

Statistical Analysis Plan

TRIAL FULL TITLE	Values and options in cancer care (VOICE)
ClinicalTrials.gov ID:	NCT01485627
SAP VERSION	4.0
SAP VERSION DATE	Last edited on 2016March12 for clarity and grammar. Dates for key analysis decisions (as documented by time-stamped email) are indicated in sections below The first version of the SAP was drafted in November, 2014, compiling analysis decisions that originated in the NIH R01 proposal, the BMC Study protocol published in February, 2013, and investigator decisions determined during investigator conferences from March through November of that year. All decisions regarding study outcome specification were made prior to the initiation of data analysis of outcome data from Phase II (the post-randomization phase of the RCT).
TRIAL STATISTICIANS	Daniel J. Tancredi, PhD (djtancredi@ucdavis.edu); Guibo Xing, PhD (gbxing@ucdavis.edu);
TRIAL + INVESTIGATORS	Ronald Epstein, MD (Ronald_Epstein@URMC.Rochester.edu); Richard L Kravitz, MD (rlkravitz@ucdavis.edu)
SAP AUTHOR	Daniel J. Tancredi, PhD (djtancredi@ucdavis.edu)

1 Table of Contents

2 Abbreviations and Key Sources	3
3 Introduction	4
3.1 Preface	4
3.2 Purpose of the analyses	5
3.3 Study Objectives	5

3.4	Endpoints & Derived variables	5
4	Study Methods	9
4.1	General Study Design and Plan	9
4.2	Inclusion-Exclusion Criteria and General Study Population.....	10
4.3	Randomisation and Blinding	11
4.4	Study Variables.....	12
5	Sample Size.....	17
6	General Considerations	18
6.1	Timing of Analyses	18
6.2	Analysis Population	18
6.2.1	Full Analysis Population	19
6.2.2	Per Protocol Population.....	19
6.3	Covariates and Subgroups	19
6.4	Missing Data	20
6.5	Multi-centre Studies.....	21
6.6	Multiple Testing.....	21
7	Summary of Study Data	Error! Bookmark not defined.
7.1	Subject Disposition.....	21

8	Efficacy Analyses.....	21
9	Reporting Conventions	22
10	Technical Details	22
11	Summary of Changes to the Protocol.....	22
12	Acknowledgment	23

2 Abbreviations and Key Sources

Abbreviations	
APPC	Active Patient Participation Coding
CI	Confidence Interval
CPS	Control Preferences Scale
FACT-G	Functional Assessment of Cancer Therapy – General.
FPI	Framing of Prognostic Information
HCCQ	Health Care Communication Questionnaire
IPS	Illness Preference Scale
MAR	Missing at random
MQOL	McGill Quality of Life
MUIS	Mishel Uncertainty in Illness Scale
PEACE	Peace, Equanimity, and Acceptance in the Cancer Experience
PECPI	Perceived Efficacy in Caregiver-Physician Interaction
PEPPI	Perceived Efficacy in Patient-Physician Interaction
QOL-LTC-C	Quality of Death Long-Term Care – Cognitively intact
QOLND	Quality of Life Near Death
SAP	Statistical Analysis Plan
SD	Standard Deviation

SES	Socioeconomic status
THC	The Human Connection scale
TDM	Terminal Decline Model
UCD	University of California, Davis
URoch	University of Rochester
VOICE	Values and Options in Cancer Care

Key Sources	
Prot_BMC	Protocol published in BMC Cancer Hoerger et al. BMC Cancer 2013, 13:188 http://www.biomedcentral.com/1471-2307/13/188
Prop_NIH	Research Grant Proposal to NIH (R01CA140419-01A)
Reg_CT	ClinicalTrials.gov registration (NCT01485627)
Emails on Dan's H-drive	Time-stamped emails documenting investigator decisions contained in folder and subfolders of H:\VOICE\AnalysisPlan on Daniel Tancredi's UCD Health Sciences network computer account.

3 Introduction

3.1 Preface

The Values and Options in Cancer Care (VOICE) Study will test an innovative synergistic intervention designed to improve communication between patients with advanced cancer and oncologists. Based on ecological theory, the intervention has both a patient/caregiver component and a physician component, and thus overcomes some of the shortcomings of prior communication studies. It targets patients at the time of diagnosis of incurable cancer and before the patient is critically ill, anticipating the need for information and strong patient-physician relationships as the illness progresses. It directly enables physicians, patients and caregivers to communicate more effectively – in contrast to studies that use third parties such as navigators or care managers. Furthermore, this proposal links research in communication and medical decision making – historically two separate approaches addressing common issues, but without a common frame of reference. [Source: *Prop_NIH*]

3.2 Purpose of the analyses

These analyses will assess the efficacy of the experimental intervention in comparison with the control (usual care) arm on the key patient and physician outcomes described below. A separate document will describe the statistical analyses to be used for the Caregiver Bereavement Study (PI: Paul Duberstein), which aims to assess intervention effects on caregiver outcomes.

3.3 Study Objectives

Our **Aims** are:

Aim 1 (primary): To determine whether a combined intervention for patients, caregivers and physicians improves communication regarding prognosis and treatment choices in advanced cancer.

H.1a. Primary hypothesis: Compared with control, intervention group participants will show improved communication – specifically, ENGAGING patients to participate in the consultation and decisions regarding their care; RESPONDING to patients' concerns; INFORMING patients about treatment choices; and balanced FRAMING of prognosis information.

H.1b. Secondary hypotheses: Compared with control, intervention patients will report fewer unmet needs for information and support; and there will be greater shared understanding about prognosis.

Aim 2 (secondary). To determine whether the intervention improves patient well-being.

Hypotheses: Compared to control, intervention patients will experience greater well-being as indicated by lower psychological distress, improved quality of life, greater sense of peace and improved quality of death.

Aim 3 (secondary). To determine whether the intervention affects health services utilization.

Hypotheses: Compared to control, intervention patients will receive fewer aggressive interventions in the last week of life and more palliative care and hospice consultations following the study intervention.

[Source: Prop_NIH]

3.4 Endpoints & Derived variables

(ICH E9; 2.2.2)

Aim 1a

The audio-recorded office visits for all participants are coded for each of the four domains of communication behaviors hypothesized to be affected by the intervention:

- (i) Engaging: a continuous measure of the number of active patient participation communication (APPC) behaviors
 - a. Details: Scoring will be performed by Rick Street, who will count the number of active *patient* participation communication behaviors.
- (ii) Responding: a simple count of the number of occurrences of 'providing/making/opening space' physician utterances according to Verona VR-CoDES system scoring
- (iii) Informing: outcome defined in Prognostic and Treatment Choices coding system. For items 1-9 on PTCC scale, score 1 point for preliminary exploration, 2 points for any further exploration, 2 points for any validation or teach-back, and subtract 1 point for a cutoff. Maximum score of 5 per item and minimum score of 0 per item. Report total score 0 – 45.
- (iv) Framing: For items 10 and 11 on PTCC scale, score 1 point for preliminary exploration, 2 points for any further exploration, 2 points for any validation or teach-back, and subtract 1 point for a cutoff. Maximum score of 5 per item and minimum score of 0 per item. Report total score 0 – 10.

Protocol change: The BMC Protocol specified a different measure for Framing, the Optimism/Pessimism subscale of the Framing of Prognostic Information (FPI) system, but that subscale was determined not to be suitable (according to April 16, 2014, email from Ron Epstein, H:\VOICE\AnalysisPlan\emails\Aim1_composite_2014\Aim 1 analyses.eml).

A composite outcome will be derived and used as the primary outcome for Aim 1a. To develop the composite, each of the four component scores will be Z-score transformed, using sample means and variance component estimate from all available Phase-1 data. For each observation, the composite score will be the simple weighted average of the four Z-scores, which will be analysed using mixed-effects linear regression models. Secondary analyses will be performed separately for each component score.

Protocol change: The decision to use a composite outcome as described above was finalized on May 6, 2014
(“H:\VOICE\AnalysisPlan\emails\AnalysisPlan\Re Tues. May 6th VOICE conference call data analysis group - 'Epstein, Ronald' (Ronald_Epstein@URMC.Rochester.edu) - 2014-05-06 1904.eml”) Rationale for decision was that a single composite would provide a more workable and powerful hypothesis testing strategy than would the original strategy described in the NIH R01 proposal. The original strategy required separate significance testing for each component and also added a nonstandard requirement that a majority of the intervention effects on individual components exceed a given threshold (0.5 SD) presumed to represent clinical significance.

Components of composite will be secondary outcomes. Additional prespecified Secondary/Exploratory outcomes:

- (i) Engaging

- a. Physician facilitation behaviors (to assess the effectiveness of the physician training) [Using doctor data, add variable PB "partnership building" and sup "supportive talk"]
- b. Patient activation in discussions about prognosis or the future (to complement the PTCC coding) [Using patient data, use variable FUTURE, which in the most recent spreadsheet of Phase II data is missing for all records.]
- c. Physician facilitation of discussions about prognosis or the future [Using physician data, use variable FUTURE, which in the most recent spreadsheet of Phase II data is missing for all records.]

Aim 1b

For the primary assessment of Aim 1b, a derived study variable will be developed as the difference in category levels for questions PT3 and DR3. **[Operationalization finalized in August, 2014. (see emails on Dan's H-drive)]**

Aim 2

Long-term effects of the intervention on quality of life, with attention to terminal decline at the end of life

Primary analysis: Quality of life including all patient surveys, considered a composite repeated measures outcome derived from the z-scores for each of 5 measures at each time point, analyzed using a terminal decline model (see Li Z et al Statistics in Medicine DOI: 10.1002/sim.5635):

- i) McGill single item
- ii) McGill existential well-being subscale
- iii) McGill psychological well-being subscale
- iv) FACT-G physical well-being subscale
- v) FACT-G social well-being subscale

Secondary analysis:

Caregiver ratings of the patient's QOL during the final month of life, obtained from the subset of patients who died and had caregiver enrolled in the study and available to do the post mortem 1-month interview

Each of the individual McGill and FACT-G subscales in all surveys, analyzed using a terminal decline model

PEACE scale using terminal decline modeling (

Analysis of short term (3-6 month) and long term (all measures) effects using standard modeling

Protocol change: Decision to use Terminal Decline Model as primary analysis was made in August, 2014. (see emails on Dan's H-drive) Rationale for decision is that TDM is better suited for modeling quality of life trajectories in intervention studies involving terminally ill patients because it can account for potential between-arm differential mortality as well as the phenomenon of "terminal decline" in quality of life in the months before death.

Aim 3

The primary outcomes are:

1. An index of aggressive care at the end of life, scored from 0-6, and consisting of 3 components:
 - a. Chemotherapy (cytoreductive therapy) in the last month of life -- exclude hormone based therapies
 - b. Medical services emblematic of aggressive EOL care ("acts of commission")
 - c. Potentially burdensome inpatient and emergency care

Scoring details are provided below.

a. Chemotherapy: 0-2 points

- 2 points: Any chemo last 14 days of life
- 1 point: Any chemo between days 15 and 30
- 0 points: No chemo in last 30 days of life

b. Medical services emblematic of aggressive EOL care: 0-2 points

2 points: Any one of the following services initiated within last 14 days of life (CPR, mechanical ventilation, tracheostomy, G-tube or J-tube placement, dialysis)

- 1 point: Any of the above initiated between 15 and 30 days
- 0 points: None of the above initiated in the last 30 days

c. Potentially burdensome inpatient and emergency care: 0-2 points

2 points: Sum of freestanding ED visits (not leading to admission) plus inpatient admissions (with or without an associated ED visit) last 30 days of life >2 (2 points);

1 point: sum of above=2
0 points: sum of above= ≤ 1

2. Hospice use prior to death
 - a. use/no use of hospice in last 30 days before death
 - b. time from date of enrollment in hospice until death (truncated at 30 days, so that enrollment 31 or more days before death would count as 30)

In addition we propose examining the following **pre-specified secondary outcomes**:

3. # days in hospital, last 30 days of life (truncated at 30 days)
4. # days in ICU, last 30 days of life (truncated at 30 days; excludes “step down” unit admissions)
5. Death in hospital
6. Death in ICU
7. Referral to hospice

We also specify one **exploratory outcome**:

8. Chemo use in the last 30 days of life
 - a. use/no use of chemo in last 30 days before death
 - b. time from date of last chemo administration (truncated at 30 days)

4 Study Methods

4.1 General Study Design and Plan

The research method of the study is a cluster randomized control trial (RCT) of a 2-component intervention designed to improve communication about prognosis and treatment choices in advanced cancer and promote patient participation in discussions regarding their care. The project has three phases.

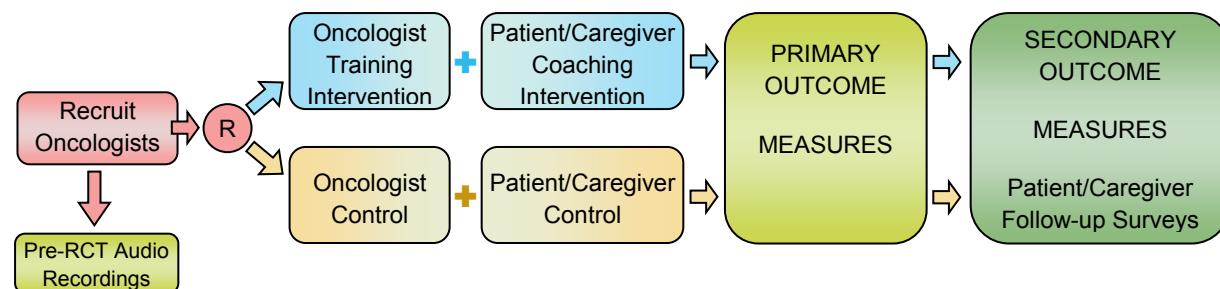
Phase 1 is for preparation, recruitment, and randomization.

Phase 2 is for the intervention: oncologists will receive two in-office tailored communication skills coaching sessions, after which patients of intervention oncologists and their caregivers will receive pre-visit coaching and question prompt lists.

Phase 3 is for data analysis.

Study subjects will include oncologists ($N \geq 40$), pre-RCT patients ($N \geq 120$) and caregivers (when available) for baseline physician communication measures (Phase 1), RCT patients ($N \geq 280$) and up to one caregiver identified by each RCT patient ($N = \text{up to } 280$) who actually participate in the clinical trial (Phase 2) for a possible total of 400 patients along with their caregivers. To establish communication patterns prior to oncologist randomization, each enrolled oncologist will have ~ 3 pre-RCT (Phase 1) audio-recorded patient visits at study entry. Randomization at the physician level to intervention or control (stratified by site and cancer subspecialty) will be performed by a statistician blinded to oncologist identities. Oncologists will be randomized to the control arm (usual care) or the intervention arm: two in-office sessions providing individual tailored feedback on their communication skills. Patients with advanced cancer and their caregivers will be assigned to the control (usual care) or the intervention: pre-visit coaching and question-prompt lists (QPLs) to address the same communication goals as oncologists. Because this is a cluster RCT, patients will be assigned to the control or intervention based on the patient's oncologist assignment. For each oncologist-patient dyad, only one oncologist-patient study visit will be audio-recorded.

The chart below was added to help facilitate your review of the study procedures.



[Sources: UCD IRB application (6/22/2011 version) and Prot_BMC]

4.2 Inclusion-Exclusion Criteria and General Study Population

(ICH E3;9.3. ICH E9;2.2.1)

Table 1. Inclusion and Exclusion Criteria for Oncologists, Patients, and Caregivers

Participant	Inclusion Criteria	Exclusion Criteria
Oncologist	<ul style="list-style-type: none"> • Currently in clinical practice at participating institutions • Oncologist that cares for patients with solid tumors • Not planning to leave the practice during the next 6 months 	<ul style="list-style-type: none"> • Non-physicians and physicians who are not oncologists • Oncologists who exclusively care for patients with hematological malignancies
Patient	<ul style="list-style-type: none"> • Currently a patient of an enrolled physician • Age 21 years or older • Diagnosis of Stage III or IV solid (non-hematological) cancer • Able to understand spoken English (study personnel to read materials to low literacy patients) • Oncologist, when asked, “would not be surprised” if the patient died within 12 months ^a 	<ul style="list-style-type: none"> • Anticipating bone marrow transplantation or diagnosed with leukemia or lymphoma • Unable to complete orally-administered surveys in English • Hospitalized or in hospice care at the time of recruitment
Caregiver	<ul style="list-style-type: none"> • Caregiver of a patient currently enrolled in the study • Age 21 years or older • Able to understand spoken English (study personnel to read materials to low literacy patients) 	<ul style="list-style-type: none"> • Unable to complete orally-administered surveys in English • Supported the patient primarily through a professional role (e.g., clergy)

^a This question is asked in order to exclude patients from the study who have stage III and IV cancers with potential for cure, such as stage IV testicular cancer or stage IIIa colon cancer.

4.3 Randomisation and Blinding

(ICH E3; 9.4.3, 9.4.6. ICH E9; 2.3.1, 2.3.2)

A stratified blocked-randomization scheme was used to assure balanced assignment by clinic site and cancer focus (breast cancer vs. other). In terms of clinic site, oncologists are grouped according to their health center, clinic, or practice of employment, and within each site, oncologists were randomly assigned approximately evenly across the treatment and control conditions. Clinic sites with a single oncologist were grouped with a similar site for randomization purposes. Because very few men are affected by breast cancer and because breast cancer patients may be more “activated” than patients with other cancers, oncologists are categorized as either a breast–cancer oncologist ($\geq 50\%$ of patients have diagnoses of breast cancer) or not, with the two groups of oncologists randomly assigned approximately evenly across the treatment and control conditions. For each clinic site/focus combination in the study, separate sequences of random numbers have been generated for use in assigning oncologists.

Randomization was implemented by the original study statistician, Paul C. Winters, MS, using SAS Statistical Software.

Blinding. To preserve blinding, assignment to the treatment or control conditions is maintained by the study statisticians and project manager, and not explicitly revealed to research assistants, transcriptionists, or coders of the audio-recorded office visits.

4.4 Study Variables

(ICH E3; 9.5.1. ICH E9; 2.2.2)

Table 2. Survey Measures Completed by Patients in the RCT

Domain	Measures	Study Entry	Post-Visit	2-4 Day Follow-up	Quarterly Follow-up
Demographics	Gender, age, race/ethnicity, SES, relationship status, religion	X			

Physician-Patient Relationship	THC	X	X	X ^a
Physician-Patient Interaction	HCCQ + MUIS + IPS		X	
Patient Communicational Self-Efficacy	PEPPI	X	X	X ^a
Physician-Patient Conversations	Topics discussed in recent medical encounters	X		X
Preferred Decision Role	CPS	X	X	
Actual Decision Role	Modified CPS	X		X
Treatment Preferences	Preferences for experimental treatments, life support, palliative care	X	X	
Illness Acceptance	PEACE	X		X
Well-being	MQOL + FACT-G	X		X
Prognostic Forecasting	Estimate of prognosis	X	X	

Note. SES = socioeconomic status; THC = The Human Connection scale; HCCQ = Health Care Communication Questionnaire; MUIS = Mishel Uncertainty in Illness Scale; IPS = Illness Preference Scale; PEPPI = Perceived Efficacy in Patient-Physician Interaction; CPS = Control

Preferences Scale; PEACE = Peace, Equanimity, and Acceptance in the Cancer Experience; MQOL = McGill Quality of Life; FACT-G = Functional Assessment of Cancer Therapy – General.

^a Administered at only the first quarterly follow-up

Table 3. Survey Measures Completed by Caregivers in the RCT

Domain	Measures	Study Entry	2-4 Day Follow-up	Quarterly Follow-up	Post-Mortem
Demographics	Gender, age, race/ethnicity, SES, X relationship with patient				
Physician-Caregiver Relationship	THC		X		X ^b
Physician-Caregiver Interaction	HCCQ		X		
	MUIS		X	X ^a	
Caregiver Communicational Self-Efficacy	PECPI		X		X ^a
Patient Treatment Preferences	Caregiver's beliefs about patient preferences for experimental treatments, life support, palliative care	X		X ^a	
Patient Illness Acceptance	PEACE		X		X ^b
Patient Well-being	MQOL + FACT-G		X		X ^b

Prognostic Forecasting	Estimate of patient's prognosis	X	X
Patient Death	Qualitative questions + QOLND + QOD-LTC-C		X ^b

Note. PECPI = Perceived Efficacy in Caregiver-Physician Interaction; QOLND = Quality of Life Near Death; QOD-LTC-C = Quality of Death Long-Term Care – Cognitively intact; other acronyms defined previously (see Table 3).

^a Administered at only the first quarterly follow-up

^b Retrospective rating of the death/dying experience, rather than the current moment

Table 4. Survey Measures Completed by Physicians

Domain	Measures	Study Entry	Post-Visit	Study Conclusion
Demographics	Age, race/ethnicity, gender, training background, practice characteristics	X		
Communication Skills	Skills in discussing prognosis and end-of-life care	X		X
Decision Making Skills	Comfort with decision making across varying levels of patient involvement	X		X
Patient Disease Status	Cancer site, progression, treatment planning		X	
Patient Treatment Preferences	Physician's beliefs about patient preferences for experimental treatments, life support, palliative care		X	
Patient Illness Acceptance	PEACE		X	
Prognostic Forecasting	Physician's estimate of prognosis, and beliefs about patient's estimate of prognosis		X	

Note. PEACE = Peace, Equanimity, and Acceptance in the Cancer Experience.

5 Sample Size

"This is a stratified cluster randomized study, with the physician as the unit of randomization. Based on prior studies, we have made the following assumptions: physician attrition 0-3%, patient attrition <5% for audio-recordings, and 10%, 30% and 35% for the 2-4 day, 3-month, and 6-month post-visit patient surveys, 80% patient mortality during 3-year follow up, availability of 85% of caregivers for post-death interviews, availability of 90% of medical records for audit, no differential attrition between the intervention and control groups, and an intraclass correlation coefficient (ICC) of .10 or less for within-physician clustering on patient and caregiver survey measures.

Prior work found that activation training increases physician communication behaviors two to three fold. This equates to an effect size of 2.0 standard deviation (**SD**) units. Thus, for our primary outcome (Aim 1a), we expect a 2.0 **SD** improvement across each communication measure; 0.5 **SD** is clinically significant. Power is based on a single measurement of communication during the oncology office visit. With at least 19 oncologists in each group, 7 patients per oncologist, and an ICC of .10, the minimum detectable effect size is 0.50 **SD**.

For our survey outcomes (Aims 1b & 2), we rely on observational data indicating effect sizes of 0.40 to 0.70 **SD** for the relationship between communication and patient well-being. With an intervention we would expect these differences to be larger. The power analysis takes advantage of a repeated-measures design. For patient data, we assume one pre-intervention measure and two post-intervention measures, and an average correlation among repeated measures of .50. Our proposed sample size is sufficient to detect an effect size of 0.40 **SD** with power of .80. For our utilization outcomes, we rely on data that suggest 3- to 8-fold differences in use of aggressive treatments during the final week of life between patients who have had discussions compared to those who have not. Thus, the study will be adequately powered, even considering attrition."

[Source: Prot_BMC]

6 General Considerations

6.1 Timing of Analyses

For each study Aim, the final analysis will be performed based on a finalized SAP and at the conclusion of Phase 2 data quality assurance activities (range checks, verification/correction of outlying values for study variables, coding of derived study variables and determination of inclusion/exclusion status for each study subject with regard to analysis populations) for all applicable data. Data quality assurance activities will be conducted in a preliminary round of data analysis that will be conducted without the use of treatment assignment information, to preserve the integrity of the analysis. The finalization of the SAP and the formal conclusion of data quality assurance will be decided by the principal investigators, in consultation with the study statistician.

6.2 Analysis Population

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 5.2)

6.2.1 Full Analysis Population

- All available data from patients, caregivers and physicians collected after formal consent has been obtained and for which formal withdrawal has not been requested. Treatment assignment is based on the randomized treatment assignment of the physician at the time of the index physician/patient visit. The randomized treatment assignment for the physician is non time-varying.

6.2.2 Per Protocol Population

- All observations from patient-physician dyads where both participants received the assigned intervention

Each study subject's inclusion/exclusion status will be established as part of the preliminary analyses

6.3 Covariates and Subgroups

(ICH E3; 9.7.1, 11.4.2.1. ICH E9; 5.7)

All regression analyses will include terms to control for the randomization stratifiers, Study Site and Oncologist Type. In addition, a parsimonious set of covariates will be used to control for measured patient-level covariates (sex, patient age, race/ethnicity). In case a key covariate is found to be unbalanced between study arms, it will be included in the model as a potential confounder.

Formal subgroup analyses will be based on formal testing of interaction terms (using an inflated significance threshold, $p < 0.05$) and will only be conducted for the following terms:

Main analysis

Study Site (UCD vs. URoch)

Oncologist type (predominantly breast cancer vs. other)

Common Cancer Site (the 5 most common sites vs. all others)

Oncologist 2 year survival estimate at baseline (Good (90-100%) vs. Poor (<75%))

Whether caregiver is enrolled

Coach rating of patient engagement in the intervention, based on 9-item post-intervention scale.

Exploratory analysis

Patient Race

Patient Gender

Cancer System or Site

Caregiver education

Number of coaching session completed (among 1 in-person and up to 3 telephone sessions)

Oncologist rating of usefulness of educational session

[Protocol change: The finalization of subgroups for analysis was made in November, 2014]

6.4 Missing Data

(ICH E3; 9.7.1, 11.4.2.2. ICH E9;5.3. EMA Guideline on Missing Data in Confirmatory Clinical Trials)

Our analysis strategy will primarily rely on mixed-effects models of all available observations from eligible participants, an approach that makes efficient use of nonmissing data under the same ignorable missing data assumption ("Missing at Random") that is necessary to ensure the validity of standard multiple imputation procedures. The MAR assumption is not testable with study data. To assess the sensitivity of treatment effect estimates to the MAR assumption, sensitivity analyses will be conducted using multiple imputation models that incorporate varying levels of nonignorable missingness. In particular, nonignorable missingness models will make imputations for missing data that assume that outcome distributions have more favourable location parameters for the control condition than for the intervention condition to permit pessimistically attenuated ("worst case") treatment effect estimates.

For intermittent missingness on scale items when more than half of the items on that scale are nonmissing for a particular observation, we will follow our usual custom of imputing the mean value of that observation's nonmissing items.

6.5 Multi-centre Studies

(ICH E3;9.7.1, 11.4.2.4. ICH E9; 3.2)

This multi-centre study is intended to be analysed as a whole. All regression models will include stratifiers for study site (UCD vs. URoch). For each outcome, modification of treatment effects by study site will be formally tested using interaction terms. In case the interaction term is significant at $p < 0.05$, study site specific treatment effects will be reported.

6.6 Multiple Testing

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 2.2.5)

The primary concern with regard to multiple testing arises from the separate analyses for the four components of the composite outcome used for Aim 1a. For these secondary analyses, no formal adjustments for multiple testing will be made. However, study reports will clearly describe these analyses as secondary and will include information on the total number of secondary outcomes that were assessed, to permit readers to perform their own adjustments.

7 Description of Subject Disposition in Trial Report

7.1 Subject Disposition

Subject disposition will be described in CONSORT flow diagrams for patients, caregivers and providers.

8 Efficacy Analyses

As described in the BMC Protocol, we will primarily rely on regression models for clustered data to account for the stratified cluster randomized longitudinal study design. The primary analysis population will be based on the assigned treatment.

Treatment effect estimates will be presented with 95% Confidence Intervals.

Protocol change: To account for pre-randomization data while permitting adjustment for patient-level covariates, the decision was made on April 16 2014, to use a differences-in-differences framework for estimating treatment effects on communication outcomes. For this analysis, main effects for study phase (0 'pre-randomization' vs. 1 'post-randomization'), study arm (0

‘control’ vs. 1 ‘intervention’) and their interaction are included in the model. The regression coefficient associated with the interaction represents the adjusted mean between-arm (intervention vs. control) difference in over-time (post vs. pre) changes. (See “H:\VOICE\AnalysisPlan\emails\Aim1_composite_2014Re Phone call Dan Rich Josh Ron April 16 2014 – ‘Richard L. Kravitz’ (rlkravitz@ucdavis.edu) – 2014-04-16 1522.eml” on Dan’s H-drive)

9 Reporting Conventions

Reporting conventions will accord with the most recent edition of the AMA Manual of Style. In particular, P-values ≥ 0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as “ <0.001 ”. The mean, standard deviation, and any other statistics other than quantiles, will be reported to no more than 2 decimal places greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

Trial reports will follow relevant CONSORT statements, including extensions for cluster-randomized trials and non-pharmacological interventions.

10 Technical Details

Drs. Xing and Tancredi will conduct the statistical analysis at UC Davis, using SAS software. All SAS code will be stored on a secure network drive available only to designated VOICE research personnel or on private personal network drives. To facilitate replication and to promote quality, SAS programming code will be developed according to standard operating procedures developed earlier for the Ca-HELP study. In particular, standard SAS code modules will be developed for study variable definition and for dataset assembly.

11 Summary of Changes to the Protocol

Important changes to the analysis plan made after publication of the BMC protocol are indicated above in boldfaced font. The most notable decisions were to use a composite outcome for the primary study outcome, to use a difference-in-differences framework for estimating intervention effects on communication outcomes, and to use a terminal decline model for modelling quality of life trajectories.

12 Acknowledgment

This SAP was prepared by modifying the SAP Template downloaded from the Cambridge University Hospitals,
http://www.cuh.org.uk/sites/default/files/research/CCTU_TPL007_SAP_template.doc