

A New Standard for Advance Care Planning (ACP) Conversations in the Hospital: Results from a Delphi Panel



Deepika Mohan, MD, MPH¹ , Olivia A. Sacks, MD², James O'Malley, PhD², Mark Rudolph, MD³, Julie Bynum, MD⁴, Megan Murphy, MS², and Amber E. Barnato, MD, MPH, MS²

¹Department of Critical Care Medicine, University of Pittsburgh, Pittsburgh, PA, USA; ²The Dartmouth Institute for Health Policy & Clinical Practice, Geisel School of Medicine at Dartmouth, Lebanon, NH, USA; ³Sound Physicians, Tacoma, WA, USA; ⁴Department of Medicine, University of Michigan, Ann Arbor, MI, USA.

BACKGROUND: Fewer than half of the US population has an advance healthcare directive. Hospitalizations offer a key opportunity for clinicians to engage patients in advance care planning (ACP) conversations. Guidelines suggest screening for the presence of “serious illness” but do not further specify how to prioritize the 12.4 million patients hospitalized each year.

OBJECTIVE: To establish a normative standard for prioritizing hospitalized patients for ACP conversations.

DESIGN AND SETTING: A modified Delphi study, with three iterative rounds of online surveys.

PARTICIPANTS: Multi-disciplinary group of US-based clinicians with research and practical expertise in ACP.

MAIN MEASURES: Indirect and direct elicitation of short-term and 1-year risk of mortality that prompt experts to prioritize ACP conversations for hospitalized adults.

MAIN RESULTS: Fifty-seven of 108 (52%) candidate panelists completed round 1, and 47 completed rounds 2 and 3. Panelists were primarily physicians (84%), with significant experience (mean years 23 [SD 9.8]), who either taught (55%) and/or performed research about ACP (55%). In round 1, > 70% of panelists agreed that all hospitalized adults \geq 65 years should have an ACP conversation before discharge, but disagreed about the timing and content of the conversation. By round 3, > 70% of participants agreed that patients with either high (> 10%) short-term or high (\geq 34%) 1-year risk of mortality should have a goals of care conversation (i.e., focused on preferences for near-term treatment), while patients with low (\leq 10%) short-term and low (< 19%) 1-year risk of mortality warranted an ACP conversation (i.e., focused on preferences for future care) before discharge.

LIMITATIONS: Use of case vignettes to elicit clinician judgment; response rate.

CONCLUSIONS: Panelists agreed that clinicians should have an ACP conversation with all hospitalized adults over 65 years in an ACP conversation, adjusting the

content and timing of the conversation conditional on the patient's risk of short-term and 1-year mortality.

KEY WORDS: advance care planning; Delphi survey; physician performance.

J Gen Intern Med 36(1):69–76

DOI: 10.1007/s11606-020-06150-0

© Society of General Internal Medicine 2020

Numerous stakeholders recommend advance care planning (ACP) to improve the quality of care that patients receive as they approach the end-of-life.^{1–3} Hospitalizations offer one opportunity for clinicians to initiate ACP conversations with patients.⁴ However, high-quality conversations, which allow patients to reveal (or potentially construct) their preferences, require clinicians to have the communication skills, the willingness to engage in emotionally complex interactions, and the time necessary to facilitate this process.^{5–8} Guidelines suggest screening patients to prioritize those with near-term mortality or morbidity risk based on the presence of “serious illness,” defined as the presence of a condition that carries a high risk of mortality or impacts quality of life.⁴ In the absence of a quantifiable definition of this term, the surprise question (which requires the treating clinician to consider whether or not he/she would be surprised if the patient died in the next year) has been widely promoted.^{9, 10} Pooled results of two different meta-analyses, however, suggest poor to modest accuracy of the surprise question for predicting death at 12 months.^{10, 11} Efforts to improve the quality of care for patients at the end-of-life therefore require better strategies to screen and prioritize patients for ACP conversations.

The objective of this study was to establish a consensus-based normative standard for risk of mortality that should prompt hospitalists to have an ACP conversation with their patients. Recognizing that people, even experts, struggle with probability-based judgments, we embedded a behavioral experiment within a Delphi process, sequentially presenting experts with cases selected from across the distribution of

The work was performed at the Dartmouth Institute for Health Policy and Clinical Practice.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-020-06150-0>) contains supplementary material, which is available to authorized users.

Received July 22, 2020

Revised July 22, 2020

Accepted August 12, 2020

Published online August 20, 2020

mortality risk and observing their judgments as the sampling frame changed. We hypothesized that experts would be more likely to recommend an immediate ACP conversation as the risk of mortality increased.

MATERIALS AND METHODS

Study Design

We used a modified Delphi method to establish an empirical standard for the threshold of mortality at which hospitalized patients should have ACP conversations. In an area that lacks certainty, the Delphi method uses multiple rounds of structured feedback to achieve consensus.^{12–15} We began with the consensus definition of ACP developed by Sudore et al., as “a process that supports adults, at any age or stage of health in understanding and sharing their personal values, life goals, and preferences regarding future medical care.”¹⁶ We conducted three rounds of surveys from March 2019 to June 2019, using *Qualtrics Survey Software* (Provo, UT) to collect anonymous responses to an iterative series of questions. Questions focused on establishing a threshold for the risk of mortality (short-term and 1-year) at which > 70% of clinicians agreed that hospitalized patients would benefit from ACP conversations, and used both case vignettes as well as explicit questions about prognosis to elicit these judgments.¹⁷ The Dartmouth Institutional Review Board reviewed and approved the study (IRB# 31186).

Expert Panel

We identified a multi-disciplinary group of clinicians with academic and practical expertise in ACP to participate in the Delphi process (see [Appendix](#) for details of sampling frame). We decided not to include patients, surrogates, or non-expert physicians on the panel because our objective was to establish to a clinical standard for screening patients.

We contacted candidate panelists through email, explained the purpose and design of the study, and offered a wage-based, fixed, monetary honorarium (paid through a link, embedded in the email invitation, to a \$25 Amazon gift card). Clinicians, who completed a round, were invited to participate in subsequent rounds. We provided an additional \$25 honorarium to those who completed all three rounds of surveys.

Instrument Design

We used three methods to generate consensus among panelists. First, we selected cases for each round to probe areas of controversy. Second, we provided aggregated feedback to panelists after each round so that they could consider and either incorporate or reject that information into their personal judgment. Third, we became increasingly transparent about the design of the study with each round. Once we developed the survey, we beta-tested it with 13 clinicians and made iterative modifications to minimize respondent burden and maximize clarity.

Demographic and Personal Characteristics. The instrument included 11 items designed to capture the demographic and personal characteristics (e.g., age, sex, training) of the respondent. It also included items specific to the respondent’s ACP practices, including their source of expertise and the number of ACP conversations held per week.

Elicitation of Mortality Threshold for Patient Selection.

Given the observation that people, even experts, lack insight into their own cognitive processes,¹⁷ we used two methods to establish the mortality threshold at which expert clinicians would recommend an ACP conversation: case vignettes to observe the use of the threshold in practice (see [Table 1](#) and the [Appendix](#) for details of development) and direct elicitation of judgment.

Indirect Elicitation Via Case Vignettes. After presenting each vignette, we asked panelists whether or not they believed an ACP conversation was indicated since no prior planning had occurred. Those who replied affirmatively were then asked to rate the priority of the conversation: (1) the patient’s primary care physician should have it in the outpatient setting; (2) the hospitalist should have it before discharge; (3) the hospitalist should have it that day. In response to feedback from panelists, after round 1, we further refined the description of the Likert options by changing the description of the current day option to “the treating hospitalist should have a “goals of care” conversation that day” (defined as discussions around decisions about near-term treatment choices and the intensity of care) with the other response options unchanged.

Round 1. Round 1 included 9 vignettes: 8 experimental stimuli plus one distractor case that we included as a check on the validity of the task and the panelists’ attention. We categorized mortality as low or high (short-term: $\leq 10\%$ or $> 10\%$; 1-year mortality: $< 34\%$ or $\geq 34\%$), and used a 3×2 factorial design to select cases, systematically varying the short-term and 1-year mortality as well as the presence or absence of dementia. Finally, we asked panelists an open-ended question about what drove their decision making for each case.

Round 2. Round 2 included 4 vignettes, selected to represent mortalities in the middle of each distribution. Given concerns about the burden imposed on respondents by the 3×2 factorial design, we dropped dementia as predictor of panelist recommendations.

Round 3. Round 3 included 5 vignettes. Based on responses in round 2, we further stratified the lower 1-year mortality group into two categories: low mortality: $< 19\%$; intermediate mortality: 19–33%. We also dropped high mortality risk from the factorial design.

Direct Elicitation. In rounds 2 and 3, we asked panelists to state the short-term and 1-year risk of mortality and the risk of failure

Table 1 Description of Case Vignettes, Stratified by Risk of Short-term, 1-Year Mortality (with Risk Calculators), and Presence or Absence of Dementia

Round	Case description	Short-term mortality	Risk calculator	1-year mortality	Risk calculator	Dementia
1	54-year-old female aerobics instructor presenting with left lower quadrant pain and fever	Low 0%	SOFA	Low < 1%	Actuarial table	No
3	65-year-old female presenting with epigastric burning, weakness, fatigue, and tarry bowel movements	1.2%	AIMS-65	4% [< 17.5%]	Walter [Fischer]	No
1	80-year-old female presenting with altered mental status, lethargy, and urinary incontinence	Low < 10%	SOFA	Intermediate 19%	Walter	Yes
1	66-year-old female with a history of breast cancer, presenting with fever, shortness of breath, and increased leg edema	8–9%	PORT/PSI	19% [≥ 17.5–49%]	Walter [Fischer]	No
2	69-year-old female with COPD on home oxygen, presenting with worsening shortness of breath and increased O ₂ requirement	8.4%	DECAF	19% [≥ 17.5–49%]	Walter [Fischer]	No
3	75-year-old female who underwent hip surgery recently, presenting with right lower extremity swelling, chest pressure, and shortness of breath	3.2–7.1%	PESI	19%	Walter	No
1	81-year-old female presenting from nursing home with worsening weakness, lethargy, and nausea	Low < 10%	Luo/Liano	High 34%	Walter	Yes
1	65-year-old female with stage IV colorectal cancer, presenting with bright red blood per rectum	7%	ACS-NSQIP	64% [> 49%]	Walter [Fischer]	No
2	82-year-old male with a history of stage IV CKD2 and hypertension, presenting with worsening shortness of breath and edema	1–5%	GWTC-HF	34%	Walter	No
3	80-year-old male presenting with fever, shortness of breath, and hypoxia	8–9%	PORT/PSI	34%	Walter	No
3	70-year-old male with recent hospitalization for osteomyelitis of the spine, presenting with weakness and lethargy	High 27–29%	PORT/PSI	Low 4%	Walter	No
1	75-year-old female with a history of chronic urinary tract infections, who recently finished a Medrol dose pack for a neck sprain, presenting with vomiting and agitation	High 40%	SOFA	Intermediate 19%	Walter	Yes
1	70-year-old female with a complicated cardiac history, presenting with chest pain	40%	TIMI	19%	Walter	No
2	75-year-old female with a history of hypertension and atrial fibrillation, presenting with chest pain, shortness of breath, and vomiting	16.5%	AIMS-65	19%	Walter	No
3	71-year-old male s/p CABG and with CKD on dialysis, presenting with left-sided chest pain radiating to both arms	20%	TIMI	19%	Walter	No
1	86-year-old wheelchair-confined male, presenting with aphasia and weakness	High 20%	MSOAR	High 64%	Walter	Yes
2	72-year-old male with a history of COPD, on home oxygen, s/p PEG tube, presenting with weakness and productive cough	27–29%	PORT/PSI	34%	Walter	No

COPD, chronic obstructive pulmonary disease; CKD, chronic kidney disease; CABG, coronary artery bypass graft; PEG, percutaneous endoscopic gastrostomy

to return to baseline physical and cognitive functioning that would prompt them to engage in a goals of care conversation with a hospitalized patient, using a slider to select the exact value between 0 and 100%. Finally, we included the surprise question, “would you be surprised if the patient died in the next year?” after each case, to further elucidate the calibration of clinician prognostication about mortality. We did not include these questions in round 1 to avoid priming panelists.

Other Determinants of ACP Conversations. To ensure that we captured other potential determinants of panelists’ screening decisions, we included an open-ended question

about advice the respondent would offer to trainees on how to prioritize ACP conversations in the hospital in round 1.

Statistical Analyses

We summarized demographic and personal characteristics of panelists using counts (percentages) and means (standard deviations) as appropriate. We summarized responses to the case vignettes, using proportions, and the direct elicitation slider questions, using medians (inter-quartile ranges). We used Spearman’s correlation coefficient to test the correlation between responses to the surprise question and the predicted probability of 1-year mortality.

Predicted probability of 1-year mortality = the probability of inpatient mortality (based on disease-specific risk calculator) * probability of 1-year post-hospitalization mortality conditional on survival to discharge (based on the Walter score).(1)

To assess other determinants of panelists' screening decisions, one author (OS) coded all the responses to the open-ended questions, using content analysis to identify dominant themes.

RESULTS

Baseline Characteristics

Of the 108 clinicians who were successfully emailed with an invitation to participate in the study, 57 (52%) completed round 1 of the Delphi process (Fig. 1). Of these initial respondents, 47 completed rounds 2 and 3 (82%). The mean age of study participants was 48.5 (SD 9.9) and 31 were male (54%). The majority were physicians (84%), working in the acute care setting (73%), with 23 (SD 9.8) years since the completion of medical school. Of the clinicians who participated in the Delphi process, the majority (55%) conducted research in ACP and were educators (55%) in the field. A summary of the participants' characteristics is provided in Table 2.

Table 2 Characteristics of Delphi Panelists

Characteristic	Total (N = 57)
Age—years, mean (SD)	48.5 (9.9)
Sex—male, n (%)	31 (54)
Race/ethnicity, n (%)	
Asian	10 (18)
Black or African American	1 (2)
Hispanic or Latino	1 (2)
White, non-Hispanic	45 (79)
Credentials, n (%)	
Physician	51 (90)
Specialty, n (%)	
Internal medicine	44 (86)
Family practice	3 (6)
General surgery	2 (4)
Neurology	1 (2)
Anesthesiology	1 (2)
Years since completion of medical school, mean (SD)	23 (9.8)
Nurse practitioner	3 (5)
Registered nurse	3 (5)
Practice setting, n (%)	
Acute care hospital (inpatient)	43 (75)
Clinic (outpatient)	9 (16)
Other	5 (9)
Source of expertise in advance care planning, n (%)	
Research	32 (56)
Quality improvement	20 (35)
Educator	32 (56)
Additional training	27 (47)
Number of advance care planning conversations performed in an average week of clinical work, n (%)	
0–5	21 (37)
6–10	14 (25)
11–15	17 (30)
> 15	6 (10)

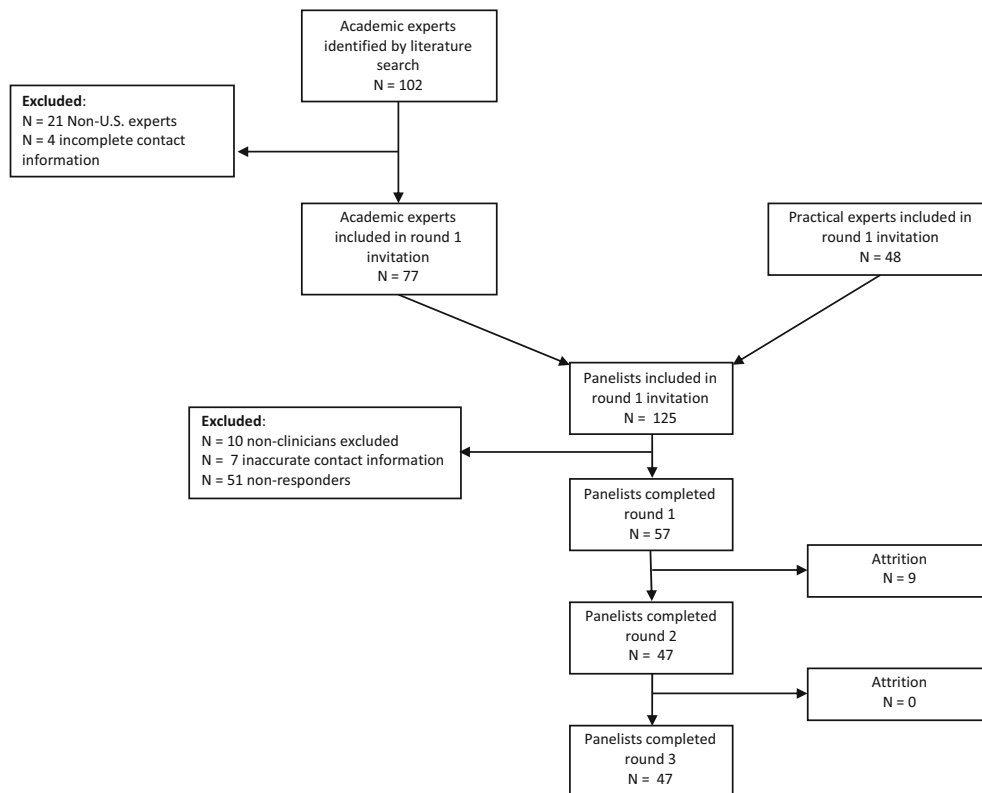


Figure 1 Sampling frame for the study. We show recruitment and retention of panelists by survey round.

Elicitation of Mortality Threshold for ACP Conversations

Indirect Elicitation Via Case Vignettes. Panelists did not reach consensus on how to treat the case of the 54-year-old otherwise healthy patient with diverticulitis, included to check both their attention and the validity of the task. However, the heterogeneity of responses suggested that (a) panelists were paying attention and (b) norms around the desirability of ACP did not extend indiscriminately to all patients, affirming the validity of the task.

Among cases of patients 65 years or older, panelists immediately agreed that ACP should occur in the hospital. They did not set a threshold for mortality at which hospitalists should engage in ACP conversations. However, they did vary in their recommendations about the timing and content of those conversations (see Fig. 2), which therefore became the focus of the analysis.

Round 1. Consensus on the timing of ACP conversations occurred only for the sickest patients, where panelists felt conversations should occur immediately. For the remainder, one-third to one-half of panelists opted to have the conversation immediately, while the others thought it could occur before discharge. The presence of dementia increased the priority that panelists assigned to having these conversations (see Fig. 2). Many participants noted potential ambiguity in the term “advance care planning,” wondering if a distinction existed between “advance care planning” and “goals of care” conversations.

Responses to the open-ended questions following each of case in round 1 about clinical characteristics that informed panelists’ decisions about whether or not to recommend an

ACP conversation further affirmed the validity of the instrument (see Table 3). Panelists repeatedly referenced three themes: the age of the patient, the presence of co-morbidities influencing long-term function (dementia) and survival (cancer), the risk of imminent death or decompensation (e.g., sepsis, respiratory failure).

Round 2. In round 2, we explicitly differentiated between “goals of care” conversations, conducted immediately to inform near-term treatment decisions, ACP conversations conducted as part of discharge planning to inform post-acute care, and ACP conversations conducted by the primary care physician to establish preferences for future medical care. There was continued agreement that a conversation (either ACP or goals of care) needed to happen during the admission regardless of mortality risk.

Round 3. Based on responses to round 2, we further probed the role that short-term and 1-year mortality played in determining the timing of the ACP conversation, sampling cases to highlight differences among cases. In round 3, > 70% of panelists agreed that patients with high (> 10%) short-term mortality risk or high (≥ 34%) long-term risk of mortality should have a goals of care conversation. Panelists also agreed that patients ≥ 65 years old who had low (≤ 10%) short-term and low (< 19%) 1-year risk of mortality warranted an ACP rather than a goals of care conversation. However, they disagreed about the case where the patient had a low risk of short-term mortality and an intermediate (19–33%) risk of 1-year mortality (see Fig. 2).

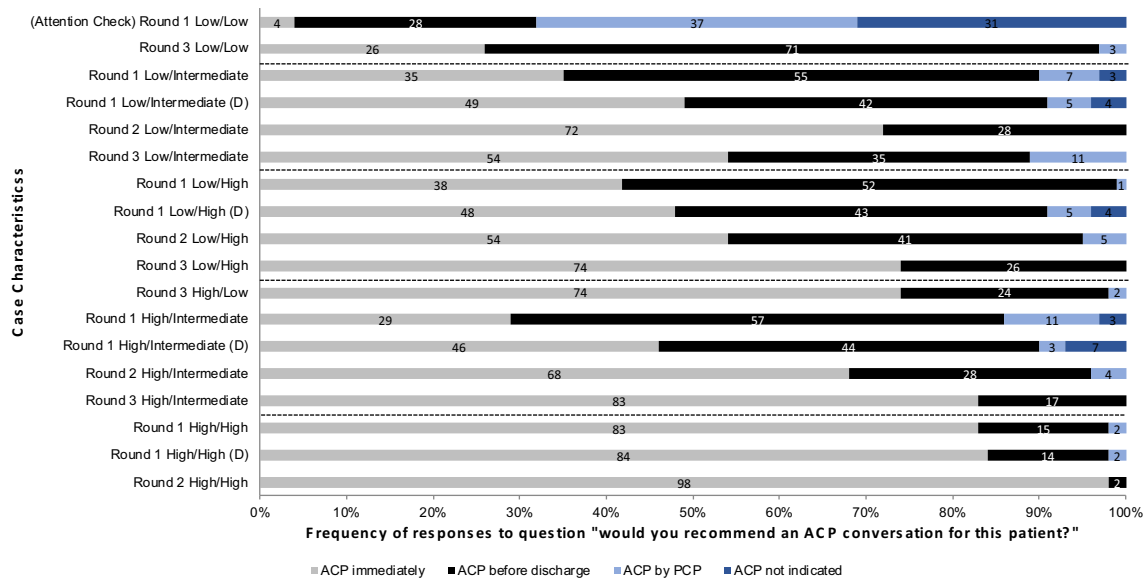


Figure 2 Frequency of ACP recommendations by case. We show the frequency of the responses to the case vignettes, organized by the probability of mortality (short-term/1-year) and by round. We categorized short-term mortality as low (≤ 10%) or high (> 10%), and 1-year mortality as low (< 19%), intermediate (19–33%), and high (> 33%). Cases describing patients with dementia are denoted by a (D). Dashed lines indicate similar groups of cases. As the survey progressed, consensus on the timing and type of conversation required was achieved for all except cases with low/intermediate risk of death.

Table 3 Determinants of Panelists' ACP Recommendations for Case Vignettes

Case vignettes	Top themes	N (%)
54-year-old female aerobics instructor presenting with left lower quadrant pain and fever. (Diverticulitis)	Age	21 (36.2)
	Comorbidities	19 (32.8)
	Risk of decompensation	13 (22.4)
	Clinical cues	9 (15.5)
80-year-old female presenting with altered mental status, lethargy, and urinary incontinence. (Urinary tract infection)	Comorbidities	29 (50.0)
	Dementia	23 (39.7)
	Clinical cues	25 (43.1)
	Age	20 (34.5)
66-year-old female with a history of breast cancer presenting with fever, shortness of breath, and increased leg edema. (Pneumonia)	Comorbidities	49 (84.4)
	Dementia	25 (43.1)
	Cancer	16 (27.6)
	Clinical cues	14 (24.1)
	Age	8 (13.8)
	Comorbidities	27 (46.6)
81-year-old female presenting from nursing home with worsening weakness, lethargy and nausea. (Acute renal failure)	Dementia	27 (46.6)
	Clinical cues	19 (32.8)
	Age	14 (24.1)
	Functional status	14 (24.1)
	Comorbidities	49 (84.5)
	Cancer	49 (84.5)
65-year-old female with stage IV colorectal cancer presenting with bright red blood per rectum. (Gastrointestinal hemorrhage)	Clinical cues	12 (20.7)
	Age	7 (12.1)
	Clinical cues	23 (39.7)
	Comorbidities	21 (36.2)
	Dementia	21 (36.2)
75-year-old female with a history of chronic urinary tract infections who recently finished a Medrol dose pack for a neck sprain presenting with vomiting and agitation. (Pyelonephritis)	Age	13 (22.4)
	Clinical cues	20 (34.5)
	Comorbidities	20 (34.5)
	Risk of decompensation	11 (19.0)
70-year-old female with a complicated cardiac history presenting with chest pain. (Non-ST segment elevation myocardial infarction)	Clinical cues	38 (65.5)
	Comorbidities	20 (34.5)
	Dementia	20 (34.5)
	Functional status	11 (19.0)
	Age	8 (13.8)
86-year-old wheelchair-bound male presenting with aphasia and weakness. (Stroke)	Clinical cues	40 (69.0)
	Comorbidities	25 (43.1)
	Risk of decompensation	14 (24.1)
	Age	8 (13.8)
77-year-old male with end-stage COPD presenting with lethargy, shortness of breath and oxygen desaturations on exertion. (Sepsis)	Clinical cues	40 (69.0)
	Comorbidities	25 (43.1)
	Risk of decompensation	14 (24.1)

*To categorize clinician responses to the free-text question about their decision making for each case vignette, we collapsed terms into broad categories:

Age = young, old, older, geriatric, elderly, age, aged

Comorbidities = comorbidities, comorbid, chronic illness, chronic conditions

Dementia = dementia, demented, Alzheimer's

Cancer = cancer, stage IV, malignancy, CA

Risk of decompensation = need for intubation, risk of decompensation, possibility of clinical decline, risk of deterioration, risk of imminent decline, stable, not at risk of decompensation/deterioration

Clinical cues = disease process, labs, vitals, physical exam (including altered mental status), imaging, any information included in HPI

Functional status = functional status, dependent on family members, functional impairment, impaired, ADL dependence

Direct Elicitation. In rounds 2 and 3, we also probed panelists for their opinion on the exact probability of short-term and 1-year mortality that should prompt a goals of care conversation, using sliders to elicit a value between 0 and 100%. In round 2, panelists stated that they would engage in a goals of care conversation if the risk of short-term mortality was higher than 38% (IQR 20–50%), or the risk of 1-year mortality was higher than 25% (IQR 16–50%). In round 3, those assessments shifted to 30% (IQR 20–40%) or 30% (IQR 20–50%) respectively. Similarly, panelists recommended engaging in a goals of care conversation immediately only if the risk of failure to return to baseline function was relatively high (physical function: 40% [IQR 25–50%]; cognitive function: 30% [IQR 20–50%]). Assessments provided in round 3 were similar (physical function: 35% [IQR 30–50%]; cognitive function: 30% [IQR 20–40%]).

There was heterogeneity in physician responses to the surprise question (i.e., would you be surprised if the patient were dead at 1 year), with only moderate correlation ($r = 0.34$) to the predicted probability that the patient would be dead at 1 year: 58% of physicians would be surprised if patients with low (< 19% mortality) were dead at 1 year; 31% if patients with intermediate (19–33% mortality) were dead at 1 year; and 9% if patients with high ($\geq 34\%$) mortality were dead at 1 year.

DISCUSSION

We conducted a modified Delphi study with an embedded behavioral experiment to establish a normative standard for the risk of mortality that should prompt clinicians to prioritize ACP conversations in the hospital. Instead of setting a threshold for the mortality that characterizes the “seriously ill,” as we

had expected, a multi-disciplinary group of experts instead recommended that clinicians engage in ACP conversations with *all* hospitalized patients over the age of 65. For those with low risk of short-term and 1-year mortality, they recommended having the conversation before discharge, focusing on preferences for future medical care. However, for those with either high short-term or 1-year mortality, they recommended having the conversation immediately to ensure alignment between treatment and goals of care.

These observations have important policy implications. ACP is an integral part of the National Academy of Medicine's objective of ensuring that patients receive person-centered, family-oriented, and evidence-based care, particularly at the end-of-life.¹ Existing guidelines have therefore advocated that clinicians use the opportunity of hospitalization to initiate these conversations, screening for the presence of "serious illness" to decide whether or not to have an ACP conversation.⁴ Mortality risk is central to the definition of "serious illness."¹⁸ However, our results demonstrate the difficulty that clinicians, even experts, have prioritizing based on the risk of mortality, except for the very highest risk patients. We found only moderate correlation between responses to the surprise question and calculator-based measures of 1-year mortality. Additionally, panelists explicitly recommended a much higher risk of short-term death when deciding whether or not to have a goals of care conversation than they implicitly used in practice when responding to the case vignettes.

Based on these findings, we conclude that strategies to increase ACP conversations by improving the calibration of clinicians' predictions about mortality (i.e., helping them to select the sickest patients) may be ineffective and inefficient. Indeed, our results may obviate the need for a screening process altogether, since our experts agree that all hospitalized patients over 65 should have an ACP conversation. Although theoretically simpler to implement, and less variable than the surprise question, this new normative standard has its own set of barriers. A total of 12.4 million patients over the age of 65 require hospitalization each year.¹⁹ At community hospitals, with few specialist services, the onus of having ACP conversations would fall almost exclusively on hospitalists, of whom 40% already report unsafe workloads.²⁰ From a societal perspective, submission of all these additional claims would increase physician billing charges and associated patient co-payments.^{21, 22}

When designing the study, we speculated that factors other than mortality would influence experts' recommendations, and specifically hypothesized that the presence of dementia might affect the urgency with which they advocated for hospitalists to have ACP conversations. Preliminary evidence confirmed our hypothesis. We believe this finding warrants additional investigation as we lacked the sample size to test the association between specific case characteristics and physician recommendations quantitatively (see Future Directions in the Appendix).

Finally, of interest, the Delphi process also highlighted the ongoing controversy over the content of ACP conversations. Consistent with Sudore's work, panelists struggled to specify when to recommend ACP (focused on long-term goals and preferences) rather than goals of care conversations about current or near-term treatment.¹⁶ After several rounds, they came to consensus for all except the most ambiguous case—patients with intermediate (< 19–33%) 1-year risk of mortality—recommending goals of care conversations for patients with either high ($\geq 10\%$) short-term or high ($\geq 34\%$) 1-year mortality. Feasibility and budgetary constraints required that we limit the number of rounds in our survey. Additional rounds may have allowed us to establish a more specific recommendation about how to manage those intermediate cases.

Our study had three limitations. First, we used case vignettes as a method of observing clinician judgment in practice so that we could make inferences about the risk of mortality that influenced their decision making. Case vignettes present a static view of the patient and may fail to elicit truly representative judgments.²³ However, in other clinical contexts, responses to case vignettes have proven a reasonable facsimile of standardized patients, the gold standard in simulation.²⁴ Second, less than 60% of invited panelists completed the first rounds of questionnaires, raising concerns about response bias. Our response rate of 50% matches those seen in other studies using survey methods.^{25, 26} Moreover, panels of 5 participants in Delphi studies have been shown to have equal validity and reliability to panels of 60, suggesting that we had sufficient numbers of participants to establish consensus.¹⁴ Third, our decision to recruit panelists with academic and practical expertise in ACP may have influenced the recommendations. A panel of community hospitalists, cognizant of the barriers to ACP in practice, may have advocated a less liberal policy for screening patients. Future work will need to address potential discordance between expert and "lay" perspectives on the topic.

CONCLUSIONS

Experts recommend that hospitalists engage in ACP with all patients 65 years or older, instead of screening for those with serious illness. Implementation of this new standard will require interventions to motivate physicians to initiate these conversations with all patients.

Acknowledgment: We thank the clinicians who beta-tested the instrument, and the panelists who participated in the Delphi process.

Corresponding Author: Deepika Mohan, MD, MPH; Department of Critical Care Medicine, University of Pittsburgh, Pittsburgh, PA, USA (e-mail: mohand@upmc.edu).

Authorship Role Study concept, design, analysis, interpretation: oas, dm, mr, aeb.
Drafting of the manuscript: oas and dm.
Critical revision of the manuscript for important intellectual content: oas, mr, jb, mm, jom.
All authors read and approved the final manuscript.

Funding Information This work was supported by the National Institutes of Health through the grant P01 AG019783 (Barnato, Mohan, O'Malley, Murphy, Bynum). The funding agency reviewed the study but played no role in its design, collection, analysis, or interpretation.

Compliance with Ethical Standards:

The Dartmouth Institutional Review Board reviewed and approved the study (IRB# 31186).

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

- Institute of Medicine 2015. *Dying in America: Improving Quality and Honoring Individual Preferences Near the End of Life*. Washington, DC: The National Academies Press.
- Benson WF and Aldrich N.** 2012. *Advance care planning: ensuring your wishes are known and honored if you are unable to speak for yourself, critical issue brief*. Atlanta: CDC. <https://www.cdc.gov/aging/pdf/advanced-care-planning-critical-issue-brief.pdf> (accessed Dec 3, 2019).
- Catholic Bishops of New York State. 2011. *Now and at the hour of our death: a catholic guide to end-of-life decision making*. <http://www.nyscatholic.org/wp-content/uploads/2011/11/End-of-Life-booklet-final.pdf> (accessed Dec 3, 2019).
- Anderson WG, Berlinger N, Ragland J et al.** Hospital-based prognosis and goals of care discussions with seriously ill patients: a pathway to integrate a key primary palliative care process into the workflow of hospitalist physicians and their teams. *Society of Hospital Medicine and The Hastings Center*, 2017.
- Howard M, Bernard C, Klein D et al.** Barriers to and enablers of advance care planning with patients in primary care. *Canadian Family Physician* 2018; 64: e190-8.
- Lund S, Richardson A, and May C.** Barriers to advance care planning at the end of life: an explanatory systematic review of implementation studies. *PLoS One* 2015; 10: e0116629.
- Heyland DK, Barwich D, Pichora D et al.** Failure to engage hospitalized elderly patients and their families in advance care planning. *JAMA Intern Med* 2013; 173: 778-787.
- Royal College of General Practitioners. The GSF Prognostic Indicator Guidance. <http://www.goldstandardsframework.org.uk/cd-content/uploads/files/General%20Files/Prognostic%20Indicator%20Guidance%20October%202011.pdf> Accessed Dec 3, 2019.
- Haydar SA, Almeder L, Michalakes L et al.** Using the surprise question to identify those with unmet palliative care needs in emergency and inpatient settings: what do clinicians think? *J Palliat Med* 2017; 20: 729-735.
- Downar J, Goldman R, Pinto R et al.** The 'surprise question' for predicting death in seriously ill patients: a systematic review and meta-analysis. *CMAJ* 2017; 189: E484-93
- White N, Kupeli N, Vickerstaff V et al.** How accurate is the 'surprise question' at identifying patients at the end of life? A systematic review and meta-analysis. *BMC Medicine* 2017; 15: 139.
- Boulkedid R, Abdoul H, Loustau M et al.** Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 2011; 6: e20476.
- Avella JR.** Delphi Panels: Research design, procedures, advantages and challenges. *Journal of Doctoral Studies*. 2016;11:305-321.
- Powell C.** The Delphi technique: myths and realities. *Journal of advanced nursing*. 2003;41(4):376-382.
- Clayton MJ.** Delphi: a technique to harness expert opinion for critical decision-making tasks in education. *Educational Psychology*. 1997;17(4):373-386.
- 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.
- Nisbett, RE and Wilson, TD.** (1977) Telling more than we can know: verbal reports on mental processes. *Psychol Rev*. 84(3): 231-259.
- Kelley AS.** Defining 'serious illness' *J Pall Med* 2014; 17: 985.
- McDermott KW, Elixhauser A, Sun R.** Trends in hospital inpatient stays in the US, 2005-2014. Statistical brief #225. In *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*. Rockville, MD, Agency for Healthcare Research and Quality, 2017.
- Michtalki HJ, Yeh H, Pronovost PJ et al.** Impact of attending physician workload on patient care: a survey of hospitalists. *JAMA Internal Medicine* 2013; 173: 375-377.
- Pelland K, Morphis B, Harris D, Gardner R.** Assessment of first-year use of Medicare's advance care planning billing codes. *JAMA Internal Medicine*. 2019.179: 827-829.
- Barnato AE, O'Malley AJ, Skinner JS, Birkmeyer JD.** Use of advance care planning billing codes for hospitalized older adults at high risk of dying: a national observational study. *Journal of hospital medicine*. 2019;14(4):229-231.
- Dhami MK, Hertwig R, and Hoffrage U.** The role of representative design in an ecological approach to cognition. *Psychol Bull* 2004; 130: 959-988.
- Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M.** Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000;283:1715-1722.
- Halpern SD, Kohn R, Dornbrand-Lo A, Metkus T, Asch DA, Volpp KG.** Lottery-based versus fixed incentives to increase clinicians' response to surveys. *Health services research*. 2011;46(5):1663-1674.
- Edwards P, Roberts I, Clarke M, DiGiuseppi C, Pratap S, Wentz R, and Kwan I.** Increasing response rates to postal questionnaires: systematic review. *BMJ* 2002; 324: 1183-5.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.