CADTH Horizon Scan

Artificial Intelligence Decision Support Tools for End-of-Life Care Planning Conversations

Weiyi Xie
Robyn Butcher
Why Is This an Issue?

- End-of-life care provides support for patients and their families during the last stage of life. End-of-life conversations aim to help people better understand their disease prognosis and expected survival, enabling them to make informed decisions regarding end-of-life care.
- Palliative care focuses on relieving symptoms and improving quality of life for patients with serious or life-threatening diseases. Approximately 89% of patients with life-limiting diseases, such as cancer, can benefit from palliative care. However, not all patients receive it in a timely manner.
- Due in part to prognostic uncertainty and optimism bias, end-of-life planning conversations and palliative care decisions do not occur early enough to have maximum benefit. Interventions that aim to prompt or help identify those patients who can benefit from palliative and/or end-of-life planning could improve the quality of care.

What Is the Technology?

- An artificial intelligence (AI)–based “nudge” is a decision-making support tool that uses prompts and alerts to aid clinicians in deciding whether and when to discuss end-of-life planning with patients.
- The nudge sends alerts and/or reminders to clinicians to prompt end-of-life conversations with patients who are at high risk of short-term mortality. These patients are identified by machine learning mortality prediction algorithms incorporated in the electronic health record (EHR) system.
- Two AI-based nudges designed for patients with cancer were identified. Both tools were developed and internally validated in the US.

What Is the Potential Impact?

- AI-based nudges have the potential to increase the number of end-of-life planning conversations between clinicians and patients as well as the number of referrals to end-of-life services.
- Implementing the nudges into clinical workflows could also help clinicians more easily identify patients with palliative care needs.

What Else Do We Need to Know?

- No AI-based nudges have been approved for use in Canada at the time of this writing nor have there been validation studies using Canadian data.
As with many AI algorithms, there is uncertainty about the validity and generalizability of the mortality predictive algorithms used in the nudges.

The acceptance of AI-based nudges by clinicians is unclear due to varying clinician attitudes and experiences with nudges and because we did not identify any studies that reported the experience of AI-based nudges from the patient perspective.
How It Works
An AI-based behavioural intervention, referred to as a “nudge,” uses alerts and/or reminders to prompt clinicians to initiate end-of-life conversations with patients under their care who are at high risk of short-term (such as 30-day or 180-day) mortality. These patients are identified by machine learning mortality prediction tools incorporated in the EHR system. Through end-of-life conversations, patients can better understand their disease prognosis and expected survival, enabling them to make informed decisions regarding future treatment options. For example, they could decide whether to choose disease-directed treatment or palliative care, which focuses on relieving symptoms and improving quality of life for patients with serious or life-threatening diseases.

Two machine learning AI-based nudges, 1 of which has been commercialized, have been developed to assist oncologists in the US to decide whether and when to discuss palliative care and end-of-life planning with patients. Machine learning is a domain of AI that trains computer algorithms to provide accurate predictions based on identified patterns from historical data. Clinicians often overestimate patient survival. Machine learning–based tools use data from EHRs and/or other sources to provide a more precise estimate of short-term mortality. Moreover, machine learning AI-based nudges can create a list of patients at high risk of 30-day or 180-day mortality, with thresholds tailored to clinician capacity for end-of-life planning conversations. For example, a nudge developed by Manz et al. uses structured electronic record data to identify the 10% of patients at the highest risk of 180-day mortality.

AI-based nudges generally offer weekly notifications within a secure platform, identifying patients at high risk for short-term mortality and supporting the identification of those in need of advance care planning.

For example, the nudge designed by Manz et al. delivers prompts through a secure dashboard, which encourage oncologists to initiate “serious illness conversations” (SICs) through 3 approaches:

- a weekly personalized list of up to 6 highest-risk patients for 180-day mortality, including patient identifiers, the health professional and date of any prior SICs, and a default checkbox for opting out of reminder texts
- an opt-out message sent on the patient’s appointment day to prompt clinicians to consider initiating the SIC
- a weekly email regarding the number of SICs clinicians conducted in the preceding 4 weeks and a peer comparison message tailored to clinicians who performed fewer or more than 8 SICs or ranked among the top 10 performers of SICs during the previous 4 weeks.
Who Might Benefit?

AI-based nudges developed for oncologists\(^1\)\(^-\)\(^6\) may improve early access to palliative care for patients with cancer. In 2022, it was estimated that there were 85,100 cancer-related deaths in Canada.\(^1\)\(^9\) According to a 2023 report from the Canadian Institute for Health Information (CIHI), 89% of people with life-limiting diseases, such as cancer and kidney diseases, can benefit from palliative care.\(^2\)\(^0\) In 2021 to 2022, 58% of people who died (89,000 people) in Canada were described as palliative and had received some form of palliative care.\(^2\)\(^0\)

However, CIHI noted that not everyone in Canada receives palliative care early enough.\(^2\)\(^0\) It reported that, regardless of care setting, half of patients died within 22 days after being identified as palliative.\(^2\)\(^0\) Half of patients receiving care in hospital with identified palliative care needs lived 11 days or fewer following the identification.\(^2\)\(^0\) People with cancer often receive aggressive treatments near the end of life.\(^2\)\(^1\)\(^,\)\(^2\)\(^2\) These patients may experience severe physical and psychosocial symptoms before death. Palliative care can reduce these symptoms and enhance patients’ quality of life.\(^2\)\(^0\)

AI-based nudges have the potential to help clinicians identify patients who could benefit from palliative and end-of-life care planning and facilitate earlier conversations about their treatment goals and end-of-life preferences.

Availability in Canada

Neither of the 2 nudges\(^1\)\(^-\)\(^6\) developed in the US have been approved for use in Canada.

The nudge developed by Manz et al.\(^4\) has not been commercialized, whereas another nudge\(^1\)\(^,\)\(^2\) has been commercialized for, and is available in, the US market.

What Does It Cost?

Although no specific cost information is available, the price of an AI-based tool may include tool usage, training, and ongoing subscriptions for system updates and user support.\(^2\)\(^3\) Implementation training could help clinicians better interpret prediction results, address disagreements with AI predictions, and maintain responsiveness as the volume of system notification grows. Further, evidence suggests that clinicians often feel unprepared to initiate SIC conversations,\(^2\)\(^4\) highlighting training is needed to develop essential communication skills for these difficult conversations. For health care services without an EHR system, additional costs include implementing, maintaining, and training for the EHR software.
Current Practice

In Canada, decisions to initiate palliative care and/or end-of-life care are made by assessing the patient’s condition using tools such as the Palliative Performance Scale. However, many patients are not eligible until they approach the end of their lives.

The Ontario Palliative Care Network recommends the Hospital One-year Mortality Risk (HOMR) tool for early identification of palliative care needs across different care settings. HOMR uses statistical methods to estimate mortality risk and is not AI enhanced. Refer to the Related Developments section for more details.

What Is the Evidence?

Predicting Performance of AI Algorithms Used in the Nudges

The prompt designed by Manz et al. was prospectively validated using new data from the original EHR systems, whereas the AI algorithm of the commercial nudge was validated within the same practice where it was developed. Both tools lack external validation with structurally different data from other health systems, so their performance in different populations is unknown. A summary of prediction performance of the 2 tools is provided in Table 1. In general, AI algorithms of both nudges demonstrated:

- good overall performance in discriminating patients with cancer who had high risk or low risk of 30-day or 180-day mortality, with both receiver operating characteristic curves exceeding 0.8
- accurate identification of patients at low risk of 30-day or 180-day mortality, as indicated by specificities and negative predictive values all exceeding 0.95
- limited ability to identify high-risk patients, with sensitivity values dropping below 0.3, showing that they can capture fewer than 30% of high-risk patients.

In addition, Manz et al. had a positive predictive value of 0.45. This means that for patients identified by the algorithm as having a high risk of 180-day mortality, there is a 45% chance of passing away within 180 days. The positive predictive value of the commercial nudge was unreported, suggesting signs of selective reporting in the validation study.

Improving Palliative and End-of-Life Care Planning in Patients With Cancer

A stepped-wedge cluster randomized trial compared the nudge developed by Manz et al. with usual care in patients with cancer. The study involved 20,506 patients and 41,021 patient encounters, including 5,520 (13.5%) high-risk patient encounters.

For all patient encounters, the study found:

- a significant increase in SIC rates compared to the control period (4.4% vs. 1.3%)
- a significant decrease in end-of-life systemic therapy compared to the control period (7.5% vs. 10.4%)
- no impact on hospice enrolment or length of stay, inpatient death, or end-of-life intensive care unit use.
Among high-risk patient encounters, the study found:

- a significant increase in SIC rates compared to the control period (13.5% vs. 3.4%).

A real-world before-after study reported that the commercial AI-based nudge significantly increased palliative care consults and hospice referrals for patients with cancer. In high- or medium-risk patients, the study found:

- a nearly 2-fold increase in the average number of palliative care consults (after controlling for a 6 month adaptation period)
- a nearly 13-fold increase in hospice referrals compared to baseline.

Issues to Consider

Predicting Performance Concerns
Both AI nudges demonstrated sensitivity less than 0.3, indicating they can correctly identify fewer than 30% of high-risk patients for mortality. Both nudges had a specificity exceeding 0.95, which means they can accurately identify more than 95% of patients without high mortality risk. The high specificity values suggest a potential trade-off between specificity and sensitivity. Sensitivity is generally prioritized for rare event prediction, such as mortality prediction. For example, some researchers preferred high sensitivity over specificity when designing a machine learning mortality predictive model for older adults. One systematic review noted that the low sensitivity of the nudge designed by Manz et al. could lead to its poor performance in identifying patients at high mortality risk. Nevertheless, the emphasis on sensitivity or specificity depends on the models’ purposes. In a prediction model designed to identify patients at high risk of short-term mortality, researchers may prioritize high specificity to minimize the misclassification of low-risk patients as high risk. This is likely particularly important for large clinical practices, where a high number of false-positives would either require clinicians to spend considerable time reviewing patient records and speaking with people who are not at high mortality risk or lead to the nudge being ignored due to the time required for review.

Manz et al. also employed a 40% threshold for high-risk patient identification during validation. This meant that the nudge would flag the top 40% of patients with the greatest risk of 180-day mortality as high risk, while the other patients were classified as low risk. But in real-world applications, researchers used a 10% threshold without assessing the nudge's performance at this threshold. Consequently, it is unknown whether the nudge could predict mortality with a 10% threshold as precisely as it did using a 40% threshold.

Generalizability and Equity Concerns
There is uncertainty regarding the generalizability of AI-based nudges due to the lack of external validation. Researchers have raised concerns about potential data bias in algorithm training datasets, which influences the performance of AI models and may cause health inequities. For example, EHR-based AI models have demonstrated decreased performance and calibration across various geographic locations and
over time.\textsuperscript{31} AI models also showed poor performance in historically underrepresented groups,\textsuperscript{31} potentially reinforcing existing inequities.

External validation using data from other health systems can test whether the algorithms have site bias or centre bias,\textsuperscript{32} which may result from variations in treatment protocols. It can also detect sampling bias and whether the models’ predictive performance changes when applied to new cohorts with different demographics.\textsuperscript{32} However, current AI-based nudges lack external validation. Therefore, it is unclear whether they can predict mortality as accurately in Canadian health care systems as they did in the development setting. For example, although the nudge developed by Manz et al.\textsuperscript{4} was prospectively validated using new data in the original system to detect algorithm overfitting to the development dataset, it cannot evaluate the nudge’s performance in other EHR systems. Therefore, this prospective internal validation cannot guarantee safe application of the model in other populations.

**Clinician Attitudes and Responses**

The effect of the AI-based nudges in clinical practice can be influenced by clinician acceptance. Evidence showed that clinicians varied in their attitudes toward machine learning–based clinical support tools.\textsuperscript{27}

> Regarding using AI–based nudges in routine care of patients with cancer, some clinicians believed it could aid in validating their own prognostic estimates and prompt conversations about treatments patients prefer to receive near the end of their lives.\textsuperscript{33}

However, other clinicians have concerns about algorithm accuracy, over-reliance on algorithm predictions, and the ethical implications related to disclosure of algorithm predictions.\textsuperscript{33}

In addition, clinician response rates to the alerts can affect the effectiveness of these nudges because clinicians are responsible for making clinical decisions, with the nudges serving as supportive tools. Behavioural studies revealed that, over time, the impact of EHR alerts decreased due to alert fatigue.\textsuperscript{34,35} Health care providers will pay less attention to alerts or reminders when responding to them requires a lot of time or a lot of effort.\textsuperscript{36} In a secondary analysis of a randomized controlled trial, the nudge developed by Manz et al.\textsuperscript{4} was applied in clinical practice and found that specialists with low patient volumes were more responsive than general oncologists and specialists with higher patient volumes.\textsuperscript{37} Therefore, low response rates from clinicians due to limited capacity or alert fatigue may be an important barrier to implementation, especially for nudges sending multiple prompts.

**Related Developments**

HOMR is an EHR electronic medical record–integrated application that automatically calculates 12-month mortality risk for newly admitted patients.\textsuperscript{38,39} Based on multivariable binomial logistic regression modelling, it predicts mortality risk using patient demographic variables and health administrative data.\textsuperscript{39} HOMR was internally validated using population-based health administrative databases in Ontario\textsuperscript{39} and externally
validated in adult, nonpsychiatric, inpatient acute care settings in Ontario and Alberta as well as Boston, Massachusetts. HOMR can notify the admitting medical team of high-risk patients using an electronic sign-out tool within the EHR. It can also send email alerts to the admitting physician for consideration of palliative and end-of-life care needs. The Ontario Palliative Care Network recommends using HOMR in acute care settings. Some hospitals in Ontario have implemented HOMR either in practice or within clinical studies. In addition, the latest version, HOMR Now!, calculates risk immediately upon patient admission, which was internally validated using data from a tertiary-care teaching hospital in Ontario.

Researchers at Washington University created and internally validated an AI-based nudge for the general inpatient population in a community setting. It notified physicians when a patient was identified as having a high risk of 30-day mortality to prompt the discussion about care goals before discharge. The AI model can predict short-term mortality or hospice outcomes on the second day of a patient's admission. However, details about the delivery method and frequency of these alerts were unknown.

An AI-based behavioural nudge triggering palliative care intervention for hospitalized patients was developed and internally validated in the US. Unlike nudges using mortality as a proxy for potential palliative care needs, this tool identified patients at low risk of short-term mortality who may still benefit from palliative care. It was developed by learning the associations between variables in the electronic medical record and palliative care consultation. The nudge was delivered through a clinical Control Tower, where Control Tower Operators and palliative care service clinicians evaluated potential palliative care needs of patients flagged by the nudge.

CADTH released a 2022 report introducing clinical applications of AI. The report included developments and application of predictive AI models that can estimate disease progression, patient outcomes, and overall survival across various medical conditions.

Looking Ahead

AI-based behavioural nudges have the potential to assist clinicians in identifying high-risk patients for short-term mortality and initiating end-of-life planning conversations, thus improving timely referrals to palliative care or hospice services. Although there are nudges for patients with cancer that have been tested in real-world settings, the predictive performance and generalizability are uncertain due to the lack of external validation. In addition, the acceptance of AI prognostic prediction tools is unclear due to varying clinician attitudes and unknown patient perspectives. More validation studies are required to confirm the robustness and generalizability of machine learning AI models in various populations.

More real-world research is needed to investigate clinician and patient acceptance of these nudges.

Additional research that examines the use of AI-based nudges for patients with nonmalignant diseases and those who have reduced access to palliative care compared with patients with cancer would be helpful to
understand whether the AI-based tools are helpful for conditions other than cancer. A 2023 report from CIHI indicated that among all patients with life-limiting illness in Canada, individuals with end-stage cancer were 3 times more likely to receive palliative care than those with other conditions. Developing AI-based nudges and/or other interventions for patients with nonmalignant diseases may facilitate equitable access to palliative care services.
References


Appendix 1: Methods

Note that this appendix has not been copy-edited.

Literature Search Strategy

An information specialist conducted a literature search on key resources including MEDLINE, Embase, the Cochrane Database of Systematic Reviews, the International HTA Database, the websites of Canadian and major international health technology agencies, as well as a focused internet search. The search approach was customized to retrieve a limited set of results, balancing comprehensiveness with relevancy. The search strategy comprised both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. Search concepts were developed based on the elements of the research questions and selection criteria. The main search concepts were artificial intelligence and end of life. The search was completed on September 6, 2023, and limited to English-language documents published since January 1, 2018.

Selection Criteria

One author screened the literature search results and reviewed the full text of all potentially relevant studies. Studies were considered for inclusion if the intervention was a combination of AI-based mortality predictive models and behavioural interventions targeting clinicians to initiate end-of-life conversations with patients. Conference abstracts and grey literature were included when they provided additional information to that available in the published studies.
Appendix 2: Additional Information

Note that this appendix has not been copy-edited.

Table 1: Characteristics and Predicting Performance of the AI Algorithms of the Included Behavioural Nudges

<table>
<thead>
<tr>
<th>AI-based behavioural nudge</th>
<th>A commercial nudge developed for the US market(^1,2)</th>
<th>AI-based nudge developed by researchers at University of Pennsylvania(^4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>US</td>
<td>US</td>
</tr>
<tr>
<td>Type of disease</td>
<td>All cancer</td>
<td>All cancer</td>
</tr>
<tr>
<td>Algorithm</td>
<td>N-dimensional eigenspace</td>
<td>Gradient-boosted tree</td>
</tr>
<tr>
<td>Input features</td>
<td>Clinical data and billing information (e.g., diagnosis codes, assessments, laboratories, medications, cancer staging information, vitals, and screenings) from the EHR</td>
<td>Structured EHR data including demographic, clinicopathologic, laboratory, comorbidity, and electrocardiogram data</td>
</tr>
<tr>
<td></td>
<td>Socioeconomic data(^a) (e.g., income, household size, transportation) from publicly available resources including US Census Bureau, US Department of Agriculture and the National Oceanic and Atmospheric Administration</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Behavioural data(^a) (e.g., history of internet searches on health conditions, purchasing channels and life stage) from third-party data vendors such as Acxiom, Experian, and Transunion</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>30-day mortality</td>
<td>180-day mortality</td>
</tr>
<tr>
<td>Validation</td>
<td>Internally validated in 3,671 patients in a community practice in the Pacific Northwest of the US</td>
<td>Internally validated in 24,582 patients in a tertiary practice within the University of Pennsylvania Health System</td>
</tr>
<tr>
<td></td>
<td>No external validation</td>
<td>No external validation</td>
</tr>
<tr>
<td></td>
<td>Threshold for high-risk patients: 5%</td>
<td>Threshold for high-risk patients: 40%</td>
</tr>
<tr>
<td>Area under the receiver operating characteristic curve</td>
<td>0.86</td>
<td>0.89</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.28</td>
<td>0.27</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.95</td>
<td>0.99</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>Not reported</td>
<td>0.45</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.99</td>
<td>0.97</td>
</tr>
</tbody>
</table>

EHR = electronic health record.

\(^a\)Socioeconomic and behavioural data were used during model development. It is unclear whether the commercial nudge retrieves patients' behavioural and socioeconomic data out of the EHR system to predict individual mortality risk in practical use.