The Impact of lung cancer
Geerse, Olaf

DOI:
10.33612/diss.94412905

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 11-12-2019
Appendix I

Impact of the Serious Illness Care Program: Results of a randomized controlled trial in outpatient oncology


ABSTRACT

Importance: High-quality conversations between clinicians and seriously ill patients about values and goals are associated with improved outcomes but occur infrequently.

Objective: To examine feasibility, acceptability, and effectiveness of a communication quality-improvement intervention (Serious Illness Care Program, SICP) on patient outcomes.

Design, Setting, Participants: Cluster-randomized controlled trial of SICP in outpatient oncology. Patients with advanced cancer and oncology clinicians participated between September 2012 and June 2016.

Intervention: The intervention included tools, training, and system changes.

Main outcomes and Measures: The co-primary outcomes included goal-concordant care (Life Priorities) and peacefulness (PEACE) at the end of life. Secondary outcomes included therapeutic alliance (Human Connection Scale), anxiety (GAD-7), depression (PHQ-9), and survival. We evaluated uptake and effectiveness of clinician training, clinician adoption of the conversation tool, and conversation duration.

Results: We analyzed data from 91 clinicians in 41 clusters (72% participation, 48 intervention, 43 control) and 278 patients (46% participation, 134 intervention, 144 control). Clinicians (47/48) rated the training as effective (4.3/5, SD=0.7); of those who received a reminder, 87% completed at least one conversation (median duration 19 minutes, range 5-70). Peacefulness, therapeutic alliance, anxiety, and depression did not differ at baseline. We were only able to evaluate the co-primary outcomes in 64 patients; no differences were found between the intervention and control groups. However, the trial demonstrated significant reductions in the proportion of patients with moderate-severe anxiety (10% vs. 5% (p=0.05)) and depression symptoms (21% vs. 11% (p=0.04) in the intervention group at 14 weeks after baseline. Anxiety reduction was sustained at 24 weeks (10.4% vs. 4.2%, p=0.02), while depression reduction was not sustained (18% vs. 13%, p=0.31). Survival and therapeutic alliance did not differ between groups.

Conclusions: The results of this cluster-randomized trial were null with respect to the co-primary outcomes of goal-concordant care and peacefulness at the end of life. Methodologic challenges for the primary outcomes, including measure selection and sample size, limited the conclusions that can be drawn from the study. However, the significant reductions in anxiety and depression in the intervention group are clinically meaningful and require further study.

Trial Registration: Registered with ClinicalTrials.gov (NCT01786811), available at www.clinicaltrials.gov.
BACKGROUND

In the last year of life, patients with serious illness suffer with physical and emotional distress, inadequate communication with clinicians, and medical interventions inconsistent with patient priorities and preferences.1-7 Patients who discuss end-of-life care with their clinicians, especially earlier in the disease trajectory, are more likely to have positive outcomes, including better quality of life, less distress, and a higher likelihood of receiving care consistent with their preferences.8-10 However, evidence indicates gaps in the frequency, timing, and quality of such conversations.11-14 To address these deficiencies, national medical organizations have called for improved communication about patients’ values, goals, and care preferences15,16 (“Serious Illness conversations”).

While palliative care clinicians train for this task, a limited palliative care workforce suggests the need for other clinicians to effectively lead serious illness conversations.17,18-20 However, interventions seeking to equip non-palliative-care specialists to better communicate with patients about end-of-life concerns have not improved patient outcomes, such as psychological symptoms or quality of life. In fact, in one trial of trainee physicians, a communication training program was associated with an increase in patients’ depressive symptoms, raising concerns that end-of-life conversations may worsen psychological symptoms.25 Additionally, clinicians cite concerns about harming patients as a barrier to initiating these conversations.26,27 Because non-palliative care clinicians must fill this gap in communication with patients, we systematically developed and extensively pilot-tested the Serious Illness Care Program (SICP) with clinicians and patients.28

This trial evaluated the feasibility and acceptability of our intervention (SICP) - including uptake and effectiveness of training, adoption of the conversation guide, and duration of conversations—and its impact on patient outcomes: goal-concordant care and peacefulness at the end of life (co-primary outcomes) and therapeutic alliance, anxiety, depression, and survival for the total population (secondary outcomes). We chose the measures of peacefulness and goal-concordant care because they are patient-centered, important to patients and caregivers, and do not make assumptions about patients’ care preferences (life-prolonging vs. comfort-focused).
METHODS

Context
Dana Farber Cancer Institute (DFCI), an NCI-designated cancer center, in Boston, Massachusetts was the study site.

Intervention Description
The intervention included tools, training, and system changes. Clinical tools included a clinician-facing Serious Illness Conversation Guide (SICG), a patient letter introducing the SICG, and a Family Guide after the discussion. Clinician training included a 2.5-hour interactive, skills-based training on the SICG delivered by palliative care experts who offered follow-up coaching. System changes included routine identification of patients at high risk of death, email reminders to initiate conversations (“reminders”), and a novel structured template in the electronic medical record (EMR) for SICG documentation. Control clinicians provided usual care; control patients did not receive supporting tools. Each clinician received a $150 gift card for participation. Patients and caregivers received no compensation for participation.

Trial Design
We employed a cluster randomized-controlled trial design from September 2012 to June 2016.

Participants
We invited clinicians (physicians (MD), physician assistants (PA), and nurse practitioners (NP)) from ten disease centers and two satellite clinics to enroll. We excluded gynecology-oncology clinicians (participating in a concurrent study on end-of-life care) and melanoma clinicians (pilot subjects). We defined clusters as units of clinicians within a disease center based on clinical workflow; a typical cluster included 1 NP/PA and 2-3 physicians. Cluster sizes varied. Enrolled oncology clinicians identified eligible patients by reviewing patient lists at regular intervals and answering the surprise question- “Would I be surprised if this patient died in the next year?” Patients for whom clinicians responded “no” were eligible for participation. We excluded patients with cognitive impairment, non-English-speaking patients, and patients unable to identify a caregiver.

Outcomes

Patient Measures for Decedents
All patients completed a baseline survey at enrollment and follow-up surveys approximately every two months for two years or until death. Lacking a ‘gold standard’ for measuring the co-primary outcome of concordance between patient goals and care provided at the end of
life, we developed two novel, patient-centered tools that did not pre-judge patient values: Life Priorities survey for patients and the Family Perceptions survey. Based on a list of nine common goals drawn from an extensive literature review and patient interviews, we asked patients to select and rank five goals in order of importance and following patient death, we asked caregivers whether each goal was fulfilled in the patient’s final week and final three months of life.

We assessed the co-primary outcome of peacefulness in decedents through the PEACE (Peace, Equanimity and Acceptance in the Cancer Experience) questionnaire, a validated tool yielding two subscales: Struggle with Illness, measuring feelings of upset, worry, unfairness, shame, and anger at diagnosis (7 questions, score range 7 to 28, Cronbach's alpha 0.81) and Peaceful Acceptance, measuring acceptance of diagnosis, inner calm, and feelings of being well-loved (5 questions, score range 5 to 20, Cronbach's alpha 0.78).

Patient Measures for Total Population

We measured therapeutic alliance (secondary outcome) with a modified version of The Human Connection (THC) scale, which evaluates patients' sense of mutual understanding, caring, and trust with their physicians. To decrease patient burden and avoid redundancy, we included 7 of the original 16 items (Cronbach's alpha 0.90), a reduction supported by the tool developer (Mack and Bernacki, personal communication). Scores on this shortened THC range from 7 to 28 (Cronbach’s alpha of 0.83 in this trial data). We assessed anxiety and depressive symptoms (secondary outcomes) using the Generalized Anxiety Disorder-7 Scale (GAD-7, 7 items, range 0 to 21, Cronbach’s alpha 0.92) and the Patient Health Questionnaire 9 (PHQ-9, 9 items, range 0 to 27, Cronbach’s alpha 0.86 to 0.89). We defined scores in the moderate or severe category on both scales (10 or higher) as clinically significant.

We identified patient deaths from the Dana-Farber Clinical Operational and Research Information System database.

Clinician Measures

At baseline, we surveyed clinicians in both arms about profession (MD, PA, NP), gender, years in practice, percentage clinical time, and disease center. After training, we surveyed clinicians for training effectiveness (Likert scale range 0-5). After sending a reminder, we surveyed clinicians for conversation occurrence and duration.

Randomization

We stratified clinician clusters by disease center or satellite facility, and within strata, randomized one-half of the clusters to the intervention (20) and one-half to control (21).
**Blinding**

Enrolled clinicians were not blinded to study arm. Patients were blinded to the study arm of their clinicians.

**Sample Size**

We performed power calculations for the study’s primary outcomes. To ensure an overall 5% type I error rate, we used a 2.5% type I error rate for each of the two primary hypotheses. We based the power calculations on having 200 evaluable patients per study arm. We allowed for 6% un-evaluability due to patient dropout; however, we had significantly fewer evaluable patients than expected, prompting us to conduct a post-hoc power calculation. With 38 intervention patients and 26 control patients with goal-concordant care outcomes, a post-hoc power calculation found 25% power to detect at least a 0.6 higher mean on the intervention arm (a-priori specified clinically important increase). With 47 intervention and 47 control patients with the PEACE outcome, a post-hoc power calculation found 29% power to detect at least a 1.3 point higher mean score on the intervention arm (a-priori specified clinically important increase).

**Statistical Analysis**

We performed statistical analyses with SAS software, version 9.4 (SAS Institute, Cary NC). We used proportions for categorical variables and means/medians for continuous variables. All comparisons across study arms accounted for clustering of patients within clinician teams. We considered a p-value ≤ 0.05 as statistically significant. All analyses were conducted based on intention-to-treat.

**Clinician and Patient Characteristics**

When comparing baseline clinician and patient characteristics between arms, we used generalized estimating equations (GEE)\(^{39}\), chi-square tests for categorical variables, and t-tests for continuous variables.

**Patient Measures for Decedents**

**Goal-Concordant Care**

We evaluated goal-concordant care by matching each decedent’s final Life Priorities\(^{28}\) survey (within three months of death) with their caregiver’s Family Perception\(^{28}\) survey. We scored each of the patient’s three highest ranking goals as “concordant” if the caregiver indicated the goal had been achieved “to a large extent” resulting in a score of 0, 1, 2, or 3 goals met. We compared the arms using a GEE Wilcoxon rank-sum type score test for ordinal categorical data.\(^{39}\)
Peacefulness
Using GEE chi-square tests for ordinal data, we compared both PEACE\textsuperscript{35} subscales for decedents at baseline and at 3 months before death across the study arms.

Patient Measures for Total Population

Therapeutic Alliance, Anxiety, and Depression
We created a separate model for each outcome of interest, using a continuous score for therapeutic alliance, and dichotomizing anxiety\textsuperscript{37} and depression\textsuperscript{38} as moderate/severe versus none/mild. Due to variation in timing (patients did not complete surveys at the same fixed points), we fit repeated-measures models via GEE.\textsuperscript{39,40} We calculated the mean therapeutic alliance score and logits of the probabilities of moderate/severe anxiety and moderate/severe depression as a cubic spline of time of survey, using all data on all patients from all time points in an intention-to-treat repeated-measures model.\textsuperscript{41} We modeled the correlation between outcomes on the same patient at a pair of times as auto-regressive.\textsuperscript{42} Because conversations occurred on average at 12 weeks after baseline in the intervention arm, we compared patient outcomes across study arms at 14 and 24 weeks after baseline (the average completion time for the next two surveys) using the estimated means and probabilities at these two time points from the repeated-measures models.

We fit separate spline models for control and intervention arms, allowing the trends to vary over time differently in each arm, and we used inverse propensity weighting to balance the three outcomes between the two arms at baseline to ensure that differences at later time points were not due to baseline differences (even though differences were non-significant at baseline). For each outcome, we modeled the propensity score (probability of being in the intervention arm) via logistic regression with baseline outcome (therapeutic alliance scores, anxiety, or depression), and patient characteristics as predictors. Although dropout and survival did not differ between arms, the models\textsuperscript{40} protected against potential biases arising from patients in one arm being followed for longer time periods.

Survival
We obtained Kaplan-Meier 2-year survival estimates from date of baseline and used a log-rank test to compare survival differences between all enrolled intervention and control patients.

Ethics
The DFCI Institutional Review Board reviewed and approved the study; it is registered with clinicaltrials.gov (NCT01786811). All clinicians and patients provided written informed consent.
RESULTS

Sample Recruitment and Demographics

We enrolled 91 oncology clinicians, grouped into 41 randomized clusters (73% participation, Figure 1).

---

FIGURE 1: CONSORT Diagram
We enrolled and consented 379 patients (46% participation, Figure 1), 278 of whom had analyzable data (Figure 1). Over 75% (n=209) of patients completed at least one post-baseline survey. Patients who did not participate were significantly older (p<0.001) and less likely to have breast cancer (p=0.04) than participants, although there were no gender differences.

Patients with analyzable data were significantly more likely to be married (p=0.002) and have higher incomes (p=0.03) than those with non-analyzable data; no other demographic differences were significant. Neither baseline clinician (Table 1) nor patient characteristics (Table 2) demonstrated significant differences between arms.

**TABLE 1.** Baseline Characteristics of the Clinician Population<sup>a,b,c</sup>

<table>
<thead>
<tr>
<th></th>
<th>Intervention</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=48 (20 clusters)</td>
<td>n=43 (21 clusters)</td>
</tr>
<tr>
<td>Female sex – no. (%)</td>
<td>30 (63)</td>
<td>22 (51)</td>
</tr>
<tr>
<td>Discipline – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD</td>
<td>36 (75)</td>
<td>30 (70)</td>
</tr>
<tr>
<td>NP</td>
<td>11 (23)</td>
<td>11 (26)</td>
</tr>
<tr>
<td>PA</td>
<td>1 (2.1)</td>
<td>2 (4.7)</td>
</tr>
<tr>
<td>Cluster size – mean (95% CI)</td>
<td>3.3 (2.9-3.8)</td>
<td>2.8 (2.3-3.2)</td>
</tr>
<tr>
<td>Years of practice – mean (95% CI)</td>
<td>12.8 (9.7-16.0)</td>
<td>10.2 (7.4-12.9)</td>
</tr>
<tr>
<td>Disease center&lt;sup&gt;d&lt;/sup&gt; – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast oncology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal, Genitourinary, Head &amp; Neck, Neurology,</td>
<td>11 (23)</td>
<td>10 (23)</td>
</tr>
<tr>
<td>Sarcoma, Thoracic, other</td>
<td>27 (56)</td>
<td>22 (51)</td>
</tr>
<tr>
<td>Hematologic Malignancies, Lymphoma</td>
<td>6 (13)</td>
<td>6 (14)</td>
</tr>
<tr>
<td>Community-based clinics</td>
<td>4 (8.3)</td>
<td>5 (11.6)</td>
</tr>
<tr>
<td>Percentage of screened panel patients identified as eligible by surprise question – mean (95% CI)</td>
<td>23 (16-30)</td>
<td>27 (19-36)</td>
</tr>
</tbody>
</table>

<sup>a</sup> P values between arms are all > 0.07
<sup>b</sup> Percentages may not sum to exactly 100 due to rounding.
<sup>c</sup> 13% are missing years of practice, 2% missing pts eligible by surprise and all else have no missings. Calculations for percentages were based on non-missing data.
<sup>d</sup> Disease center does not include gynecologic oncology due to a concurrent trial being conducted at that center

**Intervention Measures**

We trained 47 of 48 intervention clinicians, and clinicians rated the training as effective (4.3/5, SD=0.7). Of those trained, 83% (n=39) received at least one reminder to conduct a serious illness conversation, and of those reminded, 87% (n=34) completed at least one conversation. Clinicians reported a median conversation duration of 19 minutes (range 5-70).
### TABLE 2. Baseline Characteristics of the Patient Population\(^{a,b,c}\)

<table>
<thead>
<tr>
<th>Category</th>
<th>Intervention n=134</th>
<th>Control n=144</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years - mean (95% CI)</td>
<td>62 (58-65)</td>
<td>62 (58-66)</td>
</tr>
<tr>
<td>Female sex – no. (%)</td>
<td>72 (54)</td>
<td>76 (53)</td>
</tr>
<tr>
<td>Race(^d) – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>124 (93)</td>
<td>127 (93)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>2 (1.5)</td>
<td>3 (2.2)</td>
</tr>
<tr>
<td>Other</td>
<td>7 (5.3)</td>
<td>7 (5.1)</td>
</tr>
<tr>
<td>Hispanic – no. (%)</td>
<td>3 (2.3)</td>
<td>4 (2.9)</td>
</tr>
<tr>
<td>Married/partnered – no. (%)</td>
<td>107 (80)</td>
<td>115 (80)</td>
</tr>
<tr>
<td>Income ≥$75,000 – no. (%)</td>
<td>77 (60)</td>
<td>65 (50)</td>
</tr>
<tr>
<td>Disease center (^e) – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast oncology</td>
<td>32 (24)</td>
<td>37 (26)</td>
</tr>
<tr>
<td>Gastrointestinal, Genitourinary, Head &amp; Neck, Neurology</td>
<td>93 (69)</td>
<td>91 (63)</td>
</tr>
<tr>
<td>Sarcoma, Thoracic, other</td>
<td>9 (6.7)</td>
<td>16 (11.1)</td>
</tr>
<tr>
<td>Health insurance type – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td>65 (49)</td>
<td>60 (43)</td>
</tr>
<tr>
<td>Medicaid/Mass Health</td>
<td>9 (6.8)</td>
<td>11 (8.0)</td>
</tr>
<tr>
<td>Private</td>
<td>58 (44)</td>
<td>65 (47)</td>
</tr>
<tr>
<td>No insurance</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Patient-reported health status – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relatively healthy and not seriously ill</td>
<td>21 (16)</td>
<td>24 (17)</td>
</tr>
<tr>
<td>Relatively healthy and terminally ill</td>
<td>77 (58)</td>
<td>73 (53)</td>
</tr>
<tr>
<td>Seriously but not terminally ill</td>
<td>26 (20)</td>
<td>28 (20)</td>
</tr>
<tr>
<td>Seriously and terminally ill</td>
<td>9 (6.8)</td>
<td>14 (10.1)</td>
</tr>
<tr>
<td>College, graduate or professional school – no. (%)</td>
<td>112 (84)</td>
<td>112 (80)(^e)</td>
</tr>
</tbody>
</table>

\(^{a}\) P values between arms are all >0.21.

\(^{b}\) Percentages will not sum to exactly 100 due to rounding.

\(^{c}\) Since the percent missing for any variable was less than 7%, missing data are not shown in this table. Calculations for percentages were based on non-missing data.

\(^{d}\) Race or ethnic group was self-reported.

\(^{e}\) Disease center does not include gynecologic oncology due to a concurrent trial being conducted at that center.

---

### Patient Measures for Decedents

#### Goal-Concordant Care

We matched a Family Perception\(^{28}\) survey to an appropriately-timed Life Priorities\(^{28}\) survey for 64 decedents (38 intervention, 26 control). There was no significant difference in the median number of top-three goals met between study arms (Table 3).
Peacefulness
There were no significant differences between study arms for decedents in the PEACE\textsuperscript{35} subscales at baseline or within three months before death (Table 3).

Patient Measures for Total Population

Therapeutic Alliance
Among all patients, mean scores of the Human Connection Scale\textsuperscript{36} did not differ significantly between arms at baseline (25.3 (CI 24.8-25.8) intervention vs. 25.5 (CI 25.0-26.0) control, p=0.60), at 14 weeks after baseline (25.5 (CI 24.8-26.2) intervention vs. 25.7 (CI 25.1-26.2) control, p=0.65), or at 24 weeks after baseline (25.5 (CI 25.0-26.1) intervention and 25.4 (CI 24.8-26.0) vs. control, p=0.71) (Figure 2).

Anxiety
Among all patients, the proportion of patients reporting moderate or severe anxiety symptoms did not differ significantly between arms at baseline (9.6% control vs. 9.0% intervention (p=0.85)). At 14 weeks after baseline, the proportion of patients reporting moderate or severe anxiety symptoms was significantly lower in the intervention arm (10% vs. 5%, p=0.05). At 24 weeks after baseline, intervention patients remained less likely than control patients to report moderate or severe anxiety symptoms (10.4% vs. 4.2%, p=0.02) (Figure 2).

Depression
Among all patients, the proportion of patients reporting moderate or severe depression symptoms did not differ significantly between arms at baseline (20% control vs. 19% intervention (p=0.84). At 14 weeks after baseline, the proportion of patients reporting moderate or severe depression symptoms was significantly lower in the intervention arm (21% vs. 11%, p=0.04). At 24 weeks after baseline, the proportion of patients reporting moderate or severe depression symptoms did not differ significantly between arms (18% vs. 13%, p=0.31) (Figure 2).

Survival
Median 2-year survival did not differ between study arms (13.9 months intervention, 13.6 months control, log-rank p=0.91).
Several explanations for the lack of impact of our intervention on the primary outcomes are possible. First, because of smaller number of expected deaths and poor survey response, our study was under-powered. Second, because of the absence of a strong patient-centered measure of goal-concordant care, we used an unvalidated survey. Third, our measurements of goal-concordant care were dependent on patient responses late in the illness and family responses early in bereavement, which may have been too burdensome. Fourth, although
we measured peacefulness with a validated scale, this measure may have been inadequate to capture elements of peacefulness that respond to improved communication. As a result, we are uncertain whether our intervention was ineffective at improving these outcomes, if our outcome measures were not appropriate or feasible, or if we lacked sufficient numbers to detect meaningful differences. Our challenges reflect the urgent need in our field for patient-centered measures of communication that are agreed upon, validated, and demonstrably sensitive to communication interventions.

This trial demonstrated significant improvements in the secondary outcomes of moderate-to-severe anxiety and depression symptoms that regularly burden patients with cancer. In contrast to prior research, this study, using well-validated and widely-used measures, demonstrated significantly decreased rates of anxiety and depression symptoms within two weeks of the conversation in the intervention group, and the reduction in anxiety symptoms lasted until at least 24 weeks after baseline, suggesting that trained oncologists can discuss important and difficult topics without causing harm and with potential benefit. To our knowledge, this is the first study to identify a clinically meaningful benefit to psychological symptoms from a structured communication approach, suggesting that psychological outcomes be considered primary outcomes in future communication studies in oncology. This finding also highlights the need for measurement of communication and outcomes over the illness trajectory, not just at the end of life, which may help to better understand how to improve patients’ well-being as they live with serious illness.

We found that intervention clinicians readily adopted the program; they attended the training and rated it as effective. They conducted serious illness conversations in a feasible timeframe with respect to the constraints of a typical oncology practice. We expect these findings to be transferrable to other clinical contexts that treat advanced cancer patients while also recognizing that these intervention components require substantial organizational resources.

Among several study limitations was insufficient power for the primary outcomes. Our patient participation rate in the trial, while low, is consistent with other population-level trials of seriously ill patients. Due to lower patient accrual rates, fewer deaths than expected, longitudinal design, and difficulties obtaining surveys from patients and bereaved caregivers, a relatively large number of patients were not included in the primary outcomes analysis. However, nonparticipants and unanalyzed participants were not meaningfully different from analyzed participants, and randomization still produced comparable groups between study arms. The variation in timing of outcome assessment may also be a limitation; however, we found that dropout and timing of measurement were similar across arms. Use of the surprise question by all clinicians and frequent survey completion by all patients may have prompted conversations in the control arm, attenuating potential between-arm differences. Additionally,
findings may not be generalizable because the study was conducted at a single oncology institution with a fairly small number of participants that were relatively white, college-educated, and affluent. Finally, the multi-component nature of the intervention prevents assessment of which components contributed to the outcomes.

The results of this cluster-randomized trial were null with respect to the co-primary outcomes of goal-concordant care and peacefulness for decedents (but were significantly under-powered for the primary outcomes) but demonstrated significant reductions in the secondary outcomes of anxiety and depression symptoms immediately after the conversation and a sustained reduction in anxiety among intervention patients in the total population. This study showed that a feasible system-level communication intervention with high clinician adoption may improve some meaningful patient outcomes among advanced cancer patients. Advancements in serious illness communication interventions will require more reliable and well-accepted patient-centered outcome measures and additional testing of the impact on patients throughout their illness trajectory.
REFERENCES

Appendix I


Appendix I


