Default Options in Advance Directives for Seriously Ill Patients: A Randomized Clinical Trial

Study Protocol and Statistical Analysis Plan

A prospective randomized controlled trial to examine whether structuring advance directives to request comfort-oriented goals of care by default improves patients' quality of life and reduces resource utilization without reducing the number of days that patients are alive and living outside of an acute-care hospital.

Supported by:

The Gordon and Betty Moore Foundation and the Otto Haas Charitable Trust

Table of Contents

I. Original Protocol

1. Abstract

Although most seriously ill Americans wish to avoid burdensome therapies near life's end, aggressive care is provided unless or until patients or their family members actively request that it is stopped. Advance directives (ADs) hold great promise for combating this societal default of aggressive end-of-life care, but to date this promise has been largely unrealized. This study will test the premise that ADs can better align the end-of-life care patients receive with the care they want if the ADs are restructured such that comfort-oriented care is provided as the default, rather than forcing patients to make emotionally and existentially challenging choices to receive it. In this study, we will determine whether this simple and readily scalable intervention can improve patients' quality of life and reduce resource utilization without reducing the number of days that patients are alive and living outside of an acute-care hospital.

2. Background and Significance

Most Americans wish to die at home and to avoid aggressive care and life support when terminally ill. Yet the opposite commonly happens: one in five Americans dies in or shortly following a stay in an intensive care unit (ICU), roughly half of U.S. deaths occur in a hospital one third of elderly patients undergo an inpatient surgical procedure during their last year of life, one half of elderly Americans visit emergency departments in the last month of life, and more than one quarter of Medicare dollars are spent on patients in their final year. Perhaps even more concerning are recent observations that aggressive treatment of patients with serious illnesses is associated with reduced quality and perhaps quantity of life near its end. When such care culminates in ICU-based deaths, it also produces long-lasting pathological bereavement among family members contravening most patients' strong desires not to burden their loved ones.

Despite past failures, written advance directives (ADs) hold great promise. A recent study highlights a key reason for the discrepancy between the care we want and the care we receive near life's end: critical healthcare decisions must be made for 43% of older Americans near the times of their deaths, but 70% of these patients cannot participate in making these decisions. The cumulative result – that 30% of older Americans cannot choose their care when such choices are needed – highlights the potential benefits of improving the quality of advance care planning, including written advance directives (ADs).

ADs include living wills, in which patients can choose to receive or avoid life-sustaining therapies if they lose capacity to make such decisions, and designation of a durable power of attorney for healthcare to serve as the patient's decision-maker in similar circumstances. Many

experts have bemoaned the shortcomings of ADs, particularly for the living will component . Such concerns have spawned a broader focus on advance care planning that seeks to prepare patients and family members for difficult decisions. Sound in principle, this approach is difficult in process. For the right patient, surrounded by the right family, and cared for by the right clinicians, such coordinated communication may prove optimal. But this approach may be difficult to implement across diverse populations with differential access to longitudinal care.

By contrast, fixing the problems with ADs may yield more scalable ways to improve end-of-life care for all Americans. Recent evidence provides substantial motivation to try. Observational studies in the United States show that elderly patients who complete ADs less commonly die in a hospital, more often receive care consistent with their preferences, and receive less costly care.

Despite these recent studies showing the promise of ADs, none provide sufficient evidence that completing ADs, or certain types of ADs, will cause changes in clinical, economic, or patient-centered outcomes. Studies noting improved patient-centered and economic outcomes among patients completing ADs were all observational in nature, preventing conclusions about whether AD completion caused these benefits or was a marker for people likely to attain them anyway. Thus, given federal policies promoting AD completion, and evidence that completion rates are increasing in the U.S., an RCT is desperately needed to determine how best to design ADs to improve patient outcomes without increasing resource utilization.

3. Objectives

3.1 Overall objectives

This study will test the premise that ADs can better align the end-of-life care patients receive with the care they want if the ADs are restructured such that comfort-oriented care is provided as the default, rather than forcing patients to make emotionally and existentially challenging choices to receive it.

3.2 Primary outcome variable(s)

The primary outcome is “Hospital-Free Days” (HFDs), a measure that PI Halpern has been developing in collaboration with Dr. Jeffrey Silber at Penn’s Center for Outcomes Research. As the name describes, HFDs represent the number of days alive and not in an acute care facility. Although this is a simple concept and provides an outcome measure of obvious importance to patients, the use of HFDs as a primary outcome in an RCT is highly innovative. To bolster confidence in the results, we will evaluate two key variations on the theme. First, we will explore “Healthcare Facility-Free Days,” which represents the number of days alive where a patient is neither in an acute care facility, a chronic care facility, nor a nursing home. We will also evaluate HFDs within a defined period of follow-up – 6 months in this case. This is analogous to the established outcome of ventilator-free days used commonly in RCTs among

ICU patients.

3.3 Secondary outcome variable(s)

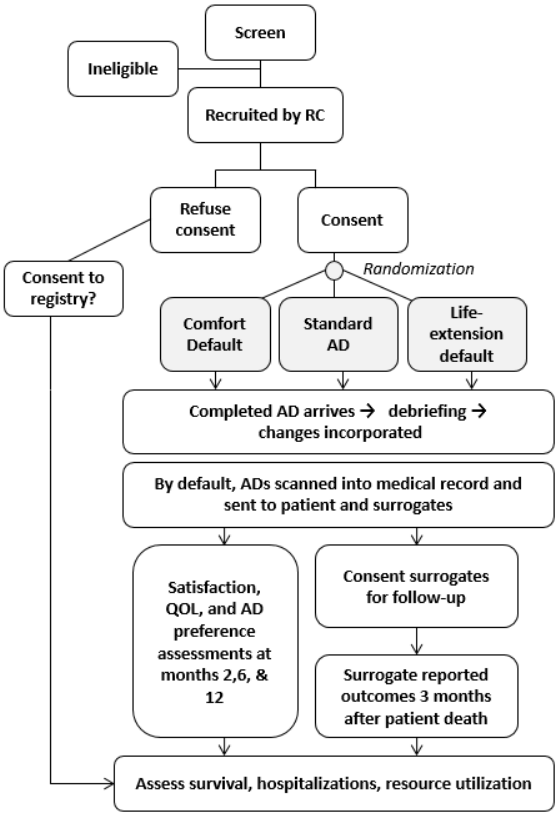
1. Hospital and ICU admissions: The numbers of admissions will be analyzed as count data. From the dates of hospital and ICU admissions, we will calculate the proportion of each patient's total survival time during study follow-up that was spent in the hospital or ICU.
2. Costs of care: We will combine all costs of inpatient and outpatient hospice, hospital stays, and life-sustaining procedures. The perspective will be that of all potential payers. Costs will be inflated to the date on which analyses are performed using the U.S. gross domestic product deflator.
3. Hospice utilization: We will analyze hospice utilization in 2 ways: (a) time from AD completion to hospice enrollment; and (b) duration of hospice utilization prior to death.
4. Choices to receive 4 potentially life-sustaining interventions, and the concordance of these choices with whether the interventions were actually received: The outcomes databases we will use contain codes for each of the 4 interventions, enabling us to determine which patients received each. Thus, we will be able to reliably evaluate the proportions of patients who received unwanted interventions. Because we cannot determine the denominator of patients with indications for these interventions, we will not evaluate the proportions of patients who went without desired services.
5. Choices regarding post-hospitalization care, and the concordance of these choices with the care actually received.
6. Decision regret and satisfaction: Decision regret will be measured using the 5-item decision regret scale that has previously been shown to have good internal consistency and strong inverse associations with decision satisfaction. Satisfaction will also be measured more specifically with the CANHELP instrument's global satisfaction with end-of-life care question.
7. Quality of life, using the McGill Quality of Life (MQOL) instrument. The MQOL is a well-validated and widely used scale designed specifically for patients with serious illnesses. The MQOL can be completed by family members on behalf of patients who have lost the capacity to complete it. Thus, we will have surrogates complete the MQOL for incapacitated patients to minimize missing data.
8. Surrogates' Perception of the quality of death and dying: We will assessed this outcome with surrogates of deceased patients using the quality of dying and death (QODD) instrument.
9. Bereavement outcomes: The risk of post-traumatic stress disorder in surrogates among deceased patients will be assessed using the Impact of Events Scale (IES). The IES is a valid and reliable scale that has been used frequently to assess PTSD risk among family members of critically ill patients. Finally, complicated grief will be assessed using

Prigerson's Inventory of Complicated Grief to distinguish pathologic grieving from normal bereavement.

4. Study Design

4.1 Schema

This is a prospective, randomized, controlled trial.



4.2 Duration

The study period is two years. Subjects will be accrued over a period of 18 months starting in January 2014. The total time it will take for the research coordinator to explain the study, obtain consent and for a subject to complete the advance directive will, conservatively, take no more than two hours. The debriefing discussion and follow up interviews will take approximately 15 – 25 minutes each. The total time spent on research activities for patients should be no more than 4 hours.

5. Subject recruitment

We will recruit 270 patients with severe respiratory, oncological, neuromuscular, or cardiovascular diseases and limited life expectancy from the Perelman Center for Advance Care Medicine, Penn Presbyterian Medical Center, Pennsylvania Hospital, and the University of Pittsburgh Medical Center. Each week the research coordinators will screen the electronic medical records of patients scheduled for routine visits to determine their study eligibility using the eligibility criteria outlined above.

Once eligible patients have been identified, research coordinators will email eligible patients' providers to 1) alert them to their patients' eligibility for participation 2) inform them their patients will be recruited for enrollment 3) provide them an opportunity to decline or defer any given patient's enrollment by responding to the email. Research coordinators will approach potential study participants while they are in the waiting areas, chemotherapy infusion areas, or in exam rooms waiting to see their doctor on the day of their visit.

5.1 Accrual

During our pilot study we were able to recruit approximately six patients per month with one full-time research coordinator. We anticipate that with the equivalent of 3.5 full-time research coordinators and an additional site (University of Pittsburgh), we will be able to recruit approximately 18 patients per month.

5.2 Key inclusion criteria

The eligibility criteria, all of which must be met, are:

1. Age 18 or older
2. Speaks and reads fluent English
3. Has seen current physician at least once prior to current visit
4. Resident of Pennsylvania or New Jersey
5. One or more of the following diagnoses:

- Amyotrophic lateral sclerosis
- Stage IIIB or IV non-small cell lung cancer, pancreatic cancer, or cholangiocarcinoma
- Stage IV colorectal, esophageal, gastric (including GIST), pancreatic, prostate, or urothelial cancer; paraganglioma, or pheochromocytoma
- Stage C or D hepatocellular carcinoma
- Stage IV renal cell carcinoma
- Stage IV or V chronic kidney disease
- Mesothelioma or any malignancy metastatic to the pleura
- Other incurable interstitial lung diseases with at least severe restriction on most recent pulmonary function tests or eligible for long-term oxygen therapy

- Chronic obstructive pulmonary disease with at least severe airflow obstruction on most recent spirometry or eligible for long-term oxygen therapy
- Congestive heart failure with NYHA Class IV status or Class III plus 1 heart failure related hospitalization in the past 12 months or ACC stage D or C classification with 1 heart failure related hospitalization in the past 12 months
- Stage IV breast cancer except patients whose only metastases are to the bones or who are receiving endocrine therapy without receiving concomitant traditional chemotherapy

5.3 Key exclusion criteria

Patients will be excluded if they are currently listed for or being considered for solid organ transplant and if they have a previously signed advance directive or living will. Cognitively impaired patients will be excluded from the study to avoid the necessity of proxy consent.

5.4 Subject Remuneration

Patients will be compensated with $20 at the day of enrollment in cash. In order to enhance study retention and participation in follow-up assessments, $20 will also be paid to subjects at the completion of the two, six, and twelve month follow-ups. Surrogates will also be compensated $20 after they consent to participate.

6. Randomization

6.1 Groups

Subjects enrolled in this RCT will be randomized into three groups. Depending on which group they've been assigned, subjects will be given one of three AD forms. The three AD forms have been created with different default treatment options. Form 1 (life-extension default) will state that 4 specific life-extending interventions (cardiopulmonary resuscitation, mechanical ventilation, hemodialysis, and feeding tube insertion) will be provided unless patients specifically opt-out from such selections. Form 2 (comfort default) will state that the 4 specific life-extending interventions will not routinely be provided unless patients elect to receive such measures. Finally, Form 3 (standard advance directive) will use the standard approach of requiring patients to actively choose whether or not they wish to receive each intervention, as they would if completing an AD outside of a research setting. In this case, if they do not make a selection, decision making would default to their surrogates as in usual practice.

Because patients may focus on an overall plan of care rather than the receipt of specific interventions, all AD forms will also include a general question regarding treatment priorities. The response to this question, is modeled on one used in a Study to Understand Prognoses and

Preferences for Outcomes and Risks of Treatments (SUPPORT) study. The question acknowledges that while, in general, most people wish to both live as long as possible and avoid pain and suffering, in some situations, choosing between these two goals may be necessary. It then asks patients, if they are in a situation where such a choice is needed, whether they prefer a plan of care that focuses on extending life as much as possible even if it means having more pain and suffering, or a plan of care that focuses on relieving pain and suffering even if that means not living as long. The default framing of this general question will be in accord with that used for the specific interventions in each AD form, and all patients will be able to select a "no" option in response to this question.

Finally, we will include a specific question about the care patients wish to receive upon discharge from the hospital, defaulting to hospice-based care (in the comfort-default group), long-term care (in the life-extension-default group), or no option pre-selected. In the standard AD group, although no options will be pre-selected, we will randomly assign whether the comfort-oriented option or the life-extending-oriented option is presented first so as to mitigate ordering effects. In all cases, the option of not deciding will be presented last.

6.2 Assignment

Eligible patients will be approached about participation by the research coordinators in the outpatient clinics at the Perelman Center for Advanced Medicine, Pennsylvania Hospital, and Presbyterian Hospital. Consenting subjects will be randomized with a 33.3% probability to each trial arm (life extension default, comfort default, standard AD) using electronic procedures monitored by the Data Management Unit within the Biostatistics Analysis Center. We will stratify the randomization by recruiter/research coordinator, and will use variable block sizes of 3 and 6 patients to promote balance of follow-up duration among the 3 trial arms.

Each research coordinator will go to his or her clinics each day with a sealed envelope in which there is a pre-determined sequence of the 3 trial packets. The research coordinator will become unblinded to the patient's allocation at the time of consent, but with variable block sizes, can never predict with certainty what the next packet will be.

7. Study Procedures

7.1 Screening for Eligibility

The research coordinators will screen electronic medical records of patients visiting pulmonary, renal, heart failure, movement disorder, and oncology clinics at the Perelman Center for

Advance Care Medicine, Penn Presbyterian Medical Center, Pennsylvania Hospital and the University of Pittsburgh Medical Center for eligibility. Patient's eligibility status will be entered into the eligibility database. We will record ICD9 and ICD 10 codes, staging information, relevant provider name, clinic location, and upcoming appointments for eligible patients.

7.2 Recruitment

Eligible patients will be approached by a research coordinator in the clinics who will seek patients' consent to participate in a study comparing different types of ADs. Of note, while some providers may be more proactive than others in engaging their patients in conversations about advance care planning, it is generally not standard-of-care that patients are approached about completing ADs. The research coordinator will specify that the ADs in this study are intended to be real ADs and that they will be included in patients' outpatient medical records, but that, as with all ADs, patients retain the right to change their selections at later dates. The research coordinator will also specify that, like all real ADs, they are most useful if copies are shared with their loved ones and physicians.

7.3 Informed Consent

Following discussion of the study, research coordinators will obtain written consent from patients. The consent forms will contain HIPAA statements of authorization of release of medical records, thus facilitating our collection of data from medical and billing records during the study. The consent includes clear explanations that different types of ADs will be assigned by chance, but that patients in all groups may select or decline any intervention or treatment goal, and may revise their choices at any time. The research coordinators will explain who will be enrolled, how many patients are being targeted for enrollment, the specific components of patient follow-up, patients' rights to withdraw from the study at any time and for any reason, and what the outcomes of interest are (e.g., utilization of healthcare services, AD selections).

Patients who do not wish to complete an AD and decline consent will be asked to instead sign a limited consent form that would provide authorization to assess their long term health outcomes from electronic health records, but would not entail any further direct patient contact. The goal of this research would be to compare the outcomes of patients who have an advance directive against those who do not. Consenting patients would agree to participate in a registry by providing their social security number for purposes of merging with state maintained datasets described in detail below. Patients who agree to participate in the registry would also provide their age, race, ethnicity, and gender.

7.4 Enrollment

After patient consent is obtained, the research coordinator will ask subjects to complete the demographics survey and walk subjects through the process of filling out the AD. Along with their AD forms, consenting subjects will be given a copy of their consent form, an informational brochure about advance care planning, contact information for research study staff, instructions for mailing back their completed AD forms, and a stamped and addressed envelope. To enhance retention, patients will also be given $20 at the point of consent.

Subject IDs will be assigned at the point of consent. Subject ID numbers, demographic information and group assignments will be entered into the analytic database. Subject contact information, including social security number, will be entered into a subject tracking database.

If completed ADs are not returned within 10 days, staff will call patients weekly to remind them to return their ADs, to schedule special clinic visits for AD completion if patients desire, and to answer any questions. If research staff members are unable to reach patients over the phone after three attempts, a letter will be sent to patients to remind them to return their ADs and encourage them to contact research staff if they have any questions or difficulties. If we are still unable to reach patients, they will be approached by the research coordinator in their next clinic visit.

7.5 Subject Debriefing

After patients complete their assigned AD, there will be a structured debriefing session conducted over the phone by a research team member in which a standardized explanation of all three ADs will be given. This debriefing will be held to alert patients to exactly how the three ADs used in the study differ. As in the pilot study, patients will not be alerted to the different default framings up front because patients in clinical settings (and indeed in this study) are only asked to complete a single AD. Explaining non-relevant ADs prior to completion of the relevant one could influence decisions in ways that would not reflect actual clinical settings, thereby biasing the results. However, because this is a research study and AD assignment is at random, it is appropriate to debrief patients afterwards to grant them such broader information. Once patients are fully informed about the variations in the ADs used in the study, they will be asked if they wish to change any of their AD selections prior to finalizing the documents as a part of the medical record. Patient ADs will not be considered “complete” until the debriefing session

has taken place. After the debriefing call, patients' AD selections will be entered into the analytic database.

During the debriefing call, we will tell subjects that we will scan their AD forms into their medical records for them, unless they do not desire this (it is optional, not a requirement of the study). Similarly, we will also inform subjects that a copy of their completed AD will be sent back to them, and that, if they wish, a copy will be sent to their appointed healthcare agent/surrogate. Completed ADs will be sent to patients and surrogates along with letters explaining that ADs can be changed at any time and they can contact the research team with questions.

The research team will contact the appointed health care agent/surrogate identified in completed AD forms to seek the surrogate's consent to a) contact him or her in the event that we are unable to reach the patient for follow-up, and b) participate in an interview related to surrogate outcomes participate in follow-up and surrogate interviews. We will also ask surrogates to notify the research team if patients die during follow-up. The research team will also contact the appointed health care agent/surrogate identified in completed AD forms to seek the surrogate's consent to a) contact him or her in the event that we are unable to reach the patient for follow-up b) participate in an interview related to surrogate outcomes participate in follow-up and surrogate interviews. We will also ask surrogates to notify the research team if patients die during follow-up. This consent process will take place over the phone and, thus, we are requesting a waiver of documentation of informed consent.

7.6 Subject Follow-up

Two, six, and twelve months after AD completion, subjects will be contacted for participation in follow-up interviews. The follow-up interviews will take place over the phone with a research associate blinded to the subject's study arm. The research associates will attempt to contact patients up to three times over the phone. If the research associates are unable to reach the patients, they will inform the research coordinator. The research coordinator will contact the patients in person the next time they arrive in clinic to ensure that patients do not have any questions or concerns about their participation in the study and set-up a time for the follow-up call. If patients are unavailable to participate in follow-up calls because they are deceased or otherwise incapacitated, we will interview their surrogates.

In the event of a patient's death, a research associate will contact the patient's surrogate between 2-3 months after the death for a telephone interview. During the telephone interview the research associate will assess quality of death and dying and bereavement outcomes using

the Quality of Dying and Death (QODD) instrument, the Impact of Events Scale (IES), and Prigerson's Inventory of Complicated Grief.

7.7 Assessment of Health Outcomes

We will assess hospitalizations, ICU admissions, costs of inpatient care, and utilization of life-sustaining therapies by querying state-run databases that capture all admissions and inpatient procedures in Pennsylvania and New Jersey. The Pennsylvania Health Care Cost Containment Council (PHC4) is an independent state agency that maintains a database of inpatient hospital discharge and outpatient procedure records from all hospitals and ambulatory surgery centers in Pennsylvania. These data include specific treatment information including costs. As roughly one-third of Penn's outpatient population resides in New Jersey, we will obtain comparable data from the New Jersey Discharge Data Collection System (NJDDCS) managed by the New Jersey Department of Health and Senior Services within their Department of Health Care Quality and Assessment (HCQA). We will establish data use agreements with both of these entities and be subject to IRB approval by HCQA. Linkages with both PHC4 and NJDDCS will be performed by the respective database administrators after we provide lists of included social security numbers and subject IDs. PHC4 and NJDHSS will send our team a report in which patients are identified by subject ID only. Identical processes have been used seamlessly and with high fidelity by many Penn investigators.

We will collect data on hospice utilization and costs via data use agreements with Penn Wissahickon Hospice and Family Hospice and Palliative Care. These organizations provide hospice services for 80% of Penn and Pitt patients, respectively.

8. Data Management

8.1 Data Confidentiality

Only authorized project personnel will have access to the data. All study data will be stored behind firewalls on Center for Clinical Epidemiology and Biostatistics (CCEB) servers; none will be stored on stand-alone PCs or laptops. All study personnel who work with these data will have undergone required human subjects training. To ensure that participant confidentiality is preserved, individual identifiers (such as social security number) will only be used to link patient records (e.g., linking subject database to PHC4 data). Once linkages between databases have been achieved, all linkage-identifiers will be dropped from all datasets. Throughout the study duration, we will maintain one master list that will link study identification numbers to patient identifiers. This list will be maintained by the principal investigator in a locked file drawer in his locked private office to ensure file security. This file will be made available to other research staff on a need-to-know basis only, and, in that case, only temporarily. The study ID will be used exclusively in all analytical files.

All datasets and computer files with study ID numbers will be further secured as follows. The University of Pennsylvania (Penn) Data Management Unit (DMU) is an arm of the broader Biostatistics and Epidemiology Consulting Center (BECC), all of which is housed within the CCEB. The DMU will be the hub for the database infrastructure that will support the project. The DMU provides a secure computing environment for a large volume of highly sensitive data, including clinical, genetic, socioeconomic, and financial information. We will implement multiple, redundant protective measures to guarantee the privacy and security of the participant data. All data for this project will be stored on the secure/firewalled servers of the CCEB in data files that will be protected by multiple password layers. These data servers are maintained in a guarded facility behind several locked doors, with very limited physical access rights. They are also cyber-protected by extensive firewalls and multiple layers of communication encryption. Electronic access rights are carefully controlled by Penn system managers. We will use highly secure methods of data encryption for all transactions involving participants' financial information using a level of security comparable to what is used in commercial financial transactions. This multi-layer system of data security, identical to the system protecting the University of Pennsylvania Health System's medical records, greatly minimizes privacy risks.

8.2 Subject Confidentiality

Steps will be taken to ensure that all information will be kept confidential and secure. Unique patient identifiers numbers will be assigned to each subject locally and kept in a secure encrypted file. Records with patient social security numbers will be maintained, used, and destroyed in a way that is consistent with Penn policy. All paper records will be kept in locked files; all computers will be password protected and kept in locked rooms; all databases will be password protected and maintained on encrypted hard-drives behind the CCEB firewall. All study data will be stored behind firewalls on Center for Clinical Epidemiology and Biostatistics (CCEB) servers; none will be stored on stand-alone PCs or laptops. All data will be destroyed after 7 years.

8.3 Subject Privacy

Individual-level data for participants will be kept confidential and will only be stored on highly secure servers available for patient-level data. Only authorized project personnel will have access to the data and the data will be stored on servers only and not stand-alone PCs or laptops. All study personnel who work with subject identifiers and contact information will have undergone all required human subjects training. They will work with the data in password protected files and once enrollment and follow up are complete, all identifying information will be removed. Personally identifiable information will NOT be included in the analytic database.

Potential subjects will be approached, in clinics, by highly trained research staff members who understand the importance of subject privacy. In most cases, the initial encounter with patients

will take place in private exam rooms or infusion suites. Potential subjects may be approached in waiting areas, but it will be done in a way that is sensitive to maintaining privacy.

Follow-up phone calls will be conducted by trained research staff who will be calling, primarily, from their offices in Blockley Hall. Efforts will be made to ensure that phone calls will not be overheard by anyone who is not directly involved with the research. In the event that research staff member needs to leave a voicemail message for a subject, they will do so in a way that maintains subject privacy.

9. Data and Safety Monitoring

9.1 Monitoring Plan

The data and safety monitoring plan will have 3 parts. First, the BECC will implement methods of validating entered data, as they have done for numerous trials before, thereby ensuring the quality of our data. Second, the PI will be directly responsible for identifying and reporting all serious adverse events, protocol deviations/violations, and unanticipated events to the IRBs and funding agency promptly, as appropriate. He will also report all adverse events, accrual rates, retention rates, mortality/survival data and all other logistical issues to the DSMB at least biannually (and more frequently as requested or needed). Third, we have convened a DSMB that will be responsible for monitoring the trial and making decisions about the termination of individual study arms or the study itself.

The DSMB will consist of individuals with considerable expertise in human subjects research, vulnerable populations, bioethics, clinical trials, decision making, palliative care, and biostatistics. The PI (Dr. Halpern), the project manager (Elizabeth Cooney), and the lead statistician (Dr. Troxel), will participate in all DSMB meetings as non-voting members. The PI, assisted by the project manager, will be responsible for maintaining communication between the DSMB and the individual project staff.

The DSMB will perform several duties. First, they will review and approve the research protocol and plans for data and safety monitoring. Second, they will evaluate the progress of the trial. This will include assessment of data quality, participant recruitment, accrual and retention, participant risk versus benefit, and study outcomes. This assessment will be performed at meetings every 6 months during the study and more frequently if needed. They will be paying particularly close attention to patient survival as well as selections made on advance directive forms. Third, they will make recommendations to ensure that all of the issues above are appropriately addressed. Dr. Halpern, as the study PI, will be responsible for responding to all recommendations of the DSMB and submitting DSMB reports to the Penn and Pitt IRBs.

9.2 Data Safety Monitoring Board Members

The DSMB has been constituted and includes the following members:

1. David Wendler, PhD: expertise in research with vulnerable populations and research ethics, including the role of debriefing in RCTs.
2. Vicki Jackson, MD, MPH: expertise in palliative care for dying patients, and physician-patient-family communication regarding end-of-life decisions.
3. Manisha Desai, PhD: expertise in statistical methods for the analysis of clinical trials, including the implementation of stopping rules.

The DSMB will also be responsible for reviewing the provided data at the 6 month and 1 year interim analyses, determining the scientific validity and safety to determine whether the study should be continued, and will advise the PI regarding whether to bring the project to a close. The project manager, Elizabeth Cooney, and staff analyst, Dr. Nicole Gabler, will assist Drs. Halpern and Troxel in providing the DSMB with any additional information on request.

10. Human Subjects Protection

10.1 Risk / Benefit Assessment

This study presents no more than minimal risk. Many precautions will be taken to protect subjects against the most likely risk which is breach of confidentiality. In addition, the ADs are not legally binding and therefore are unlikely to erect barriers to patients receiving desired care. Instead, the ADs may merely help them avoid unwanted treatments. As a result, the potential benefits to individual subjects in terms of learning about ADs and to society from learning about a scalable intervention to improve the uptake, patient-centeredness, and effectiveness of advance directives far exceed the potential risk.

The potential risks to human subjects in this research include (1) risks of breach of confidentiality of personal health information (PHI), (2) risks of emotional distress brought on by being asked to contemplate end-of-life care, and (3) risks that the interventions could have untoward impacts on patients or their family members. Potential untoward impacts include unfavorable changes in quality of life, duration of life, satisfaction with end-of-life care planning, surrogate perceptions of the quality of dying and death, surrogate bereavement and psychiatric disturbance following deaths of loved ones, or altering (increasing or decreasing) utilization of interventions at the end of life in ways that patients would not prefer. Of note, we anticipate favorable – or at worst neutral – impacts on each of these outcomes, but are

designing our study to detect and respond quickly to unforeseen negative impacts in any of these domains.

Participants in this study may benefit directly from the opportunities to discuss and clarify their end-of-life care preferences with experienced personnel who can facilitate inclusion of these preferences into their future clinical care. Participants also may benefit from the knowledge that their surrogates have clear direction on their wishes and thus, may experience fewer burdens with difficult decision-making, perhaps alleviating subsequent stress or depression. However, participants will be instructed that this is research, and like all research, it is being conducted with the primary goal of producing generalizable knowledge. Thus, the primary benefits to be gained are those related to the knowledge to be gained.

The knowledge to be gained in this study may be of considerable importance. Given the widespread dissatisfaction with the quality of end-of-life care in the U.S., this randomized trial a readily scalable intervention to improve the uptake, patient-centeredness, and effectiveness of advance directives, which stems from a novel and innovative conceptual framework, holds great promise for improving public health. The simple and inexpensive methods to be tested may go a long way towards narrowing the gap between the care patients prefer near the end of life and the care they actually receive.

10.2 Protective Measures

The first safeguard for protection of human subjects includes an experienced and well-trained study team. Dr. Scott Halpern (PI) is the Principal investigator. He has substantial experience conducting RCTs of behavioral economic interventions to modify health-related behaviors, in the ethics of applying behavioral economics to health decisions, and in the design, ethics, and recruitment barriers of RCTs. As Principal Investigator for the proposed trial, Dr. Halpern will be primarily responsible for the completion of all aspects of this RCT including study design, underlying data infrastructure, compliance with IRB requests and requirements, participant recruitment, data collection and management, data analysis, adherence to all policies and procedures for clinical research.

Collaborating with Dr. Halpern as co-investigators and overseeing recruitment at the University of Pittsburgh are Drs. Cindy Bryce and Doug White. Dr. Bryce is a health services researcher who has spent considerable time investigating the use of decision science to improve medical decision-making in the context of critical illness. In addition to overseeing the implementation of this study at Pitt, she brings her expertise as an investigator in preference-based assessment of quality-of-life, cost effectiveness analysis, and behavioral decision theory for understanding patient and surrogate decision making. Dr. White directs the University of Pittsburgh Program

on Ethics and Decision-Making in Critical Illness, which encompasses both empirical research on, and normative ethical analysis of decision-making for, patients with life-threatening illness. He will work with Dr. Bryce in coordinating the logistics and oversight of the study at Pitt and will assist Dr. Halpern's team at Penn in interpreting results and preparing manuscripts related to his area of particular expertise – surrogate decision-making.

All study team members have completed training in HIPAA regulations and human subjects research.

The debriefing process is an important element of human subjects protection. It will ensure that patients (1) understand their selections on their AD forms; (2) do not simply go with the default options because they failed to recognize that a choice was to be made or that a default was being used; (3) have multiple opportunities to withdraw their participation or data; and (4) are actively engaged in the research and comfortable with the research process.

Additional layers of protection for human subjects include the robust informed consent process (section 7.3), exceptional data security (sections 9.1, 9.2 & 9.3), and the empowered Data Safety and Monitoring Board (sections 10.1 & 10.2), all described in detail in this protocol.

This original protocol was finalized on January 9, 2014.

II. Final Protocol

1. Abstract

Although most seriously ill Americans wish to avoid burdensome therapies near life's end, aggressive care is provided unless or until patients or their family members actively request that it is stopped. Advance directives (ADs) hold great promise for combating this societal default of aggressive end-of-life care, but to date this promise has been largely unrealized. This study will test the premise that ADs can better align the end-of-life care patients receive with the care they want if the ADs are restructured such that comfort-oriented care is provided as the default, rather than forcing patients to make emotionally and existentially challenging choices to receive it. In this study, we will determine whether this simple and readily scalable intervention can improve patients' quality of life and reduce resource utilization without reducing the number of days that patients are alive and living outside of an acute-care hospital.

2. Background and Significance

Most Americans wish to die at home and to avoid aggressive care and life support when terminally ill. Yet the opposite commonly happens: one in five Americans dies in or shortly following a stay in an intensive care unit (ICU), roughly half of U.S. deaths occur in a hospital one third of elderly patients undergo an inpatient surgical procedure during their last year of life, one half of elderly Americans visit emergency departments in the last month of life, and more than one quarter of Medicare dollars are spent on patients in their final year. Perhaps even more concerning are recent observations that aggressive treatment of patients with serious illnesses is associated with reduced quality and perhaps quantity of life near its end. When such care culminates in ICU-based deaths, it also produces long-lasting pathological bereavement among family members contravening most patients' strong desires not to burden their loved ones.

Despite past failures, written advance directives (ADs) hold great promise. A recent study highlights a key reason for the discrepancy between the care we want and the care we receive near life's end: critical healthcare decisions must be made for 43% of older Americans near the times of their deaths, but 70% of these patients cannot participate in making these decisions. The cumulative result – that 30% of older Americans cannot choose their care when such choices are needed – highlights the potential benefits of improving the quality of advance care planning, including written advance directives (ADs).

ADs include living wills, in which patients can choose to receive or avoid life-sustaining therapies if they lose capacity to make such decisions, and designation of a durable power of attorney for healthcare to serve as the patient's decision-maker in similar circumstances. Many experts have bemoaned the shortcomings of ADs, particularly for the living will component. Such concerns have spawned a broader focus on advance care planning that seeks to prepare patients and family members for difficult decisions. Sound in principle, this approach is difficult in process. For the right patient, surrounded by the right family, and cared for by the right clinicians, such coordinated communication may prove optimal. But this approach may be difficult to implement across diverse populations with differential access to longitudinal care.

By contrast, fixing the problems with ADs may yield more scalable ways to improve end-of-life care for all Americans. Recent evidence provides substantial motivation to try. Observational studies in the United States show that elderly patients who complete ADs less commonly die in a hospital, more often receive care consistent with their preferences, and receive less costly care.

Despite these recent studies showing the promise of ADs, none provide sufficient evidence that completing ADs, or certain types of ADs, will cause changes in clinical, economic, or patient-centered outcomes. Studies noting improved patient-centered and economic outcomes among

patients completing ADs were all observational in nature, preventing conclusions about whether AD completion caused these benefits or was a marker for people likely to attain them anyway. Thus, given federal policies promoting AD completion, and evidence that completion rates are increasing in the U.S., an RCT is desperately needed to determine how best to design ADs to improve patient outcomes without increasing resource utilization.[20,34]

3. Objectives

3.1 Overall objectives

This study will test the premise that ADs can better align the end-of-life care patients receive with the care they want if the ADs are restructured such that comfort-oriented care is provided as the default, rather than forcing patients to make emotionally and existentially challenging choices to receive it.

3.2 Primary outcome variable(s)

The primary outcome is "Hospital-Free Days" (HFDs), a measure that PI Halpern has been developing in collaboration with Dr. Jeffrey Silber at Penn's Center for Outcomes Research. As the name describes, HFDs represent the number of days alive and not in an acute care facility. Although this is a simple concept and provides an outcome measure of obvious importance to patients, the use of HFDs as a primary outcome in an RCT is highly innovative. To bolster confidence in the results, we will evaluate two key variations on the theme. First, we will explore "Healthcare Facility-Free Days," which represents the number of days alive where a patient is neither in an acute care facility, a chronic care facility, nor a nursing home. We will also evaluate HFDs within a defined period of follow-up – 6 months in this case. This is analogous to the established outcome of ventilator-free days used commonly in RCTs among ICU patients.[35]

3.3 Secondary outcome variable(s)

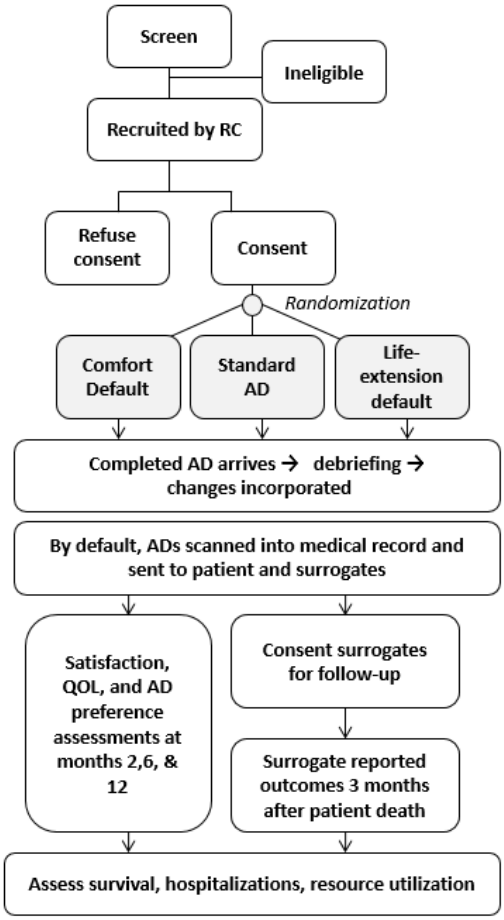
1. Hospital and ICU admissions: The numbers of admissions will be analyzed as count data. From the dates of hospital and ICU admissions, we will calculate the proportion of each patient's total survival time during study follow-up that was spent in the hospital or ICU.
2. Costs of care: We will combine all costs of inpatient and outpatient hospice, hospital stays, and life-sustaining procedures. The perspective will be that of all potential payers. Costs will be inflated to the date on which analyses are performed using the U.S. gross domestic product deflator.
3. Choices to receive 4 potentially life-sustaining interventions, and the concordance of these choices with whether the interventions were actually received: The outcomes databases we will use contain codes for each of the 4 interventions, enabling us to determine which patients received each. Thus, we will be able to reliably evaluate the proportions of patients who received unwanted interventions. Because we cannot

792 determine the denominator of patients with indications for these interventions, we will
793 not evaluate the proportions of patients who went without desired services.
794 4. Choices regarding post-hospitalization care, and the concordance of these choices with
795 the care actually received.
796 5. Decision conflict and satisfaction: The Decision Conflict Scale is a well-validated
797 instrument used to assess patients' certainty in making healthcare decisions.
798 Satisfaction will also be measured more specifically with the CANHELP instrument's
799 global satisfaction with end-of-life care question.
800 6. Quality of life, using the McGill Quality of Life (MQOL) instrument. The MQOL is a well-
801 validated and widely used scale designed specifically for patients with serious illnesses.
802 The MQOL can be completed by family members on behalf of patients who have lost the
803 capacity to complete it themselves. Thus, we will have surrogates complete the MQOL
804 for incapacitated patients to minimize missing data.
805 7. Surrogates' Perception of the quality of death and dying: We will assess this outcome
806 with surrogates of deceased patients using Prigerson's Quality of Death measures.
807 8. Bereavement outcomes: The risk of post-traumatic stress disorder in surrogates among
808 deceased patients will be assessed using the Impact of Events Scale (IES). The IES is a
809 valid and reliable scale that has been used frequently to assess PTSD risk among family
810 members of critically ill patients.
811 9. Healthcare system distrust: The Healthcare System Distrust Scale will be used to assess
812 two primary domains of distrust in healthcare (values and competence). This scale will
813 be used to explore if distrust of the healthcare system has a mediating effect on
814 surrogate outcomes, such as their perceptions of quality of death and dying and post-
815 traumatic stress.
816

817 ## *4. Study Design*

818 ### 4.1 Schema

819 This is a prospective, randomized, controlled trial.



820

821 **4.2 Duration**

822 The study period extended to 34 months. Subjects were accrued over a period of 27 months
823 starting in February 2014. The total time it will take for the research coordinator to explain the
824 study, obtain consent and for a subject to complete the advance directive will, conservatively,
825 take no more than two hours. The debriefing discussion and follow up interviews will take
826 approximately 15 – 25 minutes each. The total time spent on research activities for patients
827 should be no more than 4 hours.

828 ## *5. Subject recruitment*

829

830 We will recruit 270 patients with severe respiratory, oncological, neuromuscular, or
831 cardiovascular diseases and limited life expectancy from the Perelman Center for Advance Care
832 Medicine, Penn Presbyterian Medical Center, Pennsylvania Hospital, and the University of
833 Pittsburgh Medical Center. Each week the research coordinators will screen the electronic
834 medical records of patients scheduled for routine visits to determine their study eligibility using
835 the eligibility criteria outlined above.

Once eligible patients have been identified, research coordinators will email eligible patients' providers to 1) alert them to their patients' eligibility for participation 2) inform them their patients will be recruited for enrollment 3) provide them an opportunity to decline or defer any given patient's enrollment by responding to the email. Research coordinators will approach potential study participants while they are in the waiting areas, chemotherapy infusion areas, or in exam rooms waiting to see their doctor on the day of their visit.

5.1 Accrual

During our pilot study we were able to recruit approximately six patients per month with one full-time research coordinator. We anticipate that with the equivalent of 3.5 full-time research coordinators and an additional site (University of Pittsburgh), we will be able to recruit approximately 18 patients per month.

5.2 Key inclusion criteria

The eligibility criteria, all of which must be met, are:

1. Age 18 or older
2. Speaks and reads fluent English
3. Has seen current physician at least once prior to current visit
4. Resident of Pennsylvania or New Jersey
5. One or more of the following diagnoses:

- Amyotrophic lateral sclerosis
- Stage IIIB or IV non-small cell lung cancer or cholangiocarcinoma
- Stage IV colorectal, esophageal, gastric (including GIST), pancreatic, prostate, uterine, cervical, ovarian or urothelial cancer; paraganglioma, or pheochromocytoma
- Stage C or D hepatocellular carcinoma
- Stage IV renal cell carcinoma
- Stage IV or V chronic kidney disease
- Mesothelioma or any malignancy metastatic to the pleura
- Other incurable interstitial lung diseases with at least severe restriction on most recent pulmonary function tests or eligible for long-term oxygen therapy
- Chronic obstructive pulmonary disease with at least severe airflow obstruction on most recent spirometry or eligible for long-term oxygen therapy
- Congestive heart failure with NYHA Class IV status or Class III plus 1 heart failure-related hospitalization in the past 12 months or ACC stage D or C classification with 1 heart failure-related hospitalization in the past 12 months

- Stage IV breast cancer except patients whose only metastases are to the bones or who are receiving endocrine therapy without receiving concomitant traditional chemotherapy

5.3 Key exclusion criteria

Patients will be excluded if they are currently listed for or being considered for solid organ transplant and if they have a previously signed advance directive or living will. Cognitively impaired patients will be excluded from the study to avoid the necessity of proxy consent.

5.4 Subject Remuneration

Patients will be compensated with a $20 Amazon.com gift card following completion of the debriefing session. In order to enhance study retention and participation in follow-up assessments, a $20 Amazon.com gift card will also be given to subjects at the completion of the two, six, and twelve month follow-ups. Surrogates will also be compensated with a $20 Amazon.com gift card after they consent to participate.

6. Randomization

6.1 Groups

Subjects enrolled in this RCT will be randomized into three groups. Depending on which group they've been assigned, subjects will be given one of three AD forms. The three AD forms have been created with different default treatment options. Form 1 (life-extension default) will state that 4 specific life-extending interventions (cardiopulmonary resuscitation, mechanical ventilation, hemodialysis, and feeding tube insertion) will be provided unless patients specifically opt-out from such selections. Form 2 (comfort default) will state that the 4 specific life-extending interventions will not routinely be provided unless patients elect to receive such measures. Finally, Form 3 (standard advance directive) will use the standard approach of requiring patients to actively choose whether or not they wish to receive each intervention, as they would if completing an AD outside of a research setting. In this case, if they do not make a selection, decision making would default to their surrogates as in usual practice.

Because patients may focus on an overall plan of care rather than the receipt of specific interventions, all AD forms will also include a general question regarding treatment priorities. The response to this question, is modeled on one used in a Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT) study. The question acknowledges that while, in general, most people wish to both live as long as possible and avoid

pain and suffering, in some situations, choosing between these two goals may be necessary. It then asks patients, if they are in a situation where such a choice is needed, whether they prefer a plan of care that focuses on extending life as much as possible even if it means having more pain and suffering, or a plan of care that focuses on relieving pain and suffering even if that means not living as long. The default framing of this general question will be in accord with that used for the specific interventions in each AD form, and all patients will be able to select a "no" option in response to this question.

Finally, we will include a specific question about the care patients wish to receive upon discharge from the hospital, defaulting to hospice-based care (in the comfort-default group), long-term care (in the life-extension-default group), or no option pre-selected. In the standard AD group, although no options will be pre-selected, we will randomly assign whether the comfort-oriented option or the life-extending-oriented option is presented first so as to mitigate ordering effects. In all cases, the option of not deciding will be presented last.

6.2 Assignment

Eligible patients will be approached about participation by the research coordinators in the outpatient clinics at the Perelman Center for Advanced Medicine, Pennsylvania Hospital, and Presbyterian Hospital. Consenting subjects will be randomized with a 33.3% probability to each trial arm (life extension default, comfort default, standard AD) using electronic procedures monitored by the Data Management Unit within the Biostatistics Analysis Center. We will stratify the randomization by recruiter/research coordinator, and will use variable block sizes of 3 and 6 patients to promote balance of follow-up duration among the 3 trial arms.

Each research coordinator will go to his or her clinics each day with a sealed envelope in which there is a pre-determined sequence of the 3 trial packets. The research coordinator will become unblinded to the patient's allocation at the time of consent, but with variable block sizes, can never predict with certainty what the next packet will be.

7. Study Procedures

7.1 Screening for Eligibility

The research coordinators will screen electronic medical records of patients visiting pulmonary, renal, heart failure, movement disorder, and oncology clinics at the Perelman Center for Advance Care Medicine, Penn Presbyterian Medical Center, Pennsylvania Hospital and the University of Pittsburgh Medical Center for eligibility. Patient's eligibility status will be entered

into the eligibility database. We will record ICD9 and ICD 10 codes, staging information, relevant provider name, clinic location, and upcoming appointments for eligible patients.

7.2 Recruitment

Eligible patients will be approached by a research coordinator in the clinics who will seek patients' consent to participate in a study comparing different types of ADs. Of note, while some providers may be more proactive than others in engaging their patients in conversations about advance care planning, it is generally not standard-of-care that patients are approached about completing ADs. The research coordinator will specify that the ADs in this study are intended to be real ADs and that they will be included in patients' outpatient medical records, but that, as with all ADs, patients retain the right to change their selections at later dates. The research coordinator will also specify that, like all real ADs, they are most useful if copies are shared with their loved ones and physicians.

7.3 Informed Consent

Following discussion of the study, research coordinators will obtain written consent from patients. The consent forms will contain HIPAA statements of authorization of release of medical records, thus facilitating our collection of data from medical and billing records during the study. The consent includes clear explanations that different types of ADs will be assigned by chance, but that patients in all groups may select or decline any intervention or treatment goal, and may revise their choices at any time. The research coordinators will explain who will be enrolled, how many patients are being targeted for enrollment, the specific components of patient follow-up, patients' rights to withdraw from the study at any time and for any reason, and what the outcomes of interest are (e.g., utilization of healthcare services, AD selections).

7.4 Enrollment

After patient consent is obtained, the research coordinator will ask subjects to complete the demographics survey, indicate whether they prefer to be contacted by phone or email, and walk subjects through the process of filling out the AD. Along with their AD forms, consenting subjects will be given a copy of their consent form, an informational brochure about advance care planning, contact information for research study staff, instructions for mailing back their completed AD forms, the decision conflict scale, and a stamped and addressed envelope. The DCS will be sent home with consenting patients to complete and return, along with an instruction sheet explaining to patients that they should complete their AD first, followed by the DCS.

Subject IDs will be assigned at the point of consent. Subject ID numbers, demographic information and group assignments will be entered into the analytic database. Subject contact information, including social security number, will be entered into a subject tracking database.

If completed ADs are not returned within 10 days, staff will call or email patients weekly to remind them to return their ADs, to schedule special clinic visits for AD completion if patients desire, and to answer any questions. If research staff members are unable to reach patients over the phone or email after three attempts, a letter will be sent to patients to remind them to return their ADs and encourage them to contact research staff if they have any questions or difficulties. If we are still unable to reach patients, they will be approached by the research coordinator at their next clinic visit.

7.5 Subject Debriefing

After patients complete their assigned AD, there will be a structured debriefing session conducted over the phone, regardless of patients' preferred contact method, by a research team member in which a standardized explanation of all three ADs will be given. If the patient is unable to be reached by phone, they will be approached at their next clinic visit to complete the debriefing in person. This debriefing will be held to alert patients to exactly how the three ADs used in the study differ. As in the pilot study, patients will not be alerted to the different default framings up front because patients in clinical settings (and indeed in this study) are only asked to complete a single AD. Explaining non-relevant ADs prior to completion of the relevant one could influence decisions in ways that would not reflect actual clinical settings, thereby biasing the results. However, because this is a research study and AD assignment is at random, it is appropriate to debrief patients afterwards to grant them such broader information. Once patients are fully informed about the variations in the ADs used in the study, they will be asked if they wish to change any of their AD selections prior to finalizing the documents as a part of the medical record. Patients who choose to make changes to their AD's during the debriefing can choose to have their original AD sent back to them along with a blank AD to complete and return, or the research team will make the changes directly on the AD forms and send a letter back to the patient indicating the changes have been made and instructing the patients to call the research team if they do not approve of the changes and/or would like additional changes. If we do not hear from the patients within 10 days, the study team will consider the revised AD complete. Patient ADs will not be considered "complete" until the debriefing session has taken place. After the debriefing call, patients' AD selections will be entered into the analytic database.

During the debriefing call, we will tell subjects that we will scan their AD forms into their medical records for them, unless they do not desire this (it is optional, not a requirement of the study). Similarly, we will also inform subjects that a copy of their completed AD will be sent back to them along with a $20 Amazon.com gift card as compensation for their time, and that, if they wish, a copy will be sent to their appointed healthcare agent/surrogate. Completed ADs will be sent to patients and surrogates along with letters explaining that ADs can be changed at any time and they can contact the research team with questions.

Research coordinators will help facilitate the scanning of patients' completed ADs into their medical records, for patients who wish to do so. Completed ADs will be given to clinic administrative staff along with a step by step instruction sheet explaining that we are asking that the AD be scanned into the patient's medical record and where, in the medical record, the ADs should be placed. Two weeks after the completed ADs have been delivered to clinic staff, the research coordinators will review the medical record in Epic to confirm the successful upload of the documents. In addition, a confirmation email will be sent to patients' physicians informing them that their patients have active ADs as part of their medical record.

7.6 Subject Follow-up

Two, six, and twelve and months after AD completion, subjects will be contacted for participation in follow-up interviews. The follow-up interviews will take place over the phone with a research associate blinded to the subject's study arm, or online through REDCap, depending on the patient's preferred method of communication. The research associates will attempt to contact patients up to two times using their preferred method of communication. If the two first attempts are unsuccessful, we will attempt to contact the patient using the alternate method. If we are unable to reach the patient following the third attempt, we will scan EPIC for the patient's next in-clinic appointment, during which a research coordinator, blinded to the patient's study arm will attempt to complete the follow-up interview in person. In advance of this in-person meeting, we will send a letter to the patient notifying them of our efforts to reach them, and indicate a member of our study team would like to meet with them during their next clinic visit. If patients are unavailable to participate in follow-up calls because they are deceased or otherwise incapacitated, we will interview their surrogates.

Prior to contacting patients for follow-up assessments, we will screen their EPIC medical records to check patient mortality. EPIC will capture the vast majority of deaths within 2-3 weeks, as mortality data are entered by clinic staff in regular contact with seriously ill patients.

7.7 Assessment of Health Outcomes

We will assess hospitalizations, ICU admissions, costs of inpatient care, and utilization of life-sustaining therapies by querying state-run databases that capture all admissions and inpatient procedures in Pennsylvania and New Jersey. The Pennsylvania Health Care Cost Containment Council (PHC4) is an independent state agency that maintains a database of inpatient hospital discharge and outpatient procedure records from all hospitals and ambulatory surgery centers in Pennsylvania. These data include specific treatment information including costs. As roughly one-third of Penn's outpatient population resides in New Jersey, we will obtain comparable data from the New Jersey Discharge Data Collection System (NJDDCS) managed by the New Jersey Department of Health and Senior Services within their Department of Health Care Quality and Assessment (HCQA). We will establish data use agreements with both of these entities and be subject to IRB approval by HCQA. Linkages with both PHC4 and NJDDCS will be performed by the respective database administrators after we provide lists of included social security numbers and subject IDs. PHC4 and NJDHSS will send our team a report in which patients are identified by subject ID only. Identical processes have been used seamlessly and with high fidelity by many Penn investigators.

8. Data Management

8.1 Data Confidentiality

Only authorized project personnel will have access to the data. All study data will be stored behind firewalls on Center for Clinical Epidemiology and Biostatistics (CCEB) servers; none will be stored on stand-alone PCs or laptops. All study personnel who work with these data will have undergone required human subjects training. To ensure that participant confidentiality is preserved, individual identifiers (such as social security number) will only be used to link patient records (e.g., linking subject database to PHC4 data). Once linkages between databases have been achieved, all linkage-identifiers will be dropped from all datasets. Throughout the study duration, we will maintain one master list that will link study identification numbers to patient identifiers. This list will be maintained by the principal investigator in a locked file drawer in his locked private office to ensure file security. This file will be made available to other research staff on a need-to-know basis only, and, in that case, only temporarily. The study ID will be used exclusively in all analytical files.

We will implement multiple, redundant protective measures to guarantee the privacy and security of the participant data. All data for this project will be stored on the secure/firewalled servers of the CCEB in data files that will be protected by multiple password layers. These data

servers are maintained in a guarded facility behind several locked doors, with very limited physical access rights. They are also cyber-protected by extensive firewalls and multiple layers of communication encryption. Electronic access rights are carefully controlled by Penn system managers. We will use highly secure methods of data encryption for all transactions involving participants' financial information using a level of security comparable to what is used in commercial financial transactions. This multi-layer system of data security, identical to the system protecting the University of Pennsylvania Health System's medical records, greatly minimizes privacy risks.

8.2 Subject Confidentiality

Steps will be taken to ensure that all information will be kept confidential and secure. Unique patient identifiers numbers will be assigned to each subject locally and kept in a secure encrypted file. Records with patient social security numbers will be maintained, used, and destroyed in a way that is consistent with Penn policy. All paper records will be kept in locked files; all computers will be password protected and kept in locked rooms; all databases will be password protected and maintained on encrypted hard-drives behind the CCEB firewall. All study data will be stored behind firewalls on Center for Clinical Epidemiology and Biostatistics (CCEB) servers; none will be stored on stand-alone PCs or laptops. All data will be destroyed after 7 years.

8.3 Subject Privacy

Individual-level data for participants will be kept confidential and will only be stored on highly secure servers available for patient-level data. Only authorized project personnel will have access to the data and the data will be stored on servers only and not stand-alone PCs or laptops. All study personnel who work with subject identifiers and contact information will have undergone all required human subjects training. They will work with the data in password protected files and once enrollment and follow up are complete, all identifying information will be removed. Personally identifiable information will NOT be included in the analytic database.

Potential subjects will be approached, in clinics, by highly trained research staff members who understand the importance of subject privacy. In most cases, the initial encounter with patients will take place in private exam rooms or infusion suites. Potential subjects may be approached in waiting areas, but it will be done in a way that is sensitive to maintaining privacy.

Follow-up phone calls will be conducted by trained research staff who will be calling, primarily, from their offices in Blockley Hall. Efforts will be made to ensure that phone calls will not be overheard by anyone who is not directly involved with the research. In the event that research

staff member needs to leave a voicemail message for a subject, they will do so in a way that maintains subject privacy.

9. Data and Safety Monitoring

9.1 Monitoring Plan

The data and safety monitoring plan will have 3 parts. First, the BECC will implement methods of validating entered data, as they have done for numerous trials before, thereby ensuring the quality of our data. Second, the PI will be directly responsible for identifying and reporting all serious adverse events, protocol deviations/violations, and unanticipated events to the IRBs and funding agency promptly, as appropriate. He will also report all adverse events, accrual rates, retention rates, mortality/survival data and all other logistical issues to the DSMB at least biannually (and more frequently as requested or needed). Third, we have convened a DSMB that will be responsible for monitoring the trial and making decisions about the termination of individual study arms or the study itself.

The DSMB will consist of individuals with considerable expertise in human subjects research, vulnerable populations, bioethics, clinical trials, decision making, palliative care, and biostatistics. The PI (Dr. Halpern), the project manager (Elizabeth Cooney), and the lead statistician (Dr. Troxel), will participate in all DSMB meetings as non-voting members. The PI, assisted by the project manager, will be responsible for maintaining communication between the DSMB and the individual project staff.

The DSMB will perform several duties. First, they will review and approve the research protocol and plans for data and safety monitoring. Second, they will evaluate the progress of the trial. This will include assessment of data quality, participant recruitment, accrual and retention, participant risk versus benefit, and study outcomes. This assessment will be performed at meetings every 6 months during the study and more frequently if needed. They will be paying particularly close attention to patient survival as well as selections made on advance directive forms. Third, they will make recommendations to ensure that all of the issues above are appropriately addressed. Dr. Halpern, as the study PI, will be responsible for responding to all recommendations of the DSMB and submitting DSMB reports to the Penn and Pitt IRBs.

9.2 Data Safety Monitoring Board Members

The DSMB has been constituted and includes the following members:

1. David Wendler, PhD: expertise in research with vulnerable populations and research ethics, including the role of debriefing in RCTs.

2. Vicki Jackson, MD, MPH: expertise in palliative care for dying patients, and physician-patient-family communication regarding end-of-life decisions.
3. Manisha Desai, PhD: expertise in statistical methods for the analysis of clinical trials, including the implementation of stopping rules.

The DSMB will also be responsible for reviewing the provided data at the 6 month and 1 year interim analyses, determining the scientific validity and safety to determine whether the study should be continued, and will advise the PI regarding whether to bring the project to a close. The project manager, Elizabeth Cooney, and the staff analyst will assist Drs. Halpern and Troxel in providing the DSMB with any additional information on request.

10. Human Subjects Protection

10.1 Risk / Benefit Assessment

This study presents no more than minimal risk. Many precautions will be taken to protect subjects against the most likely risk which is breach of confidentiality. In addition, the ADs are not legally binding and therefore are unlikely to erect barriers to patients receiving desired care. Instead, the ADs may merely help them avoid unwanted treatments. As a result, the potential benefits to individual subjects in terms of learning about ADs and to society from learning about a scalable intervention to improve the uptake, patient-centeredness, and effectiveness of advance directives far exceed the potential risk.

The potential risks to human subjects in this research include (1) risks of breach of confidentiality of personal health information (PHI), (2) risks of emotional distress brought on by being asked to contemplate end-of-life care, and (3) risks that the interventions could have untoward impacts on patients or their family members. Potential untoward impacts include unfavorable changes in quality of life, duration of life, satisfaction with end-of-life care planning, surrogate perceptions of the quality of dying and death, surrogate bereavement and psychiatric disturbance following deaths of loved ones, or altering (increasing or decreasing) utilization of interventions at the end of life in ways that patients would not prefer. Of note, we anticipate favorable – or at worst neutral – impacts on each of these outcomes, but are designing our study to detect and respond quickly to unforeseen negative impacts in any of these domains.

Participants in this study may benefit directly from the opportunities to discuss and clarify their end-of-life care preferences with experienced personnel who can facilitate inclusion of these preferences into their future clinical care. Participants also may benefit from the knowledge that their surrogates have clear direction on their wishes and thus, may experience fewer

burdens with difficult decision-making, perhaps alleviating subsequent stress or depression. However, participants will be instructed that this is research, and like all research, it is being conducted with the primary goal of producing generalizable knowledge. Thus, the primary benefits to be gained are those related to the knowledge to be gained.

The knowledge to be gained in this study may be of considerable importance. Given the widespread dissatisfaction with the quality of end-of-life care in the U.S., this randomized trial a readily scalable intervention to improve the uptake, patient-centeredness, and effectiveness of advance directives, which stems from a novel and innovative conceptual framework, holds great promise for improving public health. The simple and inexpensive methods to be tested may go a long way towards narrowing the gap between the care patients prefer near the end of life and the care they actually receive.

10.2 Protective Measures

The first safeguard for protection of human subjects includes an experienced and well-trained study team. Dr. Scott Halpern (PI) is the Principal investigator. He has substantial experience conducting RCTs of behavioral economic interventions to modify health-related behaviors, in the ethics of applying behavioral economics to health decisions, and in the design, ethics, and recruitment barriers of RCTs. As Principal Investigator for the proposed trial, Dr. Halpern will be primarily responsible for the completion of all aspects of this RCT including study design, underlying data infrastructure, compliance with IRB requests and requirements, participant recruitment, data collection and management, data analysis, adherence to all policies and procedures for clinical research.

Collaborating with Dr. Halpern as co-investigators and overseeing recruitment at the University of Pittsburgh are Drs. Cindy Bryce and Doug White. Dr. Bryce is a health services researcher who has spent considerable time investigating the use of decision science to improve medical decision-making in the context of critical illness. In addition to overseeing the implementation of this study at Pitt, she brings her expertise as an investigator in preference-based assessment of quality-of-life, cost effectiveness analysis, and behavioral decision theory for understanding patient and surrogate decision making. Dr. White directs the University of Pittsburgh Program on Ethics and Decision-Making in Critical Illness, which encompasses both empirical research on, and normative ethical analysis of decision-making for, patients with life-threatening illness. He will work with Dr. Bryce in coordinating the logistics and oversight of the study at Pitt and will assist Dr. Halpern's team at Penn in interpreting results and preparing manuscripts related to his area of particular expertise – surrogate decision-making.

All study team members have completed training in HIPAA regulations and human subjects research.

The debriefing process is an important element of human subjects protection. It will ensure that patients (1) understand their selections on their AD forms; (2) do not simply go with the default options because they failed to recognize that a choice was to be made or that a default was being used; (3) have multiple opportunities to withdraw their participation or data; and (4) are actively engaged in the research and comfortable with the research process.

Additional layers of protection for human subjects include the robust informed consent process (section 7.3), exceptional data security (sections 8.1, 8.2 & 8.3), and the empowered Data Safety and Monitoring Board (sections 9.1 & 9.2), all described in detail in this protocol.

III. Summary of Changes

The following changes to the protocol were made after the original protocol had been finalized on January 09, 2014 and patient enrollment had begun on February 6, 2014.

1. The study's enrollment period was originally planned for 18 months. Due to slower than expected accrual, regulatory delays, and turnover of research staff the enrollment period lasted 27 months (February 2014 – April 2016)

2. We abandoned efforts to enroll patients who declined participation into the registry after roughly 20% of the sample had been enrolled due to low interest in the registry from patients. The purpose of the registry was to enabled outcomes to be collected among a broader group of patients who did not complete ADs, thereby enabling complier average treatment effect analyses of patients' quality of life. This goal would only be enabled with nearly complete accrual of non-enrolled patients into the registry. This proved infeasible early on.

3. Submitted to the IRB 03.31.14 – For ease of use and risk management reasons we changed patient remuneration for completion of follow-up assessments to amazon.com gift cards instead of cash.

4. Submitted 04.23.14 – In order to ease the burden of a lengthy assessment for bereaved family members we eliminated the use of Prigerson's Complicated Grief Inventory and the Quality of Death and Dying Instrument. We replaced these instruments with Prigerson's Quality of Death measure.

5. Submitted 09.08.13 – We added specificity around the timing and frequency of follow-up calls to patients to encourage them to return their completed AD within 30 days. We also expanded eligibility criteria to include patients with Stage IV uterine, cervical, and ovarian cancer.

6. Submitted 04.28.15 – We added procedures to ensure and confirm AD upload in patients' medical records including 1) an instruction sheet for clinical staff indicating where ADs should go in the medical record and 2) a protocol to confirm presence of an AD in the record within 2 weeks.

7. Submitted 05.22.15 – We modified our demographics form to collect patients' email addresses and ask patients if they prefer email vs. phone call follow-up. Additionally, we modified follow-up procedures to allow for electronic survey completion of follow-up measures. We also added in-person completion of follow-up measures in outpatient clinics. We also modified the timing of

patient remuneration such that patients would receive their first $20 gift card after AD completion instead of at the point of consent.

8. Submitted 12.22.15 – Due to low response rate and resource constraints we eliminated collection of follow-up measures for patients who did not complete ADs.

9. Submitted 07.20.16 – Due to observation, in preparation for a DSMB meeting, of missing demographic data, we obtained permission from the IRB to manually search the electronic health records to improve demographic data completeness.

No further changes to the protocol were made after this ninth modification. Thus, the protocol was considered finalized after receiving IRB approval for the final modification on August 10, 2016.

IV. Original Statistical Analytic Plan

1. Analytic Methods

To assess balance among groups achieved by randomization, we will compare baseline values of all variables across arms using ANOVA and chi-square tests for continuous and binary data, respectively. We will use Poisson models to assess the number of hospital free days (HFDs) from the time of randomization. We will use logistic, linear, or quantile (1) regression, as appropriate based on outcome parameterizations and distributions, for all secondary outcomes. In all analyses, we will model the clinic from which patients are recruited as a random effect to adjust for potential clustering within clinics and to mitigate confounding by clinic (2). We will employ standard covariate-selection procedures for etiologic models to assess, and potentially adjust for, chance covariate imbalance among arms. Specifically, patient-level covariates (e.g., gender, race, diagnosis category) will be included in multivariable models based on pre-specified hypotheses or if their inclusion – singly or jointly – modifies the coefficient for the randomized exposure by ≥ 15% (3).

All analyses will be conducted using the intention-to-treat approach to avoid selection bias. Some patients who consent to participate and receive their assigned AD may not return the AD. In our pilot work we developed several interventions that successfully mitigated this possibility. However, any patients who do not return the AD will be retained in the primary analyses, and will be classified as having not specified preferences for goals of care or specific interventions.

2. Specific aims and hypothesis

(a) ADs with preselected comfort care options, compared with those defaulting to life-extension or standard ADs, will produce an increase in hospital-free days (HFDs), a measure that represents the number of days alive and not in an acute care facility.

(b) Compared with standard ADs or ADs defaulting to life-extension, ADs defaulting to comfort care will:

1. produce no change in survival
2. reduce hospital and ICU admissions
3. reduce costs of inpatient care
4. improve patients' quality of life
5. improve patients' satisfaction with end-of-life care and decision making
6. improve surrogates' perceptions of the quality of dying and death
7. decrease the incidence of symptoms of post-traumatic stress among surrogates following their loved ones' death

3. Exposure

Intervention group (standard AD, life-extension default, or comfort care default)

4. Outcomes

4.1 Primary

Primary outcome is Hospital-free days (HFD). This metric represents the number of days alive and not in an acute care facility following the date of consent. We chose the date of consent as day 0 so that all enrolled participants, including those who do not return ADs, are eligible for ITT analyses.

The choice of HFDs as the primary outcome reflects the desire to choose a measure that is patient-centered, readily measured and analyzed, and reflects a patient's holistic state rather than a specific symptom. HFDs have many attractive properties: they are continuous, enhancing power; they can be analyzed reliably and flexibly, to account for different values patients may place on avoiding hospitalization; and in nearly all cases, they are unidirectional, in the sense that nearly all patients prefer longer lives to shorter ones, and to have more of those days spent outside a hospital than within.

4.2 Secondary

Secondary outcomes include several clinical, economic and patient-reported measures including:

1. Survival – Patient deaths will be captured via medical records and verified by the Pennsylvania and New Jersey Departments of Health vital statistics
2. Hospital and ICU admission – Captured by querying state-run databases that capture all admission and inpatient procedures in Pennsylvania and New Jersey.
3. Inpatient care charges – Captured via the database detailed in (2).
4. Hospice utilization – Captured via data use agreements with Wissahickon hospice and Family Hospice and Medical Care, organizations the provide care for 80% of eligible patients at Penn and Pitt.
5. Receipt of life-sustaining therapies – Captured via the database detailed in (2).
6. Concordance between patients' expressed desires in ADs regarding four potentially life-sustaining therapies (CPR, mechanical ventilation, dialysis, and feeding tube) and care received
7. Quality of life – Measured at 2, 6, and 12 months following AD completion with the McGill quality of life (MQOL), which can be completed by family members on behalf of patients who are unable to complete it themselves. MQOL during follow-up is missing for a high number of participants. In examining the data over time, we were able to determine that MQOL does not change over time and that time to follow-up is not significantly related to MQOL values. Therefore, we will only report one MQOL score per patient (in the per protocol analysis), and this score will be the one closest to the 6 month follow-up period. Also, per-protocol patients who die are assigned a value of 0 and the remainder are imputed.
8. Satisfaction with advance care planning – Measured at 2, 6, and 12 months following AD completion with the Canadian Healthcare Evaluation Project (CANHELP) instrument's global satisfaction and end-of-life care question.
9. Satisfaction/conflict with decision-making – Measured immediately following AD completion using the validated decision conflict scale (DCS).

10. Surrogates' perceptions of the quality of dying and death – Measured using Prigerson's quality of death measures.
11. Symptoms of post-traumatic stress among surrogates following their loved ones' death – Measured using the Impact of Events Scale.

5. Analysis

We aim to answer two primary questions in this study:

(1) What is the overall effectiveness of offering people the opportunity to complete advance directives with different embedded default options?
(2) What are the specific effects of making certain choices within ADs on patient and caregiver outcomes?

The primary way we'll answer question (1) is through the modified ITT analysis; question (2) will be answered using a complier average treatment effect (CATE) analysis.

5.1 Modified ITT

The unit of analysis for the primary outcome (HFDs) will be the individual patient. mITT analyses include all patients except for (1) post-randomization ineligibles; (2) withdraws; and (3) patients who died within 30 days of randomization. The rationale for these exclusions is that none of these patients were fully eligible to complete the assigned intervention in a way that would be accessible to the investigators. Further, as expected, these losses are evenly distributed across the 3 arms (see CONSORT diagram) such that their exclusion could not affect the results. In primary analyses, only patients who return an AD and are debriefed will be counted as having returned an AD. In secondary analyses, all patients who return ADs, regardless of debriefing status, will be included. mITT analyses will be conducted using linear regression, adjusting for center, to compare the effects of assignment to complete ADs with different default options on HFDs. This approach will use data from all randomized patients and will provide the truest test of the overall effectiveness of the intervention among those randomly assigned to receive it.

5.2 CATE analysis

This analysis examines the effects of making certain choices within ADs on outcomes and accounts for tendencies to not complete ADs. The CATE analysis surmounts the selection effects inherent in per-protocol analyses, as well as the inability of ITT analyses to provide specific tests of the effects of choices made in ADs because these effects will be diluted by the fact that many randomized patients will not complete their assigned ADs. The CATE approach entails a two-stage least-squares regression in which the randomization arm is modelled as an instrumental variable (IV) in complier average treatment effect (CATE) analysis. This analysis will also be adjusted for center. Like the ITT analysis, these analyses use data on all randomized patients to estimate the effects of specifying any treatment choice in ADs regardless of group assignment, and after accounting for the possibility that AD completion rates may differ among the three arms by using the randomization arm as the IV. Thus, the estimated effect of the choices patients make is adjusted for the percentage of assigned patients who complete an AD at all, and the percentage who opt out from their assigned default option. This IV uses data on all randomized patients and then adjusts for AD completion rates, thereby attenuating the selection effects.

This analysis also requires the use of principal stratification methods to formulate the causal quantities of interest and determine the proportions of patients in each arm who would choose comfort care if they were assigned to complete each version of the AD. The analysis assumes that all patients who would choose comfort care in a standard AD would also choose it in an AD that defaults to comfort care, and that all patients who would choose comfort care in an AD that defaults to aggressive care would also choose it in a standard AD or an AD the defaults to comfort care. Coupled with the possibilities that some participants would never return an AD, and that others would return an AD but not choose comfort care regardless of group assignment, this creates five compliance classes (principal strata) of participants. These classes are:

i. Patients would not complete an AD regardless of group assignment
ii. Patients would complete an AD but not choose comfort care regardless of group assignment
iii. Patients would complete an AD and only choose comfort care if assigned to the comfort-default AD
iv. Patients would complete an AD and choose comfort care if assigned to the comfort-default AD or standard AD
v. Patients would complete an AD and choose comfort care regardless of group assignment

Each patient has three potential outcomes (see below). Only one of the potential outcomes can be observed, the outcome corresponding to the actual intervention the patient received. This is represented by a binary endpoint – whether or not patients would have a high quality of life in the future:

Y_i^A = whether patient i would have high quality of life if assigned to complete an aggressive-default AD
Y_i^S = whether patient i would have high quality of life if assigned to complete an standard AD
Y_i^C = whether patient i would have high quality of life if assigned to complete an comfort-default AD

Our approach assumes the exclusion restriction that AD assignment only influences the potential outcomes through the causal pathway of determining which type of care the patient chooses through the AD. However, this assumption is likely to hold in this case, because the randomly assigned IV – which of three versions of the AD is offered – would not influence outcomes unless it modified the probability of AD completion or the choices made in the ADs.

5.3 Secondary analyses

Per-protocol analysis: The per-protocol analysis will compare patients who choose comfort care on their ADs with patients who do not choose comfort care. Again, the main per-protocol analysis will only include patients who return an AD and are debriefed, but an additional secondary analysis will be performed that includes patients who return ADs and are not debriefed. This analysis will assess the efficacy of an intervention among those who choose to accept it. However, it is important to recognize that this analysis will likely be biased by selection effects because patients who complete ADs and choose comfort care are likely different from those who do not complete ADs or make other choices in completed ADs. These underlying differences may influence outcomes such as quality of life. We will assess the magnitude of such selection effects by comparing results between the per-protocol, mITT, and CATE analyses.

Secondary outcomes analyses: Secondary outcomes will be analyzed using logistic, linear, or quantile regression, as appropriate. The number of hospital and ICU admissions will be analyzed as count data. Charges will be inflated to the date on which analyses are performed using the US gross domestic product deflator.

In all models, center will be entered as a random effect to adjust for potential clustering within centers and to mitigate confounding by clinic. Gender, race, and diagnosis category will be included in all multivariable models based on pre-specified hypotheses, and others will be added if their inclusion – singly or jointly – modifies the coefficient for the randomized exposure by ≥ 15%.

5.4 Sensitivity analyses modifying the HFD calculation

a) We will recode the outcome as "Healthcare facility-free days", which represent the number of days alive where a patient is not in an acute care facility, a chronic care facility or a nursing home will be evaluated as an alternative to hospital-free days
b) We will also analyze effects on the original "Hospital-free days" but only up through six months of follow-up

5.5 Subgroup analyses

Planned subgroup analyses will be conducted across groups defined by gender (male vs. female), age (analyzed as a continuous variable), race (White vs. Black, excluding all other races), religion (Christian vs. not Christian), diagnostic category (cancer vs. non-cancer), and the three prior experience questions.

5.6 Mediator analysis

We will conduct three mediation analyses. First, the presence of an AD in the medical record (i.e., the successful uploading of the completed AD to the patient's medical record) will be examined as a mediator variable for (1) the primary analysis examining the relationship between randomization group and HFDs and (2) also for the secondary outcome of concordance of care. In addition, we will examine (3) surrogates' distrust of the healthcare system, measured by the Health System Distrust Scale, as a mediating variable in the relationship between randomization group and surrogates' perceptions of the quality of death and dying.

In order to establish a variable as a mediator, we will first confirm that the proposed mediating variable precedes the outcome in time, and then conduct a series of regressions to evaluate the following four hypotheses (presented for the primary analysis, below). Rejection of all four hypotheses is necessary to establish the presence of an AD in the medical record as a mediator. These four hypotheses are:

1. Randomization group has a significant effect on the presence of an AD in the medical record
2. Having an AD in the medical record has a significant effect on HFDs
3. Randomization group has a significant effect on HFDs
4. The effect of randomization group on HFDs is attenuated when the presence of an AD in the medical record is added to the model

Each hypothesis will be examined using linear or logistic regression, as appropriate, and will be adjusted for center to account for any center differences.

If the null hypothesis is rejected for the above four hypotheses, we will determine the proportion of variability explained by the presence of an AD in the medical record by quantifying the change in the treatment assignment coefficient between the reduced (#3 above) and full model (#4 above).

6. Sample Size and Power

We calculate our sample size as that required to rule out a significant reduction in HFDs attributable to random assignment to a default AD. This approach entails non-inferiority tests of data from a Poisson distribution, such that we seek to reject the hypothesis of a rate ratio (RR) for HFDs that is significantly >1.0. By enrolling 270 patients who complete ADs—90 in each of the three arms—we will obtain at least 80% power to demonstrate non-inferiority up to a margin of an RR for HFDs ≥1.18 associated with use of a default AD. This calculation is based on: (1) a one-sided α of 0.05, yielding an upper confidence limit on the observed RR that falls entirely below an RR of 1.18; (2) a mean number of HFDs in the control group of 100, such that a RR of 1.18 would correspond to 15 (15%) fewer HFDs in a given AD group (100/85=1.18); (3) an allowance for considerable dispersion in the distribution of HFDs; (4) no loss to follow-up because all deaths and hospitalizations will be checked against the Social Security Death Index and Pennsylvania Health Care Cost Containment Council (PHC4), respectively; (5) an allowance for two primary hypotheses tests (comparing both the comfort-default and life-extension default arms to the control arm) and (6) a true RR of 1.0. This final choice reflects our hypothesis that assignment to all three ADs will produce equivalent numbers of HFDs. If the true RR is below 1.0 (eg, the comfort default increases HFDs), power would increase considerably. Further, because simulations used to generate these sample size estimates included scenarios with extreme assumptions of data dispersion, and the proposed sample sizes incorporate this conservative assumption, our observed power is likely to be higher than stated.

This original Statistical Analysis Plan was finalized on March 18, 2014, after review and approval by the DSMB during its first meeting.

V. Final Statistical Analytic Plan

1. Analytic Methods

To assess balance among groups achieved by randomization, we will compare baseline values of all variables across arms using ANOVA and chi-square tests for continuous and binary data, respectively. We will use Poisson models to assess the number of hospital free days (HFDs) from the time of randomization. We will use logistic, linear, or quantile (1) regression, as appropriate based on outcome parameterizations and distributions, for all secondary outcomes. In all analyses, we will model the clinic from which patients are recruited as a random effect to adjust for potential clustering within clinics and

to mitigate confounding by clinic (2). We will employ standard covariate-selection procedures for etiologic models to assess, and potentially adjust for, chance covariate imbalance among arms. Specifically, patient-level covariates (e.g., gender, race, diagnosis category) will be included in multivariable models based on pre-specified hypotheses or if their inclusion – singly or jointly – modifies the coefficient for the randomized exposure by ≥ 15% (3).

All analyses will be conducted using the intention-to-treat approach to avoid selection bias. Some patients who consent to participate and receive their assigned AD may not return the AD. In our pilot work we developed several interventions that successfully mitigated this possibility. However, any patients who do not return the AD will be retained in the primary analyses, and will be classified as having not specified preferences for goals of care or specific interventions.

2. Specific aims and hypothesis

a) Compared with standard ADs, neither ADs with preselected comfort care options nor ADs with preselected options intended to promote life extension will reduce patients' hospital-free days (HFDs), a measure that represents the number of days alive and not in an acute care facility.
b) Compared with standard ADs or ADs defaulting to life-extension, ADs defaulting to comfort care will:
    1. produce no change in survival
    2. reduce hospital and ICU admissions
    3. reduce costs of inpatient care
    4. improve patients' quality of life
    5. improve patients' satisfaction with end-of-life care and decision making
    6. reduce the receipt of life-sustaining therapies

3. Exposure

Intervention group (standard AD, life-extension default, or comfort care default)

4. Outcomes

4.1 Primary

Primary outcome is Hospital-free days. This metric represents the number of days alive and not in an acute care facility following the date of consent. We chose the date of consent as day 0 so that all enrolled participants, including those who do not return ADs, are eligible for ITT analyses.

The choice of HFDs as the primary outcome reflects the desire to choose a measure that is patient-centered, readily measured and analyzed, and reflects a patient's holistic state rather than a specific symptom. HFDs have many attractive properties: they are continuous, enhancing power; they can be analyzed reliably and flexibly, to account for different values patients may place on avoiding hospitalization; and in nearly all cases, they are unidirectional, in the sense that nearly all patients prefer longer lives to shorter ones, and to have more of those days spent outside a hospital than within.

4.2 Secondary

Secondary outcomes include several clinical, economic and patient-reported measures including:

1. Survival – Patient deaths will be captured via medical records and verified by the Pennsylvania and New Jersey Departments of Health vital statistics
2. Hospital and ICU admission – Captured by querying state-run databases that capture all admission and inpatient procedures in Pennsylvania and New Jersey.
3. Total Inpatient care charges – Captured via the database detailed in (2).
4. Inpatient care charges per day per visit- Captured via the database detailed in (2).
5. Receipt of life-sustaining therapies – Captured via the database detailed in (2).
6. Concordance between patients' expressed desires in ADs regarding four potentially life-sustaining therapies (CPR, mechanical ventilation, dialysis, and feeding tube) and care received
7. Quality of life – Measured at around 6 months following AD completion with the McGill quality of life (MQOL), which can be completed by family members on behalf of patients who are unable to complete it themselves. MQOL during follow-up is missing for a high number of participants. In examining the data over time, we were able to determine that MQOL does not change over time and that time to follow-up is not significantly related to MQOL values. Therefore, we will only report one MQOL score per patient (in the per protocol analysis), and this score will be the one closest to the 6 month follow-up period. Also, per-protocol patients who die are assigned a value of 0 and the remainder are imputed.
8. Satisfaction with advance care planning – Measured at 2 months following AD completion with the Canadian Healthcare Evaluation Project (CANHELP) instrument's global satisfaction and end-of-life care question.
9. Satisfaction/conflict with decision-making – Measured immediately following AD completion using the validated decision conflict scale (DCS).

5. Analysis

We aim to answer two primary questions in this study:

(1) What is the overall effectiveness of offering people the opportunity to complete advance directives with different embedded default options?
(2) What are the specific effects of making certain choices within ADs on patient outcomes?

The primary way we'll answer question (1) is through the modified ITT analysis; question (2) will be answered using a complier average treatment effect (CATE) analysis.

5.1 Modified ITT

The unit of analysis for the primary outcome (HFDs) will be the individual patient. mITT analyses include all patients except for (1) post-randomization ineligibles; and (2) withdraws. The rationale for these exclusions is that none of these patients were fully eligible to complete the assigned intervention in a way that would be accessible to the investigators. Further, as expected, these losses are evenly distributed across the 3 arms (see CONSORT diagram) such that their exclusion could not affect the results. In primary analyses we include all patients regardless of AD return or debriefing.

mITT analyses will be conducted using count regression, adjusting for center, to compare the effects of assignment to complete ADs with different default options on HFDs. We found center and diagnosis are

highly correlated and used diagnosis in the model building. This approach will use data from all randomized patients and will provide the truest test of the overall effectiveness of the intervention among those randomly assigned to receive it.

5.2 CATE analysis

This analysis examines the effects of making certain choices within ADs on outcomes and accounts for tendencies to not complete ADs. The CATE analysis surmounts the selection effects inherent in per-protocol analyses, as well as the inability of ITT analyses to provide specific tests of the effects of choices made in ADs because these effects will be diluted by the fact that many randomized patients will not complete their assigned ADs. The CATE approach entails a two-stage least-squares regression in which the randomization arm is modelled as an instrumental variable (IV) in complier average treatment effect (CATE) analysis. This analysis will also be adjusted for center. Like the ITT analysis, these analyses use data on all randomized patients to estimate the effects of specifying any treatment choice in ADs regardless of group assignment, and after accounting for the possibility that AD completion rates may differ among the three arms by using the randomization arm as the IV. Thus, the estimated effect of the choices patients make is adjusted for the percentage of assigned patients who complete an AD at all, and the percentage who opt out from their assigned default option. This IV uses data on all randomized patients and then adjusts for AD completion rates, thereby attenuating the selection effects.

This analysis also requires the use of principal stratification methods to formulate the causal quantities of interest and determine the proportions of patients in each arm who would choose comfort care if they were assigned to complete each version of the AD. The analysis assumes that all patients who would choose comfort care in a standard AD would also choose it in an AD that defaults to comfort care, and that all patients who would choose comfort care in an AD that defaults to aggressive care would also choose it in a standard AD or an AD the defaults to comfort care. Coupled with the possibilities that some participants would never return an AD, and that others would return an AD but not choose comfort care regardless of group assignment, this creates five compliance classes (principal strata) of participants. These classes are:

I. Patients would not complete an AD regardless of group assignment
II. Patients would complete an AD but not choose comfort care regardless of group assignment
III. Patients would complete an AD and only choose comfort care if assigned to the comfort-default AD
IV. Patients would complete an AD and choose comfort care if assigned to the comfort-default AD or standard AD
V. Patients would complete an AD and choose comfort care regardless of group assignment

Each patient has three potential outcomes (see below). Only one of the potential outcomes can be observed, the outcome corresponding to the actual intervention the patient received. This is represented by a binary endpoint – whether or not patients would have a high quality of life in the future:

Y_i^A = whether patient i would have high quality of life if assigned to complete an aggressive-default AD
Y_i^S = whether patient i would have high quality of life if assigned to complete an standard AD

Y_i^C = whether patient i would have high quality of life if assigned to complete an comfort-default AD

Our approach assumes the exclusion restriction that AD assignment only influences the potential outcomes through the causal pathway of determining which type of care the patient chooses through the AD. However, this assumption is likely to hold in this case, because the randomly assigned IV – which of three versions of the AD is offered – would not influence outcomes unless it modified the probability of AD completion or the choices made in the ADs.

5.3 Secondary analyses

Secondary outcomes analyses: Secondary outcomes will be analyzed using logistic, linear, or quantile regression, as appropriate. The number of hospital and ICU admissions will be analyzed as count data. Charges will be inflated to the date on which analyses are performed using the US gross domestic product deflator.

In all models, center will be entered as a random effect to adjust for potential clustering within centers and to mitigate confounding by clinic. Since center and diagnosis are strongly correlated as we mentioned above, we will only include diagnosis as fixed effect in the models. Gender, race (categorical) and age (continuous) will be included in all multivariable models based on pre-specified hypotheses, and others will be added if their inclusion – singly or jointly – modifies the coefficient for the randomized exposure by ≥ 15%.

5.4 Sensitivity analyses modifying the HFD calculation

(1) We will impute the HFD for the patients with invalid SSNs. The imputation method will be model based multiple imputation approach and we will report the pooled estimates.

6. Sample Size and Power

We calculate our sample size as that required to rule out a significant reduction in HFDs attributable to random assignment to a default AD. This approach entails non-inferiority tests of data from a Poisson distribution, such that we seek to reject the hypothesis of a rate ratio (RR) for HFDs that is significantly >1.0. By enrolling 270 patients who complete ADs—90 in each of the three arms—we will obtain at least 80% power to demonstrate non-inferiority up to a margin of an RR for HFDs ≥1.18 associated with use of a default AD. This calculation is based on: (1) a one-sided α of 0.05, yielding an upper confidence limit on the observed RR that falls entirely below an RR of 1.18; (2) a mean number of HFDs in the control group of 100, such that a RR of 1.18 would correspond to 15 (15%) fewer HFDs in a given AD group (100/85=1.18); (3) an allowance for considerable dispersion in the distribution of HFDs; (4) no loss to follow-up because all deaths and hospitalizations will be checked against the Social Security Death Index and Pennsylvania Health Care Cost Containment Council (PHC4), respectively; (5) an allowance for two primary hypotheses tests (comparing both the comfort-default and life-extension default arms to the control arm) and (6) a true RR of 1.0. This final choice reflects our hypothesis that assignment to all three ADs will produce equivalent numbers of HFDs. If the true RR is below 1.0 (eg, the comfort default increases HFDs), power would increase considerably. Further, because simulations used to generate these sample size estimates included scenarios with extreme assumptions of data dispersion, and the

proposed sample sizes incorporate this conservative assumption, our observed power is likely to be higher than stated.

VI. Summary of Changes to the Statistical Analytic Plan

1. We redefined the mITT sample to now include who died within 30 days of randomization. This change was recommended by our DSMB during our July 19, 2016 meeting, well before any trial data were reviewed even in cumulative form, let alone unblinded. Thus, the only exclusion criteria were (1) post-randomization ineligibles and (2) patients who withdrew.
2. We had considerable difficulty obtaining responses from surrogates after patients died. The low response rate was discussed with the DSMB at our March 1, 2017 meeting. After reviewing the cumulative data (not stratified by arm) on April 20, 2018, we elected to forgo analyses of surrogate-reported outcomes.
3. We were not able to evaluate hospice utilization because these data were unexpectedly missing from the NJ and PA databases. We discussed this with the DSMB at our March 1, 2017 meeting. During the Spring of 2018, we pursued other hospice-specific databases and spoke with hospice organizations at both participating health systems. However, because patients from both participating health systems may end up in multiple different hospice systems, we were concerned that this approach would yield incomplete data. Thus, we abandoned the plan to analyze hospice utilization on May 25, 2018. This decision was made by the PI (Dr. Halpern), who was still blinded to trial data.
4. We only analyzed satisfaction with advance care planning at 2 months following AD completion because the data available to analyze the 6 and 12 months measures were frequently missing. This choice was similarly made by Dr. Halpern while blinded to arm-specific data.
5. We elected not to conduct the proposed sensitivity analysis in which the primary outcome was changed to "Healthcare facility-free days," which would represent the number of days that a patient spent alive and outside an acute care facility, a chronic care facility, or a nursing home. We abandoned this plan because we could not obtain reliable data on days spent in the latter two types of facilities.

All of the above modifications were made prior to unblinding of trial data to anyone other than the Data Manager, Brian Bayes. Mr. Bayes had no role in making the foregoing decisions. Unblinded analyses were then prepared by Dr. Chowdury, in collaboration with Drs. Halpern, Small, and Troxel.

6. Because the hospital-free days distribution was observed to be highly left skewed regardless of duration of follow-up, we chose not to perform the planned sensitivity analysis using different time cut-offs. This decision was made on August 1, 2018, by Drs. Halpern, Chowdhury, and Troxel.
7. Due to unplanned missing data on the primary outcome measure, we used multiple imputation to impute missing HFD data for the 55 patients with invalid SSNs. We elected, on August 8, 2018, to report these analyses among all patients in the mITT sample using imputation, and among the 88.8% of patients with observed outcomes.

8. Also on August 8, 2018, we elected to compare patient-level characteristics between the 55 patients who did not provide valid SSNs and the 437 patients who did provide valid SSNs. We made this decision so as to assess the possibility of selection effects stemming from this form of non-response in the analyses without imputed data.
9. Because both the mITT and CATE analyses were null, we elected not to perform the per-protocol analysis as had originally been planned, because inferences from such an analysis would have yielded ambiguous conclusions. This choice was made by Drs. Halpern, Chowdhury, Small, and Troxel on September 17, 2018.
10. Also on September 17, 2018, we elected not to perform the proposed subgroup analyses for purposes of this first manuscript due to space considerations, and to instead report these in a subsequent brief manuscript.
11. We also modified the plans for mediator analyses on September 17, 2018. We elected not to examine mediation of the primary outcome by uploading of AD into the EHR because the primary comparison of the randomization group on this outcome was null. We elected to pursue the second proposed mediator analysis, on the outcome of goal-concordant care, in a subsequent report.

The Statistical Analysis Plan was considered final at close of business on September 17, 2018.

12. Afterwards, during preparation of our manuscript for submission, we elected to pursue per-protocol analyses among the 186 patients who returned ADs, were debriefed, and had their ADs uploaded into the EHR. We reasoned that this would assist in interpretation of a trial reporting no differences across arms in any clinical outcomes. In reporting this analysis, we clearly specify that it was a post-hoc analysis.
13. During peer-review of our submitted manuscript, protocol, and SAP, reviewers and editors correctly noted an error in Hypothesis 2a in the original SAP, which stated that we hypothesized that ADs with comfort-oriented defaults would *increase* the number of hospital-free days. This hypothesis was inconsistent with what we stated in our trial protocol (in which we state that "we will determine whether this simple and readily scalable intervention can improve patients' quality of life and reduce resource utilization *without reducing* the number of days that patients are alive and living outside of an acute-care hospital.") Indeed, this language of testing the noninferiority of comfort-oriented defaults on the outcome of hospital-free days is also present in our original grant application and our original posting of the trial protocol on ClinicalTrials.gov on December 16, 2013. We regret this error in the original SAP, and have corrected it in the submitted final SAP such that hypothesis 2a now correctly reads: "Compared with standard ADs, neither ADs with preselected comfort care options nor ADs with preselected options intended to promote life extension will reduce patients' hospital-free days (HFDs), a measure that represents the number of days alive and not in an acute care facility."