Long-term Effect of Machine Learning–Triggered Behavioral Nudges on Serious Illness Conversations and End-of-Life Outcomes Among Patients With Cancer
A Randomized Clinical Trial

Christopher R. Manz, MD; Yichen Zhang, PhD; Kan Chen, MA; Qi Long, PhD; Dylan S. Small, PhD; Chalanda N. Evans, BS; Corey Chivers, PhD; Susan H. Regli, PhD; C. William Hanson, MD; Justin E. Bekelman, MD; Jennifer Braun, MHA, RN, BSN; Charles A. L. Rareshide, MS; Nina O’Connor, MD; Pallavi Kumar, MD, MPH; Lynn M. Schuchter, MD; Lawrence N. Shulman, MD; Mitesh S. Patel, MD, MBA; Ravi B. Parikh, MD, MPP

IMPORTANCE Serious illness conversations (SICs) between oncology clinicians and patients are associated with improved quality of life and may reduce aggressive end-of-life care. However, most patients with cancer die without a documented SIC.

OBJECTIVE To test the impact of behavioral nudges to clinicians to prompt SICs on the SIC rate and end-of-life outcomes among patients at high risk of death within 180 days (high-risk patients) as identified by a machine learning algorithm.

DESIGN, SETTING, AND PARTICIPANTS This prespecified 40-week analysis of a stepped-wedge randomized clinical trial conducted between June 17, 2019, and April 20, 2020 (including 16 weeks of intervention rollout and 24 weeks of follow-up), included 20,506 patients with cancer representing 41,021 encounters at 9 tertiary or community-based medical oncology clinics in a large academic health system. The current analyses were conducted from June 1, 2021, to May 31, 2022.

INTERVENTION High-risk patients were identified using a validated electronic health record machine learning algorithm to predict 6-month mortality. The intervention consisted of (1) weekly emails to clinicians comparing their SIC rates for all patients against peers’ rates, (2) weekly lists of high-risk patients, and (3) opt-out text messages to prompt SICs before encounters with high-risk patients.

MAIN OUTCOMES AND MEASURES The primary outcome was SIC rates for all and high-risk patient encounters; secondary end-of-life outcomes among decedents included inpatient death, hospice enrollment and length of stay, and intensive care unit admission and systemic therapy close to death. Intention-to-treat analyses were adjusted for clinic and wedge fixed effects and clustered at the oncologist level.

RESULTS The study included 20,506 patients (mean [SD] age, 60.0 [14.0] years) and 41,021 patient encounters: 22,259 (54%) encounters with female patients, 28,907 (70.5%) with non-Hispanic White patients, and 5520 (13.5%) with high-risk patients; 1417 patients (6.9%) died by the end of follow-up. There were no meaningful differences in demographic characteristics in the control and intervention periods. Among high-risk patient encounters, the unadjusted SIC rates were 3.4% (59 of 1754 encounters) in the control period and 13.5% (510 of 3765 encounters) in the intervention period. In adjusted analyses, the intervention was associated with increased SICs for all patients (adjusted odds ratio, 2.09 [95% CI, 1.53-2.87]; P < .001) and decreased end-of-life systemic therapy (7.5% [72 of 957 patients] vs 10.4% [24 of 231 patients]; adjusted odds ratio, 0.25 [95% CI, 0.11-0.57]; P = .001) relative to controls, but there was no effect on hospice enrollment or length of stay, inpatient death, or end-of-life ICU use.

CONCLUSIONS AND RELEVANCE In this randomized clinical trial, a machine learning–based behavioral intervention and behavioral nudges to clinicians led to an increase in SICs and reduction in end-of-life systemic therapy but no changes in other end-of-life outcomes among outpatients with cancer. These results suggest that machine learning and behavioral nudges can lead to long-lasting improvements in cancer care delivery.

TRIAL REGISTRATION ClinicalTrials.gov Identifier: NCT03984773

Published online January 12, 2023.
Serious illness conversations (SICs) between oncology clinicians and patients with cancer focus on patients’ understanding of and priorities around their illness.1 These conversations are appropriate at any stage of cancer, improve mood and quality of life, and may reduce end-of-life systemic therapy use.2–3 However, most patients die without an SIC.4 Clinician factors, including inaccurate prognostication and perceptions that SICs are appropriate only close to death, contribute to this dearth of SICs.5 To address these factors, we conducted a stepped-wedge, cluster-randomized clinical trial that evaluated behavioral nudges to clinicians to prompt SICs among patients at high risk for death within 180 days (high-risk patients) identified by a machine learning algorithm.6 A prior analysis after 16 weeks showed significant increases in SICs among all and high-risk patients. Here we report the intervention impact on pre-specified 40-week SIC and end-of-life outcomes.

Methods

The 16-week results of this trial were previously published.6 The trial protocol was approved by the University of Pennsylvania institutional review board (Supplement 1), which granted a waiver of informed consent owing to minimal risk. Patients with encounters at 1 of 9 tertiary and community-based medical oncology clinics between June 17, 2019, and April 20, 2020, without a prior documented SIC were eligible. The current analyses were conducted from June 1, 2021, to May 31, 2022. Details of randomization are shown in Figure 1. We identified high-risk patients using a prospectively validated machine learning algorithm predicting 6-month mortality; patients with a 10% or higher absolute mortality risk were deemed high risk.6,7 The intervention consisted of (1) weekly emails to clinicians comparing their SIC rates for all patients against peers’ rates, (2) weekly lists of 6 or more forthcoming encounters with high-risk patients, and (2) opt-out reminder texts to clinicians on the morning of encounters with high-risk patients. While emails reported SICs for all patients, opt-out text messages only applied to encounters with high-risk patients. After a 4-week baseline wedge in which all groups remained in usual care (control), groups were randomly assigned to the intervention in 4-week wedges until all clinics received the intervention by the fifth wedge (week 17) and were followed up through week 40.

Primary outcomes at 40 weeks (after 16 weeks of rollout and 24 weeks of follow-up) included SIC rates for all, high-risk, and non–high-risk patients, along with subgroup analyses stratified by age, self-reported race and ethnicity, sex, insurance status, and clinic type. A specific type of note indicated SICs in the advance care planning module of the electronic medical record. End-of-life outcomes included inpatient death, intensive care unit use within 30 days of death, systemic therapy (chemotherapy or checkpoint inhibitor therapy) within 14 days of death, hospice enrollment prior to death, and hospice length of stay.5,8 End-of-life outcomes were derived from an institutional cancer registry and were available for 1188 patients with 3 or more oncology encounters in the year prior to death; hospice outcomes were available for 856 patients eligible for a health system-affiliated hospice. Decedents were assigned to control vs intervention group based on intervention status on the date of the last clinic encounter. Outcomes were analyzed at either the patient-encounter (ie, SIC) or patient (end-of-life) level using generalized estimating equations with a logit link (binary outcomes) or gamma link (continuous outcomes) and an independent correlation structure using oncologist as the clustering unit to account for between-clinician variation, with clinic-group and wedge-period fixed effects. Heterogeneity of intervention effects was evaluated using (subgroup × intervention) interaction terms. Analyses were performed using R statistical software, version 4.2.1 (R Project for Statistical Computing). A 2-tailed \( P < .05 \) was considered statistically significant.

Results

The study included 20,506 patients (mean [SD] age, 60.0 [14.0] years) and 41,021 patient encounters: 22,259 (54%) encounters with female patients, 8,175 (19.9%) with non-Hispanic Black patients, 28,907 (70.5%) with non-Hispanic White patients, and 3,939 (9.6%) with patients categorized as other (including American Indian, Asian, East Indian, Hispanic Black, Hispanic White, and Pacific Islander); 3,520 encounters (13.5%) were with high-risk patients (eTable 1 in Supplement 2). There were no meaningful baseline differences between control and intervention periods. For the overall study population, SIC rates increased during the initial rollout and plateaued at a higher baseline in the follow-up period (Figure 2). Among all patient encounters, the unadjusted SIC rates were 1.3% (158 of 12,356 encounters) and 4.4% (1,262 of 28,665 encounters) in the control and intervention periods, respectively (Table). Among encounters with high-risk patients, the unadjusted SIC rate was 3.4% (59 of 1754 encounters) and 13.5% (510 of 3,765 encounters) in the control and intervention periods, respectively. Among encounters with non–high-risk patients, the unadjusted SIC rates were 0.9% (99 of 10,602 encounters) and 3.0% (752 of 24,899 encounters) in the control and intervention periods, respectively. The intervention was associated with significant increases in SICs among all patient encounters (adjusted odds ratio [aOR], 2.09 [95% CI, 1.53–2.87]; \( P < .001 \)), those with high-risk patients (aOR, 2.62 [95% CI, 1.84–3.72]; \( P < .001 \)), and those with non–high-risk patients (aOR, 2.07 [95% CI, 1.52–2.82]; \( P < .001 \)). The intervention effect was significant across all subgroups.
interaction models revealed a greater impact on SIC rates among Medicaid beneficiaries (5.2% of encounters in the intervention group vs 0.9% in the control group; aOR, 2.19 [95% CI, 1.10-4.37]; P = .03 for interaction). No other significant interaction effects were found. There were 419 SICs (29.5% of all SICs) among patients who died; the unadjusted SIC rates were 19.5% (42 of 215 patients) and 31.4% (377 of 1202 patients) in the control and intervention periods, respectively (aOR, 2.44 [95% CI, 1.35-4.42]; P < .001).

Intervention patients were less likely than control patients to receive systemic therapy at the end of life (all decedents: 7.5% [72 of 957 patients] vs 10.4% [24 of 231 patients]; aOR, 0.25 [95% CI, 0.11-0.57]; P = .001; high-risk decedents: 7.2% [50 of 696 patients] vs 8.3% [15 of 181 patients]; aOR, 0.37 [95% CI, 0.13-1.04]; P = .06) (Table). There were no differences between control and intervention groups in other end-of-life outcomes.

**Discussion**

In this randomized clinical trial, a machine learning–based intervention increased SICs and decreased end-of-life systemic therapy among patients with cancer. Our findings suggest that machine learning–based interventions may lead to long-term reduction in low-value cancer care and offer several lessons. First, our intervention led to a sustained increase in SICs, and rates remained nearly 4 times above baseline after 6 months. A prior effort to improve serious illness communication has found
inconsistent long-term results. Our intervention's sustained effect on SICs likely stemmed from the integration of proven behavioral interventions, including peer comparisons, performance reports, and opt-out reminders. Of note, our intervention increased SICs even among patients not flagged by the algorithm (ie, as a spillover).

Second, our subgroup analyses suggest that intervention effects were as strong or stronger among Medicaid beneficiaries, who have historically lower SIC rates. Our algorithm-based intervention may have had a larger impact on disadvantaged groups by addressing clinician-level biases that contributed to lower baseline SIC rates.

Third, our intervention decreased end-of-life systemic therapy. This impact stands in contrast to a prior trial, which found no impact of SIC education alone on end-of-life outcomes. Practices participating in performance-based payment models may find that machine learning-based interventions facilitate higher-quality end-of-life care.

**Limitations**

This study has limitations. First, this single-health system trial may not be generalizable to other systems. Second, our intervention's impact on SIC and end-of-life outcomes may have been limited because it nudged time-constrained oncologists. Future machine learning-based nudges could involve special palliative care referral and nonclinician staff. Third, there are no established standards defining a clinically meaningful impact on end-of-life systemic therapy. However, the effect estimate for end-of-life systemic therapy compares similarly to a prior randomized evaluation of early palliative care. Finally, SICs have improved mental health, quality of life, and emotional distress in prior work, but our study did not measure these outcomes. Future prospective SIC studies should prioritize patient-reported outcome measurement.

**Conclusions**

In this randomized clinical trial, a machine learning-based behavioral intervention led to a sustained increase in SICs and reduction in end-of-life systemic therapy among outpatients with cancer, suggesting that machine learning-based interventions can lead to long-lasting improvements in cancer care delivery.
Effect of Machine Learning-Triggered Nudges on Serious Illness Conversations With Patients With Cancer

ARTICLE INFORMATION

Accepted for Publication: September 29, 2022.
Published Online: January 12, 2023.
Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2023 Manz CR et al. JAMA Oncology.

Author Affiliations: Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts (Manz); Harvard Medical School, Boston, Massachusetts (Manz); Division of Health Policy, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Zhang, Bekelman, Parikh); Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Chivers, O’Connor, Shulman, Parikh); Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Evans, Rareshide, O’Connor, Kumar, Schuchter, Shulman, Parikh); Penn Medicine, Philadelphia, Pennsylvania (Chivers, Regli, Hanson, Braun); Penn Center for Cancer Care Innovation, Abramson Cancer Center, University of Pennsylvania, Philadelphia (Bekelman, Schuchter, Shulman, Parikh); Corporal Michael J. Crescenz Veterans Affairs Medical Center, Philadelphia, Pennsylvania (Kumar, Parikh); Ascension Health, Veterans Affairs Medical Center, Philadelphia, Pennsylvania (Chivers, O’Connor, Kumar, Schuchter, Shulman, Parikh); Corporal Michael J. Crescenz Pennsylvania, Philadelphia (Bekelman, Schuchter, Shulman, Parikh); Corporal Michael J. Crescenz Veterans Affairs Medical Center, Philadelphia, Pennsylvania (Kumar, Parikh); Ascension Health, St Louis, Missouri (Patel).

Author Contributions: Mr Chen and Dr Parikh had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Manz, Zhang, Chivers, Regli, Hanson, Braun, O’Connor, Shulman, Patel, Parikh.

Acquisition, analysis, or interpretation of data: Manz, Zhang, Chen, Long, Small, Evans, Chivers, Bekelman, Braun, Rareshide, O’Connor, Kumar, Schuchter, Shulman, Parikh.

Drafting of the manuscript: Manz, Zhang, Evans, Bekelman, Parikh.

Critical revision of the manuscript for important intellectual content: Manz, Zhang, Chen, Long, Small, Chivers, Regli, Hanson, Braun, Bekelman, Braun, Rareshide, O’Connor, Kumar, Schuchter, Shulman, Patel, Parikh.

Statistical analysis: Manz, Zhang, Chen, Long, Small, Chivers, Rareshide, Parikh.

Obtained funding: Manz, Parikh.

Administrative, technical, or material support: Manz, Zhang, Evans, Braun, Shulman, Patel, Parikh.

Supervision: Manz; Small, Hanson, O’Connor, Schuchter, Shulman, Patel, Parikh.

Conflict of Interest Disclosures: Dr Long reported receiving grants from the National Institutes of Health (NIH) during the conduct of the study and grants from Pfizer and Bayer outside the submitted work. Dr Bekelman reported receiving personal fees from United Healthcare, grants from UnitedHealth Group, and personal fees from the National Comprehensive Cancer Network (NCCN) and Reimagine Care outside the submitted work. Dr Patel reported having part ownership of Catalyst Health and receiving personal fees from Humana outside the submitted work. Dr Parikh reported receiving grants from the NIH, Prostate Cancer Foundation, National Palliative Care Research Center, NCCN Foundation, Conquer Cancer Foundation, Humana, Emerson Collective, and Veterans Health Administration; personal fees and equity from GNS Healthcare, Thyme Care, and Onc.AI; personal fees from the Cancer Study Group, Biofourms, Humana, and Nanolory; honoraria from Flatiron and Medscape; having board membership (unpaid) at the Coalition to Transform Advanced Cancer and American Cancer Society; and serving on a leadership consortium (unpaid) at the National Quality Forum outside the submitted work. No other disclosures were reported.

Funding/Support: This study was supported by NIH grant SKO2CA263541 and the Penn Center for Precision Medicine (Drs Manz and Parikh) and NIH grant KO8CA263541 (Dr Parikh).

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

Meeting Presentation: This study was presented at the American Society of Clinical Oncology Annual Meeting; June 1, 2022; Chicago, Illinois.

Data Sharing Statement: See Supplement 3.

REFERENCES


© 2023 Manz CR et al. JAMA Oncology.