Advance Directives as Acts of Communication

A Randomized Controlled Trial

Peter H. Ditto, PhD; Joseph H. Danks, PhD; William D. Smucker, MD; Jamila Bookwala, PhD; Kristen M. Coppola, PhD; Rebecca Dresser, JD; Angela Fagerlin, PhD; R. Mitchell Gready, PhD; Renate M. Houts, PhD; Lisa K. Lockhart, PhD; Stephen Zyzanski, PhD

Background: Instructional advance directives are widely advocated as a means of preserving patient self-determination at the end of life based on the assumption that they improve surrogates' understanding of patients' life-sustaining treatment wishes. However, no research has examined whether instructional directives are effective in improving the accuracy of surrogate decisions.

Participants and Methods: A total of 401 outpatients aged 65 years or older and their self-designated surrogate decision makers (62% spouses, 29% children) were randomized to 1 of 5 experimental conditions. In the control condition, surrogates predicted patients' preferences for life-sustaining medical treatments in 9 illness scenarios without the benefit of a patient-completed advance directive. Accuracy in this condition was compared with that without the benefit of a patient-completed advance directive. Surrogates were also measured.

Results: None of the interventions produced significant improvements in the accuracy of surrogate substituted judgment in any illness scenario or for any medical treatment. Discussion interventions improved perceived surrogate understanding and comfort for patient-surrogate pairs in which the patient had not completed an advance directive prior to study participation.

Conclusions: Our results challenge current policy and law advocating instructional advance directives as a means of honoring specific patient wishes at the end of life. Future research should explore other methods of improving surrogate decision making and consider the value of other outcomes in evaluating the effectiveness of advance care planning.

Arch Intern Med. 2001;161:421-430

©2001 American Medical Association. All rights reserved.
METHODS

PARTICIPANTS
ADVANCE participants were recruited from 6 primary care practices affiliated with Summa Health System in Akron, Ohio. Randomly selected patients aged 65 years and older (N = 2544) received a letter from their physician inviting them to participate in the study. Individuals aged 75 years and older were oversampled. Potential participants were given 1 week after receipt of the letter to call the project office if they were unable or unwilling to participate (189 did so). Professional interviewers attempted to contact each remaining patient by telephone (612 potential participants could not be reached because of death, disconnected telephones, etc.). The patients contacted (n = 1704) were informed that the study involved three 1- to 2-hour in-home interviews over 2 years, additional interviews if hospitalized, and the coparticipation of a surrogate decision maker (defined as the individual they would want to make medical decisions for them if they were no longer able). Prior research27 led us to anticipate that these study requirements would result in a participation rate of approximately 30%. Accordingly, age and sex information was obtained from all individuals contacted and an attempt was made to collect additional information about patients’ plans for future medical care. A total of 447 potential participants were excluded for 1 or more of the following reasons: no available surrogate decision maker, spouse already participating, or patient judged unable to give informed consent. Of 849 refusing to participate, 510 agreed to answer questions about planning for future medical care. We randomized the remaining 408 patient-surrogate pairs to 1 of 2 AD documents. The health care directive (HCD)20 is available on the standard version. Scenarios and treatments were varying in their severity, nature of impairment, prognosis, and level of pain: (1) the patient’s current health; (2) Alzheimer disease with moderately severe cognitive impairment; (3) emphysema with severe physical limitations; (4) coma persisting 6 weeks after a stroke with no obvious cognitive abilities and physician opinion of no chance of recovery; (5) the same coma scenario as item 4 with a physician opinion of a very slight chance of recovery; (6) stroke resulting in partial paralysis, language deficits, total dependence in activities of daily living, and physician opinion of no chance of improvement; (7) the same stroke scenario as item 6 with a physician opinion of a very slight chance of improvement; (8) colon cancer with fatigue, no pain, and a life expectancy of 6 months; and (9) the same cancer scenario as item 8 with pain that requires the constant use of medication. Patients were asked to imagine themselves in each scenario and indicate their preference for receiving each of 4 life-sustaining medical treatments chosen to vary in invasiveness: (1) antibiotics for life-threatening pneumonia, (2) cardiopulmonary resuscitation (CPR) for cardiac arrest, (3) emergency gallbladder surgery for life-threatening gallbladder infection, and (4) artificial nutrition and hydration (ANH) for inability to take food or water. Preferences for ANH were not solicited in the current health scenario because patients were unlikely to require ANH in that situation. Patients indicated their treatment preferences using a 5-point Likert scale ranging from “definitely don’t want” to “definitely want” treatment.10,11,13 Surrogates were asked to imagine the patient in each illness scenario and predict the patient’s treatment preferences using the same scale.

The secondary outcomes (perceived benefits of AD completion) were measured with 5 questions using 5-point Likert scales. The questions measured patients’ and surrogates’ beliefs regarding (1) surrogates’ general understanding of patients’ life-sustaining treatment preferences, (2) confidence in surrogates’ ability to predict accurately patients’ treatment decisions, (3) surrogates’ likelihood of honoring patients’ treatment wishes, (4) surrogates’ comfort in making medical decisions on behalf of the patient, and (5) the importance of having an AD.

OUTCOME MEASURES
The primary outcome (accuracy of substituted judgment) was measured with the Life-Support Preferences/Predictions Questionnaire (LSPQ).28-30 The LSPQ (written at the eighth-grade reading level)31 describes 9 different illness scenarios chosen to capture a wide range of conditions varying in their severity, nature of impairment, prognosis, and level of pain: (1) the patient’s current health; (2) Alzheimer disease with moderately severe cognitive impairment; (3) emphysema with severe physical limitations; (4) coma persisting 6 weeks after a stroke with no obvious cognitive abilities and physician opinion of no chance of recovery; (5) the same coma scenario as item 4 with a physician opinion of a very slight chance of recovery; (6) stroke completed instructional directive improves the accuracy of substituted judgment.18 If instructional directives fail to communicate information in a way that improves the ability of surrogate decision makers to understand and predict patient preferences, then the wisdom of advocating ADs as a means of preserving patients’ ability to control specific treatment decisions near the end of life is suspect. Surrogates cannot honor a patient’s wishes if they do not understand those wishes.

A test of the assumption that instructional ADs improve the accuracy of substituted judgment requires a direct comparison of the accuracy of surrogate decision makers randomized to make treatment predictions with and without the benefit of reviewing a patient’s completed AD. A thorough test of this assumption must also consider the possibility that different methods of documenting AD information may be differentially effective in improving the quality of surrogate decisions. The most straightforward
identical. The primary difference was that we asked about the goal of medical care once after all scenarios had been completed rather than after each scenario. The valued life activities directive (VLA) is a value-based AD developed by our group.24 Respondents are asked to think about the activities “that make your life worth living” and generate a list of activities they believe are so important to their well-being that they would not want to live if they were no longer able to engage in those activities. Both ADs were read aloud by the interviewer and patients’ responses were recorded on a paper copy. After completing the AD, patients were administered the LSPQ and perceived benefit measures. The surrogate interview began by allowing the surrogate to review the patient’s AD. The format of each directive was explained and surrogates were given as much time as desired to review the document before and during completion of the LSPQ.

In the 2 discussion intervention conditions, patients completed the AD in the presence of the surrogate. Pairs were encouraged to discuss the patient’s responses to the directive through a series of structured prompts delivered by the interviewer. These prompts asked patients to explain the reasons underlying their choices and encouraged surrogates to ask patients to clarify the reasons for their choices. (Copies of both ADs with the content and schedule of discussion prompts can be obtained from the first author [P.H.D.] on request.) After completing the AD, patients and surrogates separated and the interviews proceeded as in the other conditions.

Patients and surrogates provided standard demographic information. Patients also completed the Medical Outcomes Study 36-Item Short-Form Health Survey32 and the Center for Epidemiological Studies Depression Scale 10.33

STATISTICAL ANALYSIS

Characteristics of patients agreeing and refusing to participate were compared with t tests for continuous variables and χ² tests for categorical variables. Similar analyses compared characteristics of patients and surrogates randomized to the 5 experimental conditions.

Responses on the LSPQ were dichotomized into “want treatment” (“definitely want,” “probably want,” or “unsure”) and “don’t want treatment” (“probably do not want” and “definitely do not want”) responses for each of the 35 treatment decisions. (Following past research,10,11,13 “unsure” responses were categorized with “want treatment” responses because in most instances the clinical default is to provide treatment unless specifically refused.10,11,13 Analyzing data excluding “unsure” responses or treating them as a third response category produced no significant differences in the study results.) Proportion indexes were generated for preferences and predictions in each scenario by summing the number of “want treatment” responses within each scenario and dividing by the number of decisions in that scenario. Repeated-measures analyses of variance compared preference-prediction indexes across illness scenarios and the overall proportion of want treatment responses given by patients and surrogates.

Surrogate predictions were defined as accurate if for a given treatment decision surrogates and patients gave the same dichotomized response. Inaccurate predictions were categorized into overtreatment errors (surrogate predicted patient would want treatment when patient actually did not) and undertreatment errors (surrogate predicted patient would not want treatment when patient actually did). Proportion indexes were created for each scenario by summing the number of accurate predictions within each scenario and dividing by the number of decisions in that scenario. Multivariate analysis of variance (MANOVA) was used to simultaneously compare these 9 indexes across experimental conditions. Dunnett post hoc tests further compared each of the 4 intervention conditions with the no-AD control condition. This process was repeated with overtreatment and undertreatment indexes. To confirm the results of the scenario-based analyses, an analogous set of analyses was conducted using proportion indexes calculated for each treatment (collapsed across medical scenarios) as dependent variables. Finally, to examine whether the interventions were differentially effective for different subgroups of participants, MANOVAs were reconducted using selected patient (sex, age, education, report of prior AD completion), surrogate (sex, age, education), and patient-surgeon relationship (type and length of relationship) characteristics as a second independent variable with experimental condition.

To examine the effects of the interventions on the secondary outcome measures, 2 MANOVAs were conducted. The first included the 5 patient-perceived benefit measures as dependent variables. The second included the 5 surrogate-perceived benefit measures as dependent variables. Subgroup analyses were again conducted to examine for differential effects of the interventions for different groups of participants.

POWER ANALYSIS

Setting α at .05 and statistical power at .80, our primary analyses (MANOVAs comparing 5 groups with 9 dependent variables) require 58 observations per condition to detect large effects and 105 observations per condition to detect medium-size effects.35 Our secondary analyses (MANOVAs comparing 5 groups with 4 dependent variables) require 72 observations per condition to detect medium-size effects. Accordingly, we set 80 patient-surgeon pairs per condition as our target sample size to provide us with adequate power to detect medium-size to large effects.
The Advance Directives, Values Assessment, and Communication Enhancement (ADVANCE) project was a 3-phase, longitudinal study designed to test key assumptions underlying the use of instructional ADs. This article describes the results of phase 1 of the ADVANCE project: a randomized controlled trial designed to evaluate the effectiveness of 2 instructional directives, completed with and without patient-surrogate discussion, to improve the accuracy of surrogate substituted judgment.

**RESULTS**

**SAMPLE CHARACTERISTICS**

Patients agreeing to participate in the study were significantly younger than refusers (mean ages, 73 years vs 75 years; \( P < .001 \)) and less likely to be women (56% vs 62%; \( P = .03 \)). Participating patients were also significantly more likely than refusers to report having made plans for future medical care (69% vs 57%; \( P = .001 \)) and to rate having plans for future care as important (mean, 4.51 vs 3.84; \( P < .001 \)).

No statistically significant differences across intervention conditions were found on any patient or surrogate characteristic, confirming that randomization was successful. Overall, patients and surrogates were predominantly European American and Protestant, with relatively high socioeconomic status (Table 1). A slight minority of patients reported having a living will or other AD, while a slight majority reported having a durable power of attorney for health care. Surrogates were typically spouses or adult children of patients and most (78%) reported knowing the patient for 40 years or more.

**PATIENT PREFERENCES AND SURROGATE PREDICTIONS**

Preferences for treatment varied significantly across illness scenarios (\( P = .001 \)). Patient preferences ranged from an almost unanimous desire to receive all life-sustaining treatments in their current health (mean proportion of want responses, 0.96) to high rates of treatment rejection in the “coma no chance” scenario (mean, 0.12; Figure 2). Surrogates predicted a similar pattern to patient preferences. Surrogates predicted greater desire for treatment than that expressed by patients (overall mean proportion of want, 0.54 vs 0.44, respectively, \( P = .001 \)). The interventions had no effect on preferences or predictions.

**ACCURACY OF SUBSTITUTED JUDGMENT**

Surrogates predicting patient preferences without the benefit of an AD showed only modest accuracy in their substituted judgments (Table 2). Although surrogates in the no-AD condition demonstrated relatively high accuracy in the “current health” scenario (mean proportion of correct predictions, 0.94), surrogate predictions were correct less than 70% of the time on average for the “Alzheimer disease,” “coma slight chance,” “stroke no chance,” “stroke slight chance,” and “cancer no pain” scenarios (overall predictive accuracy in the no-AD condition, 0.72). In general, accuracy levels closely mirrored the degree of variation in patients’ preferences with surrogates predicting most accurately for scenarios in which most patients either wanted (eg, current health) or did not want (eg, coma no chance) to receive treatment. The majority of errors made by surrogates in the no-AD condition were errors of overtreatment, with the ratio of overtreatment to undertreatment errors ranging from 2:1 in the “current health” scenario to greater than 3:1 in the “emphysema,” “coma no chance,” “coma slight chance,” and “stroke no chance” scenarios.

The AD interventions failed to improve the accuracy of surrogate substituted judgment (scenario MANOVA, \( F_{15,122} = 1.55, P = .25 \); treatment MANOVA, \( F_{16,127} = 0.66, P = .83 \); Table 2). Dunnett post hoc tests confirmed that neither no-discussion intervention produced improvements in surrogate predictive accuracy (in any illness scenario or for any type of medical treatment) beyond the level observed in the no-AD control condition (overall mean predictive accuracy for HCD no discussion, 0.75, VLA no discussion, 0.73). Discussion was equally ineffective in improving surrogate accu-
Table 1. Patient and Surrogate Characteristics*  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patients (n = 401)</th>
<th>Surrogates (n = 481)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SEM, y</td>
<td>73.0 ± 0.3</td>
<td>62.5 ± 0.7</td>
</tr>
<tr>
<td>Female</td>
<td>224 (56)</td>
<td>270 (67)</td>
</tr>
<tr>
<td>Race, white</td>
<td>370 (92)</td>
<td>369 (92)</td>
</tr>
<tr>
<td>Married</td>
<td>273 (68)</td>
<td>343 (86)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;High school diploma</td>
<td>74 (19)</td>
<td>38 (10)</td>
</tr>
<tr>
<td>High school diploma</td>
<td>152 (38)</td>
<td>145 (36)</td>
</tr>
<tr>
<td>&gt;High school diploma</td>
<td>173 (43)</td>
<td>218 (54)</td>
</tr>
<tr>
<td>Income, $</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤14,999</td>
<td>87 (22)</td>
<td>35 (9)</td>
</tr>
<tr>
<td>15,000-29,999</td>
<td>139 (35)</td>
<td>127 (32)</td>
</tr>
<tr>
<td>≥30,000</td>
<td>123 (31)</td>
<td>190 (47)</td>
</tr>
<tr>
<td>Refused to provide</td>
<td>52 (13)</td>
<td>49 (12)</td>
</tr>
<tr>
<td>Religious affiliation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>276 (69)</td>
<td>263 (66)</td>
</tr>
<tr>
<td>Catholic</td>
<td>93 (23)</td>
<td>94 (24)</td>
</tr>
<tr>
<td>Other</td>
<td>32 (8)</td>
<td>43 (11)</td>
</tr>
<tr>
<td>Self-reported general health†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very good or excellent</td>
<td>192 (48)</td>
<td>NA</td>
</tr>
<tr>
<td>Good</td>
<td>139 (35)</td>
<td>NA</td>
</tr>
<tr>
<td>Fair or poor</td>
<td>70 (18)</td>
<td>NA</td>
</tr>
<tr>
<td>CESD-10 score ≥ 10‡</td>
<td>59 (15)</td>
<td>NA</td>
</tr>
<tr>
<td>With AD or living will</td>
<td>184 (46)</td>
<td>NA</td>
</tr>
<tr>
<td>With DPAHC</td>
<td>208 (52)</td>
<td>NA</td>
</tr>
<tr>
<td>Patient-surrogate relationship</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse</td>
<td>246 (62)</td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td>114 (29)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>37 (9)</td>
<td></td>
</tr>
<tr>
<td>Length of relationship, mean ± SEM, y</td>
<td>46.9 ± 0.6</td>
<td></td>
</tr>
</tbody>
</table>

* Data are given as number (percentage) of participants except as otherwise noted and may not total due to missing values. NA indicates not applicable. AD, advance directive; CESD-10, Center for Epidemiological Studies Depression Scale; and DPAHC, durable power of attorney for health care.
†From Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36).
‡Scores ≥ 10 indicate depressive symptoms.

racy. Surrogates allowed to complete and discuss the ADs with the patient were also unable to predict patients’ treatment wishes at accuracy levels that exceeded those of surrogates making predictions without the benefit of an AD (overall mean predictive accuracy for HCD discussion, 0.73, VLA discussion, 0.69). Overtreatment and undertreatment errors showed no significant differences across study conditions.

Planned subgroup analyses did not identify any subgroup of patients, surrogates, or patient-surrogate relationships for which any of the AD interventions produced levels of predictive accuracy greater than those observed in the no-AD control condition.

PERCEIVED BENEFITS OF AD COMPLETION

More than 90% of patients in the no-AD condition believed that their surrogate understood their life-sustaining treatment wishes at least “pretty well” and were at least “pretty confident” that their surrogate could accurately predict their wishes (top of Table 3). Most no-AD patients also believed that their surrogate would honor their wishes (89%), but fewer believed surrogates were comfortable making medical decisions for them (69%). Surrogates’ responses to analogous questions were similar to those of patients (bottom of Table 3).

Patient measures revealed a significant intervention effect (F20,1287 = 2.73, P < .001). Dunnett post hoc tests indicate that both discussion interventions produced small but significant increases in patients’ perceptions of surrogate understanding relative to the no-AD control condition (P < .001 for all; Table 3). The VLA discussion intervention produced a similar increase in patients’ perceptions of surrogate comfort (P = .03). Subgroup analyses revealed that the effects of the discussion interventions on perceived understanding and comfort were limited to patients without a previously completed AD (intervention × AD interaction, F20,1248 = 1.62, P = .04). In the no-AD condition, patients with a prior AD believed their surrogate to have both better understanding of their wishes (mean, 4.56) and greater comfort with making end-of-life decisions (mean, 4.15) than did patients without a previous AD (means, 4.20 and 3.34, respectively, P < .02 for all). The discussion interventions did little to improve the already high level of perceived understanding and comfort in patients with a previous AD (P > .39 for all).

Both discussion interventions, however, produced significant increases in perceived understanding (HCD discussion mean, 4.89, VLA discussion mean, 4.84) and perceived comfort (HCD discussion mean, 4.03, VLA discussion mean, 4.12) for patients without a previous AD (P < .01 for all).

Surrogate measures did not differ by experimental condition (F20,1284 = 1.30, P = .17). Dunnett post hoc tests, however, reveal that surrogates in the HCD discussion condition had significantly higher scores (relative to the no-AD condition) on perceived understanding of patients’ wishes (P = .04), confidence in their own predictive accuracy (P = .02), and belief in the importance of having an AD (P = .01). The VLA discussion intervention also produced a slight increase in surrogates’ perceived comfort (P = .08). Subgroup analyses revealed no significant qualifications of these effects.

Phase 1 of the ADVANCE project revealed only modest accuracy of surrogate decisions uninformed by instructional ADs. Most surrogates were spouses or children of...
proving surrogate decisions. The results of our study session of a patient’s instructional directive would improve the accuracy of surrogate substituted judgment. Empirical ADs and patient-surrogate discussion to improve the ineffectiveness of both instructional incapacitation, the study did not require that patients complete detailed ADs or that family members or physicians discuss or even read any directives that were received the discussion interventions to improve shared understanding of patients’ treatment wishes.

The novel contribution of our study is its unequivocal demonstration of the ineffectiveness of both instructional ADs and patient-surrogate discussion to improve the accuracy of surrogate substituted judgment. Empirical data on surrogate inaccuracy have often been taken as support for ADs based on the assumption that possession of a patient’s instructional directive would improve surrogate decisions. The results of our study challenge this assumption. Our randomized trial design provided a stringent test of the ability of 4 different AD interventions to improve the accuracy of surrogate predictions. The results were clear and consistent across all 4 interventions and over every treatment decision examined. The ineffectiveness of an AD document alone to improve surrogate decisions might have been expected based on past writings skeptical of the usefulness of written directives in the absence of a broader process of advance care planning. More surprising was the finding that supplementing a written directive with guided patient-surrogate discussion was equally ineffective. This was true despite the fact that many patients and surrogates perceived the discussion interventions to improve shared understanding of patients’ treatment wishes.

The present study extends past research documenting numerous practical problems facing attempts to improve the quality of end-of-life medical care. The largest and best-known study, SUPPORT (Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments), found that a multifaceted intervention designed to facilitate advance care planning and improve patient-physician communication failed to produce significant improvements in the care and outcomes of seriously ill hospitalized patients. Our study differs from SUPPORT in several important ways. Although one goal of the SUPPORT intervention was to facilitate the documentation of treatment wishes in advance of decisional incapacitation, the study did not require that patients complete detailed ADs or that family members or physicians discuss or even read any directives that were completed. In fact, SUPPORT was crucial in documenting that patients often complete directives containing little or no clinically useful information that completed directives are often not placed or their existence even men-

<table>
<thead>
<tr>
<th>Index</th>
<th>No AD (n = 77)</th>
<th>HCD No Discussion (n = 79)</th>
<th>VLA No Discussion (n = 78)</th>
<th>HCD Discussion (n = 78)</th>
<th>VLA Discussion (n = 80)</th>
<th>Total (N = 392)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Health</td>
<td>0.94 ± 0.02</td>
<td>0.97 ± 0.01</td>
<td>0.95 ± 0.02</td>
<td>0.96 ± 0.01</td>
<td>0.94 ± 0.02</td>
<td>0.95 ± 0.01</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.04 ± 0.01</td>
<td>0.02 ± 0.01</td>
<td>0.03 ± 0.01</td>
<td>0.01 ± 0.01</td>
<td>0.05 ± 0.02</td>
<td>0.03 ± 0.01</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.02 ± 0.01</td>
<td>0.02 ± 0.01</td>
<td>0.02 ± 0.01</td>
<td>0.02 ± 0.01</td>
<td>0.04 ± 0.01</td>
<td>0.02 ± 0.01</td>
</tr>
<tr>
<td>Alzheimer Disease</td>
<td>0.65 ± 0.04</td>
<td>0.66 ± 0.04</td>
<td>0.70 ± 0.04</td>
<td>0.70 ± 0.03</td>
<td>0.63 ± 0.04</td>
<td>0.67 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.25 ± 0.04</td>
<td>0.16 ± 0.03</td>
<td>0.22 ± 0.03</td>
<td>0.21 ± 0.03</td>
<td>0.23 ± 0.04</td>
<td>0.21 ± 0.02</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.10 ± 0.03</td>
<td>0.17 ± 0.03</td>
<td>0.09 ± 0.03</td>
<td>0.09 ± 0.02</td>
<td>0.14 ± 0.03</td>
<td>0.12 ± 0.01</td>
</tr>
<tr>
<td>Emphysema</td>
<td>0.73 ± 0.04</td>
<td>0.77 ± 0.03</td>
<td>0.68 ± 0.04</td>
<td>0.70 ± 0.04</td>
<td>0.66 ± 0.04</td>
<td>0.71 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.21 ± 0.03</td>
<td>0.12 ± 0.03</td>
<td>0.26 ± 0.04</td>
<td>0.23 ± 0.04</td>
<td>0.24 ± 0.04</td>
<td>0.21 ± 0.02</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.06 ± 0.02</td>
<td>0.11 ± 0.03</td>
<td>0.06 ± 0.02</td>
<td>0.07 ± 0.02</td>
<td>0.10 ± 0.02</td>
<td>0.08 ± 0.01</td>
</tr>
<tr>
<td>Coma No Chance</td>
<td>0.78 ± 0.04</td>
<td>0.83 ± 0.03</td>
<td>0.83 ± 0.03</td>
<td>0.86 ± 0.03</td>
<td>0.81 ± 0.04</td>
<td>0.82 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.17 ± 0.04</td>
<td>0.10 ± 0.03</td>
<td>0.13 ± 0.03</td>
<td>0.12 ± 0.03</td>
<td>0.14 ± 0.03</td>
<td>0.13 ± 0.01</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.05 ± 0.02</td>
<td>0.07 ± 0.02</td>
<td>0.05 ± 0.02</td>
<td>0.06 ± 0.02</td>
<td>0.08 ± 0.02</td>
<td>0.06 ± 0.01</td>
</tr>
<tr>
<td>Coma Slight Chance</td>
<td>0.68 ± 0.04</td>
<td>0.68 ± 0.04</td>
<td>0.66 ± 0.04</td>
<td>0.63 ± 0.04</td>
<td>0.62 ± 0.05</td>
<td>0.66 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.24 ± 0.04</td>
<td>0.18 ± 0.03</td>
<td>0.23 ± 0.04</td>
<td>0.20 ± 0.04</td>
<td>0.23 ± 0.04</td>
<td>0.22 ± 0.02</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.07 ± 0.02</td>
<td>0.14 ± 0.03</td>
<td>0.12 ± 0.03</td>
<td>0.16 ± 0.03</td>
<td>0.15 ± 0.03</td>
<td>0.13 ± 0.01</td>
</tr>
<tr>
<td>Stroke No Chance</td>
<td>0.62 ± 0.04</td>
<td>0.70 ± 0.03</td>
<td>0.63 ± 0.05</td>
<td>0.68 ± 0.04</td>
<td>0.63 ± 0.04</td>
<td>0.66 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.28 ± 0.04</td>
<td>0.15 ± 0.02</td>
<td>0.25 ± 0.04</td>
<td>0.18 ± 0.04</td>
<td>0.28 ± 0.04</td>
<td>0.23 ± 0.02</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.09 ± 0.02</td>
<td>0.12 ± 0.03</td>
<td>0.12 ± 0.03</td>
<td>0.14 ± 0.03</td>
<td>0.09 ± 0.03</td>
<td>0.11 ± 0.01</td>
</tr>
<tr>
<td>Stroke Slight Chance</td>
<td>0.69 ± 0.04</td>
<td>0.69 ± 0.04</td>
<td>0.69 ± 0.04</td>
<td>0.74 ± 0.04</td>
<td>0.58 ± 0.04</td>
<td>0.68 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.21 ± 0.04</td>
<td>0.15 ± 0.03</td>
<td>0.20 ± 0.04</td>
<td>0.15 ± 0.03</td>
<td>0.30 ± 0.04</td>
<td>0.20 ± 0.02</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.09 ± 0.02</td>
<td>0.16 ± 0.03</td>
<td>0.11 ± 0.03</td>
<td>0.11 ± 0.03</td>
<td>0.11 ± 0.03</td>
<td>0.12 ± 0.01</td>
</tr>
<tr>
<td>Cancer No Pain</td>
<td>0.66 ± 0.04</td>
<td>0.70 ± 0.04</td>
<td>0.68 ± 0.04</td>
<td>0.68 ± 0.04</td>
<td>0.61 ± 0.04</td>
<td>0.66 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.25 ± 0.04</td>
<td>0.18 ± 0.03</td>
<td>0.24 ± 0.04</td>
<td>0.19 ± 0.04</td>
<td>0.31 ± 0.04</td>
<td>0.23 ± 0.02</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.09 ± 0.02</td>
<td>0.12 ± 0.03</td>
<td>0.08 ± 0.02</td>
<td>0.13 ± 0.03</td>
<td>0.08 ± 0.02</td>
<td>0.10 ± 0.01</td>
</tr>
<tr>
<td>Cancer With Pain</td>
<td>0.79 ± 0.03</td>
<td>0.74 ± 0.04</td>
<td>0.77 ± 0.04</td>
<td>0.67 ± 0.04</td>
<td>0.75 ± 0.04</td>
<td>0.74 ± 0.02</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.15 ± 0.03</td>
<td>0.14 ± 0.03</td>
<td>0.13 ± 0.03</td>
<td>0.20 ± 0.04</td>
<td>0.20 ± 0.04</td>
<td>0.16 ± 0.02</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.06 ± 0.02</td>
<td>0.12 ± 0.03</td>
<td>0.09 ± 0.03</td>
<td>0.13 ± 0.03</td>
<td>0.05 ± 0.02</td>
<td>0.09 ± 0.01</td>
</tr>
<tr>
<td>Total</td>
<td>0.72 ± 0.02</td>
<td>0.75 ± 0.02</td>
<td>0.73 ± 0.02</td>
<td>0.73 ± 0.02</td>
<td>0.69 ± 0.02</td>
<td>0.72 ± 0.01</td>
</tr>
<tr>
<td>Overtreatment</td>
<td>0.21 ± 0.02</td>
<td>0.14 ± 0.01</td>
<td>0.19 ± 0.02</td>
<td>0.17 ± 0.02</td>
<td>0.22 ± 0.02</td>
<td>0.19 ± 0.01</td>
</tr>
<tr>
<td>Under-treatment</td>
<td>0.07 ± 0.01</td>
<td>0.12 ± 0.02</td>
<td>0.09 ± 0.01</td>
<td>0.10 ± 0.01</td>
<td>0.09 ± 0.01</td>
<td>0.09 ± 0.01</td>
</tr>
</tbody>
</table>

* Data are given as mean ± SEM. AD indicates advance directive; HCD, health care directive; and VLA, valued life activities directive.
†P < .05 on Dunnett post hoc test.
The fact that instructional ADs were found to be ineffective in other studies.46-49 Because attitudes toward end-of-life care have been found to vary across demographic and ethnic groups,15,50-54 the unique characteristics of our sample must be considered in evaluating our results. Still, the fact that instruction ADs were found to be ineffective in a relatively educated sample of individuals motivated to make plans for their future medical care does not bode well for their success more generally.

It is possible that other AD documents might prove more effective in enhancing the accuracy of substituted judgment. The directives used in our study, however, included treatment instructions much more detailed than the vague expression of wishes contained in most living wills.62 The HCD is a prototypical scenario-based AD with considerable empirical evidence involved in its development.19,20,55-57 The VLA directive was developed from research showing that assessments of health-related quality of life are based on outcome and function24,38,39 as well as suggestions that identifying a threshold for medical usefulness in end-of-life decision making.26,60,61 It could be argued that the VLA directive was too general to allow easy application by surrogates to specific treatment decisions. The HCD, however, was equally ineffective in improving surrogate decisions, even in scenarios that closely resembled specific HCD scenarios (eg, the coma no chance and slight chance scenarios). The fact that 2 detailed, empirically derived, but otherwise quite different instructional directives proved equally ineffective in improving surrogate substituted judgment supports the generalizability of our findings.

It is also possible that a more elaborate or long-term discussion intervention might be effective in enhancing surrogate accuracy. Although intensive and focused, our intervention was a brief, single-session discussion without the guidance of a physician or any explicit educational component. We agree that any ben-

### Table 3. Responses to Perceived Benefit Questions by Experimental Condition

<table>
<thead>
<tr>
<th>Question/Response</th>
<th>No AD (n = 80)</th>
<th>HCD No Discussion (n = 80)</th>
<th>VLA No Discussion (n = 79)</th>
<th>HCD Discussion (n = 80)</th>
<th>VLA Discussion (n = 82)</th>
<th>Total (N = 401)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surrogate understands?</td>
<td>4.36 ± 0.08</td>
<td>4.53 ± 0.07</td>
<td>4.40 ± 0.09</td>
<td>4.84 ± 0.04†</td>
<td>4.75 ± 0.05†</td>
<td>4.58 ± 0.03</td>
</tr>
<tr>
<td>At least pretty well</td>
<td>74 (93)</td>
<td>76 (95)</td>
<td>70 (89)</td>
<td>80 (100)</td>
<td>81 (99)</td>
<td>81 (99)</td>
</tr>
<tr>
<td>Surrogate is accurate?</td>
<td>4.53 ± 0.08</td>
<td>4.55 ± 0.07</td>
<td>4.58 ± 0.08</td>
<td>4.72 ± 0.05</td>
<td>4.56 ± 0.08</td>
<td>4.59 ± 0.03</td>
</tr>
<tr>
<td>At least pretty confident</td>
<td>74 (93)</td>
<td>76 (96)</td>
<td>74 (94)</td>
<td>79 (99)</td>
<td>77 (94)</td>
<td>78 (95)</td>
</tr>
<tr>
<td>Surrogate will honor wishes?</td>
<td>4.19 ± 0.09</td>
<td>4.08 ± 0.10</td>
<td>4.22 ± 0.11</td>
<td>4.35 ± 0.09</td>
<td>4.35 ± 0.10</td>
<td>4.24 ± 0.04</td>
</tr>
<tr>
<td>At least pretty sure</td>
<td>69 (89)</td>
<td>65 (83)</td>
<td>64 (82)</td>
<td>69 (86)</td>
<td>72 (89)</td>
<td>74 (89)</td>
</tr>
<tr>
<td>Surrogate is comfortable</td>
<td>3.68 ± 0.13</td>
<td>3.82 ± 0.11</td>
<td>3.94 ± 0.11</td>
<td>3.99 ± 0.11</td>
<td>4.07 ± 0.11†</td>
<td>3.90 ± 0.05</td>
</tr>
<tr>
<td>At least pretty comfortable</td>
<td>55 (69)</td>
<td>54 (68)</td>
<td>55 (70)</td>
<td>61 (76)</td>
<td>63 (77)</td>
<td>288 (72)</td>
</tr>
<tr>
<td>Having an AD is important?</td>
<td>4.38 ± 0.09</td>
<td>4.53 ± 0.09</td>
<td>4.51 ± 0.09</td>
<td>4.48 ± 0.11</td>
<td>4.57 ± 0.09</td>
<td>4.49 ± 0.04</td>
</tr>
<tr>
<td>At least pretty important</td>
<td>67 (84)</td>
<td>67 (85)</td>
<td>71 (90)</td>
<td>69 (88)</td>
<td>77 (94)</td>
<td>351 (88)</td>
</tr>
<tr>
<td>Surrogate I understand patient’s wishes?</td>
<td>4.36 ± 0.08</td>
<td>4.45 ± 0.07</td>
<td>4.51 ± 0.07</td>
<td>4.61 ± 0.05†</td>
<td>4.54 ± 0.07</td>
<td>4.49 ± 0.03</td>
</tr>
<tr>
<td>At least pretty well</td>
<td>75 (94)</td>
<td>78 (98)</td>
<td>77 (98)</td>
<td>80 (100)</td>
<td>79 (96)</td>
<td>389 (97)</td>
</tr>
<tr>
<td>I am accurate?</td>
<td>4.32 ± 0.07</td>
<td>4.40 ± 0.07</td>
<td>4.48 ± 0.07</td>
<td>4.61 ± 0.06†</td>
<td>4.39 ± 0.09</td>
<td>4.44 ± 0.02</td>
</tr>
<tr>
<td>At least pretty confident</td>
<td>74 (93)</td>
<td>77 (96)</td>
<td>73 (92)</td>
<td>78 (98)</td>
<td>73 (95)</td>
<td>375 (94)</td>
</tr>
<tr>
<td>I will honor patient’s wishes?</td>
<td>4.18 ± 0.09</td>
<td>4.36 ± 0.08</td>
<td>4.18 ± 0.09</td>
<td>4.31 ± 0.08</td>
<td>4.27 ± 0.11</td>
<td>4.26 ± 0.04</td>
</tr>
<tr>
<td>At least pretty sure</td>
<td>71 (90)</td>
<td>70 (88)</td>
<td>64 (81)</td>
<td>72 (90)</td>
<td>69 (86)</td>
<td>346 (87)</td>
</tr>
<tr>
<td>I am comfortable?</td>
<td>3.71 ± 0.11</td>
<td>3.89 ± 0.10</td>
<td>3.84 ± 0.12</td>
<td>4.01 ± 0.10</td>
<td>4.06 ± 0.11†</td>
<td>3.90 ± 0.05</td>
</tr>
<tr>
<td>At least pretty comfortable</td>
<td>50 (63)</td>
<td>57 (71)</td>
<td>56 (71)</td>
<td>59 (74)</td>
<td>60 (73)</td>
<td>282 (71)</td>
</tr>
<tr>
<td>Having an AD is important?</td>
<td>4.35 ± 0.10</td>
<td>4.55 ± 0.09</td>
<td>4.53 ± 0.08</td>
<td>4.71 ± 0.06†</td>
<td>4.59 ± 0.10</td>
<td>4.55 ± 0.04</td>
</tr>
<tr>
<td>At least pretty important</td>
<td>68 (85)</td>
<td>70 (88)</td>
<td>74 (94)</td>
<td>77 (96)</td>
<td>74 (91)</td>
<td>363 (91)</td>
</tr>
</tbody>
</table>

*Data for questions are given as mean ± SEM, and data for responses are given as number (percentage). AD indicates advance directive; HCD, health care directive; and VLA, valued life activities directive.
†P < .05 on Dunnett post hoc test.
‡P < .10 on Dunnett post hoc test.
ecial effects of discussing end-of-life issues will likely
depend on the content and quality of that discussion. Past
correlational research, for example, has produced only
inconsistent evidence of any association between natu-
really occurring patient-surrogate discussion and surro-
gate accuracy. What our results clearly show is
that despite the faith placed by many physicians and
most of our study participants) in the beneficial effects
discussing end-of-life issues, a structured discussion almost certainly more rigorous than most naturally oc-
curring end-of-life discussions produced no discernible
improvements in patient-surrogate understanding.

Another possible limitation of this study is its reli-
ance on a particular set of hypothetical illness sce-
narios. Because the LSPQ was constructed based on an
extensive review of previous surrogate decision-making
research to include a broad range of realistic life-
sustaining treatment decisions, it is unlikely that inclu-
sion of other hypothetical treatment choices would change
the observed results. Still, future research should exam-
ine the effectiveness of ADs with patients facing chronic
or progressive diseases for whom life-sustaining treat-
ment decisions are less hypothetical and more likely to be
discussed end-of-life discussions produced no discernible
improvements in patient-surrogate understanding.

Finally, we adopted a sample size and analysis strat-
ogy that maximized the probability of finding inter-
vention effects with the consequent risk of inflating our
likelihood of type I errors. Analyses were conducted to
confirm the lack of intervention effects across specific sce-
narios and treatments and within different subgroups of
patients, surrogates, and patient-surrogate relation-
ships. The consistency of negative results across analyses makes it unlikely that any clinically important effect
of our interventions was missed.

IMPLICATIONS FOR POLICY AND RESEARCH

The increasing institutionalization of ADs in US law
and medical practice stands in stark contrast to a grow-
ing body of research challenging the effectiveness of
advance care planning to produce specific improvements
in end-of-life medical care. This inconsistency raises
questions about whether policy and law should con-
tinue to encourage the use of ADs and what directions
empirical and ethical analysis should take to help ad-
ress this issue.

One direction for future research is to attempt to de-
velop more effective methods of improving the accu-

racy of surrogate decision making. This research must
move beyond simple documents to long-term interven-
tion strategies and should be guided by basic research in
communication and human decision making rather
than the mixture of common sense, legal analysis, and
political consensus underlying many current ap-
proaches (eg, state-specific living will forms). More-
over, additional consideration needs to be given to the
difficult ethical question of what constitutes an “accept-
able” level of surrogate accuracy. Interpretation of the
present findings differs depending on the level of con-
cordance considered indicative of adequate surrogate
understanding.

Alternatively, a recent review of the cumulative find-
ings of SUPPORT concludes that the initial guiding
assumption of that study—that enhanced patient-level
decision making is the key to quality end-of-life care—
was fundamentally flawed. Research may show that it is
either impossible or prohibitively difficult to improve the
ability of surrogates to predict patient preferences for spe-
cific treatments in specific medical circumstances. Sur-
rogate decision making is likely subject to the same limi-
tations as other forms of “clinical” judgment. The
accuracy of surrogate decisions may also be limited by
the tendency of patients’ life-sustaining treatment pref-
erences to change over time. Surrogates cannot be ex-
pected to predict patients’ future treatment wishes bet-
ter than patients can predict their own.

Even if the accuracy of surrogate decision making has some theoretical or practical limit, however, ADs
may still have an important role in end-of-life decision
making. Although instructional directives may not im-
prove decisions made by family members, they may be
more effective in improving the poorer accuracy ex-
pected from decision makers with little or no past rela-
tionship with the patient such as an emergency depart-
ment physician (see article by Coppola et al in this
issue). Advanced directives are also likely to have a
number of psychological benefits for patients and their
families, particularly if their completion serves to stimu-
late discussion of end-of-life issues. Our discussion inter-
ventions were found to produce a sense of mutual under-
standing and comfort with end-of-life decision making.
These effects were minimal for pairs in which the patient
had previously completed an AD but were more pro-
ounced for pairs in which the discussion of end-of-life
issues was presumably more novel. Improved satisfac-
tion with decision making was also the only identifiable
positive effect of the SUPPORT intervention.

The question that remains is how to evaluate these
psychological benefits in the absence of actual improve-
ments in the accuracy of substituted judgment. Because
law and policy advocating ADs has always been deeply
rooted in the preservation of patient self-determina-
tion, the goal of improving the accuracy of surrogate
decision making is not easily abandoned. At the same
time, however, patients and families facing real decisions about
end-of-life treatment often seem less concerned with the
ability to predict specific treatment decisions than they
are with gaining a general sense of control over the
dying process and reducing the level of burden on sur-
rogate decision makers.

The results of the present study clearly challenge the
effectiveness of instructional ADs as a means of preserv-
ing patients’ ability to control specific treatment deci-
sions near the end of life. What is less clear is the extent
to which the majority of patients and surrogates desire
this level of control and the relative value to assign to the
goals of accurate surrogate decision making vs psy-
chological benefits in future policy development. A fi-
nal evaluation of the wisdom of advocating instruc-
tional directives as a means of improving end-of-life
medical care requires additional debate, and eventually
a broader consensus, regarding the specific outcomes we
hope ADs to achieve.
References


