Quality of Life Research
https://doi.org/10.1007/s11136-024-03675-3

Optimization of alert notifications in electronic patient-reported outcome (ePRO) remote symptom monitoring systems (AFT-39)

Gina L. Mazza[1,2] · Amylou C. Dueck[1,2] · Brenda Ginos[1,2] · Jennifer Jansen[3] · Allison M. Deal[3] · Philip Carr[3] · Victoria S. Blinder[4] · Gita Thanarajasingam[5] · Mattias Jonsson[3] · Minji K. Lee[6,7] · Lauren J. Rogak[2] · Gita N. Mody[3] · Deborah Schrag[4] · Ethan Basch[3]

Accepted: 29 April 2024

Abstract

Purpose Clinical benefits result from electronic patient-reported outcome (ePRO) systems that enable remote symptom monitoring. Although clinically useful, real-time alert notifications for severe or worsening symptoms can overburden nurses. Thus, we aimed to algorithmically identify likely non-urgent alerts that could be suppressed.

Methods We evaluated alerts from the PRO-TECT trial (Alliance AFT-39) in which oncology practices implemented remote symptom monitoring. Patients completed weekly at-home ePRO symptom surveys, and nurses received real-time alert notifications for severe or worsening symptoms. During parts of the trial, patients and nurses each indicated whether alerts were urgent or could wait until the next visit. We developed an algorithm for suppressing alerts based on patient assessment of urgency and model-based predictions of nurse assessment of urgency.

Results 593 patients participated (median age = 64 years, 61% female, 80% white, 10% reported never using computers/tablets/smartphones). Patients completed 91% of expected weekly surveys. 34% of surveys generated an alert, and 59% of alerts prompted immediate nurse actions. Patients considered 10% of alerts urgent. Of the remaining cases, nurses considered alerts urgent more often when patients reported any worsening symptom compared to the prior week (33% of alerts with versus 26% without any worsening symptom, $p = 0.009$). The algorithm identified 38% of alerts as likely non-urgent that could be suppressed with acceptable discrimination (sensitivity = 80%, 95% CI [76%, 84%]; specificity = 52%, 95% CI [49%, 55%]).

Conclusion An algorithm can identify remote symptom monitoring alerts likely to be considered non-urgent by nurses, and may assist in fostering nurse acceptance and implementation feasibility of ePRO systems.

Keywords Symptom · Remote monitoring · Alert notification · Supportive care · Symptom report · Electronic patient-reported outcome

✉ Gina L. Mazza
Mazza.Gina@mayo.edu

1 Alliance Foundation Trials Statistics and Data Center, Mayo Clinic, 13400 East Shea Boulevard, Scottsdale, AZ 85259, USA

2 Department of Quantitative Health Sciences, Mayo Clinic, Scottsdale, AZ, USA

3 Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

4 Memorial Sloan Kettering Cancer Center, New York, NY, USA

5 Division of Hematology, Mayo Clinic, Rochester, MN, USA

6 Alliance Foundation Trials Statistics and Data Center, Mayo Clinic, Rochester, MN, USA

7 Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA

Plain English summary

Systematic monitoring of symptoms using surveys completed by patients between scheduled visits with their care team has been shown to benefit patients receiving cancer treatment and is of substantial and increasing interest for use in routine oncology practice. Severe or worsening symptoms reported via these surveys generate automated alert notifications to the care team (generally a nurse), enabling them to react and manage concerning symptoms in near real time. In a trial in which oncology practices implemented remote symptom monitoring, participating nurses indicated that while they found alerts to be helpful, about half of nurses felt they received too many alerts. Nurses opted not to act immediately on 41% of alerts, suggesting the potential

Published online: 21 May 2024

to unburden nurses by identifying and suppressing alerts upfront that are likely non-urgent. Therefore, we sought to optimize alerts based on patient assessment of urgency and model-based predictions of nurse assessment of urgency, and we identified 38% of alerts from the trial that could have been suppressed while missing few urgent alerts. Our results suggest that nurses generally perceive new or worsening symptoms to be clinically urgent warranting immediate action, while stable ongoing symptoms, even if severe, are more often regarded as appropriate for discussion at the next visit. This finding seems driven by persistently high symptoms already being known by the care team. Reducing the number of alert notifications can encourage the use of remote symptom monitoring in routine oncology practice by fostering nurse acceptance and feasibility of reviewing and responding to alerts.

Introduction

Symptoms are common among patients receiving treatment for advanced cancers, and are a major source of complications and morbidity [1–4]. Systematic monitoring of symptoms using electronic patient-reported outcomes (ePROs) has been shown to yield clinical benefits and is of substantial and increasing interest for deployment in routine oncology practice [5, 6].

In general, remote symptom monitoring involves an electronic data capture system that regularly administers surveys to patients regarding their symptoms (Fig. 1). Patients complete the surveys between clinic visits using a computer, smart device, or interactive voice response system (i.e., automated telephone call). Severe or worsening symptoms generate automated alert notifications to the care team (generally a nurse), enabling the care team to react and manage concerning symptoms in near real time. Longitudinal symptom reports are available during scheduled visits to guide discussions and care.

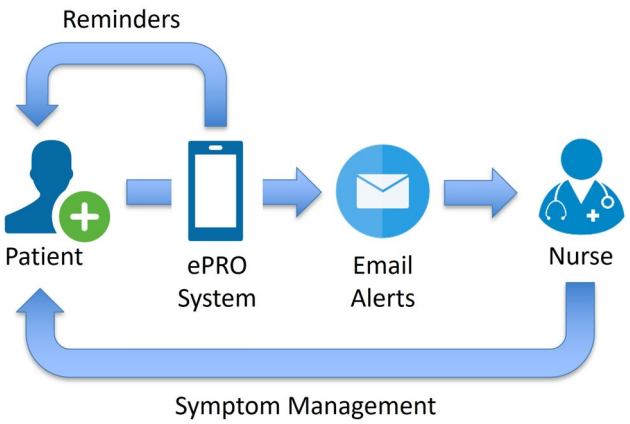

Fig. 1 Information flow with integration of electronic patient-reported outcomes for symptom monitoring in routine oncology care

PRO-TECT (Alliance AFT-39) was a U.S. national trial evaluating ePRO remote symptom monitoring versus usual care in patients receiving metastatic cancer treatment (ClinicalTrials.gov: NCT03249090) [7]. As previously described, patients receiving the intervention reported greater physical function, symptom control, and health-related quality of life at 3 months compared to patients receiving usual care. Most patients said that remote symptom monitoring made them feel in greater control of their care (77%, 387/504) and improved discussions with their care team (72%, 365/504). Among participating nurses, surveys and qualitative interviews indicated that while nurses found alerts to be helpful (44/58, 76%), about half of nurses felt they received too many alerts (28/57, 49%) [7]. Qualitative interviews confirmed that although nurses valued the alerts in caring for patients, reviewing and responding to alerts were perceived as clinically burdensome, thus warranting strategies for reducing the number of alert notifications. Notably, only 59% (4122/6947) of alerts prompted immediate nurse actions in this trial, suggesting the potential to algorithmically identify alerts that are likely to be considered non-urgent by nurses—and thereby could be suppressed from generating a real-time notification (i.e., could wait for discussion at the next visit). Therefore, using data from the PRO-TECT trial, we sought to develop a model-based algorithm to optimize alerts, based on the hypothesis that patient and symptom characteristics could contribute to predicting which events nurses considered urgent or non-urgent.

Methods

Trial design and patients

PRO-TECT (Alliance AFT-39) was a multicenter, cluster-randomized trial evaluating remote symptom monitoring with weekly ePRO surveys, conducted in U.S. community oncology practices, and has previously been described [7]. Each practice could enroll up to 50 adult patients with metastatic cancer of any type (except leukemia or indolent lymphoma) receiving chemotherapy, targeted oral therapy, and/or immunotherapy. Central and local institutional review boards provided approval, and patients provided written informed consent.

Patients at PRO-TECT intervention practices completed weekly surveys for 1 year or until they discontinued all cancer treatment. At baseline, patients selected a day of the week and time to receive an email or automated telephone call prompting survey completion. Patients also selected the mode of administration (i.e., online versus interactive voice response system). Patients who did not complete the weekly

survey after 24 h received an automated reminder. If needed, after 72 h, a research staff member called patients to give a reminder or verbally administer the survey. Caregivers and research staff could assist patients with survey completion.

As described elsewhere [7], the weekly survey included items from the National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) [8, 9] for pain (frequency, severity, and interference), nausea (frequency and severity), vomiting (frequency), diarrhea (frequency), dyspnea (severity and interference), and insomnia (severity) and additional items assessing food/fluid intake, performance status (per Eastern Cooperative Oncology Group criteria [10]), depression, falls, and (every four weeks) financial toxicity. Higher scores indicated greater symptom burden for all items (Table 1). A final item asked patients to indicate whether they had any other symptoms, with the online system providing a "free text" box for patients to enter that information.

Figure 1 illustrates the information flow and clinical workflow for symptom monitoring, alert notifications, and longitudinal reports. The study team conducted a literature review and engaged patients, clinicians, health services researchers, scientific advisory board members, and Alliance for Clinical Trials in Oncology committee members to develop the weekly surveys and set criteria for sending alert notifications [11]. Whenever a patient reported a concerning symptom, defined by a prespecified high score or worsened score compared to the prior week (Table 1), the practice's clinical research assistant received an automated email alert notification to forward to the responsible clinical nurse. The alert included the patient's survey response(s) for the concerning symptom(s) and links to standardized symptom management pathways (Supplementary Fig. S1). Nurses had discretion for whether and how to address alerts, in accordance with their judgment regarding whether an immediate action was clinically indicated. Within 72 h, the clinical research assistant documented what action(s), if any, the nurse took in response to the alert. An immediate nurse action was defined as contacting the patient or caregiver by telephone, email, patient portal, or in person; prescribing

Table 1 Criteria for sending alert notifications to nurses for patient-reported outcomes in the PRO-TECT trial

Symptom	Response options	Criteria for alerting care team
Decreased food/fluid intake	Not at All, A Little Bit, Somewhat, Quite a Bit, Very Much	"Quite a Bit" or "Very Much" or 2-Point Worsening
Pain frequency	Never, Rarely, Occasionally, Frequently, Almost Constantly	"Frequently" or "Almost Constantly" or 2-Point Worsening
Pain severity	None, Mild, Moderate, Severe, Very Severe	"Severe" or "Very Severe" or 2-Point Worsening
Pain interference	Not at All, A Little Bit, Somewhat, Quite a Bit, Very Much	"Quite a Bit" or "Very Much" or 2-Point Worsening
Performance status	Normal With No Limitations (0) to Pretty Much Bedridden, Rarely Out of Bed (4)	Patient-Reported ECOG Performance Status of 3 or 4 or 2-Point Worsening[a]
Depression	Not at All, Several Days, More Than Half the Days, Nearly Every Day	"More Than Half the Days" or "Nearly Every Day"
Nausea frequency	Never, Rarely, Occasionally, Frequently, Almost Constantly	"Frequently" or "Almost Constantly" or 2-Point Worsening
Nausea severity	None, Mild, Moderate, Severe, Very Severe	"Severe" or "Very Severe" or 2-Point Worsening
Vomiting frequency	Never, Rarely, Occasionally, Frequently, Almost Constantly	"Frequently" or "Almost Constantly" or 2-Point Worsening
Diarrhea frequency	Never, Rarely, Occasionally, Frequently, Almost Constantly	"Frequently" or "Almost Constantly" or 2-Point Worsening
Constipation severity	None, Mild, Moderate, Severe, Very Severe	"Severe" or "Very Severe" or 2-Point Worsening
Dyspnea severity	None, Mild, Moderate, Severe, Very Severe	"Severe" or "Very Severe" or 2-Point Worsening
Dyspnea interference	Not at All, A Little Bit, Somewhat, Quite a Bit, Very Much	"Quite a Bit" or "Very Much" or 2-Point Worsening
Insomnia severity	None, Mild, Moderate, Severe, Very Severe	"Severe" or "Very Severe" or 2-Point Worsening
Have fallen	No, Yes	"Yes"
Financial toxicity	Not at All, A Little Bit, Somewhat, Quite a Bit, Very Much	"Quite a Bit" or "Very Much"
Any other symptoms	No, Yes	"Yes"

[a]For patient-reported Eastern Cooperative Oncology Group (ECOG) performance status, alert notifications were initially sent if a 1-point worsening occurred. Due to a large number of resulting alerts, this criterion was changed to a 2-point worsening on March 21, 2018 based on feedback from nurses, evidence suggesting that nurses did not act upon these alerts, and deliberations with the PRO-TECT advisory board and patient representatives

or modifying supportive medications; modifying or holding chemotherapy; providing advice for self-management; scheduling an appointment for evaluation in clinic; ordering an imaging or laboratory test; referring the patient to another clinic; or sending the patient to the emergency room, urgent care, or hospital for admission. During scheduled visits, information from the weekly surveys, including each symptom's trajectory over (up to) 10 weeks, were summarized in a report shared with the care team.

To evaluate whether contextual information from patients would be helpful in developing an algorithm for prioritizing alerts, during two testing periods of the trial (May 30, 2019 to August 19, 2019 and December 18, 2019 to January 27, 2021), for research purposes only, patients were asked the following question whenever they reported a concerning symptom that would generate an alert: "You have reported one or more concerning symptoms; would you like a message sent to your cancer care team about this now, or can it wait until your next appointment?" Patients could select "yes, report now" or "no, wait until next appointment." Regardless of the patient's selection, the nurse received an alert notification, and the patient's response to this question was not shared with the nurse.

Outcome

During two testing periods of the trial (December 17, 2019 to April 15, 2020 and April 27, 2020 to February 1, 2021), nurses were asked the following question regarding the urgency of each alert: "In the opinion of the clinical nurse, could the symptom(s) reported in this alert likely have waited to be addressed until this patient's next appointment?" Nurses could select "no, likely could not have waited" or "yes, likely could have waited."

Statistical analysis

The number and reasons for alerts, patient and nurse assessments of symptom urgency, and immediate nurse actions in response to alerts were summarized descriptively. Fisher's exact test was used to evaluate whether nurses were more likely to act or were more likely to consider an alert urgent if the patient's weekly survey included any high score, any worsened score, or any other (free text) symptom.

When attempting to identify likely non-urgent alerts, we felt that nurses should always receive notifications if requested by the patient (i.e., a "yes, report now" response when asked if they would like a message sent to their care team). Thus, the analysis aimed at reducing alert notifications included all weekly surveys that generated an alert where (1) patients indicated the alert could wait until their next appointment and (2) nurses provided their assessment of the alert's urgency. Generalized linear mixed modeling with a logit link was conducted to predict the nurse's assessment of alert urgency. Fixed effects included baseline demographic and clinical characteristics (i.e., sex, age, race, ethnicity, cancer type, months since diagnosis, two or more hospitalizations in the past year, line of systemic cancer treatment, and intravenous [rather than oral] delivery of cancer treatment), and information from the weekly surveys (i.e., patient-reported symptoms, worsening of each symptom from the prior week, and alert type for each symptom [no alert, first alert for that symptom during the study period, second or later alert]). A random patient intercept accounted for nonindependence of alerts due to patients generating multiple alerts over the study period, and a random practice intercept accounted for nonindependence within practices. Predicted probabilities were calculated solely based on fixed effects [12]. To understand the added benefit of information from the weekly surveys, results were compared to those from a generalized linear mixed model with baseline demographic and clinical characteristics (only) as fixed effects. Statistical testing of each fixed effect was two-sided, with p-values < 0.05 considered statistically significant. For each model, the receiver operating characteristic (ROC) curve was plotted, and the area under the ROC curve was calculated. A cutpoint based on the predicted probability of an alert being urgent was selected to achieve 80% sensitivity. Sensitivity, specificity, positive predictive value, and negative predictive value were estimated along with 95% Wald confidence intervals. Model results were used to calculate how many alerts could be suppressed while maintaining 80% sensitivity. All statistical analyses were conducted by the Alliance Foundation Trials Statistics and Data Management Center in SAS 9.4 (SAS Institute Inc., Cary, NC) on the study database frozen on July 27, 2022.

Results

At 26 oncology practices, 746 patients were assessed for eligibility, and 593 were included (enrolled October 30, 2017 to March 16, 2020; Fig. 2) [7]. Median age was 64 years (range 29–89), 359/593 (61%) patients were female, 473/588 (80%) were white, 218/592 (37%) had a high school education or less, 154/593 (26%) were recruited from a rural site, and 62/593 (10%) reported never using a computer, tablet, or smartphone. Patients were diverse in cancer type (Table 2). Approximately one-third of patients (215/593, 36%) opted to use the interactive voice response system, with the remaining patients completing the weekly surveys online.

Patients remained on the study intervention for a median of 11.3 months (interquartile range 5.1–12.0). During this time, patients completed 91% (20,565/22,486) of expected surveys, with 34% (6979/20,565) of completed surveys generating an alert. Almost all patients (566/593, 95%) had at

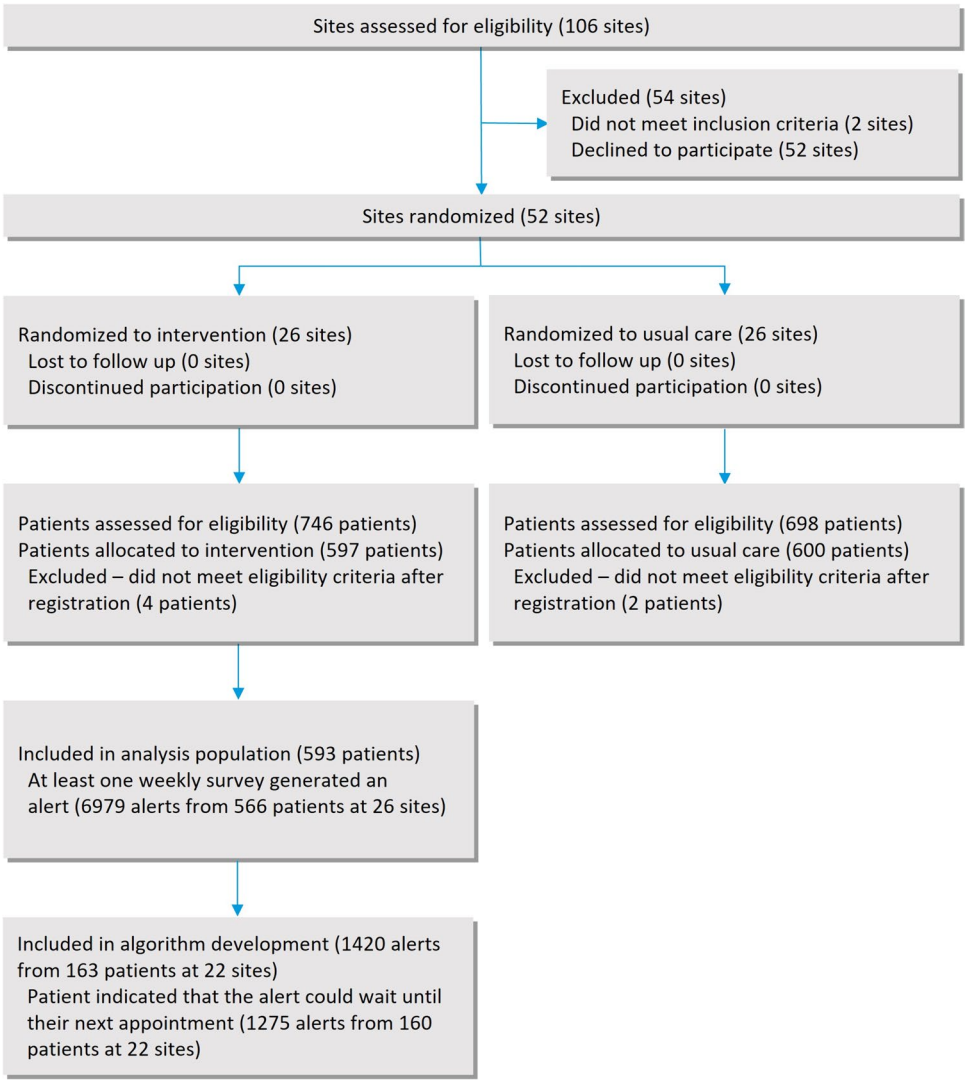

Fig. 2 CONSORT flow diagram for sites, patients, and alerts in the PRO-TECT trial

least one weekly survey generate an alert. Figure 3 shows the number of alerts due to high scores, worsened scores, and/or any other (free text) symptoms. Almost half of alerts (3311/6979, 47%) were solely due to symptoms with high scores. Pain had the greatest prevalence of high scores (3249/6979, 47% for frequency, severity, and/or interference) and worsened scores (783/6979, 11%; Supplementary Table S1).

Documentation on whether an immediate nurse action was taken was available for > 99% (6947/6979) of alerts. Of these 6947 cases, 4122 (59%) prompted immediate nurse actions, most commonly a telephone discussion with the patient or caregiver (3097/4122, 75%), advice for self-management (944/4122, 23%), and/or supportive medications prescribed or modified (868/4122, 21%). When they did not act upon alerts, nurses often reported already being aware of the patient's symptoms (2075/2819, 74% of alerts with no immediate nurse action). Nurses were more likely to act upon an alert if the patient's weekly survey included any worsened score compared to the prior week (Fisher's exact $p < 0.001$, 1663/2429 [68%] alerts with versus 2459/4518 [54%] without any worsened score). Nurses were not more likely to act upon an alert if the patient's weekly survey included any high score (Fisher's exact $p = 0.20$, 3210/5447 [59%] alerts with versus 912/1500 [61%] without any high score) or any other (free text) symptom (Fisher's exact $p = 0.84$, 1039/1745 [60%] alerts with versus 3083/5202 [59%] without any other symptom).

In the overlapping testing periods of the trial during which patients were asked if they would like a message sent to their care team and nurses were asked if the symptom(s) reported in the alert likely could have waited to be addressed, 1420 alerts occurred from 163 patients. Among these 1420 alerts, there were 1275 (90%) cases for which the patient indicated the alert could wait until their next appointment, encompassing 160 patients whose baseline demographic and clinical characteristics were similar to those of the full sample (Table 2). For these 1275 cases, despite the patient

Table 2 Patient demographic and clinical characteristics at baseline

Characteristic	All patients ($n=593$)	Subsample included in testing periods[a] ($n=160$)
Age in years, median (range)	64 (29–89)	64 (29–87)
Sex, *n* (%)		
Female	359 (61%)	101 (63%)
Male	234 (39%)	59 (37%)
Race, *n* (%)		
American Indian or Alaska Native	11 (2%)	1 (0.6%)
Asian	2 (0.3%)	1 (0.6%)
Black or African American	99 (17%)	24 (15%)
Native Hawaiian or Pacific Islander	2 (0.3%)	1 (0.6%)
White	473 (80%)	132 (83%)
Multiple races	1 (0.2%)	0 (0%)
Not reported	5	1
Hispanic ethnicity, *n* (%)		
Yes	14 (2%)	5 (3%)
No	577 (98%)	155 (97%)
Not reported	2	0
Education, *n* (%)		
1st to 8th grade	10 (2%)	0 (0%)
9th to 11th grade	35 (6%)	5 (3%)
High school graduate/GED	173 (29%)	36 (23%)
Some college, associate degree, or other certification	218 (37%)	72 (45%)
College degree	91 (15%)	25 (16%)
Advanced degree	65 (11%)	22 (14%)
Not reported	1	0
Recruited from rural site, *n* (%)	154 (26%)	24 (15%)
Weekly survey mode of administration, *n* (%)		
Internet	378 (64%)	116 (73%)
Automated telephone	215 (36%)	44 (28%)
Cancer type, *n* (%)		
Colorectal, anal	100 (17%)	27 (17%)
Thoracic (lung, thymus)	118 (20%)	35 (22%)
Breast	97 (16%)	24 (15%)
Gynecologic (ovarian, cervix, uterine, vaginal)	64 (11%)	19 (12%)
Pancreas, hepatobiliary	48 (8%)	13 (8%)
Gastro-esophageal, small bowel	25 (4%)	9 (6%)
Genitourinary non-prostate (bladder, kidney, testicular, penile)	36 (6%)	6 (4%)
Myeloma, lymphoma	31 (5%)	8 (5%)
Prostate	33 (6%)	7 (4%)
Melanoma	11 (2%)	6 (4%)
Other (brain, sarcoma, other soft tissue, head/neck, thyroid, unknown primary)	30 (5%)	6 (4%)
Months since diagnosis, median (range)	21 (0–454)	19 (0–221)
Line of systemic cancer treatment, *n* (%)		
1st	206 (35%)	60 (38%)
2nd	176 (30%)	46 (29%)
3rd	102 (17%)	26 (16%)
4th or later	109 (18%)	28 (18%)
Intravenous delivery of cancer treatment, *n* (%)	526 (89%)	148 (93%)
2 + hospitalizations in past year, *n* (%)	114 (19%)	26 (16%)

[a]Development of the algorithm used data from two testing periods during which patients were asked to indicate if the alert could wait until their next appointment and nurses were asked to document the urgency of the alert (1420 alerts from 163 patients). During these testing periods, 160 patients generated 1275 alerts for which they indicated that the alert could wait until their next appointment

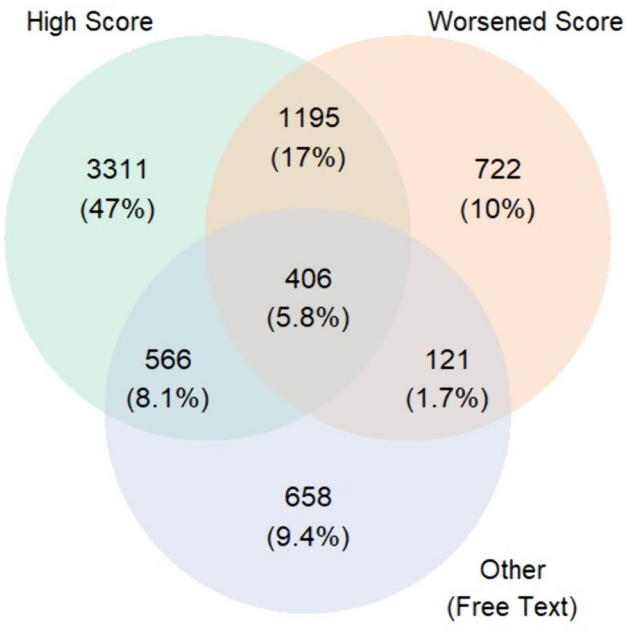

Fig. 3 Venn diagram of the number of alerts due to symptoms with a high score, symptoms with a worsened score compared to the prior week, and/or other (free text) symptoms (total $N = 6979$ alerts)

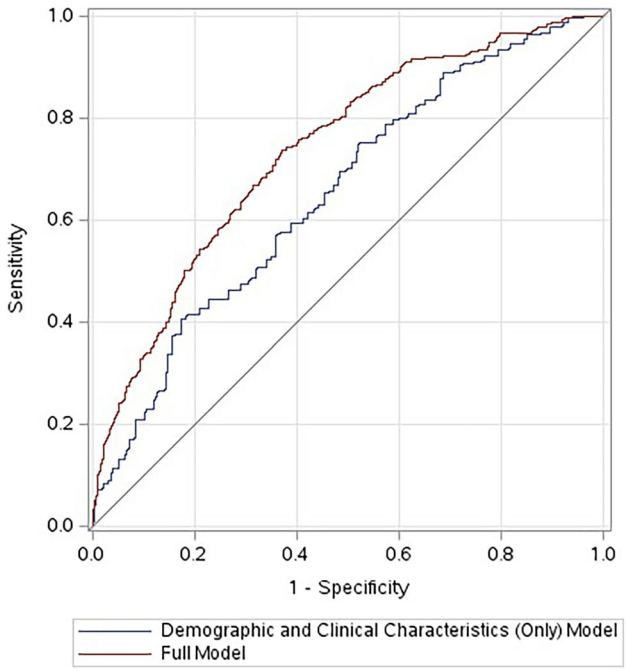

Fig. 4 Comparison of receiver operating characteristic (ROC) curves of models predicting nurse opinions regarding the urgency of an alert. The first model (in blue) included baseline demographic and clinical characteristics (only) as fixed effects (area under the ROC curve = 0.66, 95% CI [0.63, 0.69]), whereas the second model (in red) additionally included information from weekly surveys (i.e., patient-reported symptoms, worsening of each symptom from the prior week, and alert type for each symptom) as fixed effects (area under the ROC curve = 0.74, 95% CI [0.71, 0.77]). (Color figure online)

indication that the symptom(s) could wait, an alert was sent to nurses nonetheless (and the nurses were unaware of the patient assessment of symptom urgency). Upon receipt of these alerts, nurses indicated that the alert was urgent and likely could not have waited until the patient's next appointment for 28% (361/1275) of cases. Nurses were more likely to consider an alert urgent if the patient's weekly survey included any worsened score compared to the prior week (Fisher's exact $p = 0.009$, 131/393 [33%] alerts with versus 230/882 [26%] without any worsened score) or any other (free text) symptom (Fisher's exact $p < 0.001$, 109/240 [45%] alerts with versus 252/1035 [24%] without any other symptom). Nurses were less likely to consider an alert urgent if the patient's weekly survey included any high score (Fisher's exact $p = 0.001$, 268/1022 [26%] alerts with versus 93/253 [37%] without any high score).

Generalized linear mixed models were estimated based on the 1275 alerts where (1) patients indicated the alert could wait until their next appointment and (2) nurses provided their assessment of the alert's urgency. The initial generalized linear mixed model including only baseline demographic and clinical characteristics as fixed effects yielded an area under the ROC curve of 0.66 (95% CI [0.63, 0.69]) for predicting whether nurses considered an alert urgent. Adding information from the weekly surveys as fixed effects yielded an improved area under the ROC curve of 0.74 (95% CI [0.71, 0.77]; Fig. 4). Based on this model, nurses were more likely to consider an alert urgent when patients had certain cancer types (i.e., thoracic, melanoma, or other relative to breast), had less time since diagnosis, were receiving their third or later line of systemic cancer treatment (relative to first), were receiving cancer treatment intravenously, reported worsening of pain severity, reported any other (free text) symptoms, generated an alert for constipation for the first time, or reported a fall for the second or later time (Table 3). Nurses were less likely to consider an alert urgent when patients generated an alert for pain for the second or later time.

Model results were used to develop an algorithm for classifying alerts as urgent versus non-urgent based on the predicted probability of the nurse designating that the alert was urgent (≥ 0.19 versus < 0.19). Based on this classification, sensitivity equaled 80% (95% CI [76%, 84%]), specificity equaled 52% (95% CI [49%, 55%]), positive predictive value equaled 40% (95% CI [36%, 43%]), and negative predictive value equaled 87% (95% CI [84%, 90%]). By sending alerts to the care team only if (1) requested by the patient or (2) indicated by the model-based algorithm, 542 (38%) of the 1420 alerts from the two testing periods of the PRO-TECT trial could have been suppressed. Care teams would have received 878 alerts (instead of 1420) and deemed 358 (41%)

Table 3 Odds ratio estimates from the generalized linear mixed model predicting the nurse's opinion regarding the urgency of an alert ($N = 1275$ alerts from 160 patients at 22 sites)

Fixed effect	Odds ratio (95% CI)	*p*-value
Female sex (reference: male)	1.420 (0.846, 2.383)	0.1842
Age in years	0.997 (0.977, 1.017)	0.7736
Race (reference: white)		
Black	1.251 (0.637, 2.456)	0.5144
Other race	2.641 (0.599, 11.649)	0.1994
Hispanic/Latino (reference: no)	1.554 (0.334, 7.224)	0.5737
Cancer type (reference: breast)		
Colorectal, anal	1.419 (0.693, 2.906)	0.3378
Thoracic (lung, thymus)	2.872 (1.352, 6.099)	0.0061
Gynecologic (ovarian, cervix, uterine, vaginal)	1.918 (0.845, 4.355)	0.1195
Pancreas, hepatobiliary	2.134 (0.880, 5.177)	0.0936
Gastro-esophageal, small bowel	2.090 (0.758, 5.765)	0.1544
Genitourinary non-prostate (bladder, kidney, testicular, penile)	2.090 (0.533, 8.200)	0.2904
Myeloma, lymphoma	2.696 (0.814, 8.927)	0.1043
Prostate	1.747 (0.509, 5.993)	0.3749
Melanoma	8.801 (2.676, 28.946)	0.0004
Other (brain, sarcoma, other soft tissue, head/neck, thyroid, unknown primary)	4.185 (1.450, 12.080)	0.0082
Months since diagnosis	0.993 (0.987, 0.999)	0.0151
2 + hospitalizations in past year (reference: no)	0.761 (0.435, 1.330)	0.3374
Line of systemic cancer treatment (reference: 1st)		
2nd	1.583 (0.940, 2.667)	0.0841
3rd	1.965 (1.016, 3.802)	0.0449
4th or later	3.100 (1.609, 5.973)	0.0007
Intravenous delivery of cancer treatment (reference: no)	4.049 (1.864, 8.795)	0.0004
Decreased food/fluid intake	1.090 (0.828, 1.435)	0.5370
Pain frequency	1.018 (0.754, 1.374)	0.9084
Pain severity	0.871 (0.626, 1.214)	0.4153
Pain interference	1.251 (0.949, 1.650)	0.1125
Performance status	0.953 (0.693, 1.310)	0.7647
Depression	1.146 (0.786, 1.671)	0.4782
Nausea frequency	1.328 (0.878, 2.008)	0.1785
Nausea severity	0.720 (0.450, 1.154)	0.1718
Vomiting frequency	0.791 (0.526, 1.190)	0.2600
Diarrhea frequency	1.255 (0.966, 1.630)	0.0886
Constipation severity	1.090 (0.831, 1.429)	0.5338
Dyspnea severity	1.000 (0.702, 1.424)	0.9993
Dyspnea interference	0.814 (0.561, 1.181)	0.2778
Insomnia severity	0.921 (0.705, 1.203)	0.5448
Severe financial toxicity (reference: no)	1.483 (0.678, 3.241)	0.3233
Any other symptoms (reference: no)	2.137 (1.410, 3.237)	0.0004
Worsened decreased food/fluid intake (reference: no)	0.544 (0.188, 1.575)	0.2618
Worsened pain frequency (reference: no)	1.139 (0.556, 2.332)	0.7218
Worsened pain severity (reference: no)	4.037 (1.270, 12.828)	0.0181
Worsened pain interference (reference: no)	1.097 (0.479, 2.515)	0.8262
Worsened performance status (reference: no)	1.102 (0.392, 3.099)	0.8534
Worsened nausea frequency (reference: no)	0.555 (0.190, 1.618)	0.2802
Worsened nausea severity (reference: no)	3.726 (0.677, 20.498)	0.1303
Worsened vomiting frequency (reference: no)	0.319 (0.032, 3.152)	0.3279
Worsened diarrhea frequency (reference: no)	1.539 (0.642, 3.686)	0.3334
Worsened constipation severity (reference: no)	0.388 (0.119, 1.258)	0.1144

Table 3 (continued)

Fixed effect	Odds ratio (95% CI)	*p*-value
Worsened dyspnea severity (reference: no)	1.320 (0.310, 5.627)	0.7074
Worsened dyspnea interference (reference: no)	2.445 (0.713, 8.380)	0.1548
Worsened insomnia severity (reference: no)	2.166 (0.668, 7.027)	0.1976
Decreased food/fluid intake alert type (reference: 1st alert)		
No alert	0.416 (0.113, 1.535)	0.1875
2nd or later alert	0.487 (0.162, 1.460)	0.1986
Pain alert type (reference: 1st alert)		
No alert	0.592 (0.237, 1.478)	0.2609
2nd or later alert	0.439 (0.194, 0.991)	0.0475
Performance status alert type (reference: 1st alert)		
No alert	0.602 (0.183, 1.977)	0.4024
2nd or later alert	0.653 (0.223, 1.912)	0.4366
Depression alert type (reference: 1st alert)		
No alert	1.804 (0.409, 7.958)	0.4354
2nd or later alert	1.695 (0.371, 7.743)	0.4956
Nausea alert type (reference: 1st alert)		
No alert	0.582 (0.149, 2.273)	0.4354
2nd or later alert	0.747 (0.237, 2.360)	0.6190
Vomiting alert type (reference: 1st alert)		
No alert	0.081 (0.005, 1.314)	0.0770
2nd or later alert	0.320 (0.039, 2.602)	0.2860
Diarrhea alert type (reference: 1st alert)		
No alert	1.644 (0.435, 6.208)	0.4630
2nd or later alert	1.027 (0.337, 3.132)	0.9627
Constipation alert type (reference: 1st alert)		
No alert	0.128 (0.032, 0.515)	0.0038
2nd or later alert	0.312 (0.094, 1.036)	0.0572
Dyspnea alert type (reference: 1st alert)		
No alert	0.687 (0.193, 2.448)	0.5627
2nd or later alert	1.018 (0.301, 3.442)	0.9775
Insomnia alert type (reference: 1st alert)		
No alert	1.648 (0.386, 7.034)	0.4995
2nd or later alert	2.022 (0.509, 8.024)	0.3165
Have fallen alert type (reference: 1st alert)		
No alert	0.838 (0.318, 2.209)	0.7203
2nd or later alert	3.952 (1.147, 13.614)	0.0294

of them urgent. This reduction by 542 alerts would be at a cost of 71 missed potentially urgent alerts.

Discussion

In this study, we found that a model-based algorithm could inform the reduction of real-time alert notifications from ePRO remote symptom monitoring by 38%. Reducing the number of alert notifications would foster nurse acceptance and implementation feasibility of ePRO systems because while nurses find ePRO systems valuable in clinical care, they also feel that they receive too many alert notifications and that many alerts pertain to symptoms that could reasonably and safely be deferred until the next visit [13]. In the setting of nursing staffing shortages, automated strategies to prioritize symptom reports have the potential to enhance the value of ePRO systems. Indeed, in this trial nurses opted not to act immediately on 41% of alerts. Therefore, an opportunity exists to identify and suppress alerts upfront that are likely non-urgent, thereby unburdening nurses and enabling them to address the most concerning issues. Notably, even if some potentially urgent alert notifications were suppressed, the ePRO information in those alerts would still be available for review in symptom reports during scheduled visits.

Our results highlight that nurses generally perceive new or worsening symptoms to be clinically urgent warranting immediate action, while stable ongoing symptoms, even if severe, are more often regarded as appropriate for discussion at the next visit. This finding seems driven by persistently high symptoms already being known by the care team.

A potential concern in applying an algorithm to identify and suppress non-urgent alerts is that urgent and/or actionable alerts could also be suppressed. Such algorithms can be calibrated to dial up or down the number of suppressed alert notifications depending on available resources and clinical priorities of an organization. Such adjustments are no different than any clinical or population health management program that is resource-intensive, such as care coordination, ambulatory care access, community outreach, or inpatient bed capacity. In all these examples, clinical coverage versus available resources are inherent tradeoffs, albeit often less systematically traded off than in an algorithm. The model-based algorithm we developed has an area under the ROC curve of 0.74, which is considered acceptable discrimination [14]. Rather than suppressing alert notifications, another option would be to send notifications with annotation regarding whether the patient feels the symptom can wait, and a designation of potential urgency as determined by an algorithm. Nurses could use this additional information to decide whether to act immediately or defer management.

This study has limitations. First, we used nurse perceived urgency and actions, which are subject to variability in practice patterns and clinical judgment, as outcomes for our analyses. However, arguably clinician decision-making is the gold standard for assessment of symptom urgency. An alternative approach would be central arbitration of clinical assessments of symptom urgency, although such an approach would also be subjective and would lose the real-world grounding of the current analyses. Second, because patients were diverse in cancer type, conclusions about any given cancer type are limited and require replication in a larger sample of patients with the cancer type of interest. More generally, the sample may not allow for conclusions about all patient subgroups of interest. However, the PRO-TECT trial aimed to achieve greater representation by enrolling patients from urban and rural community oncology practices across the U.S. and by using purposive sampling of Black and African American patients. Third, alert notifications were generated only for symptoms that were severe or worsened by two points, and it is unclear if the algorithm would perform differently in an ePRO system with different alert notification thresholds. However, the thresholds were determined based on clinical consensus and engagement of patients, clinicians, and researchers [11], and therefore likely captured most potentially urgent symptoms. Finally, we did not test the performance of this algorithm in an independent sample, though surveys and qualitative interviews with participating nurses supported our conclusions regarding the limited utility of alerts for ongoing symptoms already known to the care team.

This algorithm has not been implemented prospectively to evaluate its impact on the number of alert notifications when deployed, nurse perceptions of its acceptability and appropriateness, or most importantly its impact on patients' clinical outcomes. Future work should include conducting a randomized trial to prospectively compare the performance of different algorithms for generating ePRO remote symptom monitoring alerts, including an algorithm that only alerts nurses about new or worsening symptoms.

Conclusion

A model-based algorithm can identify ePRO remote symptom monitoring alerts likely to be considered non-urgent, and may assist in reducing the burden of alert notifications on nurses.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11136-024-03675-3.

Acknowledgements The sponsor, Alliance Foundation Trials, LLC, coordinated contracting with participating study practice sites and submission of regulatory documentation. The sponsor was not involved in the design or conduct of the study; collection, analysis, or interpretation of the data; preparation of the manuscript; or decision to submit the manuscript for publication. The Alliance Publications Committee reviewed and approved this manuscript for publication and did not have the right to veto publication or to control the decision regarding to which journal the manuscript was submitted. All statements in this publication, including its findings, are solely those of the authors and do not necessarily represent the views of PCORI or its Board of Governors or Methodology Committee. Dr. Mazza presented results from this article at the International Society for Quality of Life Research Conference in Calgary, AB, Canada on October 20, 2023.

Author contributions GLM and EB formulated the research question and designed the study. EB and ACD provided supervision for the research activity planning and execution. BG, ACD, and GLM prepared the data for analysis. GLM and BG analyzed the data. GLM led the writing of the manuscript. GLM created the tables and figures. GLM, ACD, BG, JJ, AMD, PC, VSB, GT, MJ, LJR, DS, and EB contributed to the conduct and primary analysis of the PRO-TECT trial. All authors interpreted the results as well as reviewed and approved the manuscript for publication.

Funding This work was funded by the Patient-Centered Outcomes Research Institute (PCORI) award IHS-1511-33392. The trial made use of technology systems provided by the Patient-Reported Outcomes Core (PRO Core; https://pro.unc.edu) at the Lineberger Comprehensive Cancer Center of the University of North Carolina, funded by NCI Cancer Center Core Support grant P30CA016086 and the University Cancer Research Fund of North Carolina. The trial was sponsored by Alliance Foundation Trials, LLC (https://acknowledgments.alliancefound.org).

Data availability Investigators can submit a request for data from this trial by completing an Alliance Data Sharing Request Form, found at

https://www.allianceforclinicaltrialsinoncology.org/main/public/standard.xhtml?path=%2FPublic%2FDatasharing, and submitting it to DataSharing@AllianceNCTN.org.

Declarations

Conflict of interest Dr. Basch reported receipt of personal fees from Navigating Cancer, Sivan, AstraZeneca, Resilience, and the Research Triangle Institute for serving as a scientific advisor; receipt of research funding from the National Cancer Institute (NCI); being owner of Carolina Informatics; and employment by the University of North Carolina. Dr. Schrag reported receipt of personal fees from Pfizer, nonfinancial support from Grail for serving as a site principal investigator, and grants from the American Association for Cancer Research awarded to Memorial Sloan Kettering Cancer Center. Dr. Mody reported consultancy and receipt of research funding from Sivan. All remaining authors have declared no conflicts of interest.

Consent to participate Informed consent was obtained from all participants included in the PRO-TECT trial.

Ethics approval The PRO-TECT trial was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. The trial's protocol was approved by the central institutional review board and the institutional review board at each participating site.

References

1. Cleeland, C. S., Zhao, F., Chang, V. T., et al. (2013). The symptom burden of cancer: Evidence for a core set of cancer-related and treatment-related symptoms from the Eastern Cooperative Oncology Group Symptom Outcomes and Practice Patterns study. *Cancer, 119*(24), 4333–4340.
2. Reilly, C. M., Bruner, D. W., Mitchell, S. A., et al. (2013). A literature synthesis of symptom prevalence and severity in persons receiving active cancer treatment. *Supportive Care in Cancer, 21*(6), 1525–1550.
3. Panattoni, L., Fedorenko, C., Greenwood-Hickman, M. A., et al. (2018). Characterizing potentially preventable cancer- and chronic disease-related emergency department use in the year after treatment initiation: A regional study. *Journal of Oncology Practice/American Society of Clinical Oncology, 14*(3), e176–e185.
4. Henry, D. H., Viswanathan, H. N., Elkin, E. P., et al. (2008). Symptoms and treatment burden associated with cancer treatment: Results from a cross-sectional national survey in the US. *Supportive Care in Cancer, 16*(7), 791–801.
5. Schmidt, T., Valuck, T., Perkins, B., et al. (2021). Improving patient-reported measures in oncology: A payer call to action. *Journal of Managed Care and Specialty Pharmacy, 27*(1), 118–126.
6. Centers for Medicare & Medicaid Services (CMS). *Enhancing oncology model*. https://innovation.cms.gov/innovation-models/enhancing-oncology-model.
7. Basch, E., Schrag, D., Henson, S., et al. (2022). Effect of electronic symptom monitoring on patient-reported outcomes among patients with metastatic cancer: A randomized clinical trial. *JAMA, 327*(24), 2413–2422.
8. National Cancer Institute. *Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events*. https://healthcaredelivery.cancer.gov/pro-ctcae/.
9. Dueck, A. C., Mendoza, T. R., Mitchell, S. A., et al. (2015). Validity and reliability of the US National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *JAMA Oncology, 1*(8), 1051–1059.
10. Al-Rashdan, A., Sutradhar, R., Nazeri-Rad, N., et al. (2021). Comparing the ability of physician-reported versus patient-reported performance status to predict survival in a population-based cohort of newly diagnosed cancer patients. *Clinical Oncology (Royal College of Radiologists), 33*(7), 476–482.
11. Stover, A. M., Tompkins Stricker, C., Hammelef, K., et al. (2019). Using stakeholder engagement to overcome barriers to implementing patient-reported outcomes (PROs) in cancer care delivery: Approaches from 3 prospective studies. *Medical Care, 57*(5), S92–S99.
12. Kiernan, K. (2018). Insights into using the GLIMMIX procedure to model categorical outcomes with random effects. *SAS Technical Papers, 2018*, 1–23.
13. Basch, E., Stover, A. M., Schrag, D., et al. (2020). Clinical utility and user perceptions of a digital system for electronic patient-reported symptom monitoring during routine cancer care: Findings from the PRO-TECT trial. *JCO Clinical Cancer Informatics, 4*, 947–957.
14. Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology, 5*(9), 1315–1316.