Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Pain Symptom Manage. Author manuscript; available in PMC 2019 April 09.

Published in final edited form as:

J Pain Symptom Manage. 2018 September ; 56(3): 406–413.e3. doi:10.1016/j.jpainsymman.2018.06.003.

Inter-rater agreement of intensivists evaluating the goal-concordance of preference-sensitive ICU interventions

Alison E. Turnbull, DVM, MPH, PhD[1,2,3], **Sarina K. Sahetya, MD**[2], **Elizabeth Colantuoni, PhD**[1,4], **Josephine Kweku, MD, MPH**[5], **Roozbeh Nikooie, MD**[1], and **J. Randall Curtis, MD, MPH**[6,7]

[1]Outcomes After Critical Illness and Surgery Group, Johns Hopkins University, Baltimore, MD

[2]Division of Pulmonary and Critical Care Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD

[3]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

[4]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

[5]Department of Anesthesiology and Critical Care, Anne Arundel Medical Center, Annapolis, MD

[6]Division of Pulmonary and Critical Care Medicine, Harborview Medical Center, University of Washington, Seattle, WA

[7]Cambia Palliative Care Center of Excellence, University of Washington, Seattle, Washington.

Abstract

Context: Goal-concordant care has been identified as an important outcome of advance care planning and shared decision-making initiatives. However validated methods for measuring goal-concordance are needed.

Objectives: To estimate the inter-rater reliability of senior critical care fellows rating the goal-concordance of preference-sensitive interventions performed in intensive care units (ICUs) while considering patient-specific circumstances as described in a previously-proposed methodology.

Methods: We identified ICU patients receiving preference-sensitive interventions in 3 adult ICUs at Johns Hopkins Hospital. A simulated cohort was created by randomly assigning each patient 1 of 10 sets of goals and preferences about limiting life support. Critical care fellows then independently reviewed patient charts and answered two questions: 1) Is this patient's goal achievable? and 2) Will performing this intervention help achieve the patient's goal? When the answer to both questions was yes, the intervention was rated as goal-concordant. Inter-rater agreement was summarized by estimating intraclass correlation (ICC) using mixed-effects models.

Corresponding Author: Alison E. Turnbull, 1830 E. Monument St. 5th Floor, Baltimore, MD 21205, (410) 955-2190, turnbull@jhmi.edu.

Disclosure/Conflict of Interest

The authors declare no conflicts of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Results: Six raters reviewed the charts of 201 patients. Interventions were rated as goal-concordant 22% – 92% of the time depending on the patient's goal-limitation combination. Percent agreement between pairs of raters ranged from 59% – 86%. The ICC for ratings of goal concordance was 0.50 (95% CI 0.31 – 0.69) and was robust to patient age, gender, ICU, severity of illness, and length of stay.

Conclusion: Inter-rater agreement between intensivists using a standardized methodology to evaluate the goal-concordance of preference-sensitive ICU interventions was moderate. Further testing is needed before this methodology can be recommended as a clinical research outcome.

Keywords

goal-concordant care; patient-centered care; critical illness; critical care outcomes; outcome measures; inter-observer variability; psychometrics

Introduction

Researchers evaluating interventions to facilitate advance care planning, patient-family engagement, and structured communication[1] share a common struggle: we lack a good outcome measure. Process measures, such as conducting family meetings and documenting patient preferences, are important but are not an end unto themselves. Meetings are conducted and preferences are documented to help patients receive care that matches their preferences and help achieve their goals. Measures of patient and family satisfaction are problematic. For example, families of patients who die may report greater satisfaction than those who survive,[2] and physicians providing an overly optimistic prognosis may receive higher satisfaction scores than their more accurate colleagues.[3] Scales, such as The Quality of Dying and Death (QODD) and CAESAR,[4] are sophisticated measures that capture the experiences of families. However, these scales may not measure a single unidimensional construct, do not apply to patients who survive their ICU stay, and are not currently recommended as primary outcome measures.[5,6]

Faced with these challenges, stakeholders now cite "goal-concordant care" as an important outcome of both advance care planning,[7] and communication interventions for seriously ill patients and families.[8,9] A policy statement endorsed by the American Thoracic Society and American College of Critical Care Medicine calls on intensivists to engage patients and proxies in a shared decision-making process to achieve goal-concordant care when "making major treatment decisions that may be affected by personal values, goals, and preferences."[12,13] However, there are currently few objective measures of goal-concordance, making it challenging to assess whether this outcome was achieved.[7,10]

We previously proposed a methodology for measuring the incidence of goal-concordant care in the ICU setting.[11] This proposal is part of a research agenda to develop a measure for evaluating the impact of patient-engagement and communication interventions on ICU care (see Appendix Figure 1). However for the proposed methodology to be valid, trained clinicians should demonstrate strong agreement when rating concordance between patients' goals and preferences, and the care received. Therefore, we designed a simulation study using clinical data from real ICU patients to estimate the inter-rater reliability of senior

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

critical care fellows rating the goal-concordance of preference-sensitive interventions in the ICU setting. Preference-sensitive interventions were defined using the consensus of an expert panel of ICU stakeholders that was convened to identify non-emergent ICU interventions requiring consideration of a patient's goals and treatment limitations in routine critical care clinical practice.[14] The panel reached consensus on 8 procedures referred to hereafter in this study as preference-sensitive interventions.[14]

Methods

Identifying ICU patients with orders for preference-sensitive interventions

In collaboration with the Johns Hopkins Institute for Clinical and Translational Research (ICTR), we developed a screening algorithm to identify orders in the electronic medical record (EMR) for 6 of the 8 previously-identified preference-sensitive interventions at Johns Hopkins Hospital (JHH).[14] The 6 interventions identified by the screening algorithm were tracheotomy, peripherally inserted central catheter, nasogastric tube, in-hospital dialysis, percutaneous endoscopic gastrostomy, and long-term dialysis catheter. This algorithm was applied retrospectively to identify adult patients with orders for these interventions in the medical, surgical, and surgical-oncology ICUs at JHH during 2015. To estimate the inter-rater reliability using intraclass correlation (ICC) with precision of +/− 0.20, we randomly selected 40 patients with orders for each of the 6 preference-sensitive interventions creating a sample of 240 unique patients for analysis. (Figure 1)

Simulating patient goals and treatment preferences

Patient goals and preferences for life-sustaining treatments are not reliably recorded in EMRs.[15,16] Therefore, we were not able to use the actual goals or treatment limitations of study patients for analyses. Instead, we used data on patient goals and preferences for the use of life-sustaining treatments collected as part of a study conducted in the JHH medical ICU.[17] In this previous study, ICU proxies were asked about the patient's goals using a 7-item, multiple-choice question with previously validated response options.[18–20] Proxies were also asked a 5-item multiple choice question about limitations in the use of life support preferred by their loved ones avoiding medical terminology. Together, these questions created 35 potential combinations of patient goal and treatment limitation. For the current study, we used the 9 most common combinations, plus the goal/limitation combination "to be comfortable" and "Focus on keeping me as comfortable as possible, even if that means I die sooner" creating 10 goal/limitation combinations. The comfort-focused combination was included because it described a unique subset of patients whose goals and preferences are qualitatively different from the majority of ICU patients. These 10 goal-limitation combinations were randomly assigned to the patients identified by the screening algorithm. The resulting cohort consisted of 240 patients with real physiologic data, real orders for preference-sensitive treatments, and a randomly assigned goal and preference about treatment limitations. From this point forward, this cohort is described as being comprised of "simulated patients" to acknowledge the simulated nature of the preferences.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Training intensivists to rate preference-sensitive interventions as goal-concordant or goal-discordant

In January 2017, 6 critical care fellows in their 2nd-4th year of training were hired and trained to rate preference-sensitive interventions as either concordant or discordant with a patient's goals using a previously-published conceptual framework for goal-concordant care in the ICU.[11] Among the 6 raters, 3 fellows were in the Pulmonary and Critical Care Medicine (PCCM) program at JHH and had completed the majority of their clinical training in the medical ICU, and 3 fellows were in the Anesthesia and Critical Care Medicine (ACCM) program and had completed the majority of their training in JHH surgical ICUs. All fellows participated in a 90-minute in-person training, received a 25-page operations manual, and had to successfully rate 6 – 10 test patients before reviewing the charts of patients in the study cohort. The training and operations manual included the conceptual framework, guidance on reviewing patient charts, instructions for interpreting statements about patient goals and treatment limitations, and examples of preference-sensitive interventions that did and did not constitute goal-concordant care for each of the 10 assigned goal-limitation combinations.

After completing training, each rater worked independently to review the history and physical, consult notes, lab and imaging data, and progress notes up until the day of the order for a preference-sensitive intervention for 120 patients in the study cohort. After reviewing each patient's chart and their assigned goal and treatment limitations, raters answered two questions:

1. Is this patient's goal achievable?
2. Will performing this intervention, or continuing to use this intervention help achieve the patient's goal?

Response options for both questions were "Yes or maybe" vs "No." Interventions were defined as being goal-concordant if the rater selected "Yes or maybe" in response to both of these questions. Responses were entered into a data collection form in a REDCap database.[21] Raters also abstracted data on each patient's age, gender, and days of hospitalization and ICU care prior to the order for the preference-sensitive intervention. Patient's SOFA score[22–24] on the day of the intervention order were calculated separately by an investigator (R. Nikooie). The study was approved by the Institutional Review Board of Johns Hopkins University (IRB00053330).

Data analysis

Patient characteristics were summarized using frequencies for categorical variables, and medians with interquartile ranges (IQR) for continuous variables. The proportion of patient goals rated as potentially achievable and the percent of interventions rated as goal-concordant were summarized for each of the 10 assigned goal-limitation combinations, by each rater and by rater fellowship program. Percent agreement was calculated for all pairs of raters. Finally, three-level, mixed-effects models with no independent covariates were fit for the binary response to the question "Is this patient's goal achievable?" and for the definition of goal-concordance. In these models the goal-limitation combinations represent the third nesting level, individual patients represent the second nesting level, the responses of three

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

raters are nested within each patient and the models included random intercepts for both the goal-limitation combination and patient. The intraclass correlation[25] from this model estimates the proportion of variance in responses explained by the random effects for patient and the patient's assigned goal-limitation combination. To assess the robustness of this estimates, fixed effects for the patient's age, gender, ICU and SOFA score at the time of the order, days in the ICU prior to the order, and for the rater's fellowship program were included in the mixed models. As a sensitivity analysis, we also fit a two-level, mixed-effects model with the goal-limitation combination included as a fixed effect, i.e. a single random intercept for patient was included. Descriptive statistics were generated using R (Version 3.3.2; R Development Core Team, Vienna, Austria), and all models were fit using Stata (StataCorp. 2013. *Stata Statistical Software: Release 13.* College Station, TX: StataCorp LP.)

Results

We identified 203 eligible patients with 2 patients having missing data, leaving 201 patients for analysis (Figure 1). The median age, SOFA score, and ICU length of stay on the day of the order for a preference-sensitive intervention were 60 years (IQR 48, 68), 6 (IQR 3, 9), and 3 days (IQR 0, 10) respectively (Table 1). Each patient was reviewed and rated by 3 intensivists, creating 603 ratings. Intensivist ratings were strongly influenced by which goal and limitation was assigned. For example, the proportion of "Yes or maybe" responses to the question "Is this patient's goal achievable?" ranged from 22% to 100% depending on the patient's assigned combination of goal and treatment limitations (Table 2).

The difference between PCCM and ACCM raters was greatest when the patient's goal was to prioritize comfort even if it meant dying sooner (goal-limitation combination #10) and when a patient's goal was to return to a state of health that a person of the same age without any significant illness or injury is expected to experience (i.e. To be cured) and the patient's treatment limitation statement was "Try to help me get better, but don't use life support machines and if my heart stops don't do CPR" (goal-limitation combination #2). Under both of these conditions, there was unanimous agreement about goal-concordance for only 44% of patients. When comfort was prioritized, PCCM vs AACM fellows rated interventions as being goal-concordant 8% and 61% of the time respectively (Table 2). Under goal-limitation combination #2 the PCCM fellows rated patient goals as potentially achievable 4% of the time, while intensivists in the ACCM fellowship program rated patient goals as potentially achievable 38% of the time. These two goal-limitation combinations remained the most divisive even when comparing the responses of raters in the same fellowship program (Appendix Table 1).

Percent agreement about whether a patient's goal was potentially achievable ranged from 73% to 91% across the 15 rater pairings (Appendix Table 2). Percentage agreement on the goal-concordance of interventions ranged from 59% to 86%. Inter-rater reliability estimates are presented in Table 3, using the intraclass correlation coefficient (ICC). The ICC represents the fraction of the total variance that is due to variability in ratings between groups (patients) as opposed to within groups (variation in responses among raters assigned to the same patient). Perfect agreement among raters generates an ICC value of 1, while an

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ICC of 0 can occur if the raters randomly generate their response (i.e. a flip of a coin).” The intraclass correlation (ICC) for responses to the question “Is this patient’s goal achievable?” was 0.64 (95% confidence interval (CI) 0.41 – 0.82). Adding fixed effects for patient characteristics (0.62, 95% CI 0.37–0.81) or the rater’s fellowship program (0.65, 95% CI 0.43 – 0.83) to the model did not change the ICC substantially. The ICC for ratings of goal concordance was 0.50 (95% CI 0.31 – 0.69), and again remained relatively unchanged when fixed effects for patient characteristics (0.47, 95% CI 0.28 – 0.68) or rater’s fellowship program (0.53, 95% CI 0.34 – 0.72) were included in the model. The sensitivity analyses using a two-level random effects model yielded similar ICCs (Appendix Table 3).

Discussion

This study is the first to evaluate the inter-rater reliability of a measure of goal-concordance for clinical research in the ICU setting. Our findings suggest that after a single training session, the inter-rater reliability as measured by intraclass correlation lay between 0.31 and 0.69, and was robust to patient age, severity of illness, and length of stay. While further development is needed before this methodology can be recommended, these findings are encouraging.

There is currently great enthusiasm for measuring consistency between patient preferences and the care received.[26] However ICU patient goals and preferences are not consistently elicited, are often poorly documented,[15,16,27–31] and may change over time[32] creating major logistical challenges to measurement. As a result, current recommendations are to measure 1) the timing of serious illness communication, 2) patient or surrogate experience, and 3) whether bereaved caregivers report that their loved ones received care consistent with their preferences.[10] We support these recommendations, and also advocate for a direct measure of goal-concordance.

The methodology evaluated in this study in unique in using decisions, not patients, as the unit of analysis. For example, if 48 preference-sensitive interventions are ordered in an ICU and 36 are rated as goal-concordant, the incidence rate of goal-concordant treatment is 75 per 100 preference-sensitive decisions (95% CI 60 – 86). This approach allows complex or long-stay patients to contribute more data than short-stay patients, and produces more precise estimates (smaller confidence intervals) of ICUs that perform more preference-sensitive treatments. Requiring raters to consider the patient-specific circumstances and context of the intervention also avoids penalizing clinicians for responding appropriately in unusual situations. For example, a percutaneous endoscopic gastrostomy (PEG) tube is normally inappropriate for a patient prioritizing comfort at the end of life. However, a PEG tube could be rated as goal-concordant if it was placed to vent the stomach of a patient with a malignant bowel obstruction who wished to spend his final days at home.[33]

Inter-rater agreement in this study was heavily influenced by patient goal. This means that the inter-rater reliability of this measure will vary depending on the mix of goals expressed by patients in a study cohort. Therefore, a clinical researcher considering this measure as an outcome would be well-served to collect preliminary data on the goals expressed by patients or proxies in their study population.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Discordance in the ratings of fellows who train in the medical vs surgical ICU highlight the differences in how goal-concordance is interpreted across sub-specialties. This suggests the value of efforts to develop consensus across specialties in the implementation of goal-concordant care and to develop interdisciplinary educational programs to facilitate shared understanding and implementation of goal-concordant care.

Our study was limited by using simulated clinical scenarios created by randomly assigned goals and preferences about limiting treatment to actual ICU patients from a single hospital. This could have resulted in a disproportionate number of situations where patients identified unachievable goals. We also required raters to work completely independently. In the setting of a clinical study, blinded raters could confer about cases where there was initial disagreement. The process of continually discussing such cases may lead to better shared understanding of what interventions constitute goal-concordant care. Finally, all raters in this simulation were fellows at a single institution who may have demonstrated higher or lower agreement than more senior intensivists or intensivists from other sites.

In conclusion, this simulation study found moderate inter-rater agreement between intensivists using a standardized methodology to evaluate the goal-concordance of preference-sensitive interventions in ICUs. Agreement was not substantially impacted by patient age, length of stay, or severity of illness, but did vary by the patient's goals and preferences. Further testing and development is needed before this methodology can be recommended as an outcome in clinical research. This will require following ICU patients and their proxies longitudinally, documenting their changing goals and limitations over the course of a hospital stay, and allowing physicians (including attending physicians) to determine whether preference-sensitive interventions constituted goal-concordant care.

Acknowledgments

The authors wish to acknowledge the following colleagues for their assistance with data acquisition: Dale M. Needham, Pragyashree Sharma Basyal, Grace Peloquin, Lou Kirk, Rahul Raghavan, Rebecca Dezube, Earl Mantheiy, and Diana Gumas.

Funding

This work was supported by a research grant awarded by the Gordon and Betty Moore Foundation and by the National Heart, Lung, And Blood Institute of the National Institutes of Health under Award Number T32HL007534, and by the Johns Hopkins Institute for Clinical and Translational Research (ICTR) which is funded in part by Grant Number UL1 TR001079 from the National Center for Advancing Translational Sciences (NCATS) a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the Johns Hopkins ICTR, NCATS or NIH.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

APPENDIX

A. Background and Development

1. **Review of existing literature** describing how ICU patients and their families conceptualize goals of care and outcomes after critical illness.
2. **Qualitative interviews** with intensivists about facilitators and barriers to discussing goals of care and post-discharge outcomes with families of ICU patients. (Turnbull et al. 2016. *Annals of ATS,* 13 (9):1546-1551)
3. **Conceptual framework** incorporating patient goals, treatment limitations set by patients or proxies, and clinician assessment of prognosis and treatment effectiveness. (Turnbull and Hartog. 2017. *Intensive Care Med,* 43:1847-1849)

B. Feasibility and Pilot Testing

1. **Delphi process** with relevant stakeholders (ICU physicians, nurses, patients, and family-members) to identify non-emergent, potentially harmful, ICU interventions whose value is highly dependent on patient treatment goals. (Turnbull et al. 2016. *Heart & Lung,* 45(6):517-524)
2. **Evaluation of a recruitment strategy** for enrolling ICU proxies (mostly family members) in research while their loved one is critically ill. (Turnbull et al. 2017. *PLoS ONE,* 12(5):e0177741)
3. **Cross-sectional survey** to assesses the ability of ICU proxies to identify a) a goal of care and b) treatment limitations 1-8 days after ICU admission. (Under peer-review)
4. **Simulation study** to evaluate the inter-rater agreement of intensivists trained to evaluate whether the interventions identified in B1 met the criteria proposed for goal-concordant care in A3. Simulation used a sample of ICU patients randomly assigned the most common combinations of goals and treatment limitations reported in B3. (Under peer-review)
5. **Clinical pilot study** to estimate retention and missing data rates for patients and their proxies enrolled in studies estimating the incidence of goal-concordant ICU care. (Planned)

C. Validation studies (Planned)

1. **Validation study** of the correlation between goal-concordant care and recommended process measures, communication quality, shared decision-making, decisional regret, and costs.
2. **Qualitative interviews** with ICU clinicians, families and patients to understand the origins of care rated as goal-discordant.

CCAPG Critical Care Aligned with Patient Goals

Figure 1:
Methodology for Developing a Measure of Goal-concordant Care in the ICU Setting for Reaserch Purposes

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Appendix Table 1:

Chart review results by assigned patient goal, treatment limitations, and rater

			Patient goal rated as potentially achievable given treatment limitations						Interventions rated as goal-concordant[a]					
			Rater (Raters a-c are PCCM fellows; raters d-f are ACCM fellows)											
Goal-limit combination[b]	Patient Goal[c]	Treatment Limitations[d]	a n=98	b n=101	c n=98	d n=99	e n=103	f n =104	a n=98	b n=101	c n=98	d n=99	e n=103	f n=104
1: n=60	To be cured	None	30%	38%	17%	88%	56%	14%	30%	38%	17%	88%	56%	14%
2: n=54	To be cured	Some	10%	0%	0%	60%	12%	36%	10%	0%	0%	60%	12%	36%
3: n=66	To be cured	Unsure	46%	40%	0%	50%	22%	40%	46%	40%	0%	50%	22%	40%
4: n=66	To live longer	None	91%	100%	100%	80%	100%	83%	91%	67%	100%	80%	100%	83%
5: n=57	To live longer	Unsure	100%	90%	100%	100%	100%	100%	80%	80%	90%	100%	100%	100%
6: n=63	To improve health	None	80%	75%	64%	82%	82%	100%	80%	62%	64%	82%	82%	100%
7: n=60	To improve health	Some	82%	100%	93%	100%	100%	100%	82%	88%	86%	100%	100%	100%
8: n=69	To improve health	Unsure	100%	79%	100%	100%	71%	80%	100%	79%	90%	100%	71%	80%
9: n=54	To maintain health	None	83%	71%	80%	75%	88%	64%	83%	71%	80%	75%	88%	64%
10: n=54	To be comfortable	Strong	100%	100%	100%	100%	100%	100%	20%	0%	0%	57%	55%	70%

Abbreviations: ACCM refers to the subset of ratings by 3 fellows in the Anesthesia and Critical Care Medicine program who trained primarily in surgical ICUs. PCCM refers to ratings by 3 fellows in the Pulmonary and Critical Care Medicine program who trained primarily in the medical ICU

[a] Interventions are defined as being goal-concordant if the rater selected “Yes or maybe” in response to the question “Is this patient’s goal achievable?” and the question “Will performing this intervention, or continuing to use this intervention help achieve the patient’s goal?”

[b] The 10 combinations of patient goals and treatment limitations were derived from a study of ICU proxies conducted at Johns Hopkins Hospital in 2016. Examples and further instructions on how to interpret each combination were provided in the training and operations manual provided to each rater.

[c] Raters were instructed to interpret patient goal statements as follows: **To be cured -** To return to a state of health that a person of the same age without any significant illness or injury is expected to experience. **To live longer -** To live as long as possible in any health state, even if they require 24–7 care in an acute care facility, LTAC, or chronic vent facility. **To improve health -** To be well enough to live outside the hospital setting. Living in a nursing home or discharge to a rehabilitation facility is an acceptable outcome for these patients. **To maintain health -** To be well enough to return to the baseline health status and level of independence that they personally had before hospital admission. **To be comfortable -** To prioritize comfort over longevity and hope to be as free of pain and discomfort as possible.

[d] Treatment limitations were interpreted as follows: **None -** Use life support machines to keep me alive no matter what. If my heart stops, do CPR. **Some -** Try to help me get better, but don’t use life support machines and if my heart stops don’t do CPR. **Strong -** Focus on keeping me as comfortable as possible, even if that means I die sooner. **Unsure -** I (the patient’s proxy) don’t know what the patient would say.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Appendix Table 2:

Percent agreement between raters

Agreement in response to the question: "Is this patient's goal potentially achievable?"						
	Rater					
	a – PGY5	**b – PGY8**	**c – PGY6**	**d – PGY9**	**e – PGY5**	**f – PGY5**
a – PGY5	n = 98					
b – PGY8	79%	n = 101				
c – PGY6	87%	82%	n = 98			
d – PGY9	78%	73%	80%	n = 99		
e – PGY5	77%	69%	91%	80%	n = 103	
f – PGY5	83%	69%	78%	76%	77%	n = 104
Does the intervention meet the definition of goal-concordant care[a] for this patient?						
	a– PGY5	**b – PGY8**	**c – PGY6**	**d – PGY9**	**e – PGY5**	**f – PGY5**
a – PGY5	n = 98					
b – PGY8	74%	n = 101				
c – PGY6	80%	74%	n = 98			
d – PGY9	73%	59%	80%	n = 99		
e – PGY5	69%	60%	86%	80%	n = 103	
f – PGY5	75%	60%	76%	71%	70%	n = 104

[a]Interventions are defined as being goal-concordant if the rater selected "Yes or maybe" in response to the question "Is this patient's goal achievable?" and the question "Will performing this intervention, or continuing to use this intervention help achieve the patient's goal?"

Appendix Table 3:

Estimates of inter-rater agreement via two-level random effects models

Model	Model description	ICC	ICC 95% CI
Model 3	**Binary response to the question "Is this patient's goal achievable?"**		
3a.	Random effect for patient, no fixed effects	0.67	0.54 – 0.78
3b.	Random effect for patient, fixed effects for patient characteristics[b]	0.64	0.51 – 0.76
3c.	Random effect for patient, fixed effects for rater's training background	0.67	0.54 – 0.78
3d.	Random effect for patient, fixed effect for the combination[a] of patient goal & code status	0.21	0.08 – 0.45
Model 4	**Did the intervention meet the criteria for goal-concordant care?**		
4a.	Random effect for patient, no fixed effects	0.55	0.42 – 0.68
4b.	Random effect for patient, fixed effects for patient characteristics[b]	0.53	0.40 – 0.66
4c.	Random effect for patient, fixed effects for rater's training background	0.58	0.45 – 0.70
4d.	Random effect for patient, fixed effect for the combination[a] of patient goal & code status	0.20	0.08 – 0.41

Abbreviations: CI, Confidence Interval; ICC, Intraclass correlation

[a]The 10 combinations of patient goals and treatment limitations were derived from a study of ICU proxies conducted at Johns Hopkins Hospital in 2016.

[b]Patient characteristics include age, gender, ICU, sofa score, and days in the ICU at the time of the intervention

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

References

1. Oczkowski SJ, Chung HO, Hanvey L, Mbuagbaw L, You JJ. Communication tools for end-of-life decision-making in the intensive care unit: a systematic review and meta-analysis. Critical care 2016;20:97. [PubMed: 27059989]
2. Wall RJ, Curtis JR, Cooke CR, Engelberg RA. Family satisfaction in the ICU: differences between families of survivors and nonsurvivors. Chest 2007;132:1425–33. [PubMed: 17573519]
3. Weeks JC, Catalano PJ, Cronin A, et al. Patients' expectations about effects of chemotherapy for advanced cancer. The New England journal of medicine 2012;367:1616–25. [PubMed: 23094723]
4. Kentish-Barnes N, Seegers V, Legriel S, et al. CAESAR: a new tool to assess relatives' experience of dying and death in the ICU. Intensive care medicine 2016;42:995–1002. [PubMed: 26951427]
5. Curtis JR, Downey L, Engelberg RA. The quality of dying and death: is it ready for use as an outcome measure? Chest 2013;143:289–91. [PubMed: 23381306]
6. Curtis JR, Downey L, Engelberg RA. The importance and challenge of measuring family experience with end-of-life care in the ICU. Intensive care medicine 2016;42:1179–81. [PubMed: 27056551]
7. Sudore RL, Heyland DK, Lum HD, et al. Outcomes That Define Successful Advance Care Planning: A Delphi Panel Consensus. Journal of pain and symptom management 2018;55:245–55 e8. [PubMed: 28865870]
8. Narang AK, Wright AA, Nicholas LH. Trends in Advance Care Planning in Patients With Cancer: Results From a National Longitudinal Survey. JAMA oncology 2015;1:601–8. [PubMed: 26181909]
9. Cooper Z, Koritsanszky LA, Cauley CE, et al. Recommendations for Best Communication Practices to Facilitate Goal-concordant Care for Seriously Ill Older Patients With Emergency Surgical Conditions. Annals of surgery 2016;263:1–6. [PubMed: 26649587]
10. Sanders JJ, Curtis JR, Tulsky JA. Achieving Goal-Concordant Care: A Conceptual Model and Approach to Measuring Serious Illness Communication and Its Impact. Journal of palliative medicine 2018;21:S17–S27. [PubMed: 29091522]
11. Turnbull AE, Hartog CS. Goal-concordant care in the ICU: a conceptual framework for future research. Intensive care medicine 2017;43:1847–9. [PubMed: 28656453]
12. Kon AA, Davidson JE, Morrison W, Danis M, White DB. Shared Decision-Making in Intensive Care Units. Executive Summary of the American College of Critical Care Medicine and American Thoracic Society Policy Statement. American journal of respiratory and critical care medicine 2016;193:1334–6. [PubMed: 27097019]
13. Kon AA, Davidson JE, Morrison W, et al. Shared Decision Making in ICUs: An American College of Critical Care Medicine and American Thoracic Society Policy Statement. Critical care medicine 2016;44:188–201. [PubMed: 26509317]
14. Turnbull AE, Sahetya SK, Needham DM. Aligning critical care interventions with patient goals: A modified Delphi study. Heart & lung : the journal of critical care 2016;45:517–24. [PubMed: 27593494]
15. Heyland DK, Barwich D, Pichora D, et al. Failure to engage hospitalized elderly patients and their families in advance care planning. JAMA internal medicine 2013;173:778–87. [PubMed: 23545563]
16. Bear A, Thiel E. Documentation of Crucial Information Relating to Critically Ill Patients. Journal of palliative care 2018;33:5–8. [PubMed: 29260613]
17. Turnbull AE, Chessare CM, Coffin RK, Needham DM. A brief intervention for preparing ICU families to be proxies: A phase I study. PloS one 2017;12:e0185483. [PubMed: 28968409]
18. Kaldjian LC, Curtis AE, Shinkunas LA, Cannon KT. Goals of care toward the end of life: a structured literature review. The American journal of hospice & palliative care 2008;25:501–11. [PubMed: 19106284]
19. Haberle TH, Shinkunas LA, Erekson ZD, Kaldjian LC. Goals of care among hospitalized patients: a validation study. The American journal of hospice & palliative care 2011;28:335–41. [PubMed: 21097876]
20. Brandt DS, Shinkunas LA, Gehlbach TG, Kaldjian LC. Understanding Goals of Care Statements and Preferences among Patients and Their Surrogates in the Medical ICU. Journal of hospice and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

palliative nursing : JHPN : the official journal of the Hospice and Palliative Nurses Association 2012;14:126–32.

21. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. Journal of biomedical informatics 2009;42:377–81. [PubMed: 18929686]
22. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive care medicine 1996;22:707–10. [PubMed: 8844239]
23. Vasilevskis EE, Pandharipande PP, Graves AJ, et al. Validity of a Modified Sequential Organ Failure Assessment Score Using the Richmond Agitation-Sedation Scale. Critical care medicine 2016;44:138–46. [PubMed: 26457749]
24. Brown SM, Duggal A, Hou PC, et al. Nonlinear Imputation of PaO2/FIO2 From SpO2/FIO2 Among Mechanically Ventilated Patients in the ICU: A Prospective, Observational Study. Critical care medicine 2017;45:1317–24. [PubMed: 28538439]
25. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutorials in quantitative methods for psychology 2012;8:23–34. [PubMed: 22833776]
26. Unroe KT, Hickman SE, Torke AM, Group ARCW. Care Consistency With Documented Care PMethodologic Considerations for Implementing the “Measuring What Matters” Quality Indicator. Journal of pain and symptom management 2016;52:453–8. [PubMed: 27677443]
27. Holley A, Kravet SJ, Cordts G. Documentation of code status and discussion of goals of care in gravely ill hospitalized patients. Journal of critical care 2009;24:288–92. [PubMed: 19327289]
28. Spronk PE, Kuiper AV, Rommes JH, Korevaar JC, Schultz MJ. The practice of and documentation on withholding and withdrawing life support: a retrospective study in two Dutch intensive care units. Anesthesia and analgesia 2009;109:841–6. [PubMed: 19690256]
29. Duplan KL, Pirret AM. Documentation of cardiopulmonary resuscitation decisions in a New Zealand hospital: A prospective observational study. Intensive & critical care nursing 2016;37:75–81. [PubMed: 27575617]
30. Grudzen CR, Buonocore P, Steinberg J, Ortiz JM, Richardson LD, Group ARCW. Concordance of Advance Care Plans With Inpatient Directives in the Electronic Medical Record for Older Patients Admitted From the Emergency Department. Journal of pain and symptom management 2016;51:647–51. [PubMed: 26891604]
31. Robinson MT, Vickrey BG, Holloway RG, et al. The lack of documentation of preferences in a cohort of adults who died after ischemic stroke. Neurology 2016;86:2056–62. [PubMed: 27060165]
32. Berntsen GK, Gammon D, Steinsbekk A, et al. How do we deal with multiple goals for care within an individual patient trajectory? A document content analysis of health service research papers on goals for care. BMJ open 2015;5:e009403.
33. Shaw C, Bassett RL, Fox PS, et al. Palliative venting gastrostomy in patients with malignant bowel obstruction and ascites. Annals of surgical oncology 2013;20:497–505. [PubMed: 22965572]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

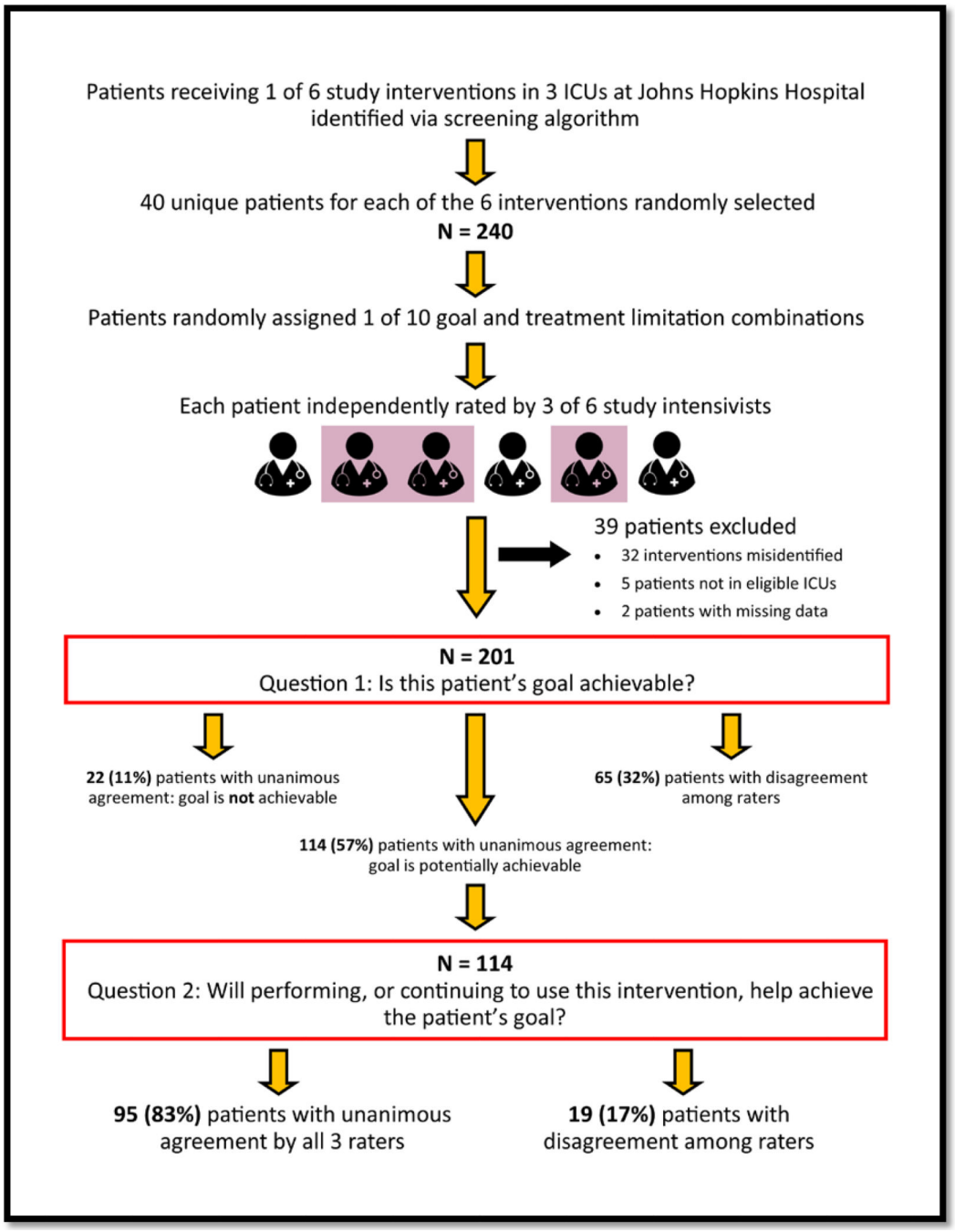

Figure 1:
Study Flow Diagram

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1:

Characteristics of study patients

Characteristic	Patients, n = 201
Age, median (IQR)	60 (48, 68)
Female, n (%)	89 (44%)
ICU, n (%)	
Medical	93 (46%)
Surgical Oncology	77 (38%)
Surgical	31 (15%)
SOFA score on the day of the order, median (IQR)	6 (3, 9)
Days in hospital prior to intervention, median (IQR)	4 (1, 14)
Days in ICU prior to intervention, median (IQR)	3 (0, 10)
Intervention	
Tracheotomy	40 (20%)
Peripherally inserted central catheter	38 (19%)
Nasogastric tube	37 (18%)
In-hospital dialysis	36 (18%)
Percutaneous endoscopic gastrostomy	34 (17%)
Long-term dialysis catheter	16 (8%)

Abbreviations: ICU, intensive care unit; IQR, interquartile range; SOFA, sequential organ failure assessment

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2:

Summary of simulation ratings by assigned patient goal and treatment limitations

			Patient goal rated as potentially achievable given treatment limitations			Interventions rated as goal-concordant[a]		
Goal-limitation combination[b]	**Patient Goal[c]**	**Treatment Limitations[d]**	**All intensivists**	**PCCM[e]**	**ACCM[e]**	**All intensivists**	**PCCM[e]**	**ACCM[e]**
1: n=60	To be cured	None	38%	31%	45%	38%	31%	45%
2: n=54	To be cured	Some	22%	4%	38%	22%	4%	38%
3: n=66	To be cured	Unsure	35%	32%	38%	35%	32%	38%
4: n=66	To live longer	None	94%	97%	91%	88%	85%	91%
5: n=57	To live longer	Unsure	98%	96%	100%	93%	84%	100%
6: n=63	To improve health	None	81%	72%	88%	79%	69%	88%
7: n=60	To improve health	Some	95%	91%	100%	92%	85%	100%
8: n=69	To improve health	Unsure	87%	92%	82%	86%	89%	82%
9: n=54	To maintain health	None	76%	78%	74%	76%	78%	74%
10: n=54	To be comfortable	Strong	100%	100%	100%	35%	8%	61%

[a] Interventions are defined as being goal-concordant if the rater selected “Yes or maybe” in response to the question “Is this patient’s goal achievable?” and the question “Will performing this intervention, or continuing to use this intervention help achieve the patient’s goal?”

[b] The 10 combinations of patient goals and treatment limitations were derived from a study of ICU proxies conducted at Johns Hopkins Hospital in 2016. Examples and instructions on how to interpret each combination were provided in the training and operations manual provided to each rater.

[c] Raters were instructed to interpret patient goal statements as follows: **To be cured -** To return to a state of health that a person of the same age without any significant illness or injury is expected to experience. **To live longer -** To live long as possible in any health state, even if they require 24–7 care in an acute care facility, LTAC, or chronic vent facility. **To improve health -** To be well enough to live outside the hospital setting. Living in a nursing home or discharge to a rehabilitation facility is an acceptable outcome for these patients. **To maintain health -** To be well enough to return to the baseline health status and level of independence that they personally had before hospital admission. **To be comfortable -** To prioritize comfort over longevity and hope to be as free of pain and discomfort as possible.

[d] Treatment limitations were interpreted as follows: **None -** Use life support machines to keep me alive no matter what. If my heart stops, do CPR. **Some -** Try to help me get better, but don’t use life support machines and if my heart stops don’t do CPR. **Strong -** Focus on keeping me as comfortable as possible, even if that means I die sooner. **Unsure -** I (the patient’s proxy) don’t know what the patient would say

[e] ACCM refers to the subset of ratings by 3 fellows in the Anesthesia and Critical Care Medicine program who trained primarily in surgical ICUs. PCCM refers to ratings by 3 fellows in the Pulmonary and Critical Care Medicine program who trained primarily in the medical ICU.

J Pain Symptom Manage. Author manuscript; available in PMC 2019 April 09.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3:

Estimated inter-rater agreement from a three-level random effects model

Model	Model description	ICC	95% CI
Model 1	**Response to the question "Is this patient's goal achievable?"**		
1a.	Random effects for patient and goal-limitation combination[a], no fixed effects	0.64	0.41 – 0.82
1b.	Random effects for patient and goal-limitation combination[a], fixed effects for patient characteristics[b]	0.62	0.37 – 0.81
1c.	Random effects for patient and goal-limitation combination[a], fixed effects for rater's fellowship program	0.65	0.43 – 0.83
Model 2	**Did the intervention meet the criteria for goal-concordant care?**		
2a.	Random effects for patient and goal-limitation combination[a], no fixed effects	0.50	0.31 – 0.69
2b.	Random effects for patient and goal-limitation combination[a], fixed effects for patient characteristics[b]	0.47	0.28 – 0.68
2c.	Random effects for patient and goal-limitation combination[a], fixed effects for rater's fellowship program	0.53	0.34 – 0.72

Abbreviations: CI, Confidence Interval; ICC, Intraclass correlation

[a]The 10 combinations of patient goals and treatment limitations were derived from a study of ICU proxies conducted at Johns Hopkins Hospital in 2016.

[b]Patient characteristics include age, gender, ICU, sofa score, and days in the ICU at the time of the intervention

J Pain Symptom Manage. Author manuscript; available in PMC 2019 April 09.