Published in final edited form as:

J Pain Symptom Manage. 2020 January ; 59(1): 1–8.e3. doi:10.1016/j.jpainsymman.2019.09.001.

Development and Validation of the ACP-CAT for Assessing the Quality of Advance Care Planning Communication

Jacqueline K. Yuen, MD,

Department of Medicine, The University of Hong Kong, Hong Kong SAR, China

Amy S. Kelley, MD, MSHS,

Brookdale Department of Geriatrics and Palliative Medicine, Icahn School of Medicine at Mount Sinai, New York, New York; Geriatric Research Education and Clinical Center, James J. Peters VA Medical Center, Bronx, New York

Laura P. Gelfman, MD, MPH,

Brookdale Department of Geriatrics and Palliative Medicine, Icahn School of Medicine at Mount Sinai, New York, New York; Geriatric Research Education and Clinical Center, James J. Peters VA Medical Center, Bronx, New York

Elizabeth E. Lindenberger, MD,

Brookdale Department of Geriatrics and Palliative Medicine, Icahn School of Medicine at Mount Sinai, New York, New York; Geriatric Research Education and Clinical Center, James J. Peters VA Medical Center, Bronx, New York

Cardinale B. Smith, MD, PhD,

Division of Hematology and Medical Oncology, Icahn School of Medicine at Mount Sinai, New York, New York

Robert M. Arnold, MD,

Division of General Internal Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania

Brook Calton, MD, MHS,

Division of Palliative Medicine, University of California, San Francisco, San Francisco, California

Jane Schell, MD,

Division of Renal-Electrolyte, University of Pittsburgh, Pittsburgh, Pennsylvania

Stephen H. Berns, MD

Division of Palliative Medicine, University of Vermont, Burlington, Vermont, USA.

Abstract

Address correspondence to: Jacqueline K. Yuen, MD, Department of Medicine, The University of Hong Kong, 4/F Professorial Block, Queen Mary Hospital, 102 Pokfulam Road, Pokfulam, Hong Kong SAR, China. jkyuen@hku.hk.
In addition, the authors would like to acknowledge Rachelle Bernacki, MD, MS, Rebecca Sudore, MD, Diane Meier, MD, and Bud Hammes, PhD for their input on the ACP-CAT items, Reena Karani MD, MHPE for her input on the OSCE design and tool development, Debby ten Hove, MS for her assistance with interrater reliability calculations, and the developers of VitalTalk for training the course faculty.

Context—High quality advance care planning (ACP) discussions are important to ensure patient receipt of goal-concordant care, however there is no existing tool for assessing ACP communication quality.

Objectives—To develop and validate a novel instrument that can be used to assess ACP communication skills of clinicians and trainees.

Methods—We developed a 20-item ACP Communication Assessment Tool (ACP-CAT) plus two summative items. Randomized rater pairs assessed residents' performances in video-recorded standardized patient encounters before and after an ACP training program using the ACP-CAT. We tested the tool for its 1) discriminating ability, 2) interrater reliability 3) concurrent validity, 4) feasibility, and 5) raters' satisfaction.

Results—Fifty-eight pre/post-training video recordings from 29 first-year internal medicine residents at Mount Sinai Hospital were evaluated. ACP-CAT reliably discriminated performance pre- and post-training (median score 6 vs. 11, $P < 0.001$). For both pre/post-training encounters, interrater reliability was high for ACP-CAT total scores (intraclass correlation coefficient or ICC = 0.83 and 0.82) and the summative items *Overall impression of ACP communication skills* (ICC = 0.73 and 0.80) and *Overall ability to respond to emotion* (ICC = 0.83 and 0.82). Concurrent validity was shown by the strong correlation between ACP-CAT total score and both summative items. Raters spent an average of 4.8 minutes to complete the ACP-CAT, found it feasible, and were satisfied with its use.

Conclusions—ACP-CAT provides a validated measure of ACP communication quality for assessing video-recorded encounters and can be further studied for its applicability with clinicians in different clinical contexts.

Keywords

Advance care planning; communication quality; assessment; tool development; validation

Introduction

Advance care planning (ACP) is an important part of delivering high-value care.¹⁻² ACP discussions enable clinicians to align care with what matters most to patients and has been shown to improve patients' satisfaction with care, decrease unwanted treatments near the end of life, and lower rates of anxiety and depression amongst patients and surrogates.³⁻⁶ To incentivize ACP discussions, the Centers for Medicare and Medicaid added payment codes to reimburse health care providers for ACP counseling in 2016.⁷

The Institute of Medicine Report *Dying in America* recommended the establishment of quality standards and trainings for clinicians to engage in high-quality ACP conversations.¹ The process for training clinicians to lead high-quality ACP discussions and monitoring its quality is not yet well defined. International consensus definitions for ACP and expert recommendations on the ACP process have only been proposed in recent years.⁸⁻⁹ Unlike routine medical consultations which focus on managing present clinical problems, ACP communication focuses on supporting patients to define their "personal values, life goals, and preferences regarding future medical care."⁸ Given the emphasis on planning for future

medical care, ACP conversations require different communication competencies and follow a different communication framework.^{1,10}

Instruments such as checklists that consist of observable communication tasks can provide a structure to support efforts to teach, assess, and improve communication skills of clinicians. In the era of the Accreditation Council for Graduate Medical Education's New Accreditation System, valid and feasible assessment tools are needed to assess communication skills competencies.¹¹ However, existing validated communication assessment instruments for routine clinical consultations do not assess the competencies specific to ACP conversations.¹²⁻¹³

For these reasons, we developed a novel Advance Care Planning Communication Assessment Tool (ACP-CAT) that reflects a values-based, patient-centered approach to ACP in line with expert-recommended best practices to facilitate teaching and evaluation of ACP communication quality.^{1-2,8-9}

The ACP-CAT is a 20-item instrument consisting of communication tasks plus two summative items to assess the quality of ACP communication in a clinician-patient encounter. This paper describes the development of this tool and its initial validation with first-year internal medicine residents before and after GOCARE, a training program on ACP communication previously reported in the literature.¹⁴ The instrument's psychometric properties, feasibility, and raters' satisfaction with the ACP-CAT are reported.

Methods

Tool development

Conceptual framework and content validity—The development of ACP-CAT was informed by an extensive literature review and consultation with content experts on ACP communication. Key communication tasks in ACP conversations were identified and a framework for ACP communication was developed. The ACP-CAT Framework consists of five sections:

Pre-Step: Clarifying health status and prognosis

Step 1: Assessing readiness and getting buy-in

Step 2: Identifying and involving health care proxy

Step 3: Exploring goals, values, and preferences

Step 4: Matching care plan to goals

Steps 1-4 describe phases in ACP conversations with adults in any stage of health. For patients with a serious or chronic progressive life-limiting condition, an additional phase *Pre-Step: Clarifying health status and prognosis* is also recommended early in the ACP process. This facilitates a shared understanding of the clinical context and allows patients to express goals and preferences with a realistic understanding of prognosis.

The ACP-CAT Framework is a flexible framework where the communication tasks need not all be completed in sequence and over one encounter by a single provider. The timing and content of discussions should be dictated by the patient/family's emotional readiness and practical considerations such as clinicians' time. The discussions are revisited when the clinical context, patient's preferences or family's needs change over time.^{1-2,8-9,15}

Item generation and pilot testing—Initially, the study investigators generated a list of 47 ACP communication tasks through a literature review of ACP communication interventions, expert opinions on ACP communication, existing validated communication assessment tools, and the investigators' own clinical and teaching experience in ACP communication^{1-2,8-10,12-16}. The majority of the communication tasks focus on content. These content items include topics to be covered in the five sections of the ACP-CAT Framework. Some communication tasks focus on process. These are communication behaviors or skills perceived to be important for the clinicians to demonstrate throughout the ACP process. Therefore, a sixth section, “*General communication skills*,” is added to categorize these skills items.

The initial list of ACP communication tasks was refined through consultations with four content experts who rated the importance of each communication task. Suggestions were also elicited to reduce redundancy, refine the wording of items, and prioritize ACP communication topics that can be discussed with adults in any stage of health. This process eliminated 27 items, leaving 20 items in the instrument.

The instrument was pilot tested with five physician raters in internal medicine, geriatrics, and palliative medicine with clinical experience in conducting ACP conversations. The raters independently evaluated video recordings of five simulated ACP encounters led by four internal medicine residents and a palliative care specialist. The group convened to identify items where the raters disagreed on the rating. The group modified the language of some items to improve clarity and consistency in marking each item through an iterative process.

The final ACP-CAT consists of 17 content items under the five sections of the ACP-CAT Framework. The sixth section consists of three general communication skills items (Appendix I). These items are *Asked permission to move forward using invitation questions* (the skill in assessing patient's readiness and information preferences in different phases of the discussion),^{1-2,17} *Avoided/explained medical jargon* (the skill in explaining medical information adapted to the patient's level of understanding), and *Used empathic responses* (the skill in expressing empathic understanding in response to emotion).^{1,17} The two summative items are *Overall impression of ACP communication skills* and *Overall ability to respond to emotion*. Both items used a 7-point rating scale where 0 = “Beginner” and 7 = “Expert” to reflect the Accreditation Council for Graduate Medical Education's milestones indicating progression from novice to expert in core clinical competencies.¹⁸

Coding rules—The first 17 content items are coded as “yes” if the task is done at least once in the encounter or “no” if not done. Items that contain the word “ANY” indicate that item is coded as “yes” if any of the listed topics for the item was discussed. Items that contain the word “AND” indicate the task is coded as “yes” only if all components of that

item are fulfilled. Each task may vary in difficulty level and amount of time required. The items may also be coded as “n/a” (not applicable) if the clinical context does not present an opportunity for the rater to assess the item. This option is provided to increase the flexibility of this instrument for assessment in different clinical contexts.

The final three communication skills items are important for the clinician to demonstrate whenever the opportunity is presented throughout the encounter. In the pilot testing phase, these items were given a scoring scheme of used “rarely,” “sometimes,” or “frequently” but this led to high variability in coding between raters. Subsequently, the scoring scheme was modified to the nominal (yes/no) scale but without an “n/a” option since these skills are considered applicable to all ACP contexts. A skill item is coded as “yes” if done at least some of the time during the encounter. It is coded as “no” when it is not done or done only once if more than one opportunity is presented.

Tool validation

Participants—To test the ACP-CAT and assess its psychometric characteristics, we applied a sample of video-recorded simulated ACP consultations in an observed structured clinical examination (OSCE) of first-year internal medicine residents before and two months after completion of the ACP training program GOCARE.¹⁵

All residents who completed the pre-training OSCE were eligible to participate in the study. The Institutional Review Board for the Icahn School of Medicine at Mount Sinai deemed this study exempt.

OSCE development—Two different cases of advanced cancer patients with poor prognosis were developed, one for the pre- and post-training OSCE respectively (Appendix II). Two standardized patients (SPs) were recruited and trained to enact in the role of the patients. The two SP encounters were designed to match in difficulty level and the number and type of opportunities to demonstrate ACP skills.

Given prognostic disclosure can be a challenging task for first-year residents, in the pre-training OSCE, the SP is trained to reveal knowledge of the poor prognosis if the resident has not disclosed the prognosis by five minutes into the encounter. This is done to allow time to assess other ACP communication tasks. In the post-training OSCE where residents would have received training in prognostic disclosure, the SP would not have knowledge of the prognosis unless informed by the resident in the encounter.

Participants were given a maximum of 15 minutes for each encounter to reflect the median conversation length for serious illness discussions in the outpatient oncology setting.¹⁹ All OSCE encounters were video-recorded.

Raters—Six physician faculty with communication skills teaching experience from three different institutions were recruited as raters. The raters underwent a 2-hour training to use the ACP-CAT. The training consisted of instruction on the ACP-CAT items and coding rules followed by practice using the tool to rate video-recorded simulated ACP encounters.

Each video was randomly assigned two raters from the pool of six. This rater arrangement was used (rather than all raters rate all videos) to reduce rater burden and to simulate a common and practical raters arrangement for OSCE assessments when large numbers of learners are assessed.

The raters were blinded to the pre/post-training status of the residents and independently assessed each video recording using the ACP-CAT. Subsequently, each rater pair compared ratings and reached consensus for items where there were disagreement. Both individual ratings and the consensus rating for each video recording were collected.

ACP-CAT total scores and section scores—To evaluate the discriminating ability of the ACP-CAT, total scores and scores for each of the six sections were generated. ACP-CAT total scores were tabulated by assigning one point for each item marked “yes” and zero points to items marked “no” and then summing the values across the 20 items. The section scores were generated by summing the points for each section. Any rating of “n/a” was assigned zero points. The rationale being the cases were standardized and the residents were given the same amount of time to demonstrate the communication tasks.

Data analysis—We used descriptive statistics for baseline characteristics. McNemar’s test was used to test for significant differences pre/post training for each individual item. Wilcoxon signed rank test was used to compare differences for ordinal measures including total scores, section scores, and the summative ratings.

We estimated interrater reliability for the ACP-CAT using intraclass correlation coefficient (ICC).²⁰ ICC was calculated using a two-way random effects model based on an absolute agreement of the average score of each pair of raters. ICC estimates above 0.40, 0.60, and 0.80 were considered as fair, moderate, and high level of agreement respectively.²¹ Markov chain Monte Carlo estimation²² was used to calculate ICCs and a 95% Bayesian credible interval (BCI). For all models, we used non-informative priors and the mode of the posterior distribution was reported.

Given there is no existing gold standard instrument for assessing ACP communication, concurrent validity was assessed by determining the Pearson correlation coefficient r between ACP total score and the summative item “*Overall impression of ACP communication skills*” to determine how well the ACP-CAT total score measured the construct of ACP communication quality. We also assessed for the correlation between ACP total score and the summative item “*Overall ability to respond to emotion*.”

Feasibility and rater satisfaction with the ACP-CAT was determined by the average time the raters took to mark each video recording and a two-item questionnaire that asked raters to rank the tool’s feasibility and satisfaction with the tool on a 4-point Likert scale.

We performed statistical analyses using SAS 9.3 (SAS Institute, Cary, North Carolina) and Stata 13 (StataCorp, College Station, TX).

Results

Participant Characteristics

Of the 39 eligible residents, 33 (85%) consented to participate. Amongst those, four were excluded who did not complete the post-training OSCE due to scheduling conflicts—leaving 29 participants (74%) and 58 pre/post training video-recordings in the final analyses (Table 1). Median age was 27 and 66% were female. Prior to GOCARE training, ninety percent of participants reported experience with leading at least one goals of care conversation in the inpatient setting. Only 14% had led an ACP discussion in the outpatient setting. Forty-one percent had received four or more hours of formal communication skills training prior to GOCARE.

Participants took on average 13 minutes (range 8-15 minutes) to complete the pretraining OSCE and 14 minutes (range 10-15 minutes) post-training.

Discriminating Ability of the ACP-CAT

The ACP-CAT differentiated performance of residents pre- and post-training. There was a significant increase in the ACP-CAT total score post-training (median score 6 vs. 11, $P < 0.001$) (Table 2). The ACP-CAT section scores also increased for all sections post-training with the exception of *Step 1: Assessing readiness and getting buy-in* which showed no significant change. The greatest section score increase was in *Step 2: Identifying and involving health care proxy* (median score 1 vs. 3, $P < 0.001$) and *Step 4: Matching care plan to goals* (median score 0 vs. 2, $P < 0.01$).

Overall impression of ACP communication skills ratings increased from a median of 1 (“Beginner”) to 3 (“Intermediate” or approaching competence), which was the goal competency level for first-year internal medicine residents. There was a significant increase in the proportion of participants who achieved a rating of 3 or higher posttraining (10% vs. 52%). In addition, there was a significant increase in *Overall ability to respond to emotion* ratings post-training (median rating 1 vs. 4).

Interrater Reliability

Interrater reliability (IRR) was determined separately for the pre- and post-training OSCEs to assess the extent to which it was affected by the case itself. The ACP-CAT total score showed high IRR and it did not differ significantly between cases (ICC = 0.82, 95% BCI [0.47-0.89] for Case 1 and ICC = 0.83, 95% BCI [0.45-0.92] for Case 2) (Table 3).

For the individual ACP-CAT items, the IRR for Case 1 was high (ICC estimates ranged from 0.80 to 0.97) with the exception of item 19 (*General communication skills: Avoid/explain jargon*), which showed poor IRR (ICC = 0.07, 95% BCI [0.01-0.78]). For Case 2, the IRR for individual items ranged from moderate to high (ICC estimates ranged from 0.67 to 0.96).

The two summative items showed high IRR for both cases (*Overall impression ACP communication skills* ICC = 0.80, 95% BCI [0.47-0.90] for Case 1 and ICC = 0.73, 95% BCI [0.32-0.86] for Case 2; *Overall ability to respond to emotion* ICC = 0.82, 95% BCI [0.54-0.91] for Case 1 and ICC = 0.83, 95% BCI [0.25-0.84] for Case 2).

Concurrent Validity

A strong positive correlation was found between ACP-CAT total score and *Overall impression of ACP communication skills* rating ($r = 0.87$, $P < 0.001$) (Table 4). Positive correlation between ACP-CAT total score and *Overall ability to respond to emotion* rating was also observed ($r = 0.72$, $P < 0.001$).

ACP-CAT Feasibility and Satisfaction

Raters spent an average of 4.8 minutes to complete the ACP-CAT. The raters found the tool feasible to use (mean rating 3.7 out of 4) and were satisfied with the tool (mean rating 3.5 out of 4).

Discussion

This study describes the development and initial validation of the ACP-CAT for assessing ACP communication using a sample of OSCE video recordings for first-year internal medicine residents before and after an ACP training program. The systematic development process of the ACP-CAT informed by research and education literature and consultation with content experts supported its content validity.

The ability of the ACP-CAT to discriminate pre/post-training performance on ACP communication was high. From a baseline median of six items on the ACP-CAT, the residents demonstrated an average increase of five additional items post-training (or 25% of the total maximum score). This observation supports the use of ACP-CAT to evaluate the impact of ACP communication training programs.

All individual ACP-CAT section scores showed significant increase except for *Step 1: Assessing readiness and getting buy-in* which showed no significant change. The *Step 1* communication tasks reflect Fried's conceptualization of engagement in ACP as a health behavior where individuals vary in their degree of readiness and require different levels of motivation²³⁻²⁴.

While the lack of improvement in *Step 1* may suggest that the training did not impact the residents' ability to perform these tasks, an alternative possibility is that the OSCEs used cases where the patient already has buy-in to discuss ACP, which obliterated the need for the residents to perform these tasks. This implies that certain clinical contexts may not allow a clinician's ability to perform particular ACP communication tasks to be fully assessed. A variety of clinical scenarios may be needed (e.g. a patient who is less ready to discuss ACP) to gauge performance on all aspects of ACP communication.

The IRR estimates were excellent for the ACP-CAT total score and the two summative items and did not differ significantly between the two cases. Furthermore, the use of randomly assigned pairs of raters from a pool of raters in our design is expected to have lower IRR compared with using the same rater pair to rate all subjects. Therefore, the high IRR estimates we found nonetheless are compelling.

For the individual ACP-CAT items, the average IRR was high for both cases. However, there was some variability in the ICC estimates between individual items and between cases for some items. One particular item (*Item 19-General communication skills: Avoid/explain medical jargon*) showed high IRR in Scenario 1 but low IRR in Scenario 2. A potential explanation may be the different interpretations of the raters on what constitutes “medical jargon.” By contrast, a similar item in the SEGUE Framework, an instrument for assessing general communication skills in the medical interview, “*Adapt to patient’s level of understanding (e.g., avoid/explain jargon)*,” showed high IRR when used for assessing both videotaped SP encounters and real-time SP encounters.¹² In the future, modification of *Item 19* in the ACP-CAT to “*Adapt to patient’s level of understanding*” may help to improve the IRR of this item by focusing the rater on the goal of this communication task rather than on judging whether or not a terminology used is medical jargon.

The strong positive correlation between the ACP-CAT total score and the summative item *Overall impression of ACP communication skills* provides initial evidence to support the criterion validity of the instrument. A significant correlation between the ACP-CAT total score and *Overall ability to respond to emotion* was also observed. While the skills in responding to emotion has been noted to be important in ACP communication,¹ further studies are needed to examine the relationship between empathic communication skills and ACP communication.

This study has several limitations. First, the study was conducted with residents in the same training level at one institution which limits its generalizability. Second, the small sample size limited the precision of interrater reliability estimates with wide credible intervals observed for some items. Third, the testing of the ACP-CAT using two clinical scenarios limits the ability to assess whether the performance of the ACP-CAT would be consistent under different clinical contexts. Finally, because different cases were used pre/post-training, the improvement in ACP performance post-training cannot exclude the possibility that potential differences in the difficulty level of the two cases played a role.

Further validation studies with larger samples of trainees and clinicians with variable ACP communication experience in both simulated and real clinical scenarios are needed. Additional research to assess the correlation between ACP communication quality and patient outcomes such as care satisfaction and goal-concordant care would also be of important value.

Conclusion

The ACP-CAT is a novel instrument for assessing the quality of ACP communication that uses a values-based, patient-centered framework for ACP. Initial testing of the tool indicates high interrater reliability and high sensitivity in discriminating residents’ ACP communication performance pre- and post-training in video-recorded SP encounters. Raters found the instrument feasible and were satisfied with its use. The ACP-CAT has the potential to improve the teaching of ACP communication skills and to serve as a quality measure of clinician-led ACP conversations. More studies are needed for additional validation as well to assess its benefits for patients in clinical settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Disclosures and Acknowledgments

Funding: This work was supported by the Arnold P. Gold Foundation; In addition, Dr. Yuen is supported in part by the John A. Hartford Foundation (J.K.Y.); Dr. Kelley by the National Institute on Aging (NIA) [1K23AG040774-01A1], the American Federation for Aging Research, and the National Palliative Care Research Center (NPCRC); Dr. Gelfman by NIA [K23AG04993], Mount Sinai Older Adults Independence Center [P30AG028741], and NPCRC; Dr. Lindenberger by the Cambia Health Foundation; Dr. Smith by the American Cancer Society [MRSF 1323201].

Role of funding sources: No funding bodies had a role in the design and conduct of the study, data collection, analysis, interpretation of the data, and writing of the manuscript. The authors declare no conflicts of interest.

References

1. Institute of Medicine. Dying in America: Improving quality and honoring individual preferences near the end of life. Washington, DC: The National Academies Press;2015.
2. Bernacki RE, Block SD. Communication about serious illness care goals: a review and synthesis of best practices. *JAMA Intern Med.* 2014;174:1994–2003. [PubMed: 25330167]
3. Detering KM, Hancock AD, Reade MC, Silvester W. The impact of advance care planning on end of life care in elderly patients: randomised controlled trial. *BMJ.* 2010;340:c1345. [PubMed: 20332506]
4. Wendler D, Rid A. Systematic review: the effect on surrogates of making treatment decisions for others. *Ann Intern Med.* 2011;154:336–346. [PubMed: 21357911]
5. Teno JM, Gruneir A, Schwartz Z, Nanda A, Wetle T. Association between advance directives and quality of end-of-life care: a national study. *J Am Geriatr Soc.* 2007;55:189–194. [PubMed: 17302654]
6. Bernacki R, Paladino J, Neville BA, et al. Effect of the Serious Illness Care Program in outpatient oncology: a cluster randomized clinical trial. *JAMA Intern Med.* 2019 3 14. doi: 10.1001/jamainternmed.2019.0077. [Epub ahead of print]
7. Center for Medicare and Medicaid Services. Proposed policy, payment, and quality provisions changes to the Medicare Physician Fee Schedule for Calendar Year 2016. <http://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2015-Fact-sheets-items/2015-10-30-2.html>. Published 10 30, 2015 Accessed June 8, 2016
8. Sudore RL, Lum HD, You JJ, et al. Defining advance care planning for adults: a consensus definition from a multidisciplinary Delphi panel. *J Pain Symptom Manage.* 2017;53:821–832.e1. [PubMed: 28062339]
9. Rietjens JA, Sudore RL, Connolly M et al. Definition and recommendations for advance care planning: an international consensus supported by the European Association for Palliative Care. *Lancet Oncol* 2017;18:e543–51. [PubMed: 28884703]
10. Fahner JC, Beunders AJM, van der Heide A et al. Interventions guiding advance care planning conversations: a systematic review. *J Am Med Dir Assoc.* 2019 3;20(3):227–248. [PubMed: 30471947]
11. Nasca TJ, Philibert I, Brigham T et al. The next GME accreditation system—rationale and benefits. *N Engl J Med.* 2012;366:1051–1056. [PubMed: 22356262]
12. Makoul G The SEGUE Framework for teaching and assessing communication skills. *Patient Educ Couns* 2001;45:23–34. [PubMed: 11602365]
13. Elwyn G, Hutchings H, Edwards A, et al. The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. *Health Expectations.* 2005;8:34–42. [PubMed: 15713169]
14. Berns SH, Camargo M, Meier DE, Yuen JK. Goals of Care Ambulatory Resident Education: training residents in advance care planning conversations in the outpatient setting.

15. Sudore RL, Fried TR. Redefining the "planning" in advance care planning: preparing for end-of-life decision making. *Ann Intern Med.* 2010;153:256–261. [PubMed: 20713793]
16. Gelfman LP, Lindenberg E, Fernandez H, et al. The effectiveness of the Geritalk communication skills course: a real-time assessment of skill acquisition and deliberate practice. *J Pain Symptom Manage.* 2014;48(4):738–744 e731-736. [PubMed: 24681183]
17. Baile WF, Buckman R, Lenzi R, et al. SPIKES-A six-step protocol for delivering bad news: application to the patient with cancer. *Oncologist.* 2000;5:302–11. [PubMed: 10964998]
18. The Accreditation Council for Graduate Medical Education. Milestones. 2016 Available from: <https://www.acgme.org/What-We-Do/Accreditation/Milestones/Overview>. Assessed May 7, 2019.
19. Geerse OP, Lamas DJ, Sanders JJ, et al. A qualitative study of serious illness conversations in patients with advanced cancer. *J Palliat Med* 2019; 26. doi:10.1089/jpm.2018.0487. [Epub ahead of print]
20. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996;1:30–46.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174. [PubMed: 843571]
22. Little TD, Jorgensen TD, Lang KM, Moore EWG. On the joys of missing data. *Journal of Pediatric Psychology* 2013;39:151–162. [PubMed: 23836191]
23. Fried TR, Redding CA, Robbins ML, et al. Stages of change for the component behaviors of advance care planning. *J Am Geriatr Soc* 2010;58:2329–2336. [PubMed: 21143441]
24. Orleans T, Cassidy EF. Health and behavior In: Kovner AR, Knickman JR eds. *Health care delivery in the United States*, 10th ed. New York: Springer Publishing Company, 2011:125–149.

Table 1.**Participant Characteristics**

Characteristic	
Age, median (range)	27 (26-35)
Female, No. (%)	19 (66)
Number of goals of care conversations in inpatient setting, No. (%)	
None	3 (10)
1-3	17 (59)
4-9	7 (24)
10 or more	2 (8)
Number of goals of care conversations in outpatient setting, No. (%)	
None	25 (86)
1-3	4 (14)
4-9	0 (0)
10 or more	0 (0)
Number of times personally delivered bad news, No. (%)	
None	2 (8)
1-3	19 (66)
4-9	7 (24)
10 or more	1 (3)
Number of hours of formal communication training	
None	3 (10)
1-3 hours	14 (48)
4-9 hours	9 (31)
10 or more hours	3 (10)

Table 2.

Comparison of Pre- and Post-Training Performance on ACP-CAT and Summative Items.

Score	Pre-training (n = 29)			Post-training (n = 29)			P value	Maximum Possible Score
	Median (Q1-Q3)	Min	Max	Median (Q1-Q3)	Min	Max		
ACP-CAT total score	6 (4-9)	2	13	11 (8-14)	5	17	<0.001	20
Pre-Step score	1 (1-1)	1	2	2 (2-2)	0	3	0.003	3
Step 1 score	1 (1-2)	0	2	1 (1-2)	0	2	0.94	2
Step 2 score	1 (0-2)	0	3	3 (2-3)	1	3	<0.001	4
Step 3 score	1 (0-1)	0	3	2 (1-3)	0	4	<0.001	5
Step 4 score	0 (0-1)	0	3	2 (1-3)	0	3	0.012	3
General communication skills score	1 (1-2)	0	3	2 (1-2)	0	3	0.006	3
Overall impression of ACP skills	1 (1-2)	0	3	3 (2-4)	1	5	<0.001	7
Overall ability to respond to emotion	1 (1-2)	1	4	2 (2-4)	1	6	<0.001	7

Table 3.

Interrater reliability of ACP-CAT items.

ACP-CAT Item	Case 1		Case 2	
	ICC ^a	95% BCI ^b	ICC ^a	95% BCI
<i>(1) Pre-Step-Assessed understanding</i>	0.96	[.66–.98]	–	
<i>(2) Pre-Step-Explained health/prognosis</i>	0.67	[.04–.93]	0.92	[.43–.97]
<i>(3) Pre-Step-Checked for comprehension</i>	0.94	[.36–.97]	0.80	[.24–.95]
<i>(4) Step 1-Gave ACP rationale</i>	0.80	[.12–.94]	0.82	[.23–.94]
<i>(5) Step 1-Asked prior ACP behaviors</i>	0.75	[.19–.90]	0.92	[.50–.97]
<i>(6) Step 2-Explained purpose of HCP</i>	0.78	[.26–.95]	0.93	[.48–.97]
<i>(7) Step 2-Discussed HCP criteria</i>	0.96	[.63–.98]	0.93	[.60–.97]
<i>(8) Step 2-Encouraged family discussions</i>	0.87	[.49–.94]	0.93	[.43–.97]
<i>(9) Step 2-Explored flexibility/leeway</i>	–		–	
<i>(10) Step 3-Asked important goals</i>	0.77	[.44–.92]	0.97	[.70–.98]
<i>(11) Step 3-Assessed prior experience</i>	0.94	[.52–.97]	0.81	[.15–.96]
<i>(12) Step 3-Explored fears/worries</i>	0.93	[.40–.97]	0.82	[.04–.93]
<i>(13) Step 3-Explored valued life activities</i>	0.91	[.39–.96]	0.94	[.54–.97]
<i>(14) Step 3-Asked critical functions</i>	0.67	[.09–.93]	–	
<i>(15) Step 4-Aligned with goals</i>	0.83	[.36–.94]	0.97	[.76–.98]
<i>(16) Step 4-Recommended care plan</i>	0.90	[.47–.96]	0.85	[.36–.94]
<i>(17) Step 4-Summarized/follow up plan</i>	0.89	[.54–.96]	0.97	[.78–.98]
<i>(18) General-Assessed readiness/invitation</i>	0.90	[.08–.97]	0.93	[.24–.97]
<i>(19) General-Avoided medical jargon</i>	0.88	[.11–.96]	0.07	[.01–.78]
<i>(20) General-Used empathic responses</i>	0.69	[.03–.91]	0.91	[.33–.96]
<i>ACP-CAT total score</i>	0.83	[0.46–0.92]	0.82	[.47–.89]
Summative items				
<i>Overall impression of ACP skills</i>	0.73	[0.32–0.86]	0.80	[.47–.90]
<i>Overall ability to respond to emotion</i>	0.83	[0.25–0.84]	0.82	[.54–.91]

^aICC = intraclass correlation coefficient^bBCI = Bayesian credible intervals

Table 4.

Correlation between ACP-CAT total score and summative items.

Items Compared	CC ^a	P value
ACP-CAT total score & Overall impression of ACP communication skills	0.868	<0.001
ACP-CAT total score & Overall ability to respond to emotion	0.722	<0.001

^aCC: Correlation coefficient (Pearson r)