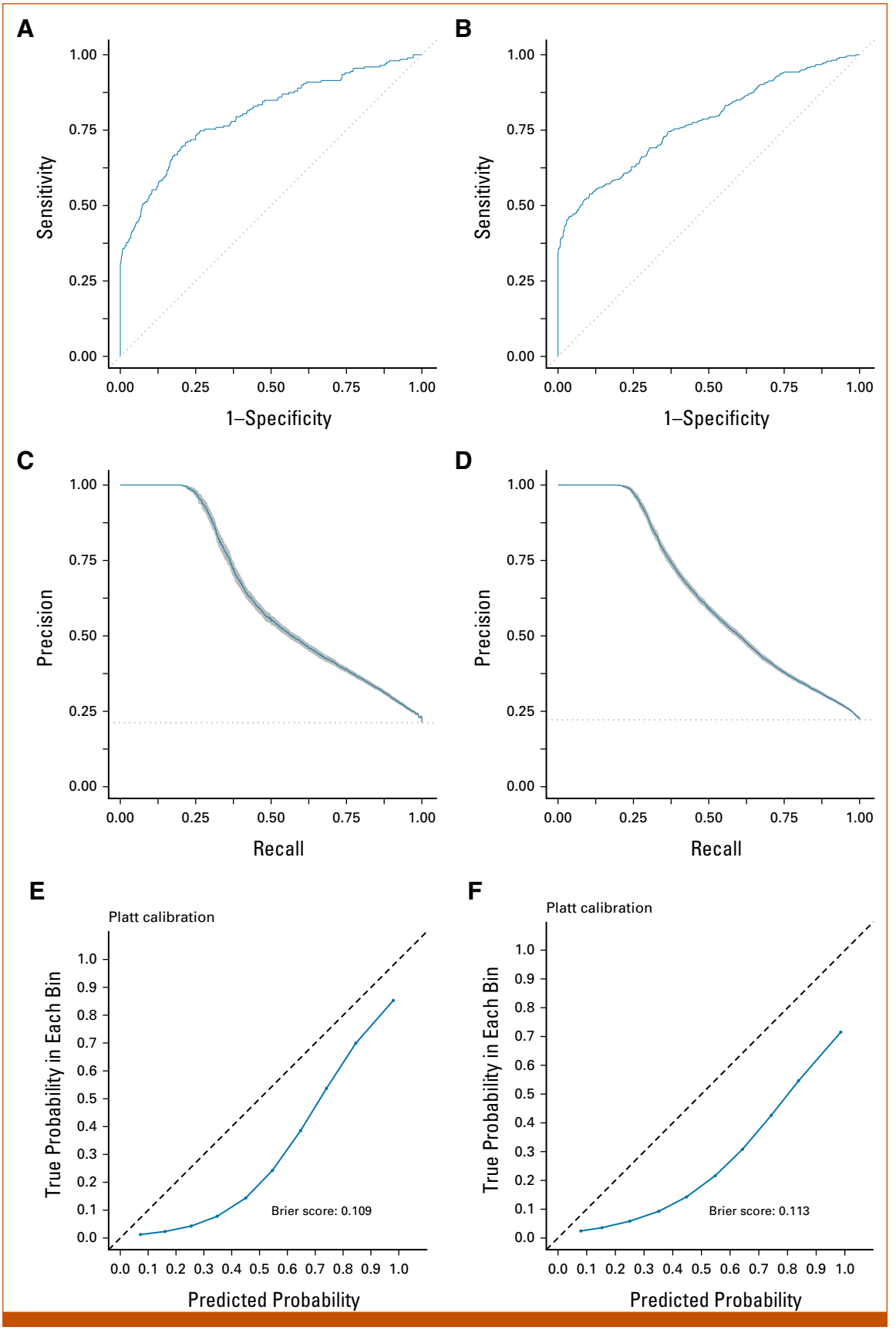


**FIG 2.** ROC, precision-recall, and calibration curves for the machine learning models developed. (A) ROC curve for the cohort without SDOH data; (B) ROC curve for the cohort with SDOH data; (C) precision-recall curve for the cohort without SDOH data; (D) precision-recall curve for the cohort with SDOH data; (E) calibration curve for the cohort without SDOH data; (F) calibration curve for the cohort with SDOH data. ROC, receiver operating characteristic curve; SDOH, social determinants of health.

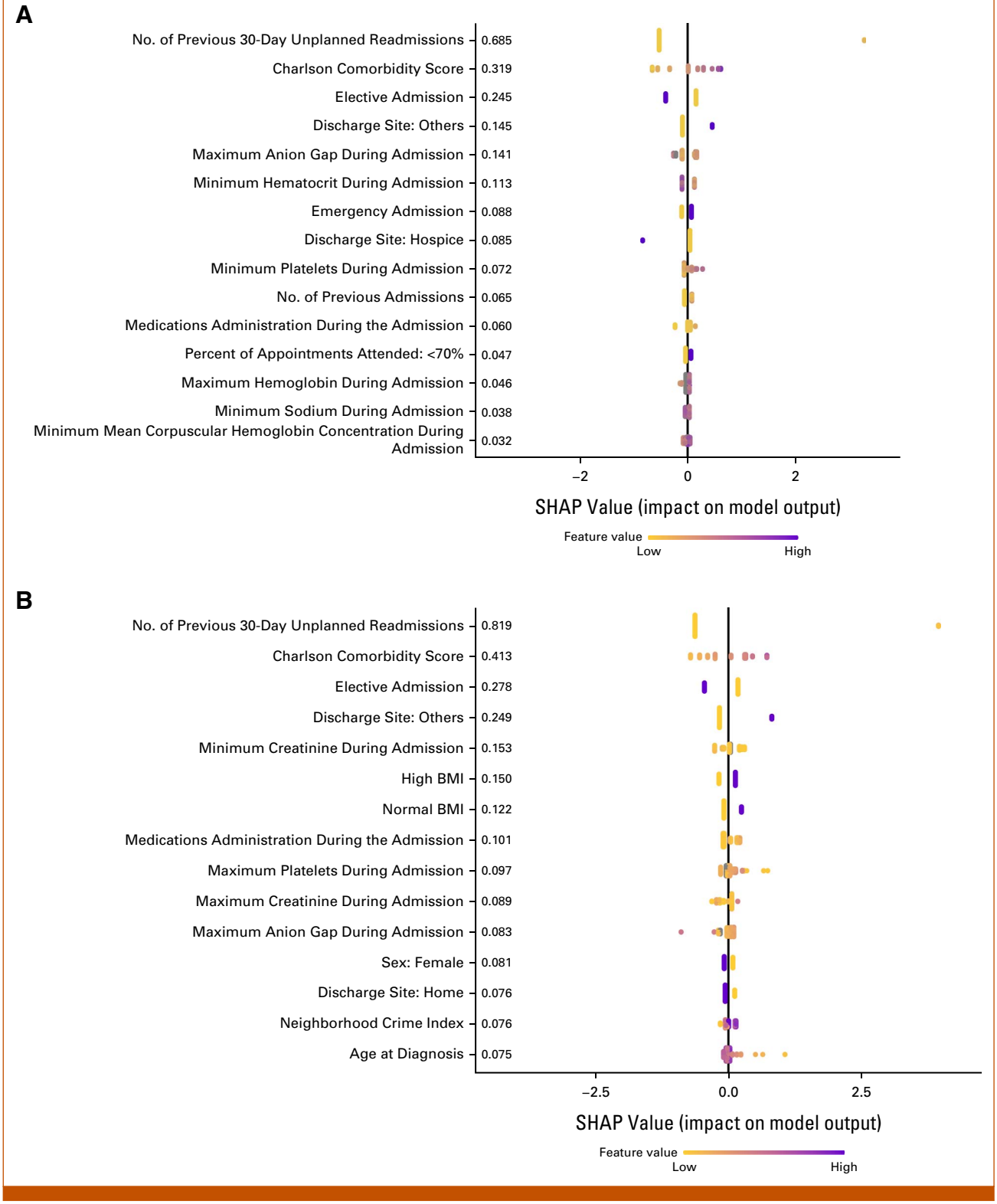


**FIG 3.** SHAP analysis from the cohorts (A) without and (B) with social determinants of health data highlighting the top 15 predictors of 30-day unplanned readmission in patients with cancer. Positive SHAP values (on the right of the line) favor the outcome. The color gradient indicates the feature values/categories. SDOH, social determinants of health; SHAP, SHapley Additive exPlanations.

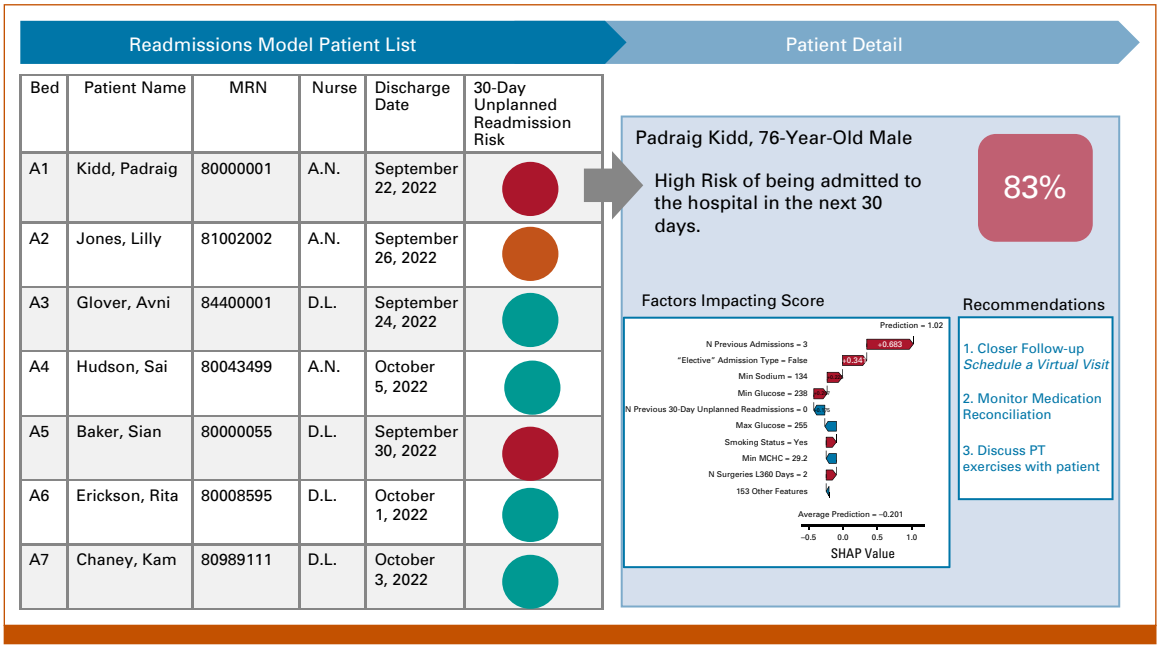


**FIG 4.** Suggested template/framework for the integration of the model with real-time EHR data. EHR, electronic health record; MCHC, mean corpuscular hemoglobin concentration; MRN, medical record number; PT, physical therapy; SHAP, SHapley Additive exPlanations.



support, lived in worse neighborhoods, and had lower economic stability.

Patients with unplanned hospital readmissions were more likely to have an advanced cancer stage and more likely to have received systemic oncologic therapy in the 30 days before the index admission.

Understanding key drivers of unplanned hospital readmissions in patients with cancer has been a topic of interest; however, the multiple factors involved have made applying traditional statistical methods difficult.[49] ML provides an alternative method for better understanding unplanned cancer readmissions, given its ability to handle multiple variables and to incorporate explainability via tree-based methods, which can offer crucial perspectives into clinical decision making applications.[20,49-51] In the published literature, one approach with respect to unplanned cancer readmissions has been to use logistic regression models.[52] Another study tested different models at various time points throughout the inpatient visit and found that the predictive performance was best at discharge, showing the importance of hospitalization-related variables.[49] More recently, a model to predict 30-day readmission among oncology patients was also developed.[53] All these, however, considered all-cause readmission. The approaches developed in our study, besides having equal or superior performance to the abovementioned models, are developed for unplanned readmissions. In addition, our model developed in the cohort without SDOH was built with commonly available variables, to facilitate the model's applicability to other health institutions.

Previous studies have shown that among patients with cancer, unplanned readmissions are dependent on multiple clinical/demographic variables including age, race, ethnicity, insurance status, underlying comorbidities, cancer type and stage (particularly metastatic pancreatic and non–small-cell lung cancers), history of depression, length of hospital stay, and specific medical/surgical interventions received.[53-55] In older patients with cancer, in addition to these factors, the number of previous chemotherapy lines and abnormal laboratory values at the time of discharge (such as sodium, albumin, and hemoglobin), admission type, and discharge disposition have also been shown to confer a higher risk of unplanned readmissions.[55,56] Moreover, the acute care literature shows that patients receiving cancer treatment (especially chemotherapy) and routine care are more likely to be hospitalized and/or to experience multiple visits and that laboratory values (including creatinine clearance, calcium level, white blood count, and platelet count) and poor coordinated care are linked with higher risk for hospitalizations.[57-59]

Our study recapitulates many of the results of these previous studies as they underscore the importance of underlying comorbidities, laboratory values, admission types, and discharge disposition. On the other hand, our findings differ from these previous studies as our results show that although demographic factors, such as age, race, and ethnicity, and cancer type, are important, in our ML model, they were not identified as the primary factors driving 30-day unplanned cancer readmissions. Our results highlight the impact of the number of medications that patients received while in the hospital, and the impact of individual patient health care utilization—both high (multiple ED visits, previous hospital admissions) and low utilization (low outpatient appointment completion rates).[60]

Our results also illustrate the value of SDOH in predicting unplanned hospital readmissions in patients with cancer. According to the WHO, SDOH are the conditions or circumstances in which people are born, grow, live, work, and age.[61] Low levels of social support, racial segregation, and education-related deaths in the United States in 2000 were comparable with the rates of lung cancer, myocardial infarction, and cerebrovascular illness, respectively.[62] With respect to cancer readmissions, previous studies have demonstrated the importance of marital and socioeconomic status.[53,54] In addition to these, we identified the importance of other SDOH, such as neighborhood and built environment factors, on the risk of unplanned 30-day cancer readmissions. Interestingly, the supplemental model without the use of SDOH data developed in the population with SDOH data available achieved better performance metrics than the model with SDOH data, probably because of the complexity added to the model when these data are added. However, this does not suggest that SDOH do not play an important role in unplanned readmissions in patients with cancer.

From a clinical perspective, identifying patients at high risk of readmissions opens up the possibility for targeted interventions, such as outpatient follow-up within 5 days of discharge (which has been shown to lower unplanned cancer readmissions and, in one study, estimated to have saved the health care system several millions of dollars), use of remote patient monitoring, or use of patient-reported outcomes guides such as Symptom Tracking and Reporting.[16,59,63,64] Our ML cancer readmission model validated in our cohort with no SDOH information available was based on commonly obtainable clinical variables available at the time of hospital discharge, which can easily estimate cancer readmission risk on an individual patient basis before discharge, and could be used by inpatient teams as a population health tool to target high-risk patients with specific interventions that could prevent unplanned hospital readmissions (Fig 4).

This study has several limitations. Our prediction models were not subjected to a third-party validation. The institutional database is EMR-based, and some of the information may be incomplete. The selection criteria for the cohort with SDOH data from LexisNexis might have led to selection bias. The outcome has a large imbalance between the readmissions and no readmissions, and the selection of the records on the basis of the outcome can inflate the model performance in the retrospective set. However, to minimize bias we applied CV methods. We also aimed to ensure that all the sets had similar rates of readmissions, as well as created variables to account for patients with a much higher than average number of readmissions (ie, high utilization rates). Some providers might have taken into account social factors (ie, marital status, social support, and type of health insurance) for admitting patients, and these cases could have biased our predictions. Nevertheless, although hospitals do not have published policies regarding social or nonmedical factors that providers should follow when making decisions on hospital admission or readmission, it is important to acknowledge that conscious and unconscious biases exist that influence these decisions. Implementation of this or another artificial intelligence (AI)–based oncology readmission model could exacerbate these biases or lead to unintended consequences, such as increasing the median length of hospital stay. Future studies before implementation of an oncology readmission tool should interview physicians using a mix of case-based scenarios and open-ended questions regarding the impact of social factors on hospital admission/readmission of high-risk patients such as distance from hospital, requirement to receive a treatment over several days, and family situations. Any implementation of an oncology-specific readmission model should also closely evaluate its impact on hospital length of stay as a balanced metric. Despite the retrospective design, for each cohort, the data were temporally split, and no patients in the validation data were in the training data, reducing the chance of data leakage. Moreover, we applied oversampling to deal with the imbalance and defined recall as the key performance metric, contrary to the commonly used AUC, which can over predict the predictive performance of imbalanced data sets.[65] However, the use of recall as a primary metric for performance evaluation might have favored models with high rates of FP.

In conclusion, according to the two different ML models developed in our study, the factors that were most predictive for unplanned hospital admission in patients with solid tumor included intensity of health care utilization, underlying comorbidities, hospitalization-related information, and specific laboratory values during the index admission, which are strong predictors of unplanned 30-day readmission risk. Furthermore, discharge planning could be optimized using cancer-specific unplanned 30-day readmission risk assessment ML algorithms and public health policies need to address social inequities related to the neighborhood context. Future studies should focus on prospective implementations, different populations, the role of SDOH in unplanned readmissions in patients with cancer, and how these algorithms could be improved with the addition of SDOH data.



## AFFILIATIONS

[1]Graduate Education Office, Case Western Reserve University School of Medicine, Cleveland, OH
[2]Department of Hematology-Oncology, University Hospitals Seidman Cancer Center, Case Western Reserve University, Cleveland, OH
[3]Faculdade Israelita de Ciências da Saúde Albert Einstein, Hospital Israelita Albert Einstein, São Paulo, Brazil
[4]Department of Population and Quantitative Health Sciences, Case Western Reserve University School of Medicine, Cleveland, OH
[5]Department of Medical Oncology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA

[6]Pandata LLC, Cleveland, OH
[7]Cancer Informatics, Seidman Cancer Center at University Hospitals of Cleveland, Cleveland, OH
[8]Oncohematology Department, Hospital Israelita Albert Einstein, São Paulo, Brazil
[9]Trans-Divisional Research Program (TDRP), Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute, National Institutes of Health, Bethesda, MD
[10]Center for Biomedical Informatics and Information Technology (CBIIT), National Cancer Institute, National Institutes of Health, Bethesda, MD

## CORRESPONDING AUTHOR

Nickolas Stabellini, BS, University Hospitals Seidman Cancer Center, Breen Pavilion, 11100 Euclid Ave, Cleveland, OH 44106; e-mail: nickolas@case.edu.

## SUPPORT

N.S. was supported through funding from the Sociedade Beneficente Israelita Brasileira Albert Einstein on the program Marcos Lottenberg & Marcos Wolosker International Fellowship for Physicians Scientist—Case Western.

## DATA SHARING STATEMENT

University Hospitals Seidman Cancer Center database is available at University Hospitals Cleveland Medical Center and has access restricted to researchers with institutional review board approval.

## AUTHOR CONTRIBUTIONS

**Conception and design:** Nickolas Stabellini, Aziz Nazha, Merilys Huhn, Kristin Waite, Jill S. Barnholtz-Sloan, Alberto J. Montero
**Administrative support:** Kristin Waite
**Provision of study materials or patients:** Jill S. Barnholtz-Sloan
**Collection and assembly of data:** Nikita Agrawal, Merilys Huhn, Alberto J. Montero
**Data analysis and interpretation:** Nickolas Stabellini, Aziz Nazha, Merilys Huhn, John Shanahan, Jill S. Barnholtz-Sloan, Alberto J. Montero
**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).

**Aziz Nazha**
This author is a member of the *JCO Clinical Cancer Informatics* Editorial Board. Journal policy recused the author from having any role in the peer review of this manuscript.
**Employment:** Incyte, Amazon
**Stock and Other Ownership Interests:** Amazon, Incyte

**Merilys Huhn**
**Stock and Other Ownership Interests:** NowRx
**Consulting or Advisory Role:** Centauri Health Solutions

**Nelson Hamerschlak**
**Consulting or Advisory Role:** Amgen, AbbVie, Sanofi
**Speakers' Bureau:** Celgene, Amgen, Jazz Pharmaceuticals

**Jill S. Barnholtz-Sloan**
**Employment:** Piedmont Health
**Leadership:** Piedmont Health
**Stock and Other Ownership Interests:** Surgical Theater
**Consulting or Advisory Role:** Monteris Medical, Medtronic
**Research Funding:** Merck, Lentigen
**Other Relationship:** National Institute of Health (NIH)

**Alberto J. Montero**
**Honoraria:** Celgene, AstraZeneca, OncoSec
**Consulting or Advisory Role:** New Century Health, Welwaze, Paragon Healthcare
**Research Funding:** F. Hoffmann-La Roche Ltd, Basel, Switzerland
**Uncompensated Relationships:** Roche
**Open Payments Link:** https://openpaymentsdata.cms.gov/physician/618396

No other potential conflicts of interest were reported.

## ACKNOWLEDGMENT

This project was performed in consultation with Pandata LLC on the development of ML algorithms.

## REFERENCES

1. Gruneir A, Dhalla IA, van Walraven C, et al: Unplanned readmissions after hospital discharge among patients identified as being at high risk for readmission using a validated predictive algorithm. Open Med 5:e104-e111, 2011
2. Anderson GF, Steinberg EP: Hospital readmissions in the Medicare population. N Engl J Med 311:1349-1353, 1984
3. Parker S: Do current discharge arrangements from inpatient hospital care for the elderly reduce readmission rates, the length of inpatient stay or mortality, or improve health status? WHO Regional Office for Europe, 2005. https://www.euro.who.int/__data/assets/pdf_file/0006/74670/E87542.pdf
4. Hospital Readmissions Reduction Program (HRRP). CMS. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Readmissions-Reduction-Program
5. Hospital discharge and readmission. UpToDate. https://www.uptodate.com/contents/hospital-discharge-and-readmission#H14
6. Jencks SF, Williams MV, Coleman EA: Rehospitalizations among patients in the Medicare fee-for-service program. N Engl J Med 360:1418-1428, 2009
7. Medicare & Medicaid statistical supplement. CMS. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Archives/MMSS
8. Bekelman JE, Halpern SD, Blankart CR, et al: Comparison of site of death, health care utilization, and hospital expenditures for patients dying with cancer in 7 developed countries. JAMA 315: 272-283, 2016
9. Graham KL, Auerbach AD, Schnipper JL, et al: Preventability of early versus late hospital readmissions in a national cohort of general medicine patients. Ann Intern Med 168:766-774, 2018
10. Manzano JGM, Lin H, Zhao H, et al: Derivation and validation of the cancer READMIT score: A readmission risk scoring system for patients with solid tumor malignancies. JCO Oncol Pract 18: e117-e128, 2022
11. Brooks GA, Li L, Sharma DB, et al: Regional variation in spending and survival for older adults with advanced cancer. J Natl Cancer Inst 105:634-642, 2013
12. Cancer facts & figures 2022. American Cancer Society. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html
13. Manzano JGM, Gadiraju S, Hiremath A, et al: Unplanned 30-day readmissions in a general internal medicine hospitalist service at a comprehensive cancer center. JCO Oncol Pract 11:410-415, 2015

14. Brooks GA, Abrams TA, Meyerhardt JA, et al: Identification of potentially avoidable hospitalizations in patients with GI cancer. J Clin Oncol 32:496-503, 2014
15. Epstein AS, Crosbie C, Martin SC, et al: 30-day-or-sooner readmissions of gastrointestinal medical oncology patients following cancer center inpatient service discharge: Characteristics and preventability. Hosp Pract 42:34-44, 2014
16. Montero AJ, Stevenson J, Guthrie AE, et al: Reducing unplanned medical oncology readmissions by improving outpatient care transitions: A process improvement project at the Cleveland Clinic. JCO Oncol Pract 12:e594-e602, 2016
17. Johnson PC, Xiao Y, Wong RL, et al: Potentially avoidable hospital readmissions in patients with advanced cancer. JCO Oncol Pract 15:e420-e427, 2019
18. Beam AL, Kohane IS: Big data and machine learning in health care. JAMA 319:1317-1318, 2018
19. Chen JH, Asch SM: Machine learning and prediction in medicine—Beyond the peak of inflated expectations. N Engl J Med 376:2507-2509, 2017
20. Huang Y, Talwar A, Chatterjee S, et al: Application of machine learning in predicting hospital readmissions: A scoping review of the literature. BMC Med Res Methodol 21:96, 2021
21. Min X, Yu B, Wang F: Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: A case study on COPD. Sci Rep 9:2362, 2019
22. Donzé J, Aujesky D, Williams D, et al: Potentially avoidable 30-day hospital readmissions in medical patients: Derivation and validation of a prediction model. JAMA Intern Med 173:632-638, 2013
23. Morgan DJ, Bame B, Zimand P, et al: Assessment of machine learning vs standard prediction rules for predicting hospital readmissions. JAMA Netw Open 2:e190348, 2019
24. Hong AS, Halm EA: Applying hospital readmissions to oncology: A square peg in a round hole? JCO Oncol Pract 18:7-10, 2022
25. Stabellini N, Bruno DS, Dmukauskas M, et al: Sex differences in lung cancer treatment and outcomes at a large hybrid academic-community practice. JTO Clin Res Rep 3:100307, 2022
26. Stabellini N, Chandar AK, Chak A, et al: Sex differences in esophageal cancer overall and by histological subtype. Sci Rep 12:5248, 2022
27. Caisis: The free & open source cancer data management system. http://www.caisis.org/
28. EMERSE: Electronic medical record search engine. https://project-emerse.org/index.html
29. PCH cancer readmissions measure overview. https://qualitynet.cms.gov/pch/measures/readmissions
30. Charlson ME, Pompei P, Ales KL, et al: A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. J Chronic Dis 40:373-383, 1987
31. Rosen RD, Sapra A: TNM classification. Treasure Island, FL, StatPearls Publishing, 2022. http://www.ncbi.nlm.nih.gov/books/NBK553187/
32. Social determinants of health. LexisNexis Risk Solutions. https://risk.lexisnexis.com/healthcare/social-determinants-of-health
33. Social determinants of health—Healthy people 2030. health.gov. https://health.gov/healthypeople/priority-areas/social-determinants-health
34. Accountable Care Organizations (ACOs). CMS. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ACO
35. Reuben DB, Tinetti ME: The hospital-dependent patient. N Engl J Med 370:694-697, 2014
36. Saravanan N, Sathish J, Balajee JM: Data wrangling and data leakage in machine learning for healthcare. Rochester, NY, 2018. https://papers.ssrn.com/abstract=3708142
37. Sapra A, Malik A, Bhandari P: Vital sign assessment. Treasure Island, FL: StatPearls Publishing, 2022. http://www.ncbi.nlm.nih.gov/books/NBK553213/
38. Weir CB, Jan A: BMI classification percentile and cut off points. Treasure Island, FL: StatPearls Publishing, 2022. http://www.ncbi.nlm.nih.gov/books/NBK541070/
39. Ke G, Meng Q, Finley T, et al: LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems. Curran Associates, Inc, 2017. https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html
40. Chen T, Guestrin C: XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). New York, NY: Association for Computing Machinery, 2016, pp 785-794. https://doi.org/10.1145/2939672.2939785
41. Lundberg SM, Lee SI: A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems. Curran Associates, Inc, 2017. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
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. BMC Med 13:1, 2015
43. RStudio: Open source & professional software for data science teams. https://www.rstudio.com/
44. Wang W, Li Y, Yan J: touch: Tools of utilization and cost in healthcare, 2022. https://CRAN.R-project.org/package=touch
45. Cooper W: icdcoder, 2022. https://github.com/wtcooper/icdcoder
46. ICD: ICD-9—International Classification of Diseases, ninth revision [Internet], 2022. https://www.cdc.gov/nchs/icd/icd9.htm
47. Lang M, Bischl B, Richter J, et al: mlr3: Machine learning in R—Next generation, 2022. https://CRAN.R-project.org/package=mlr3
48. Jones SL, Cheon O, Manzano JGM, et al: Comparison of LACE and HOSPITAL readmission risk scores for CMS target and nontarget conditions. Am J Med Qual 37:299, 2022
49. Wong CW, Chen C, Rossi LA, et al: Explainable tree-based predictions for unplanned 30-day readmission of patients with cancer using clinical embeddings. JCO Clin Cancer Inform 5:155-167, 2021
50. Azuaje F: Artificial intelligence for precision oncology: Beyond patient stratification. NPJ Precis Oncol 3:6, 2019
51. Lundberg SM, Erion G, Chen H, et al: From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2:56-67, 2020
52. Schmidt CR, Hefner J, McAlearney AS, et al: Development and prospective validation of a model estimating risk of readmission in cancer patients. J Surg Oncol 117:1113-1118, 2018
53. Hwang S, Urbanowicz R, Lynch S, et al: Towards predicting 30-day readmission among oncology patients: Identifying timely and actionable risk factors. medRxiv, 2022. https://www.medrxiv.org/content/10.1101/2022.01.05.21268065v1
54. Whitney RL, Bell JF, Tancredi DJ, et al: Hospitalization rates and predictors of rehospitalization among individuals with advanced cancer in the year after diagnosis. J Clin Oncol 35:3610-3617, 2017
55. Burhenn P, Sun CL, Scher KS, et al: Predictors of hospital readmission among older adults with cancer. J Geriatr Oncol 11:1108-1114, 2020
56. Chiang LY, Liu J, Flood KL, et al: Geriatric assessment as predictors of hospital readmission in older adults with cancer. J Geriatr Oncol 6:254-261, 2015
57. Enright K, Grunfeld E, Yun L, et al: Population-based assessment of emergency room visits and hospitalizations among women receiving adjuvant chemotherapy for early breast cancer. JCO Oncol Pract 11:126-132, 2015
58. Brooks GA, Kansagra AJ, Rao SR, et al: A clinical prediction model to assess risk for chemotherapy-related hospitalization in patients initiating palliative chemotherapy. JAMA Oncol 1:441-447, 2015
59. Handley NR, Schuchter LM, Bekelman JE: Best practices for reducing unplanned acute care for patients with cancer. JCO Oncol Pract 14:306-313, 2018
60. Wammes JJG, van der Wees PJ, Tanke MAC, et al: Systematic review of high-cost patients' characteristics and healthcare utilisation. BMJ Open 8:e023113, 2018
61. Social determinants of health. https://www.who.int/health-topics/social-determinants-of-health
62. Galea S, Tracy M, Hoggatt KJ, et al: Estimated deaths attributable to social factors in the United States. Am J Public Health 101:1456-1465, 2011
63. Azizi M, Chedid A, Oudard S: Home blood-pressure monitoring in patients receiving sunitinib. N Engl J Med 358:95-97, 2008
64. 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
65. Saito T, Rehmsmeier M: The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 10:e0118432, 2015