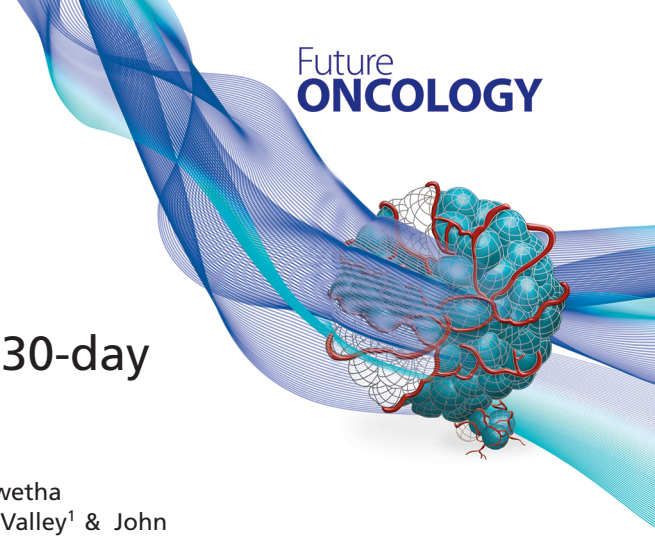
Short Communication

For reprint orders, please contact: reprints@futuremedicine.com



Augmented intelligence to predict 30-day mortality in patients with cancer

Ajeet Gajra*[,1], Marjorie E Zettler[‡,1], Kelly A Miller[2], Sibel Blau[3], Swetha S Venkateshwaran[2], Shreenath Sridharan[2], John Showalter[2], Amy W Valley[1] & John G Frownfelter[2]

[1]Cardinal Health Specialty Solutions, Dublin, OH 43017, USA

[2]Jvion, Inc., Suwanee, GA 30024, USA

[3]Rainier Hematology Oncology/Northwest Medical Specialties, Tacoma, WA 98405, USA

*Author for correspondence: Tel.: +1 847 887 2391; ajeet.gajra@cardinalhealth.com

[‡]Former employee of Cardinal Health

Aim: An augmented intelligence tool to predict short-term mortality risk among patients with cancer could help identify those in need of actionable interventions or palliative care services. **Patients & methods:** An algorithm to predict 30-day mortality risk was developed using socioeconomic and clinical data from patients in a large community hematology/oncology practice. Patients were scored weekly; algorithm performance was assessed using dates of death in patients' electronic health records. **Results:** For patients scored as highest risk for 30-day mortality, the event rate was 4.9% (vs 0.7% in patients scored as low risk; a 7.4-times greater risk). **Conclusion:** The development and validation of a decision tool to accurately identify patients with cancer who are at risk for short-term mortality is feasible.

First draft submitted: 9 March 2021; Accepted for publication: 3 June 2021; Published online: 30 June 2021

Keywords: augmented intelligence • decision tool • machine learning

In recent years, artificial intelligence (AI), the capability for machines to mimic intelligent human behavior, and its application in machine learning (ML), which allows systems to learn from data and experience without explicit programming, have demonstrated their ability to improve healthcare delivery in a myriad of ways. AI and augmented intelligence (the use of AI to aid rather than replace human decision-making) have found many applications in oncology, from improved accuracy in diagnosis through radiographic image analysis to the ability to predict cancer treatment toxicity [1,2]. Despite these advances, the application of AI and ML to identify patients at risk for short-term mortality who could benefit from either an intervention to modify actionable clinical factors or a referral to palliative care services has received little attention.

While cancer accounts for a significant proportion of deaths in the USA each year, many patients with a diagnosis of cancer die of non-cancer causes [3]. Cancer treatment itself is associated with complications that can be life threatening, including tumor lysis syndrome, febrile neutropenia and sepsis [4–6]. In addition to the risks to individual patients, emergency department visits and hospitalizations related to these complications also carry a considerable financial burden [7]. The ability to proactively identify these patients could facilitate early intervention and resolution of reversible complications.

For patients with cancer who are approaching end of life (EoL), early palliative and hospice care referrals result in improved quality of life and symptom management [8,9]. In contrast, aggressive, life-sustaining EoL care can conflict with patient preference and result in lower quality of life, family perceptions of poorer quality of care, and greater regret about treatment decisions [10]. Earlier referral also represents an opportunity to transform cancer care by reducing the potential for unnecessary, toxic and expensive treatments at EoL. The American Society of Clinical Oncology clinical practice guidelines recommend that palliative care be integrated into standard oncology care for patients with advanced cancer [11]. Additionally, the proportion of patients who died from cancer who received chemotherapy in the last 14 days of life, and the proportion who died from cancer who were admitted to hospice for fewer than 3 days, are quality measures endorsed by the National Quality Forum [12]. Many physicians

delay advance care planning and EoL conversations until patients are in the terminal phase of life [13]. The ability of physicians to predict mortality varies greatly and has been demonstrated to be largely inaccurate [14–16]. Existing mortality risk models often only include clinical factors already known to clinicians (e.g., Eastern Co-operative Oncology Group performance status, stage, cancer type and number of comorbidities) and identify patients whom clinicians are already aware of as being at risk [17,18]. However, various sociodemographic and geographic factors can also impact outcomes of patients with cancer beyond clinical factors. The impact of such nonclinical factors upon mortality in patients with cancer is an area of active research.

The aim of this study is to demonstrate that a commercially available AI tool can provide clinicians with insight into which patients with cancer are at the highest risk of mortality in the next 30 days. We developed and validated a ML model at a large community oncology practice in the Pacific Northwest of the USA utilizing clinical, environmental and behavioral data and information on social determinants of health. An objective decision support tool could assist providers in identifying patients who would benefit from a goals of care assessment and palliative and/or hospice consultation.

Patients & methods

Patient sample

A total of 3671 patients were selected from a deidentified database representing a large community-based hematology/oncology practice in the USA. The practice had 21 healthcare providers and saw >4300 unique patients per month during the period of study. Clinic locations represented both urban and rural areas of Washington State. An active patient registry was defined as distinct, living patients (≥18 years old) who had at least one office visit in the last 2 years (between 1 September 2017 and 31 August 2019; n = 14,321), had an oncology ICD-10 diagnosis code (n = 4586), had an office visit with an oncology provider during the time frame (n = 4338) and were scored by the model (n = 3671). Two subpopulations were derived from the overall sample with the aim to better identify active, current oncology patients: subpopulation A, consisting of patients ≥18 years old with an oncology diagnosis and at least two office visits in the last 180 days; and subpopulation B, consisting of patients ≥18 years old with an oncology diagnosis and at least one office visit in the past year. The data utilized in this study did not include any patient identifiers that met the standard for non-exempt research as defined by US federal regulation 45 CFR 46, and the Cardinal Health Specialty Solutions Ethics Committee ruled that no formal ethics approval was required.

Data sources

The sole data source for this study was a deidentified database that included electronic health record (EHR) data, billing data and socioeconomic determinants of care (including behavioral data). Clinical and billing data were collected directly from the practice's EHRs. Clinical data were included if they had been recorded within 12 months of the date of scoring socioeconomic determinants, and mortality events were matched to individuals prior to deidentification. Socioeconomic elements had previously been collected from publicly available sources such as the US Census Bureau, US Department of Agriculture and the National Oceanic and Atmospheric Administration. The full census dataset was considered by the model, including poverty, income, household size, transportation, employment and neighborhood characteristics. A few data elements are unclear in the government references (neighborhood in-migration, group living quarters and purchasing channel preference). Respectively, these are: the percentage of households which have moved to the neighborhood in the last 12 months; individuals living in college residence halls, residential treatment centers, skilled nursing facilities, group homes, military barracks, correctional facilities and workers' dormitories; and how individuals shop – for example, over the internet [19].

Behavioral data had been purchased from third-party data vendors such as Acxiom, Experian and Transunion. Individual behavioral data included elements such as history of internet searches on health conditions, purchasing channels and life stage. All purchased data are indexes indicating preferences and no transactional-level data was used. The purpose of the behavioral data is to enhance the model with individual preferences and behaviors. Publicly available data had been collected at the census tract level and behavioral data had been collected at the individual level. Race was defined by the race information provided by the practice in its billing data.

Outcome measures

The principal outcome measure of this study was mortality at 30 days. An individual's mortality event was pulled from the EHR and from the Social Security Administration's master death file [20]. Secondary outcomes included mortality at 60, 90 and 180 days.

Data standards

Standard claim ontologies were used for billing data. Discrete clinical data, publicly available data and purchased data were used in their native form without transformation [21].

ML approach

The AI approach is a commercially available clinical AI solution, the Jvion CORE, which uses an n-dimensional eigenspace approach. Each patient is mapped to relevant eigen clusters [22–24]. A supervised continuous learning AI layer optimizes the ML techniques used in each cluster to calculate an individual's most likely future trajectory in the eigen space. The algorithm is inclusive, and all known features are considered when the trajectory is calculated. An individual's risk is determined by the likelihood that their projected trajectory will intersect with a high-mortality cluster.

Patients positioned in the eigenspace and cluster membership are continuously adjusted as new data are added [25]. This continual fine tuning enables the localization and ongoing performance improvements that are characteristic of ML technology. The technique to identify relevance of patients is similar to identifying signal-to-noise ratio in signal processing systems [26]. The results of this approach have been shown to change clinician behavior and improve outcomes in other clinical areas, such as 30-day readmissions [27].

The algorithm was initially developed on a historic dataset of 103,468 patients from four oncology practices representing 1557 total features and validated using a 75% training, 25% testing approach. The oncology practices, one of which is part of the prospective validation described here, are located throughout the USA: two in the southeast, one in the southwest and one in the northwest. The train timeframe was 1 January through 30 June 2017, and the test timeframe was 1 July through 31 August 2017. Patients were duplicated in the test and train populations and were scored approximately 5.4× in the train portion and 2× in the test portion. The area under the curve (AUC) for the train model was 0.95 and the AUC for the test model was 0.92. To prospectively assess the accuracy of the approach, all active patients (defined as oncology patients ≥18 years old who have seen any oncology provider at one of eight clinic locations between 1 September 2017 and 31 August 2019; n = 3671) were scored for the month of September 2019. All patients meeting criteria were scored weekly during September 2019, with the highest risk during the month used for the prospective validation. Mortality and date of death were defined based on the 'date of death' field being completed in the practice's EHRs and/or the patient being listed in the Social Security Administration's master death file. Prospective validation only included patients from one practice, Northwest Medical Specialties.

Analysis

The performance of the ML algorithm to predict which patients would have a primary outcome was determined prospectively based on notation of death in the EHR. The AUC for the receiver operating characteristic was calculated for both the primary outcome of 30-day mortality and the secondary outcomes of mortality at 60, 90 and 180 days. The sensitivity and specificity were calculated in the testing set for the primary and secondary outcomes using a threshold of the top 5% highest-risk patients as the at-risk population. This threshold was set based on operational considerations of the number of patients clinicians could evaluate as well as to identify and prioritize the highest-risk patients. The highest-weighted features at the individual level were also determined and aggregated across the population to determine the features most associated with determining risk.

Results

The characteristics of the study populations are presented in Table 1. In the overall patient population (n = 3671 patients), the mean age was 68.4 years, and 54.8% of the patients were female. The majority of the patients were Caucasian (55.1%). The most prevalent cancer diagnosis was breast cancer (13.8%), followed by lung cancer (13.0%) and lymphoma (9.6%). Among the patients flagged by the algorithm as being at risk for mortality within 30 days (top 5%; n = 202 patients), the mean age was 68.4 years, 52.5% were female and 55.9% were Caucasian. The most prevalent cancer diagnosis within the at-risk group was breast cancer (21.8%), followed by lung cancer

10.2217/fon-2021-0302

Table 1. Characteristics of all active patients.

	Overall patient population (n = 3671), n (%)	Patients at highest risk (top 5%) (n = 202), n (%)
Gender		
Female	2010 (54.8)	106 (52.5)
Male	1661 (45.2)	96 (47.5)
Race		
White or Caucasian	2024 (55.1)	113 (55.9)
Black or African–American	108 (2.9)	6 (3.0)
Asian	112 (3.1)	7 (3.5)
Unknown or other	1427 (38.9)	76 (37.6)
Age group (years)		
18–25	8 (0.2)	0 (0.0)
26–40	113 (3.1)	2 (1.0)
41–50	218 (5.9)	12 (5.9)
51–65	1062 (28.9)	63 (31.2)
66–80	1632 (44.5)	91 (45.0)
>80	638 (17.4)	34 (16.8)
Primary cancer type (top 5)		
Breast	506 (13.8)	44 (21.8)
Lung	479 (13.0)	40 (19.8)
Lymphoma	353 (9.6)	Prostate cancer: 16 (7.9)
Prostate	228 (6.2)	Head and neck cancer: 10 (5.0)
Head and Neck	205 (5.6)	Kidney cancer: 9 (4.5)
Cancer staging		
Stage 0	101 (2.8)	0 (0.0)
Stage 1	437 (11.9)	4 (2.0)
Stage 2	315 (8.6)	11 (5.4)
Stage 3	392 (10.7)	21 (10.4)
Stage 4	658 (17.9)	152 (75.2)
Unknown	1768 (48.2)	14 (6.9)

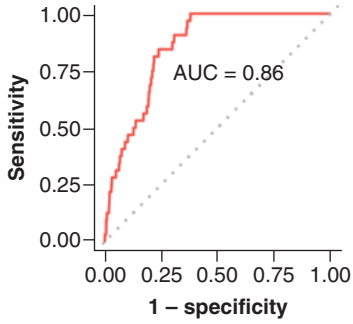


Figure 1. Overall patient population: area under the curve, 30-day mortality.
AUC: Area under the curve.

(19.8%) and prostate cancer (7.9%). Tables 2 and 3 summarize the characteristics of the two subpopulation analyses: subpopulation A and subpopulation B.

For the top 5% of patients flagged by the algorithm as being at risk for mortality within the next 30 days, the AUC for the 30-day mortality model was 0.86 for the overall population (Figure 1). Similar values were obtained in this patient population for 60-, 90- and 180-day mortality. In subpopulation A (patients ≥18 years old with an oncology diagnosis and at least two office visits in the last 180 days; n = 1732), among the top 5% of patients at risk for 30-day mortality (n = 96), the AUC was 0.78, while in subpopulation B (patients ≥18 years old with an oncology diagnosis and at least one office visit in the past year; n = 3053), among the top 5% at risk for 30-day mortality (n = 168), the AUC was 0.84 (Figures 2 & 3). For both subpopulations, the AUC was slightly increased at later time points.

Table 2. Characteristics of subpopulation A (oncology patients with at least two office visits in the last 180 days).

	All patients in subpopulation A (n = 1732), n (%)	Patients at highest risk (top 5%) in subpopulation A (n = 96), n (%)
Gender		
Female	910 (52.5)	58 (60.4)
Male	822 (47.5)	38 (39.6)
Race		
White or Caucasian	944 (54.5)	55 (57.3)
Black or African–American	56 (3.2)	4 (4.2)
Asian	48 (2.8)	4 (4.2)
Unknown or other	684 (39.5)	33 (34.4)
Age group (years)		
18–25	5 (0.3)	0 (0.0)
26–40	44 (2.5)	1 (1.0)
41–50	97 (5.6)	5 (5.2)
51–65	462 (26.7)	32 (33.3)
66–80	831 (48.0)	44 (45.8)
>80	293 (16.9)	14 (14.6)
Primary cancer type (top 5)		
Breast	282 (16.3)	36 (27.1)
Lung	242 (14.0)	27 (28.1)
Prostate	126 (7.3)	5 (5.2)
Head and neck	98 (5.7)	4 (4.2)
Lymphoma	133 (7.7)	Kidney cancer: 4 (4.2)
Cancer staging		
Stage 0	30 (1.7)	0 (0.0)
Stage 1	191 (11.0)	2 (2.1)
Stage 2	163 (9.4)	5 (5.2)
Stage 3	235 (13.6)	4 (4.2)
Stage 4	481 (27.8)	83 (86.5)
Unknown	632 (36.5)	2 (2.1)

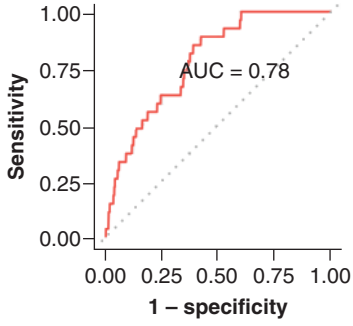


Figure 2. Subpopulation A: area under the curve, 30-day mortality.
AUC: Area under the curve.

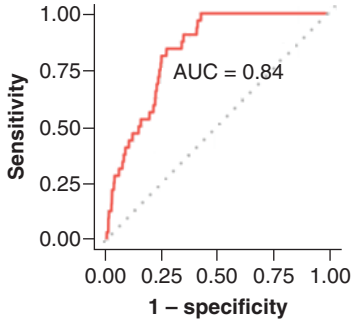


Figure 3. Subpopulation B: area under the curve, 30-day mortality.
AUC: Area under the curve.

10.2217/fon-2021-0302

Table 3. Characteristics of subpopulation B (oncology patients with at least one office visit in the last year).

	All patients in subpopulation B (n = 3053), n (%)	Patients at highest risk (top 5%) in subpopulation B (n = 168), n (%)
Gender		
Female	1669 (54.7)	94 (56.0)
Male	1384 (45.3)	74 (44.0)
Race		
White or Caucasian	1693 (55.5)	94 (56.0)
Black or African–American	86 (2.8)	5 (3.0)
Asian	96 (3.1)	6 (3.6)
Unknown or other	1178 (38.6)	63 (37.5)
Age group (years)		
18–25	7 (0.2)	0 (0.0)
26–40	85 (2.8)	1 (0.6)
41–50	178 (5.8)	10 (6.0)
51–65	875 (28.7)	54 (32.1)
66–80	1400 (45.9)	77 (45.8)
>80	508 (16.6)	26 (15.5)
Primary cancer type (top 5)		
Breast	445 (14.6)	40 (23.8)
Lung	388 (12.7)	37 (22.0)
Prostate	189 (6.2)	13 (7.7)
Lymphoma	294 (9.6)	Kidney cancer: 9 (5.4)
Head and neck	178 (5.8)	8 (4.8)
Cancer staging		
Stage 0	83 (2.7)	0 (0.0)
Stage 1	361 (11.8)	3 (1.8)
Stage 2	274 (9.0)	10 (6.0)
Stage 3	346 (11.3)	17 (10.1)
Stage 4	590 (19.3)	129 (76.8)
Unknown	1399 (45.8)	9 (5.4)

The mortality rates and risk ratios are displayed in Table 4. In the overall patient population, the mortality rate was 0.9% at 30 days. Within the 5% of the population identified by the algorithm as being at highest risk, the mortality rate was 4.9%, while in the remaining 95% the mortality rate was 0.7% (i.e., patients flagged as high risk by the algorithm had a 7.4-times greater risk for mortality). Mortality rates increased at every time point, with risk in the top 5% reaching 8.5-times at 180 days. Similarly, within subpopulation A, the mortality rate at 30 days was 1.4% (7.2% within the 5% of the population identified by the algorithm as being at highest risk, and 1.1% in the remaining 95% of the patient population; a 6.7-times greater risk). Mortality rates increased at 60, 90 and 180 days, and risk in the top 5% was comparable at each time point. Within subpopulation B, the mortality rate at 30 days was 1.0% (5.5% within the 5% of the population identified by the algorithm as being at highest risk, and 0.7% in the remaining 95% of the patient population; a 7.4-times greater risk). Mortality rates increased at 60, 90 and 180 days, and risk in the top 5% was similar at each time point.

The top features (Supplementary Table 1) include tumor staging, cancer type and chemotherapy/radiation therapy. Lifestyle was also highly relevant, with tobacco use being the second most important feature. Neighborhood, individual and household characteristics as previously described are also highly weighted [19]. The individual's clinical history and laboratory studies, and the time of year complete the top 30 features.

Discussion

In this study we found that the ML algorithm accurately predicted 30-day mortality among patients with cancer in a large oncology practice in the USA, with an AUC of 0.86. Results at 60, 90 and 180 days were similar; the AUC was stable over time (Supplementary Figures 1–9). Compared with the overall population, the AUC dropped

Table 4. Mortality rates, risk ratios and NPV across the three study populations.

Population scored and outlook	Mortality rate overall (%)	Mortality rate top 5% at risk (%)	Mortality rate remaining 95% (%)	Risk ratio	NPV
Overall patient population (n = 3671); top 5% at risk (n = 202)					
30 days	0.9	4.9	0.7	7.4	0.99
60 days	1.7	9.8	1.3	7.8	0.99
90 days	2.3	13.6	1.7	7.8	0.98
180 days	4.2	26.1	3.1	8.5	0.97
Subpopulation A (n = 1732); top 5% at risk (n = 96)†					
30 days	1.4	7.2	1.1	6.7	0.99
60 days	2.6	14.4	2.0	7.4	0.98
90 days	3.7	17.5	2.9	6.0	0.97
180 days	6.8	33.0	5.4	6.1	0.95
Subpopulation B (n = 3053); top 5% at risk (n = 168)‡					
30 days	1.0	5.5	0.7	7.4	0.99
60 days	1.9	11.0	1.5	7.6	0.99
90 days	2.7	15.2	2.1	7.4	0.98
180 days	4.8	27.4	3.6	7.7	0.96

†Subpopulation A: oncology patients only; at least two office visits in the last 180 days.
‡Subpopulation B: oncology patients only; at least one office visit in the last year.
NPV: Negative predictive value.

among the subpopulations of patients seen once in the preceding year or twice in the preceding 180 days. This observation is likely attributable to screening out those patients seen less often, who would be at lower risk for death (i.e., the negative predictive value is known). Mortality rates across the three patient populations were low at 30 days (0.9–1.4%), but markedly higher among the patients identified by the algorithm as being at highest risk (4.9–7.2%), with risk 6.7- to 7.4-times that of the patients flagged by the algorithm as being at lower risk. In addition, to better assess the discriminatory power of the ML algorithm, we conducted an exploratory analysis to study the mortality rate by risk percentile bands and the associated hazard ratios. The discrimination is good, with mortality rates in the bottom 40th percentile of less than 1% and most of the 95th–100th percentile hazard ratios over 10 (Supplementary Table 2). The findings from our study demonstrate that it is feasible to develop and validate a decision tool to aid in improving clinical care for patients with cancer, utilizing clinical, social, environmental and behavioral data.

Cancer is the second leading cause of death in the USA and was responsible for an estimated 606,520 deaths in 2020 [28,29]. However, among patients with a diagnosis of cancer, risk of noncancer death surpasses that of cancer death [3]. Complications arising from cancer treatment, particularly infections, contribute significantly to mortality [3]. The ability to identify those patients at risk for mortality due to treatable, reversible complications such as infections provides an opportunity for clinical intervention to change the outcome trajectory. Proactive interventions prompted by the AI tool also have the potential to reduce costs by decreasing unplanned admissions and emergency department visits. Physicians' estimates of prognosis are known to be subjective and overly optimistic [30]; therefore there is an unmet need for an objective tool to aid in identifying patients at risk for short-term mortality. Although validated prognostic tools exist, such as the Palliative Prognostic Score and Palliative Prognostic Indicator, they are not used routinely due to perceived complexity or inconvenience [30]. Additionally, some alternative tools such as the 'Surprise Question' have inconsistent accuracy, with AUCs ranging from 0.512 to 0.822 [31,32].

Several prediction algorithms and ML have demonstrated improved accuracy in predicting mortality risk up to 6 months [33–35]. A cancer mortality questionnaire utilizing classification trees was developed to predict the 60-, 90- and 180-day mortality risk with corresponding AUCs of 0.86, 0.84 and 0.83 [33]. Additionally, Parikh *et al.* demonstrated AUCs ranging from 0.86 to 0.88 for 6-month mortality risk for patients with cancer utilizing random forest, gradient boosting and logistic regression methodologies [34]. Finally, short-term models built to predict 30-day mortality for patients initiating chemotherapy demonstrated an even higher AUC of 0.94 [35].

To date, however, none of these tools have seen widespread adoption in clinical practice. One barrier may be the additional time and effort required to apply the AI and interpret results. In our study, the AI automatically generated patient-specific, dynamic and actionable insights with respect to patients' risk for mortality within the next 30 days on a weekly basis. Importantly, use of this AI solution required no additional documentation within the EHR, and the insights generated could be integrated back into any EHR to help inform the care plan. Because this tool is commercially available in the USA, a downstream workflow has been established at practices where it is being utilized [36]. This may include additional visits or a telemedicine visit when an 'at risk' patient is identified. At such a visit, the clinical provider assesses for any reversible causes that may be increasing the patient's risk for mortality. Such causes may be clinical (e.g., asymptomatic infection), social (e.g., missed appointments due to lack of transportation) or psychological (e.g., subclinical depression). If no reversible causes are discerned, and the patient is found appropriate for palliative care services, then a workstream to initiate a palliative care consult is initiated. The evaluation by the palliative care team can guide the decision regarding hospice services, at the appropriate time, if indicated.

Our study is subject to several limitations. First, some patients in the validation set may have been part of the training population, as both development and validation were conducted at a single practice. To address this limitation in future research, additional data are being collected at practices not involved in the training for the algorithm. Second, this study did not evaluate changes in physicians' behavior (e.g., increased referrals to palliative care or hospice) after deployment of the novel AI solution, or the impact of the AI solution upon healthcare utilization at EoL (e.g., emergency department visits, inpatient admissions). Those assessments are currently underway. Third, our study included patients from a single oncology practice and so may not be generalizable to all patients with all types of cancer; however, in a recent study of 26,525 patients receiving care at medical oncology clinics in the USA, a 180-day mortality rate of 4.0% was observed (vs 4.2% observed in our study), suggesting that our sample may be representative [34]. The AI solution has been deployed at other practices, and further work is ongoing. Additionally, there may have been gaps in acquiring mortality data as this information was derived from the EHR only; data from the US Social Security Death Index were not utilized. Delays in entering mortality events or potential omissions may have impacted the analysis. If all mortality events were received and accounted for, the model performance would likely have improved. Further study of the model's performance in clinical practice is underway. Finally, the visit information utilized professional claims data which are not 'real-time' data like the EHR. We omitted scoring 667 patients during the scoring period because they did not meet office visit criteria with the data we received at the time, but these patients subsequently would have been considered part of the active patient population.

Conclusion

The identification of patients with cancer at risk for 30-day mortality has the potential to improve outcomes for those patients with reversible clinical factors, or to ensure timely integration of palliative care in the management of patients nearing EoL. Our study found that AI can help with earlier identification of those patients at greatest risk for short-term mortality. Importantly, this novel AI solution can layer on top of existing workflows and processes to provide actionable insights for the management of patients with cancer to improve clinical care.

Future perspective

While the current use of approved augmented intelligence tools focuses on pattern recognition as in imaging and pathology, machine learning algorithms will be utilized in increasingly complex medical decision making within oncology care the near future. AI applications will utilize clinical, sociodemographic, environmental, pathologic, radiological, genetic, genomic and pharamcoepidemiological data to recommend, the treatment with greatest efficacy and least risk of toxicity, for a patient with cancer as well as offer direction for best available clinical trial, stratify risk of recurrence and implement prognostic models for a given patient as well as identify risk of disease within family members. With the rapid pace of development, such capabilities will be in clinical use within a decade as long as ethical and privacy issues are successfully navigated.

Summary points

- The application of augmented intelligence and machine learning to identify patients at risk for short-term mortality could help to improve clinical care for patients with advanced cancer, by presenting an opportunity to intervene and modify actionable clinical factors, or to refer to palliative care services.
- Socioeconomic and behavioral data, professional claims and deidentified electronic health record data from 3671 patients in a large community hematology/oncology practice were used to develop an algorithm to predict 30-day mortality risk.
- All active patients in the practice were scored using the model, and performance was determined based on the date of death as recorded in the patients' electronic health records.
- For the patients scored as highest risk for 30-day mortality (n = 202), the area under the curve for the 30-day mortality model was 0.86, with a sensitivity of 28.1% and specificity of 95.2%. Similar values were obtained for 60-, 90- and 180-day mortality with the model.
- The mortality rate in the overall patient population was low at 30 days (0.9%), but significantly higher among the patients identified by the algorithm as being at highest risk (4.9%), with risk 7.4-times that of the patients flagged by the algorithm as being at lower risk.
- Mortality rates increased at later time points, with risk in the 5% of patients flagged by the algorithm as being at highest risk reaching 8.5-times that of patients flagged at lower risk at 180 days.
- This study demonstrates that the development and validation of a decision tool to accurately identify patients with cancer who are at risk for death within 30 days is feasible. This tool may help to improve clinical care by prompting interventions, or by facilitating discussions with patients at end of life.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/fon-2021-0302

Author contributions

All authors contributed to the writing of this manuscript.

Financial & competing interests disclosure

J Frownfelter, S Venkateshwaran, S Sridharan, K Miller and J Showalter are employees of Jvion; A Gajra and M Zettler are employees of Cardinal Health; S Blau is an employee of Northwest Medical Specialties. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The Cardinal Health Specialty Solutions Ethics Committee determined that no formal ethics approval was required as the study utilized only deidentified data. The data utilized in this study did not include any patient identifiers that met the standard for non-exempt research as defined by US Federal Regulation 45 CFR 46.

Data availability

The data will not be available for sharing at this time as the data are commercially confidential. However, Cardinal Health will consider written requests to share the data on a case-by-case basis.

References

1. Ribli D, Horváth A, Unger Z, Pollner P, Csabai I. Detecting and classifying lesions in mammograms with deep learning. *Sci. Rep.* 8(1), 4165 (2018).
2. Lee S, Kerns S, Ostrer H, Rosenstein B, Deasy JO, Oh JH. Machine learning on a genome-wide association study to predict late genitourinary toxicity after prostate radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* 101(1), 128–135 (2018).
3. Zaorsky NG, Churilla TM, Egleston BL, *et al.* Causes of death among cancer patients. *Ann. Oncol.* 28(2), 400–407 (2017).
4. Durani U, Shah ND, Go RS. In-hospital outcomes of tumor lysis syndrome: a population-based study using the National Inpatient Sample. *Oncologist* 22(12), 1506–1509 (2017).
5. Lyman GH, Michels SL, Reynolds MW, Barron R, Tomic KS, Yu J. Risk of mortality in patients with cancer who experience febrile neutropenia. *Cancer* 116(23), 5555–5563 (2010).

6. Van de Louw A, Cohrs A, Leslie D. Incidence of sepsis and associated mortality within the first year after cancer diagnosis in middle aged adults: a US population based study. *PLoS ONE* 15(12), e0243449 (2020).
7. Jairam V, Lee V, Park HS, *et al.* Treatment-related complications of systemic therapy and radiotherapy. *JAMA Oncol.* 5(7), 1028–1035 (2019).
8. Haun MW, Estel S, Rücker G, *et al.* Early palliative care for adults with advanced cancer. *Cochrane Database Syst. Rev.* 6(6), CD011129 (2017).
9. Mulville AK, Widick NN, Makani NS. Timely referral to hospice care for oncology patients: a retrospective review. *Am. J. Hosp. Palliat. Care* 36(6), 466–471 (2019).
10. Starr LT, Ulrich CM, Corey KL, Meghani SH. Associations among end-of-life discussions, health-care utilization, and costs in persons with advanced cancer: a systematic review. *Am. J. Hosp. Palliat. Care* 36(10), 913–926 (2019).
11. 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(1), 96–112 (2017).
12. National Quality Forum. 2012 MAP Hospice and Palliative Care Family of Measures. (2021). www.qualityforum.org/QPS/QPSTool.aspx?p=1127
13. Mori M, Shimizu C, Ogawa A, Okusaka T, Yoshida S, Morita T. A national survey to systematically identify factors associated with oncologists' attitudes toward end-of-life discussions: what determines timing of end-of-life discussions? *Oncologist.* 20(11), 1304–1311 (2015).
14. Hui D, Kilgore K, Nguyen L, *et al.* The accuracy of probabilistic versus temporal clinician prediction of survival for patients with advanced cancer: a preliminary report. *Oncologist* 16(11), 1642–1648 (2011).
15. Glare P, Virik K, Jones M, *et al.* A systematic review of physicians' survival predictions in terminally ill cancer patients. *BMJ* 327(7408), 195–198 (2003).
16. White N, Reid F, Harris A, Harries P, Stone P. A systematic review of predictions of survival in palliative care: how accurate are clinicians and who are the experts? *PLoS ONE* 11(8), e0161407 (2016).
17. Brunello A, Fontana A, Zafferri V *et al.*Development of an oncology-multidimensional prognostic index (Onco-MPI) for mortality prediction in older cancer patients. *J. Cancer Res. Clin. Oncol.* 142, 1069–1077 (2016).
18. 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(11), 1723–1730 (2020).
19. Chen S, Bergman D, Miller K, Kavanagh A, Frownfelter J, Showalter J. Using applied machine learning to predict healthcare utilization based on socioeconomic determinants of care. *Am. J. Manag. Care* 26(1), 26–31 (2020).
20. US Social Security Administration. SSA's death information. www.ssa.gov/dataexchange/request_dmf.html
21. US Centers for Medicare & Medicaid Services. Adopted standards and operating rules. www.cms.gov/Regulations-and-Guidance/Administrative-Simplification/HIPAA-ACA/AdoptedStandardsandOperatingRules
22. Idé T, Kashima H. Eigenspace-based anomaly detection in computer systems. *Proc. Tenth ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining* 440–449 (2004).
23. Hall P, Marshall D, Martin R. Merging and splitting eigenspace models. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 042–1049 (2000).
24. Weiss Y. Segmentation using eigenvectors: a unifying view. *Proc. Seventh IEEE Int. Conf. Comput. Vision* 975–982 (1999).
25. Cai D, Zhang C, He X. Unsupervised feature selection for multi-cluster data. *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining* 333–342 (2010).
26. Lee H, Li F. An eigenvector technique for detecting the number of emitters in a cluster. *IEEE Trans. Signal Process.* 42, 2380–2388 (1994).
27. Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. Implementation of artificial intelligence-based clinical decision support to reduce hospital readmission at a regional hospital. *Appl. Clin. Inf.* 11(4), 570–577 (2020).
28. National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) Cancer Stat Facts: cancer of any site. https://seer.cancer.gov/statfacts/html/all.html
29. Centers for Disease Control and Prevention. Mortality in the United States, 2019. NCHS Data Brief No. 395, December 2020. www.cdc.gov/nchs/products/databriefs/db395.htm
30. Chu C, Anderson R, White N, Stone P. Prognosticating for adult patients with advanced incurable cancer: a needed oncologist skill. *Curr. Treat. Options Oncol.* 21(1), 5 (2020).
31. White N, Kupeli N, Vickerstaff V, Stone P. How accurate is the 'Surprise Question' at identifying patients at the end of life? A systematic review and meta-analysis. *BMC Med.* 15(1), 139 (2017).
32. Lakin JR, Robinson MG, Bernacki RE, *et al.* Estimating 1-year mortality for high-risk primary care patients using the 'surprise' question. *JAMA Intern. Med.* 176(12), 1863–1865 (2016).

33. 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).

34. Parikh RB, Manz C, Chivers C, *et al.* Machine learning approaches to predict 6-month mortality among patients with cancer. *JAMA Netw. Open* 2(10), e1915997 (2019).

35. Elfiky AA, Pany MJ, Parikh RB, Obermeyer Z. Development and application of a machine learning approach to assess short-term mortality risk among patients with cancer starting chemotherapy. *JAMA Netw. Open* 1(3), e180926 (2018).

36. Gajra A *et al.* Impact of augmented intelligence (AI) on utilization of palliative care (PC) services in oncology. *J. Clin. Oncol.* 38(Suppl.), Abstract 12015 (2020).

10.2217/fon-2021-0302