Reliable and Valid Survey-Based Measures to Assess Quality of Care in Home-Based Serious Illness Programs

Rebecca Anhang Price, PhD,[1] Melissa A. Bradley, BA,[1] Feifei Ye, PhD,[2] Danielle Schlang, MA,[3]
Maria DeYoreo, PhD,[4] Paul D. Cleary, PhD,[5] Marc N. Elliott, PhD,[4] Cheryl K. Montemayor, BS,[4]
Martha Timmer, MS,[4] Anagha Tolpadi, MS,[4] and Joan M. Teno, MD, MS[6]

Abstract

Background: There is a pressing need for standardized measures to assess the quality of home-based serious illness care. Currently, there are no validated quality measures that are specific to home-based serious illness programs (SIPs) and the unique needs of their patients.

Objective: To develop and evaluate standardized survey-based measures of serious illness care experiences for assessing and comparing quality of home-based serious illness care programs.

Methods: From October 2019 through January 2020, we administered a survey to patients who received care from 32 home-based SIPs across the United States. Using the 2263 survey responses, we assessed item performance and constructed composite measures via factor analysis, evaluated item-scale correlations, estimated reliability, and examined validity by regressing overall ratings and willingness to recommend care on each composite.

Results: The overall survey response rate was 36%. Confirmatory factor analyses supported five composite quality measures: Communication, Care Coordination, Help for Symptoms, Planning for Care, and Support for Family and Friends. Cronbach's alpha estimates for the composite measures ranged from 0.69 to 0.85, indicating adequate internal consistency in assessing their underlying constructs. Interprogram reliability ranged from 0.67 to 0.80 at 100 completed surveys per measure, meeting common standards for distinguishing between programs' performance. Together, the composites explained 45% of the variance in patients' overall care ratings. Communication, Care Coordination, and Planning for Care were the strongest predictors of overall ratings.

Conclusion: Our analyses provide evidence of the feasibility, reliability, and validity of proposed survey-based measures to assess the quality of home-based serious illness care from the perspective of patients and their families.

Keywords: patient and family care experiences; patient survey; quality measurement; serious illness care

Introduction

IN RECENT YEARS, there has been rapid growth of community-based programs that provide care for seriously ill individuals in their homes.[1–4] These serious illness programs (SIPs) are expanding at a time when both the public and private sectors are adopting more value-based payment programs for care. Value-based programs use incentives to promote the quality and efficiency of care. The use of value-based programs has particularly important implications for the seriously ill population, which is at high risk for *under*-treatment motivated by cost concerns. To that end, the Centers for Medicare & Medicaid Services has introduced a number of initiatives that test alternative models for care of high need, high cost, seriously ill populations,[5–7] and is considering others.[8]

[1]RAND Corporation, Arlington, Virginia, USA.
[2]RAND Corporation, Pittsburgh, Pennsylvania, USA.
[3]RAND Corporation, Boston, Massachusetts, USA.
[4]RAND Corporation, Santa Monica, California, USA.
[5]Yale School of Public Health, New Haven, Connecticut, USA.
[6]Oregon Health and Science University, Portland, Oregon, USA.

Quality measures are critical in such models, complementing assessments focused on utilization and cost.[9] However, to date, no standardized measures have been developed to specifically assess and monitor quality of care provided by SIPs. Measures of the degree to which care is patient- and family-centered are particularly important for seriously ill patients because of great variability across patients in both preferences for care intensity and tradeoffs between quality and length of life. Surveys of patients and their family caregivers are the main means of assessing patient- and family-centeredness of care, and survey results can be used to identify areas of patient and family experiences of care that need improvement,[10–12] monitor quality over time, and be incorporated into value-based models to allow for benchmarking and comparison of programs.

Experts have highlighted the need for standardized measures of patient and family care in SIPs.[13] To address this need, we developed a survey of care experiences of the seriously ill. We field-tested the survey in a sample of patients receiving care from SIPs in late 2019 and early 2020, immediately before the onset of the COVID-19 pandemic.

Methods

Survey

To develop the survey, we first conducted a systematic literature review of patient- and family-reported measures of serious illness care. We then conducted interviews with patients, family caregivers, and health care providers from a diverse set of SIPs nationwide, and sought input from experts in serious illness care and survey research methods. We conducted cognitive interviews with patients and family caregivers to test draft questions and questionnaires.

Guided by information from these activities and a conceptual framework of core aspects of high-quality serious illness care (Supplementary Appendix Table SA1),[2] we developed a 56-item field test version of the Serious Illness Survey. It included 28 evaluative questions about communication, emotional and spiritual support, access and responsiveness, shared decision making and advance care planning in support of patient goals, symptom management/palliation, care continuity and coordination, attention to social determinants of health (via referrals and connection to resources), attention to caregiver needs, and medication management, and two global assessments of care (overall rating of care from the program and willingness to recommend the program to friends and family).

Supplementary Appendix Table SA2 lists the field test survey questions in each of these domains. In addition, the survey included questions about patients' health status, functional status, demographic characteristics, recent visits and calls from the program, and whether and why a proxy completed the survey. Wherever possible, questions were derived or adapted from surveys for which there were published assessments of validity and reliability. Final, more concise versions of the survey instrument are available free online.[14]

Field test

Sites. The study was conducted in 32 geographically diverse SIPs that provide home-based care. Programs were recruited from a master list of 319 SIPs developed by the project team, primarily from the Center to Advance Palliative Care (CAPC) National Palliative Care Registry,[15] a report on community-based model programs for the seriously ill,[16] and responses to announcements posted by the National Hospice & Palliative Care Organization and the American Academy of Hospice and Palliative Medicine. To be eligible to participate, all programs needed to provide medical care to seriously ill patients in their homes.

Almost all included programs provide after-hours access to care either by phone or in person and have either a physician or a nurse practitioner on the team that makes home visits. Twenty-three of the programs are based out of hospices or home health agencies, 6 out of health systems or health plans, and 3 are part of medical groups. Five of the 32 programs operate in more than one state. The average program size was 203 patients actively in care at a given time (median: 116; range: 7 to 1481).

Sample. Patients within each program were eligible for the survey if they were adults (age 18 or older at the time the sample was selected), received care at a private home or assisted living facility, and had been receiving care from the program for at least 3 months and no more than 24 months at the time the sample was selected. We identified all survey-eligible patients from participating programs, for a total sample of 6456 patients. Sampled patients who indicated on the survey that they had not received visits from the program in the past three months were not considered eligible.

Survey protocol. Within each program, eligible patients were randomly assigned to one of two modes of survey administration, mail-only or mail-telephone. The mail-only mode consisted of a prenotification letter, followed by a mail survey one week later, and an additional mail survey three weeks after that if the survey had not been returned. The mail-telephone mode consisted of a prenotification letter, followed by a mail survey one week later, and up to five calls to complete the survey by phone if the mail survey was not returned after three weeks.

All cover letters and introduction scripts were addressed to the patient, but indicated that a family member or friend could assist with or complete the survey for the patient if needed. The survey was available in both English and Spanish; when indicated in the sample file, Spanish was the language for mailing and initial telephone calls. Spanish was also offered as an option by telephone interviewers. We field-tested the survey between October 2019 and January 2020. The study was approved by the RAND Corporation's Human Subjects Protection Committee, which serves as RAND's IRB.

Analyses

Like several other public reporting initiatives,[17,18] we used "top-box" scores to promote ease of understanding by consumers.[19] To calculate top-box scores, we classify the response indicating the best quality as 100 and all other responses as 0 (e.g., "always" = 100; all other responses = 0). For the overall rating item, 9–10 are classified as 100 and 0–8 as 0.[20] Item scores are not calculated for those who respond to a screening question indicating that the item is not relevant to them, or for those who select a nonapplicable response indicating that the item is not relevant to them.

Composites

The project team and technical expert panel reviewed the 28 evaluative questions administered in the field test (Supplementary Appendix SA2) and retained 19 for composite development. We used confirmatory factor analysis (CFA) to evaluate the factor structure of these 19 items and used modification indices to drop one item (about trust) that loaded on more than one factor. The final CFA analysis was performed on 18 evaluative items. We used weighted least squares means and variance adjusted (WLSMV) estimation to account for the dichotomous nature of top-box item scores.[21] We used a criterion of factor loadings ≥0.40 for inclusion within the composite,[22] and assessed overall model fit using the Comparative Fit Index (CFI), the root mean square error of approximation (RMSEA), and weighted root mean square residual (WRMR). Prior research indicates that a model with a good fit typically has a CFI >0.95, RMSEA <0.05, and WRMR <1.0, with WRMR being less critical.[23–25] The model χ^2 statistic and standard error of model estimates were adjusted to account for the clustering of patients within programs.[26,27]

For the CFA, we hypothesized five composite measures of quality of serious illness care: Communication, Care Coordination, Help for Symptoms, Planning for Care, and Support for Family and Friends. To assess the degree to which these measures assess distinct content domains, we calculated correlations between the composite scores computed as the average of top-box-scored items, adjusting for clustering within programs. Correlations exceeding 0.80 may indicate that composites are measuring aspects of care that are insufficiently distinct.[22]

Case-mix adjustment

We adjusted for differences in case mix across SIPs,[28] including patient age, education, diagnosis, proxy assistance with survey completion, self-reported ability to get out of bed or house, self-reported physical and mental health, and response percentile (a within-program rank-based measure of the time between survey administration and survey response).[29]

Reliability

To assess reliability of the proposed quality measures, we calculated the interunit (i.e., program-level) reliability of each measure, a 0–1 index of the degree to which measure scores are able to precisely distinguish between the performances of programs. We calculated the program-level reliability for each measure using intraclass correlations (ICCs) of the case-mix- and survey mode-adjusted top-box scores, excluding programs with fewer than 10 respondents. We also calculated predicted program-level reliability with 100 respondents using the Spearman Brown formula.[30] When programs are being compared, measure reliability of 0.70 or greater is commonly considered adequate.[31]

We also calculated the internal consistency reliability of the composites using Cronbach's alpha. Cronbach's alpha increases with the number of items included in a composite measure and their average correlation with one another. Larger values indicate more precise measurement of the underlying construct. Cronbach's alphas of 0.70 or higher are considered adequate for group comparisons.[31]

Validity

To assess construct validity, we evaluated the associations of each composite measure's top-box score with the top-box score of the two global measures, Overall Rating of the Program and Willingness to Recommend the Program. We estimated multivariate linear regression models with the global measures as dependent variables to highlight the unique association of each composite with those measures. All models were adjusted for the case-mix variables and mode of survey administration. We fit models that included only one composite at a time as a predictor and a model that included all composites simultaneously as predictors.

All models were estimated with the WLSMV in Mplus;[32] to correct for attenuation in regression coefficients with categorical outcomes.[33,34] We calculated the squared semi-partial correlation, or the unique r^2, associated with each composite, which indicates the proportion of variance in the outcome uniquely associated with each composite. As with the CFA models, standard errors and significance tests of regression coefficients were adjusted for clustering of patients within programs.[26,27] CFA and validity testing were performed in Mplus 8; missing data were handled using full-information likelihood estimation.

Results

The overall response rate to the survey was 36.4% (30.4% in mail only mode, and 42.5% in the mail-telephone mode).

The average age of respondents was 79; 75% were non-Hispanic white (Table 1). Forty-four percent received seven or more in-person visits from the program in three months. Fourteen percent reported that they were not able to leave the house, and 58% reported being in fair or poor health. Proxy respondents completed the survey on behalf of the patient for 33% of respondents; an additional 14% reported some other form of proxy assistance. A comparison between characteristics of respondents and nonrespondents is available in Supplementary Appendix Table SA3.

The five-factor CFA model provides an excellent fit to the data, $\chi^2_{(125)} = 269.45$; CFI = 0.992; RMSEA = 0.023; and WRMR = 1.463. Table 2 displays the factor loadings and corrected item-total correlation for the 18 evaluative items proposed for the 5 composite measures, along with Cronbach's alpha internal consistency for each composite measure. The factor loadings range between 0.71 to 0.98 and corrected item-total correlations range between 0.44 and 0.69, suggesting these items are strong indicators of the corresponding factor.

Across field test programs, on average, programs perform the best with regard to the proportion of respondents indicating that the people from the program listened carefully to them (84%), gave them the help they needed between visits (84%), and cared about them as a whole person (82%). In contrast, programs show the greatest room for improvement with regard to discussions about how to get help with everyday activities, what their health care options would be if they got sicker, or about what is important in their life, with just 48%, 50%, and 51% of respondents indicating that someone from the program "definitely" talked with them about these topics, respectively.

In addition, on average, only 53% of respondents reported that the program "definitely" gave them the help they wanted

TABLE 1. CHARACTERISTICS OF PATIENTS RESPONDING
TO SERIOUS ILLNESS SURVEY

Characteristic	Respondents (%)[a], n = 2263
Sex	
Female	58.4%
Male	41.6%
Age (mean)	78.7
18–64	13.8%
65–74	17.7%
75–84	29.6%
85–89	19.5%
90 or older	19.4%
Race/ethnicity	
Non-Hispanic white	75.4%
Non-Hispanic black or African American	10.0%
Hispanic	10.4%
Other	4.2%
Education	
Less than high school	21.2%
High school graduate or GED	36.9%
Some college or two-year degree	23.3%
College degree or more	18.6%
Language spoken at home	
English	92.7%
Spanish	4.2%
Other	3.1%
Primary payer for care	
Medicare (including FFS and Medicare Advantage)	47.9%
Medicaid	7.2%
Private	23.4%
Other (including uninsured and no payer) or Unknown	21.4%
Length of stay in program (through beginning of survey administration)	
6 weeks to 3 months	13.3%
4 to <6 months	19.3%
6 to <12 months	34.0%
12 to 24 months	33.5%
Residential setting	
Home	74.3%
Assisted living facility	5.1%
Unknown	20.6%
Primary diagnosis	
Cancer	13.6%
Alzheimer's or other dementia	9.4%
All other	77.1%
No. of in-person visits in 3 months	
1 to 2 times	12.7%
3 to 4 times	11.4%
5 to 6 times	10.4%
7 or more times	43.9%
Unknown	21.7%
Proxy assistance with survey response	
Proxy completed survey for patient	33.3%
Proxy assisted in some other way	14.0%
No proxy assistance	52.8%

(continued)

TABLE 1. (CONTINUED)

Characteristic	Respondents (%)[a], n = 2263
Self-reported functional status	
Able to leave house	85.9%
Able to get out of bed but not house	7.4%
Not able to get out of bed	6.7%
Self-reported physical health status	
Excellent	2.3%
Very good	10.0%
Good	29.5%
Fair	39.0%
Poor	19.2%
Self-reported mental health status	
Excellent	11.2%
Very good	22.4%
Good	33.4%
Fair	24.3%
Poor	8.7%

[a]Means and percentages were calculated among nonmissing values, except where large unknown categories are noted (i.e., payer for care, residential setting, and No. of in-person visits).
FFS, fee-for-service.

for feelings of anxiety or sadness. There was considerable variation across programs for many aspects of care, with the widest interquartile ranges for survey items related to getting needed help between visits, getting desired help for symptoms, planning for care, and support for family and friends (data not shown).

The five composites are moderately correlated (Table 3). Intercorrelations are highest between Care Coordination and other composites ($r = 0.446$ to $r = 0.615$), reflecting the core role that SIPs play in coordinating care through communication, planning, and assessing and managing symptoms.

Reliability

Across the 28 programs that had at least 10 respondents, there is adequate variation in measure scores, as indicated by the ICC and reliability estimates shown in Table 4. Six of the seven proposed measures exhibit acceptable program-level reliability of 0.70 or greater at 100 respondents; the remaining measure, Overall Rating, nears the threshold at reliability of 0.67 at 100 respondents.

Validity

Models including all composites account for 45% of the variance in overall rating and 45% of the variance in willingness to recommend the program, after adjusting for case-mix and survey mode. Among the models that consider each composite's association individually, the Care Coordination and Communication composites are the two strongest predictors of overall rating of care ($\beta = 0.56$ and $\beta = 0.56$, respectively) and willingness to recommend ($\beta = 0.57$ and $\beta = 0.56$, respectively) as shown in Table 5. These results indicate, for example, that compared to a respondent who selected ''Always'' (the most favorable response) to all questions in the Communication composite, a respondent

Composite and global measures and component survey items	Response options[a]	Factor loading	Corrected item-total correlation	Adjusted mean program-level top-box score[b,c]
Communication (Cronbach's alpha = 0.85)				**78.5%**
In the last three months, how often did people from this program spend enough time with you when they visited?	Never/Sometimes/Usually/**Always**	0.84	0.61	75.8%
In the last three months, how often did people from this program explain things to you in a way you could understand?	Never/Sometimes/Usually/**Always**	0.86	0.64	77.4%
In the last three months, how often did people from this program listen carefully to you?	Never/Sometimes/Usually/**Always**	0.93	0.69	83.7%
In the last three months, how often did you feel that people from this program cared about you as a whole person?	Never/Sometimes/Usually/**Always**	0.91	0.68	81.6%
In the last three months, how often did you feel heard and understood by people from this program?	Never/Sometimes/Usually/**Always**	0.97	0.66	73.9%
Care Coordination (Cronbach's alpha = 0.74)				**68.1%**
In the last three months, how often did people from this program seem to know the important information about your medical history?	Never/Sometimes/Usually/**Always**	0.78	0.51	69.4%
In the last three months, did someone from this program talk with you about the care or treatment you get from your other doctors or health care providers?	**Yes, definitely**/Yes, somewhat/No	0.71	0.53	61.2%
In the last three months, did someone from this program talk with you about all the medicines you are taking?	**Yes, definitely**/Yes, somewhat/No/I do not take any medicines	0.76	0.54	78.0%
Everyday activities include things like getting ready in the morning, getting meals, or going places in your community. In the last three months, did someone from this program talk with you about how to get help with everyday activities?	**Yes, definitely**/Yes, somewhat/No/I did not want to talk with this program about getting help with everyday activities	0.74	0.44	48.2%
In the last three months, when you contacted this program between visits, did you get the help you needed?[d]	**Yes, definitely**/Yes, somewhat/No	0.80	0.50	83.7%
Help for Symptoms (Cronbach's alpha = 0.69)				**60.0%**
In the last three months, did you get as much help as you wanted for your pain?[d]	**Yes, definitely**/Yes, somewhat/No/I did not want help for my pain	0.83	0.53	60.4%
In the last three months, did you get as much help as you wanted for your breathing?[d]	**Yes, definitely**/Yes, somewhat/No/I did not want help for my breathing	0.78	0.48	66.6%
In the last three months, did you get as much help as you wanted for your feelings of anxiety or sadness?[d]	**Yes, definitely**/Yes, somewhat/No/I did not want help for my anxiety or sadness	0.83	0.52	53.1%
Planning for Care (Cronbach's alpha = 0.73)				**55.5%**
Did someone from this program ever talk with you about what you should do during a health emergency?	**Yes, definitely**/Yes, somewhat/No	0.85	0.53	65.3%
Did someone from this program ever talk with you about what is important in your life?	**Yes, definitely**/Yes, somewhat/No	0.85	0.56	51.3%

(*continued*)

TABLE 2. (CONTINUED)

Composite and global measures and component survey items	Response options[a]	Factor loading	Corrected item-total correlation	Adjusted mean program-level top-box score[b,c]
Did someone from this program ever talk with you about what your health care options would be if you got sicker?	**Yes, definitely**/Yes, somewhat/No	0.81	0.56	50.0%
Support for Family and Friends (Cronbach's alpha = 0.70)				**70.8%**
In the last three months, did the people from the program involve your family members or friends in discussions about your health care as much as you wanted?[d]	**Yes, definitely**/Yes, somewhat/No	0.82	0.54	74.1%
In the last three months, did your family members or friends get as much emotional support as they wanted from this program?[d]	**Yes, definitely**/Yes, somewhat/No/My family members or friends did not want emotional support from this program	0.98	0.54	67.4%
Global Measure: Overall Rating of the Program				
Using any number from 0 to 10, where 0 is the worst care possible and 10 is the best care possible, what number would you use to rate your care from this program?	0, 1, 2, 3, 4, 5, 6, 7, 8, **9, 10**	N/A	N/A	75.3%
Global Measure: Willingness to Recommend the Program				
Would you recommend this program to your friends and family?	**Definitely Yes**/Probably Yes/Probably No/Definitely No	N/A	N/A	73.7%

[a]Top-box responses are noted in bold.

[b]Program level scores are adjusted for case-mix (response percentile, education, age, diagnosis, proxy use, self-reported functional status, and physical and mental health) and mode of survey administration. Adjusted program-level scores are calculated for each item assuming each program had population-average case mix and mode of survey administration. Adjusted program-level composite scores are then generated as the average of the adjusted program-level item scores for the items that compose the composite measure.

[c]Distributions of adjusted program-level scores are calculated restricting to only those programs with 10 or more respondents (28 out of 32 programs).

[d]A screening question(s) determines whether this evaluative survey question is applicable to the respondent.

who did not select "Always" in response to any of these questions would have a 56% lower chance of rating the program a 9 or 10 out of 10, or of definitely recommending the program to family and friends.

In models containing all composites, Communication is the strongest predictor of these outcomes ($\beta = 0.29$ and $\beta = 0.30$, respectively, for overall rating and willingness to recommend), with Planning for Care the second strongest predictor ($\beta = 0.23$ and $\beta = 0.19$, respectively).

Discussion

Nearly one in seven Americans is seriously ill; their care accounts for more than half of all health care expenditures.[35] Standardized, rigorously tested measures of care quality are needed to promote high-quality, person- and family-centered care for this vulnerable group. Our field test provides evidence of the reliability and validity of the proposed survey-based measures to assess the quality of home-based illness care provided by SIPs from the perspective of patients and their families. These measures address a critical gap in measurement of the quality of serious illness care,[13] particularly with regard to assessment of patients' perspectives on how well their care providers understand their priorities and support their decision making.[36]

In keeping with findings from hospice care and other settings, one of the strongest predictors of overall ratings of SIP care is Communication,[17,37–39] with Care Coordination and Planning for Care the second most important. The strong relationship between these domains and patients' overall

TABLE 3. CORRELATIONS AMONG PROPOSED SERIOUS ILLNESS CARE COMPOSITE MEASURES

	Communication	Care coordination	Help for symptoms	Planning for care	Support for family and friends
Communication	1				
Care coordination	0.615	1			
Help for symptoms	0.404	0.480	1		
Planning for care	0.392	0.579	0.433	1	
Support for family and friends	0.443	0.446	0.427	0.425	1

All correlations are significant at $p < 0.001$.

TABLE 4. INTRACLASS CORRELATION AND RELIABILITY
OF PROPOSED SERIOUS ILLNESS CARE QUALITY MEASURES

Measure	ICC	Reliability at 100 measure respondents
Composite measures		
Communication (5 items)	0.029	0.75
Care coordination (5 items)	0.034	0.78
Help for symptoms (3 items)	0.027	0.73
Planning for care (3 items)	0.036	0.79
Support for family and friends (2 items)	0.039	0.80
Global measures		
Rating of program (1 item)	0.020	0.67
Willingness to recommend (1 item)	0.032	0.77

All calculations use top-box scoring. ICCs are adjusted for case mix and mode of survey administration. Mean percentages of survey respondents completing measure are calculated as program-level averages (i.e., the average of each program's average percent of respondents completing the given measure) using all survey respondents within each included program. Reliabilities are calculated with the Spearman-Brown prophecy formula, reliability = $(k \times \text{ICC})/[(k-1)(\text{ICC}) +1]$, where k is the number of completed surveys per program. Values were calculated after restricting to programs with 10 or more respondents (28 out of 32 programs). The mean proportion of respondents responding to measures ranged from 95% to 97% for all measures with the exception of Help for Symptoms and Support for Family and Friends, for which an average of 75% and 78% of respondents responded.

ICC, intraclass correlation.

assessments of care reflect the key roles that SIPs play in helping seriously ill individuals understand the care options available to them, and acting as a central hub of information regarding medications, medical history, and supports for activities of daily living.

TABLE 5. USING SERIOUS ILLNESS SURVEY COMPOSITES
TO PREDICT GLOBAL MEASURE OF EXPERIENCE
WITH THE SERIOUS ILLNESS PROGRAM

	Models assessing composites one-by-one	Model including all composites
	Estimate	Estimate
Overall rating		
Communication	0.557***	0.292***
Care coordination	0.560***	0.154***
Help for symptoms	0.456***	0.139***
Planning for care	0.522***	0.232***
Support for family and friends	0.442***	0.085
Willingness to recommend		
Communication	0.563***	0.295***
Care coordination	0.565***	0.171***
Help for symptoms	0.438***	0.105*
Planning for care	0.501***	0.190**
Support for family and friends	0.471***	0.137**

Models are adjusted for case mix and mode of survey administration.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

A key challenge of providing high-quality home-based serious illness care is tailoring services to meet the specific needs and preferences of patients across a range of disease trajectories and functional status. For example, while most SIP patients look to their programs for help with activities of daily living, 21% of patients reported that they did not want help from the program for everyday activities. Of those respondents reporting that they had pain, trouble breathing, and anxiety or sadness, 6%, 7%, and 19%, respectively, reported that they did not want help from the program for those symptoms. These findings speak to the value of quality assessments that examine a range of care services that may be important to different groups of SIP patients.

Nearly half of our sample relied on a family member or friend to complete the survey for them or to assist them with reading the questions or writing responses. This underlines the importance of proxy assistance to represent seriously ill patients whose cognitive or other impairments interfere with survey response tasks. Family caregivers' responses have moderate-to-high agreement with patient responses regarding observable aspects of care.[40]

Although we recruited SIPs for participation in the field test from an extensive list of programs, there is no complete directory of all SIPs in the United States, so it is possible that the programs participating in our field test were not representative of all SIPs. In particular, the smallest programs did not have sufficient sample to participate.

Our overall survey response rate of 36.4% was similar to or better than that of other patient and family surveys in routine use.[41,42] Notably, the response rate to surveys administered by mail with telephone follow-up (42.5%) was substantially higher than the rate for surveys administered by mail only (30.4%). Mixed mode administration also increased the likelihood that those with Medicaid insurance responded to the survey.[43] We reduce nonresponse by allowing proxies to respond on behalf of patients who are not able to respond for themselves, and address nonresponse bias by adjusting for differences in case mix across programs,[28] which addresses nonresponse bias associated with these characteristics and allows for fair comparisons between programs.[29,44] Future research and survey efforts should continue to investigate approaches for promoting response from hard-to-reach populations.

Conclusion

We developed and tested a Serious Illness Survey that assessed of a broad range of care experiences that both seriously ill individuals and experts deem most important for high-quality serious illness care in the home. We evaluated the survey with patients receiving care in a diverse set of home-based SIPs across the United States.

We find support for the reliability and validity of the Serious Illness Survey for measuring and comparing care experiences. Results from the survey can be used to inform quality improvement efforts, monitor care quality over time, compare quality between programs, and assess the effectiveness of new initiatives that provide access to home-based serious illness care. Additional work is underway to test a version of the Serious Illness Survey to assess experiences with a broader range of providers caring for those with serious illness.

Authors' Contributions

Acknowledgments

Funding Information

Author Disclosure Statement

No competing financial interests exist.

Supplementary Material

Supplementary Appendix SA1
Supplementary Appendix SA2
Supplementary Appendix SA3

References

1. Long P, Abrams M, Milstein A, et al.: eds.: *Effective Care for High-Need Patients: Opportunities for Improving Outcomes, Value, and Health*. Washington, DC: National Academy of Medicine, 2017.

2. Cohn J, Corrigan J, Lynn J, et al.: *Community-Based Models of Care Delivery for People with Serious Illness. NAM Perspectives, Discussion Paper*. Washington, DC: National Academy of Medicine, 2017.

3. National Academies of Sciences, Engineering, and Medicine: *Models and Strategies to Integrate Palliative Care Principles into Care for People with Serious Illness: Proceedings of a Workshop*. Washington, DC: The National Academies Press, 2018.

4. Olson A, Harker M, Saunders R, et al.: *Innovations in Medicare Advantage To Improve Care for Seriously Ill Patients*. Washington, DC: Duke Margolis Center for Health Policy, 2019.

5. Centers for Medicare & Medicaid Services: VBID Model Hospice Benefit Component Overview. 2020. https://innovation.cms.gov/innovation-models/vbid-hospice-benefit-overview (Last accessed December 10, 2021).

6. Centers for Medicare & Medicaid Services: Primary Care First Model Options. 2021. https://innovation.cms.gov/innovation-models/primary-care-first-model-options (Last accessed December 10, 2021).

7. Centers for Medicare & Medicaid Services: Global and Professional Direct Contracting (GPDC) Model. 2021. https://innovation.cms.gov/innovation-models/direct-contracting-model-options (Last accessed December 10, 2021).

8. National Hospice and Palliative Care Organization: NHPCO Urges HHS, Congress to Support Community-Based Palliative Care Demo. 2020. https://www.nhpco.org/nhpco-urges-hhs-congress-to-support-community-based-palliative-care-demo (Last accessed December 10, 2021).

9. Teno JM, Price RA, Makaroun LK: Challenges of measuring quality of community-based programs for seriously ill individuals and their families. Health Affairs 2017;36:1227–1233.

10. Elliott MN, Cohea CW, Lehrman WG, et al.: Accelerating improvement and narrowing gaps: Trends in patients' experiences with hospital care reflected in HCAHPS public reporting. Health Serv Res 2015;50:1850–1867.

11. Rodriguez HP, von Glahn T, Elliott MN, et al.: The effect of performance-based financial incentives on improving patient care experiences: A statewide evaluation. J Gen Intern Med 2009;24:1281–1288.

12. Quigley DD, McCleskey SG: Improving care experiences for patients and caregivers at end of life: A systematic review. Am J Hosp Palliat Care 2020;38:84–93.

13. National Quality Forum: Issue Brief: Opportunities for Advancing Quality Measurement in Community-Based Serious Illness Care 2020.

14. RAND Corporation: RAND Serious Illness Survey. www.rand.org/Serious-Illness-Survey (Last accessed December 10, 2021).

15. Center to Advance Palliative Care: National Palliative Care Registry™. https://www.capc.org/national-palliative-care-registry (Last accessed December 10, 2021).

16. Kerr K: *Community-Based Model Programs for the Seriously Ill*. The Gordon and Betty Moore Foundation, 2017. https://www.moore.org/docs/default-source/patient-care-/report-model-programs-for-the-seriously-ill-may-2017-dls.%20pdf?sfvrsn=529b6c0c_2 (Last accessed December 10, 2021).

17. Anhang Price R, Stucky B, Parast L, et al.: Development of valid and reliable measures of patient and family experiences of hospice care for public reporting. J Palliat Med 2018;21:924–932.

18. Giordano LA, Elliott MN, Goldstein E, et al.: Development, implementation, and public reporting of the HCAHPS survey. Med Care Res Rev 2010;67:27–37.

19. Robert Wood Johnson Foundation: How to Report Results of the CAHPS Clinician & Group Survey. 2010.

20. Damiano PC, Elliott M, Tyler MC, et al.: Differential use of the CAHPS® 0–10 global rating scale by Medicaid and commercial populations. Health Serv Outcomes Res Methodol 2004;5:193–205.

21. Flora DB, Curran PJ: An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods 2004;9:466–491.

22. Brown TA: *Confirmatory Factor Analysis for Applied Research, 2nd ed*. New York, NY: Guilford Press, 2015.

23. Muthén BO: *Mplus Technical Appendices*. Los Angeles, CA: Muthén & Muthén, 1998–2004.

24. DiStefano C, Liu J, Jiang N, et al.: Examination of the weighted root mean square residual: Evidence for trustworthiness? Struct Equ Model Multidiscip J 2018;25:453–466.

25. Hu Lt, Bentler PM: Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model: A Multidiscip J 1999;6:1–55.

26. Asparouhov T: General multi-level modeling with sampling weights. Commun Stat Theory Methods 2006;35:439–460.

27. Wu J-Y, Kwok O-M: Using SEM to analyze complex survey aata: A comparison between design-based single-level and model-based multilevel approaches. Struct Equ Model Multidiscip J 2012;19:16–35.

28. DeYoreo M, Anhang Price R, Montemayor CK, et al.: Adjusting for patient characteristics to compare quality of care provided by serious illness programs. Forthcoming.

29. Elliott MN, Zaslavsky AM, Goldstein E, et al.: Effects of survey mode, patient mix, and nonresponse on CAHPS® hospital survey scores. Health Serv Res 2009;44(2 Pt 1):501–518

30. Allen MJ, Yen WM: *Introduction to Measurement Theory*. Monterey, CA: Brooks/Cole, 1979.

31. Nunnally JC, Bernstein IH: *Psychometric Theory*. New York, NY: McGraw-Hill, 1994.

32. Muthén B, du Toit SHC, Spisic D: Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes 1997.

33. Lee S-Y, Poon W-Y, Bentler PM: A three-stage estimation procedure for structural equation models with polytomous variables. Psychometrika 1990;55:45–51.

34. Muthén B: A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika 1984;49:115–132.

35. Institute of Medicine: *Dying in America: Improving Quality and Honoring Individual Preferences near the End of Life*. Washington, D.C.: The National Academies Press, 2015.

36. Sanders JJ, Paladino J, Reaves E, et al.: Quality measurement of serious illness communication: Recommendations for health systems based on findings from a symposium of national experts. J Palliat Med 2019;23:13–21.

37. Rhodes RL, Mitchell SL, Miller SC, et al.: Bereaved family members' evaluation of hospice care: What factors influence overall satisfaction with services? J Pain Symptom Manage 2008;35:365–371.

38. Paddison CA, Abel GA, Roland MO, et al.: Drivers of overall satisfaction with primary care: Evidence from the English General Practice Patient Survey. Health Expect 2015;18:1081–1092.

39. Elliott MN, Kanouse DE, Edwards CA, et al.: Components of care vary in importance for overall patient-reported experience by type of hospitalization. Med Care 2009;47:842–849.

40. McPherson CJ, Addington-Hall JM: Judging the quality of care at the end of life: Can proxies provide reliable information? Soc Sci Med 2003;56:95–109.

41. CAHPS® Hospice Survey: Scoring and Analysis. 2020. https://hospicecahpssurvey.org/en/scoring-and-analysis (Last accessed December 10, 2021).

42. Hospital CAHPS Survey: Summary Analyses. 2021. https://www.hcahpsonline.org/en/summary-analyses (Last accessed December 10, 2021).

43. DeYoreo M, Anhang Price R, Bradley M, et al.: Adding Telephone Outreach Can Increase Representativeness of Surveys of Seriously Ill People. Under Review.

44. Elliott MN, Edwards C, Angeles J, et al.: Patterns of unit and item nonresponse in the CAHPS® hospital survey. Health Serv Res 2005;40(6 Pt 2):2096–2119.

Address correspondence to:
Rebecca Anhang Price, PhD
RAND Corporation
1200 S Hayes Street
Arlington, VA 22202
USA

E-mail: ranhangp@rand.org