Accuracy of Primary Care and Hospital-Based Physicians’ Predictions of Elderly Outpatients’ Treatment Preferences With and Without Advance Directives

Kristen M. Coppola, PhD; Peter H. Ditto, PhD; Joseph H. Danks, PhD; William D. Smucker, MD

Background: Past research has documented that primary care physicians and family members are often inaccurate when making substituted judgments for patients without advance directives (ADs). This study compared the accuracy of substituted judgments made by primary care physicians, hospital-based physicians, and family surrogates on behalf of elderly outpatients and examined the effectiveness of ADs in improving the accuracy of these judgments.

Participants and Methods: Participants were 24 primary care physicians of 82 elderly outpatients, 17 emergency and critical care physicians who had no prior experience with the patients, and a baseline comparison group of family surrogates. The primary outcome was accuracy of physicians’ predictions of patients’ preferences for 4 life-sustaining treatments in 9 hypothetical illness scenarios. Physicians made substituted judgments after being provided with no patient AD, patient’s value-based AD, or patient’s scenario-based AD.

Results: Family surrogates’ judgments were more accurate than physicians’. Hospital-based physicians making predictions without ADs had the lowest accuracy. Primary care physicians’ accuracy was not improved by either AD. Accuracy and confidence in predictions of hospital-based physicians was significantly improved for some scenarios using a scenario-based AD.

Conclusions: Although ADs do not improve the accuracy of substituted judgments for primary care physicians or family surrogates, they increase the accuracy of hospital-based physicians. Primary care physicians are withdrawing from hospital-based care in growing numbers, and emergency medicine and critical care specialists most often are involved in decisions about whether to begin life-sustaining treatments. If ADs can help these physicians better understand patients’ preferences, patient autonomy more likely will be preserved when patients become incapacitated.

Arch Intern Med. 2001;161:431-440

Advance Directives (ADs) were created to ensure patient autonomy at the end of life. Ideally, patients who became decisionally incapacitated in their final weeks and days of life could still “voice” their preferences for medical care through a prerecorded document, known as the living will, or through a preappointed decision maker referred to as a proxy or surrogate decision maker. Autonomy is preserved to the extent that life-sustaining treatment decisions that physicians and family members make on behalf of the patient (based on AD information) are the same decisions that the patient would have made for himself or herself.

Considerable attention has been given to the difficulties associated with implementing policies designed to encourage completion and use of ADs. For example, surveys estimate that between 2% and 30% of Americans have actually completed an AD.1,2 Even if patients do have an AD, physicians are unlikely to be aware of patient preferences or of the existence of the document.3,4 Other researchers have reported that ADs are often unavailable when treatment decisions need to be made.5-8 Finally, even if the physician is aware of the AD, the document still may be overruled by physicians or families who must make decisions for the patient, possibly because the information contained in the document is vague or not easily applied in clinical situations.9,10 The fact that ADs are not usually completed, are not available when decisions need to be made, or are ignored or overridden by physicians or family members indicates that there is a problem with policy implementation. The difficulties of
PARTICIPANTS AND METHODS

The data for the present study were collected from physicians of patients involved in phase 1 of the Advance Directives Values Assessment and Communication Enhancement (ADVANCE) project, which examined the ability of ADs to improve the accuracy of decisions made by family surrogates. Primary care physicians and hospital-based physicians made predictions about the life-sustaining treatment preferences of a subset of phase 1 patients.

ELICITATION OF PATIENTS’ PREFERENCES

Participants in phase 1 of the ADVANCE project were recruited from a network of 6 primary care practices, which included 24 primary care physicians, affiliated with Summa Health System in Akron, Ohio. Randomly selected patients 65 years or older were initially contacted by letter introducing the study. Unless patients telephoned the project office to decline participation, trained interviewers telephoned patients to solicit participation. Interviews took place in the patients’ homes and were approximately 1 hour long. A total of 401 patients and their designated family surrogates were interviewed in the ADVANCE project.

Patients completed the Life-Support Preferences—Predictions Questionnaire (LSPQ), which measures patient treatment preferences across a broad spectrum of realistic life-sustaining treatment decisions. The LSPQ began with descriptions of 4 life-sustaining medical treatments chosen to vary in their invasiveness: (1) antibiotics, (2) cardiopulmonary resuscitation, (3) gallbladder surgery, and (4) artificial nutrition and hydration. Patients were read standard descriptions of the treatments including why the treatment is provided and general risks and benefits of the treatments. Nine different medical scenarios were described that were chosen to capture a broad range of conditions varying in their severity, nature of impairment (eg, cognitive vs physical), prognosis, and level of pain: (1) the patient’s current health (current health); (2) Alzheimer disease with moderately severe cognitive impairment and a certain but indeterminate rate of decline (Alzheimer disease); (3) emphysema with constant shortness of breath, severe physical limitations, and a certain but indeterminate rate of progression (emphysema); (4) coma persisting 6 weeks after a stroke with no obvious cognitive abilities, indeterminate rate of decline, and physician opinion of no chance of recovery (coma no chance); (5) the same coma scenario as in scenario 4 with a physician opinion of a very slight chance of recovery (coma slight chance); (6) stroke resulting in partial paralysis, language deficits, total dependence in activities of daily living, and physician opinion of no chance of improvement (stroke no chance); (7) the same stroke scenario as in scenario 6 with a physician opinion of a very slight chance of improvement (stroke slight chance); (8) terminal colon cancer with fatigue, no pain, and a life expectancy of 6 months (cancer no pain); and (9) the same cancer scenario as in scenario 8 with pain that requires the constant use of medication to control symptoms (cancer with pain).

In the patient version of the LSPQ, patients imagined themselves in each medical scenario and indicated their preference for receiving each of the 4 medical treatments (in the current health scenario, the artificial nutrition and hydration question was omitted). Patients indicated their treatment preferences using a 5-point Likert scale: “definitely don’t want treatment,” “probably don’t want treatment,” “unsure,” “probably want treatment,” and “definitely want treatment.” Surrogates were asked to imagine the patient in each medical scenario and predict the patient’s treatment preferences using the same 5-point scale.

SURROGATES’ PREDICTIONS OF PATIENTS’ PREFERENCES

Family Surrogates

A subsample (n=82) of family surrogates from the phase 1 study were used as a baseline comparison group with primary care physician and hospital-based physician accuracy. These surrogates were chosen by the patients in the phase 1 study as the individual whom they would want to make medical decisions for them if they were no longer able. The majority of family surrogates were spouses and adult children of patients. However, “family” is used in a broad sense, including friends and clergy.

As a part of the phase 1 study, patients and family surrogates were randomly assigned to either a control condition in which they did not complete an AD (no-AD) or 1 of 4 intervention conditions in which surrogates made predictions after exposure to a patient-completed AD (completed with or without discussion with the surrogate). The 4 phase 1 intervention conditions were created by the orthogonal application of 2 intervention components: (1) “type of AD,” family surrogates were provided with either a scenario-based AD (the health care directive [HCD]) or an outcome-based AD (the valued life activities directive [VLA]) and (2) “opportunity to discuss,” family surrogates were either provided with the directive without the opportunity to discuss its contents with the patient or were present when the patient completed the AD and were encouraged to discuss the directive with the patient. (Copies of both ADs, as well as a full description of all phases of the ADVANCE project can be obtained from the author, P.H.D., on request.)

The family surrogates reviewed the patient’s AD (when applicable), made predictions on the LSPQ regarding patient preferences, and rated how confident they were that they accurately predicted the wishes of the patient on a 5-point scale.

Primary Care Physicians

All of the primary care physicians (N=24) from the primary care practices affiliated with Summa Health System were contacted and agreed to participate.

Each primary care physician was to complete 5 substituted judgment tasks: predicting preferences of 1 patient who did not complete an AD (the no-AD condition), predicting preferences of 2 patients who had completed the VLA (1 patient who had completed the VLA with discussion with their family surrogate and 1 who had completed the VLA alone), and predicting preferences of 2 patients in the HCD condition (1 patient who had completed the HCD with discussion with their family surrogate and 1 who had completed the HCD alone). Primary care physicians
were told the name of the patient whom they would be making substituted judgments for and were asked to answer questions regarding their relationship with that patient (eg, “How long have you been seeing this person as a patient?”). Four patient names were not recognized by primary care physicians and were not included in the study. Physicians were then instructed to review the patient’s AD (in the VLA and HCD conditions) and to predict their patients’ preferences for life-sustaining treatment on the LSPQ. Those in the AD conditions had the option of reviewing the patient’s directive at any time during the completion of the LSPQ. After completing the LSPQ, the physicians rated how confident they were that they accurately predicted the wishes of the patient on a 5-point scale and how helpful they found the AD. This procedure was then repeated with another patient until the physician had made substituted judgments for 1 patient from each of the 5 phase 1 conditions.

Because patients were randomly assigned to a study condition in phase 1, not all primary care physicians had patients in each of the 5 study conditions. In addition, 47 (12%) patients indicated on a question from their phase 1 participation that they did not want their AD shared with their physician. Because the “discussion” conditions did not directly apply to the physicians (ie, the physicians only reviewed the patients’ ADs and were not present for discussion when the patient completed the VLA or HCD documents) having them make predictions using ADs created from discussion was not essential. To maximize the number of physician predictions, physicians without 5 unique prediction conditions completed as many unique conditions as possible, with the goal of having predictions for 1 no-AD patient, 2 HCD patients (either discussion or no discussion), and 2 VLA patients (either discussion or no discussion) (Figure 1).

Thirteen primary care physicians completed judgments for 5 patients. The following list describes the remaining patient judgments: 1 physician completed judgments for 4 patients, 1 physician completed judgments for 3 patients, 3 physicians completed judgments for 2 patients each, and 5 physicians completed judgments for 1 patient each. The total number of predictions were 83, although only 1 patient was subsequently not included in the study due to incomplete data.

**Hospital-Based Physicians**

Phase 1 of the ADVANCE project had enrolled physicians and patients from 2 hospitals associated with Summa Health System. Physicians from these hospitals who specialized in emergency or critical care medicine (n=17) and spent at least 50% of their time working in a hospital setting were contacted by letter to participate. Each hospital-based physician was yoked to 1 or more primary care physicians to complete predictions for 1 patient in each of the 5 possible AD conditions. For example, if primary care physician 1 had completed predictions for 1 patient in all 5 conditions, then hospital-based physician 1 would complete predictions for the same 3 patients. However, if primary care physician 2 only made predictions for 2 patients in the HCD condition, then hospital-based physician 2 would make predictions for those 2 patients in addition to predictions for another primary care physician’s patients who were in the no-AD and VLA conditions. Therefore, fewer hospital-based physicians were needed to complete all of the predictions.

Hospital-based physicians were provided with basic demographic information about each patient (ie, age, sex, and race) but were blinded to patients’ names. They then reviewed the patient’s AD (when applicable) and made predictions on the LSPQ regarding patient preferences. After completing the LSPQ with predictions of the patient’s preferences, the hospital-based physicians also rated how confident they were that they accurately predicted the wishes of the patient on a 5-point scale and whether they found the AD helpful.

**STATISTICAL ANALYSIS**

Patients’ and surrogates’ responses on the LSPQ were dichotomized into “want treatment” (“definitely want treatment,” “probably want treatment,” or “unsure”) and “don’t want treatment” (“probably do not want treatment” and “definitely do not want treatment”) responses for each of the 35 treatment decisions. On the basis of past research, “unsure” responses were categorized with “want treatment” responses because in most instances the clinical default is to provide treatment unless specifically refused.17,20,24,25 Analyzing data excluding “unsure” responses, treating “unsure” as a third response category, and treating definitely and probably as separate response categories produced no significant differences in study results. Proportion scores were generated for preferences and predictions made in each scenario (by summing the number of “want treatment” responses within each scenario and dividing by the number of treatment decisions in that scenario).

Three measures were generated as indicators of the accuracy of substituted judgments. Predictions were defined as accurate if, for a given treatment decision, the surrogate gave the same dichotomized response as the corresponding patient. Inaccurate predictions were further categorized into “overtreatment errors” (surrogate predicted the patient would want treatment and patient did not want treatment) and “undertreatment errors” (surrogate predicted patient would not want treatment and patient wanted treatment). Two approaches were used to compare the proportion of accurate predictions and of each type of prediction error across the 3 collapsed study conditions (ie, no-AD, HCD, VLA). First, an overall index was created by summing the number of accurate predictions (or alternatively, the number of overtreatment or undertreatment errors) for all treatments in all scenarios and dividing by the total number of predictions. Second, scenario indexes were created by summing the number of accurate predictions (or alternatively, the number of overtreatment or undertreatment errors) within each scenario and dividing by the number of treatment decisions in that scenario.

The overall ability of primary care physicians, hospital-based physicians, and families to predict patients’ treatment preferences was examined initially with repeated measures analyses of variance (ANOVAs) using the overall accuracy, overtreatment, and undertreatment indexes (ie, collapsed across all scenarios and treatments). When significant differences were found for an overall index, individual scenario indexes were then examined. Post hoc comparisons were then conducted using paired t tests with a Bonferroni correction for multiple comparisons.31
Next, accuracy by AD condition was examined for the 2 physician groups. A repeated-measures ANOVA with 1 between-subjects factor (type of physician) and 1 between-subjects factor (AD condition) was conducted on overall proportion accuracy, overall over-treatment, and overall undertreatment scores. For the between-subjects factor, Dunnett post hoc comparisons were used to examine differences between the means within each group. For the within-subjects factor, paired t tests with a Bonferroni correction were used for post hoc comparisons. When significant differences were found for an overall index, individual scenario and treatment indexes were then examined. Identical analyses and post hoc comparisons were conducted on the confidence measure for physician group by AD condition. To evaluate the extent to which significant results could be accounted for by within-physician variation, intraclass correlations were computed for those ANOVAs showing between-group differences. Since all intraclass correlations were nonsignificant, it is unlikely that within-physician variance significantly altered the present results.

Although much research has documented that physicians and family members are often inaccurate when making substituted judgments for patients without ADs, until recently, no research had examined whether ADs improve the accuracy of substituted judgment. In their article in this issue, Ditto and colleagues showed that substituted judgment decisions informed by 1 of 2 types of ADs and enhanced with patient-surrogate discussion did not improve the accuracy of family surrogates’ substituted judgments. No research has examined if ADs can improve the accuracy of primary care physicians asked to make substituted judgments for patients. Arguably, physicians have a greater need for AD information, since prior studies indicate that primary care physicians have poorer accuracy compared with family members.

In addition, AD information may be particularly valuable for physicians who do not have a close relationship with the patient or have not had the opportunity to discuss treatment preferences with the patient. For example, in the emergency department, physicians are often faced with life-sustaining treatment decisions for patients with whom they have no preexisting relationship, and for whom there is no designated or available surrogate. Indeed, as primary care physicians delegate their care of seriously ill patients to hospital-based physicians, more decisions about initiation of life-sustaining treatments will be made by physicians who do not know the patient. The previous evidence concerning accuracy of decision making for well-acquainted surrogates indicates that accuracy would likely be even lower for physicians who are not acquainted with the patient. No research has examined the accuracy of decisions made by hospital-based physicians for unfamiliar patients, how this accuracy compares with primary care physicians’ or family surrogates’ accuracy, or whether ADs play a role in informing these decisions.

This study had 2 major purposes. The first was to examine the accuracy of substituted judgments made by hospital-based physicians for unfamiliar patients in comparison to patients’ primary care physician and family surrogate. The second was to examine whether ADs can improve the accuracy of substituted judgments made by primary care physicians and hospital-based physicians.

**RESULTS**

**SAMPLE CHARACTERISTICS**

Descriptive information about primary care physicians, hospital-based physicians, and family surrogates is presented in Table 1. Primary care physicians were on average 39.8 years old and 59% were male. Most of the primary care physicians were European American (87.5%), married (87.5%), and had been trained as family practice physicians (79.2%). The majority of primary care physicians (76.4%) reported having made life-sustaining medical treatment decisions (ie, to withhold or withdraw treatment) less than 10 times in the past year. Approximately 24% of primary care physicians reported that they had an AD. The primary care physicians had seen the majority of patients in the sample (96.4% of patients) at least once in the past year and about half of the patients had been seen multiple times during the past year.

Similar to the primary care physicians, the hospital-based physicians had a mean age of 39.2 years, were European American (100%), and married (82.4%). Eighty-eight percent were male and the majority (70.6%) specialized in emergency medicine. Hospital-based physicians spend on average 94% of their time working in a hospital setting (range, 60%-100%). Compared with primary care physicians, a higher percentage of hospital-based physicians (35.7%) reported having an AD. The majority of hospital-based physicians (64.7%) reported having made life-sustaining medical treatment decisions (ie, to withhold or withdraw treatment) more than 50 times in the past year.

The demographics of the subsample of family surrogates did not significantly differ from those of the original sample of 401 family surrogates from the phase 1 study. Unlike the physician groups, the family surrogates had a mean age of 62.7 years and the majority were female (64.6%). The majority of family surrogates were European American (91.5%), married (87.8%), and of the Protestant faith (59.8%). Patients and family surrogates had known each other for an average of 45 years. Information about ADs of these surrogates was not collected.

**OVERALL ACCURACY OF SURROGATES’ PREDICTIONS**

We first examined the overall accuracy of predictions of patient preferences, as well as accuracy in predictions by
type of scenario for primary care physicians, hospital-based physicians, and family surrogates. To do this, we collapsed accuracy scores across the intervention conditions. These means are reported in Table 2. Overall, primary care physicians and hospital-based physicians were only accurate for an average of 0.66 and 0.64 of treatment decisions, respectively (Table 2), although accuracy scores differed by scenario. Accuracy scores were highest for “extreme” scenarios (ie, current health, coma with no chance of recovery, and terminal cancer with pain) when most patients wanted all or none of the treatment options.

When we compared the accuracy of the physician groups and family surrogates, we found differences for overall accuracy ($F_{2,162}=12.02$, $P<.001$), and accuracy for the emphysema, stroke with no chance of recovery, stroke with a slight chance of recovery, stroke without pain, and cancer with pain scenarios (all $P<.05$). Post hoc comparisons indicated that for overall scores, family surrogates were more accurate than both physician groups, while the 2 physician groups did not differ from each other. For emphysema, stroke with a slight chance of recovery, and cancer with no pain, family surrogates were more accurate than both physician groups. For the remaining 2 scenarios, only family surrogates were significantly more accurate than the hospital-based physicians.

We next examined the types of errors made by the primary care physicians, hospital-based physicians, and family surrogates. In Table 2, family surrogates primarily made overtreatment errors; primary care physicians consistently made undertreatment errors. Hospital-based physicians made both types of errors, with slightly more being overtreatment errors. Overtreatment errors are slightly masked in Table 2 because means are collapsed across AD conditions.

Comparing overtreatment errors across surrogate type, an overall difference was found ($F_{2,302}=6.76$, $P<.001$), as well as individual scenario differences for the coma no chance scenario, and cancer with and without pain scenarios (all $P<.05$). Post hoc comparisons indicated that for the overall score and for the coma scenario, primary care physicians made the least overtreatment errors. Hospital-based physicians made the most overtreatment errors for both cancer scenarios.

Comparing undertreatment errors across surrogate type, significant differences were found for overall scores ($F_{2,302}=24.18$, $P<.001$), as well as the Alzheimer disease, emphysema, coma slight chance, stroke no chance, stroke slight chance, cancer no pain, and cancer with pain scenarios (all $P<.001$). Post hoc analyses indicated that family surrogates made the least undertreatment errors.

![Figure 1](https://example.com/figure1.png)

**Figure 1. Study design.** ADVANCE indicates Advance Directives Values Assessment and Communication Enhancement; VLA, valued life activities directive; HCD, health care directive; and AD, advance directive.

<table>
<thead>
<tr>
<th>Table 1. Demographic Characteristics of Sample[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Age, mean ± SD, y</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>European American</td>
</tr>
<tr>
<td>African American</td>
</tr>
<tr>
<td>Marital status</td>
</tr>
<tr>
<td>Single</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Widowed</td>
</tr>
<tr>
<td>Religious affiliation</td>
</tr>
<tr>
<td>Protestant</td>
</tr>
<tr>
<td>Catholic</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>With advance directive</td>
</tr>
<tr>
<td>Living will</td>
</tr>
<tr>
<td>DPAHC</td>
</tr>
<tr>
<td>Specialty area</td>
</tr>
<tr>
<td>Family practice</td>
</tr>
<tr>
<td>Internal medicine</td>
</tr>
<tr>
<td>Emergency medicine</td>
</tr>
<tr>
<td>Times physician saw patient in the past year</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1-2</td>
</tr>
<tr>
<td>3-5</td>
</tr>
<tr>
<td>Monthly</td>
</tr>
</tbody>
</table>

[a] Data are given as number (percentage), except for age. DPAHC indicates durable power of attorney for health care. Some categories were not included due to small numbers.

©2001 American Medical Association. All rights reserved.
treatment errors overall. Primary care physicians made
the most undertreatment errors overall, and in the
stroke slight chance and both cancer scenarios. Primary
care physicians made more undertreatment errors than
family surrogates in the Alzheimer disease scenario and
both physician groups made more undertreatment
errors than the family surrogates for the remaining 3
scenarios.

The second purpose of the present study was to exam-
ine whether ADs improved predictive accuracy for pri-
mary care physicians and hospital-based physicians. Over-
all accuracy and accuracy in each scenario were compared
for decisions made without the benefit of the patient's
AD, with a patient’s VLA, and/or with a patient’s HCD.
A significant physician type
condition interaction was
found (F2,79=3.51, P
.05). Post hoc tests indicated that
for primary care physicians, neither type of AD im-
proved the accuracy of substituted judgments over not
having the patient’s AD. However, for hospital-based phy-
sicians, the HCD was effective in improving the accu-
racy of decisions compared with not having AD infor-
mation (Figure 2). As a baseline comparison, family
surrogate accuracy scores by AD condition are indi-
cated in Figure 2 with a dashed line.

At the level of the individual illness scenarios, a sig-
nificant interaction was only found for the coma with no
chance scenario (F2,79 =3.41, P < .05). Post hoc tests indicated that for primary care physicians, neither type of AD improved the accuracy of substituted judgments over not having the patient’s AD. However, for hospital-based phy-
sicians, the HCD was effective in improving the accu-
racy of decisions compared with not having AD informa-
tion (Figure 2). As a baseline comparison, family
surrogate accuracy scores by AD condition are indicated in Figure 2 with a dashed line.

ACCURACY OF SURROGATES’ PREDICTIONS
WITH AND WITHOUT ADs

The second purpose of the present study was to exam-
ine whether ADs improved predictive accuracy for pri-
mary care physicians and hospital-based physicians. Overall
accuracy and accuracy in each scenario were compared
for decisions made without the benefit of the patient's AD, with a patient’s VLA, and/or with a patient’s HCD. A significant physician type × condition interaction was found (F2,79 = 3.51, P < .05). Post hoc tests indicated that for primary care physicians, neither type of AD improved the accuracy of substituted judgments over not having the patient’s AD. However, for hospital-based phy-
sicians, the HCD was effective in improving the accu-
racy of decisions compared with not having AD informa-
tion (Figure 2). As a baseline comparison, family
surrogate accuracy scores by AD condition are indicated in Figure 2 with a dashed line.

At the level of the individual illness scenarios, a sig-
nificant interaction was only found for the coma with no
chance scenario (F2,79 = 3.41, P < .05). Like the overall
scores, primary care physician accuracy was not im-
proved by the ADs, although for the hospital-based phy-
sicians, the HCD significantly improved the accuracy of judgments over the no-AD control condition. Although no other scenario differences achieved statistical signifi-
There was a consistent trend for the HCD to improve hospital-based physicians' accuracy scores over the no-AD and VLA conditions (Table 3). We next compared the types of errors that the physicians made when using the ADs and with no AD. Overall overtreatment errors by surrogate type can be seen in Figure 3. A significant physician type × condition interaction was found for overall overtreatment scores (F(2,79) = 5.57, P < .01), as well as for the coma no chance and stroke no chance scenarios (F(2,79) = 4.19, P < .05 and F(2,79) = 3.33, P < .05, respectively). Post hoc analyses indicated that when using the HCD, hospital-based physicians made fewer overtreatment errors than when they were provided with no AD. These findings were consistent across many of the individual scenarios, although the differences did not achieve statistical significance (Table 3). No significant differences were found for primary care physicians. In addition, no differences were found for the overall undertreatment index by physician type. Therefore, differences in individual scenarios undertreatment indexes by physicians were not examined.

Finally, we examined whether ADs had any influence on the confidence that physicians had in the accuracy of their predictions of patients' preferences (Figure 4). A significant physician type × condition interaction was found (F(2,79) = 4.36, P < .05). Post hoc comparisons indicated that hospital-based physicians were more confident in their predictions using either type of AD compared with no directive. Primary care physicians had no differences in confidence using either directive compared with no directive.

### Table 3. Proportion of Accurate Predictions, Overtreatment Errors, and Undertreatment Errors by Hospital-Based Physicians*

<table>
<thead>
<tr>
<th>Scenario</th>
<th>No AD (n = 16)</th>
<th>VLA (n = 33)</th>
<th>HCD (n = 33)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>0.58 ± 0.04*</td>
<td>0.60 ± 0.03*</td>
<td>0.70 ± 0.03*</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.33 ± 0.05</td>
<td>0.25 ± 0.04</td>
<td>0.12 ± 0.02</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.10 ± 0.04</td>
<td>0.16 ± 0.03</td>
<td>0.18 ± 0.03</td>
</tr>
<tr>
<td><strong>Current Health</strong></td>
<td>0.94 ± 0.03</td>
<td>0.91 ± 0.04</td>
<td>0.97 ± 0.02</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.06 ± 0.03</td>
<td>0.04 ± 0.02</td>
<td>0.02 ± 0.02</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.00 ± 0.00</td>
<td>0.05 ± 0.03</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td><strong>Alzheimer Disease</strong></td>
<td>0.45 ± 0.08</td>
<td>0.58 ± 0.05</td>
<td>0.60 ± 0.07</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.50 ± 0.09*</td>
<td>0.26 ± 0.06*</td>
<td>0.16 ± 0.05*</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.05 ± 0.03</td>
<td>0.16 ± 0.04</td>
<td>0.24 ± 0.06</td>
</tr>
<tr>
<td><strong>Emphysema</strong></td>
<td>0.50 ± 0.10</td>
<td>0.52 ± 0.06</td>
<td>0.64 ± 0.07</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.36 ± 0.10*</td>
<td>0.29 ± 0.06*</td>
<td>0.09 ± 0.03*</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.14 ± 0.08</td>
<td>0.19 ± 0.05</td>
<td>0.27 ± 0.06</td>
</tr>
<tr>
<td><strong>Coma No Chance</strong></td>
<td>0.62 ± 0.11*</td>
<td>0.83 ± 0.06*</td>
<td>0.86 ± 0.04*</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.33 ± 0.11*</td>
<td>0.12 ± 0.05*</td>
<td>0.08 ± 0.03*</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.05 ± 0.05</td>
<td>0.05 ± 0.03</td>
<td>0.06 ± 0.03</td>
</tr>
<tr>
<td><strong>Coma Slight Chance</strong></td>
<td>0.59 ± 0.10</td>
<td>0.55 ± 0.07</td>
<td>0.69 ± 0.07</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.28 ± 0.09</td>
<td>0.16 ± 0.05</td>
<td>0.16 ± 0.05</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.12 ± 0.08</td>
<td>0.29 ± 0.07</td>
<td>0.15 ± 0.06</td>
</tr>
<tr>
<td><strong>Stroke No Chance</strong></td>
<td>0.42 ± 0.19</td>
<td>0.55 ± 0.07</td>
<td>0.68 ± 0.07</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.41 ± 0.10*</td>
<td>0.19 ± 0.06*</td>
<td>0.11 ± 0.05*</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.17 ± 0.09</td>
<td>0.26 ± 0.06</td>
<td>0.21 ± 0.06</td>
</tr>
<tr>
<td><strong>Stroke Slight Chance</strong></td>
<td>0.56 ± 0.09</td>
<td>0.55 ± 0.07</td>
<td>0.55 ± 0.07</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.30 ± 0.09</td>
<td>0.23 ± 0.06</td>
<td>0.08 ± 0.04</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.07 ± 0.07</td>
<td>0.33 ± 0.07</td>
<td>0.38 ± 0.07</td>
</tr>
<tr>
<td><strong>Cancer No Pain</strong></td>
<td>0.55 ± 0.08</td>
<td>0.48 ± 0.06</td>
<td>0.63 ± 0.07</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.38 ± 0.09</td>
<td>0.50 ± 0.06</td>
<td>0.22 ± 0.06</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.08 ± 0.05</td>
<td>0.02 ± 0.02</td>
<td>0.15 ± 0.05</td>
</tr>
<tr>
<td><strong>Cancer With Pain</strong></td>
<td>0.58 ± 0.10</td>
<td>0.55 ± 0.06</td>
<td>0.69 ± 0.06</td>
</tr>
<tr>
<td><strong>Overtreatment</strong></td>
<td>0.31 ± 0.09</td>
<td>0.38 ± 0.07</td>
<td>0.18 ± 0.05</td>
</tr>
<tr>
<td><strong>Undertreatment</strong></td>
<td>0.11 ± 0.08</td>
<td>0.07 ± 0.03</td>
<td>0.13 ± 0.04</td>
</tr>
</tbody>
</table>

*AD indicates advance directive; VLA, valued life activities directive; and HCD, health care directive. Within each scenario, pairwise post hoc comparisons were conducted. Row means ± SEs with different superscripts are significantly different at P < .02.

### COMMENT

Three noteworthy findings are reported in this study. First, without the benefit of ADs, hospital-based physicians' accuracy when predicting elderly outpatients' life-sustaining treatment preferences is significantly lower than that of primary care physicians and family surrogates. Second, although primary care physicians were significantly less accurate in their decisions than family surrogates, neither a value-based nor a scenario-based AD helped to improve the accuracy of their substituted judgments for their patients. Finally, a scenario-based AD improved the accuracy of hospital-based physicians' judgments in some scenarios, such that their level of accuracy reached similar levels of other well-acquainted surrogates.

### ACCURACY OF PHYSICIANS' PREDICTIONS OF PATIENTS' PREFERENCES

Consistent with past research, primary care physicians were not highly accurate in predicting elderly patients' preferences, there was a consistent trend for the HCD to improve hospital-based physicians' accuracy scores over the no-AD and VLA conditions (Table 3).

We next compared the types of errors that the physicians made when using the ADs and with no AD. Overall overtreatment errors by surrogate type can be seen in Figure 3. (Again, baseline scores of family surrogates are indicated with a dashed line.) A significant physicist type × condition interaction was found for overall overtreatment scores (F(2,79) = 5.57, P < .01), as well as for the coma no chance and stroke no chance scenarios (F(2,79) = 4.19, P < .05 and F(2,79) = 3.33, P < .05, respectively). Post hoc analyses indicated that when using the HCD, hospital-based physicians made fewer overtreatment errors than when they were provided with no AD. These findings were consistent across many of the individual scenarios, although the differences did not achieve statistical significance (Table 3). No significant differences were found for primary care physicians. In addition, no differences were found for the overall undertreatment index by physician type. Therefore, differences in individual scenarios undertreatment indexes by physicians were not examined.

Finally, we examined whether ADs had any influence on the confidence that physicians had in the accuracy of their predictions of patients' preferences (Figure 4). A significant physician type × condition interaction was found (F(2,79) = 4.36, P < .05). Post hoc comparisons indicated that hospital-based physicians were more confident in their predictions using either type of AD compared with no directive. Primary care physicians had no differences in confidence using either directive compared with no directive.

©2001 American Medical Association. All rights reserved.
life-sustaining treatment preferences and frequently predicted that patients would not want treatment when patients indicated that they would.15,20,26 Primary care physicians were most accurate in predictions for scenarios that can be considered extremes on a continuum of impairment (ie, current health, coma with no chance of recovery, and terminal cancer with pain). However, for scenarios between the extremes, where perhaps there may not be an overwhelming consensus for or against treatment, accuracy was considerably lower.

Although hospital-based physicians often must make life-sustaining treatment decisions for patients with whom they have no preexisting relationship, little is known about the accuracy of these decisions compared with patients’ recorded wishes. We found that without ADs, hospital-based physicians have considerable difficulty making accurate predictions of patient preferences. Unlike primary care physicians who made undertreatment errors, hospital-based physicians made more overtreatment errors in their predictions. This may be due to the default assumption that in an emergency situation everything should be done to preserve life. Therefore, without explicit directions to withhold treatment, hospital-based physicians presented with an unfamiliar patient provided life-sustaining therapies. This may be particularly problematic given that in this study, hospital-based physicians overtreated in scenarios that involved significant pain, loss of reasoning abilities, and poor prognoses for recovery—situations that many patients consider to be “fates worse than death.”32

Comparing the 3 types of surrogates (family members, primary care physicians, hospital-based physicians), our overall results are consistent with those of others who have found that family surrogates are more accurate than physicians.16,20,21 However, a closer examination of accuracy by type of scenario reveals that primary care physician accuracy is comparable to that of family surrogates in the extreme situations: current health, coma with no chance of recovery, and terminal cancer with significant pain. For these scenarios, primary care physicians were able to predict patient preferences to the same degree of accuracy as family members who have known the patient for 40 or more years. However, considering that hospital-based physicians were also able to achieve high levels of accuracy for the current health scenario and coma with no chance of recovery scenario without knowing the specific patients, it may be that primary care physicians were representing a generalized preference that many members of a community share and not responding to individual patient preferences. It should be noted that the accuracy for hospital-based physicians was slightly inflated in the scenario comparisons with the other surrogates because these means were collapsed across predictions made with and without the ADs.

Although the finding that hospital-based physicians showed low levels of accuracy is not surprising, we were surprised to find that hospital-based physicians without the benefit of ADs were as accurate as the primary care physicians for many of the scenarios. In addition to the “extreme” scenarios, hospital-based physicians’ accuracy reached comparable levels when imagining patients who had emphysema or a stroke. It may be that the hospital-based physicians develop an awareness of what medical treatments most elderly adults would want due to their direct experience with patients who have these health impairments, and their experience in making end-of-life decisions. However, without a direct relationship with the patient, their accuracy is less than that of family members.

**ACCURACY OF SURROGATES’ PREDICTIONS OF PATIENTS’ PREFERENCES USING ADs**

Ditto and colleagues27 recently reported that ADs did not help to improve the accuracy of substituted judgment for decisions made by family members of outpatients. No one has examined whether ADs can improve the accuracy of substituted judgments made by physicians. In this study, as in the Ditto et al study,27 we removed many of the confounds of testing ADs in the field. The AD was available and the physician was engaged in reviewing patients’ treatment preferences. The physicians were instructed to predict the patients’ wishes and not indicate what they would do in a clinical situation. Many physicians openly discussed differences between what they thought the patient would want and what they thought
they would be inclined to do for the patient in an actual clinical situation. Despite these highly desirable, controlled conditions, ADs still did not improve predictive accuracy of primary care physicians.

Advance directives may not aid primary care physicians or family members, but the results of this study support a role for the documents in a hospital setting. The HCD improved hospital-based physicians’ accuracy consistently across all but one of the scenarios, although not all of the differences achieved statistical significance. Decisions made with the VLA also showed this trend, with improvements in accuracy being seen in 4 of the 9 scenarios. With the use of ADs, hospital-based physician accuracy was comparable to that of family members in 4 of the 9 medical scenarios and comparable to primary care physicians’ accuracy in 8 of the 9 scenarios. These results suggest that ADs may be helpful in an emergency when no family member or physician who knows the patient is available, for patients who have no next of kin, or for patients whose primary care physician is not involved in decision making in the hospital.

The HCD, but not the VLA, improved hospital-based physicians’ accuracy by reducing the number of overtreatment errors without also increasing undertreatment errors. This would indicate that the documents did more than remove the prevailing bias to “treat” but actually communicated specific preferences to the physicians. Having either directive increased the hospital-based physicians’ confidence in the accuracy of their decisions. Perhaps the VLA directive was more problematic because the document contains values that do not translate directly into whether to use a particular treatment in a given illness scenario.

LIMITATIONS

Several limitations of the present study should be noted. First, it is possible that the use of hypothetical situations threatens external validity in real clinical situations because it is unclear whether patients’ preferences and physicians’ predictions would be different if faced with an actual illness. This problem is not unique to this study or to other studies about ADs. Indeed, the entire practice of AD completion and use in primary care offices, hospitals, and nursing homes is based on patient preferences in 4 of the 9 medical scenarios and comparable to primary care physicians’ accuracy in 8 of the 9 scenarios. These results suggest that ADs may be helpful in an emergency when no family member or physician who knows the patient is available, for patients who have no next of kin, or for patients whose primary care physician is not involved in decision making in the hospital.

The HCD, but not the VLA, improved hospital-based physicians’ accuracy by reducing the number of overtreatment errors without also increasing undertreatment errors. This would indicate that the documents did more than remove the prevailing bias to “treat” but actually communicated specific preferences to the physicians. Having either directive increased the hospital-based physicians’ confidence in the accuracy of their decisions. Perhaps the VLA directive was more problematic because the document contains values that do not translate directly into whether to use a particular treatment in a given illness scenario.

Another possible limitation was that physicians and patients did not discuss the patients’ preferences while patients completed the ADs. Physicians were unable to ask patients questions concerning the document or have information that was written in the document clarified. Although this may be a limitation of the present study, in a real-life setting most ADs are completed without physicians and physicians are often unaware that patients have ADs. In addition, one study reported poor physician understanding of patients’ preferences even after a structured physician-patient discussion.

Finally, the relatively small sample of physicians may limit the present investigation. The physicians and patients who participated in this study may not be representative of all primary care practices and may be different from other practices that did not participate in the ADVANCE project. The physicians and patients in this study may have had more interest in end-of-life issues and ADs and so were willing to be involved with the ADVANCE project. If the participants in the present study were more informed or interested in ADs, then it could be expected that accuracy would be increased. It could be argued that this would have created an environment that was optimal for agreement between patient and physician. If this study were replicated in other, randomly selected practices, it is possible that that accuracy might be even lower.

CONCLUSIONS

In Sherwin Nuland’s book, How We Die, he stated, I will not allow a specialist to decide when to let me go. I will choose my own way, or at least make the elements of my own way so clear, that the choice, should I be unable, can be made by those who know me best.

The results of this study, as well as earlier literature concerning the accuracy of surrogate decision making, clearly indicate that even those who know patients well are not highly accurate in predicting patients’ life-sustaining treatment preferences. Nuland’s quote implies that there is some way to make preferences explicit to surrogate decision makers in advance of incapacitation. Advance directives, such as the living will and durable power of attorney for health care, have been endorsed by virtually every medical organization and state and federal policy as a tool for this purpose.

Our study showed that ADs do not improve the accuracy of substituted judgments made by primary care physicians but that they do improve the accuracy of hospital-based physicians. These findings suggest several possible roles for ADs. First, ADs may be important in populations of older patients who do not have surrogates or do not have a surrogate who is readily available in an emergency situation. Second, ADs may be useful for patients who will have their care guided by a hospital-based physician rather than their primary care physician. Finally, since the life-sustaining decision process involves several parties, often including the physician providing direct care in the emergency department or intensive care unit as well as the patient’s family, the AD may be a useful tool to improve the quality of communication and increase the level of agreement between physician and family. Given these possible roles of ADs, it is clear that ADs must be available and easily accessible in an emergency situation, and, when clinically appropriate, must be reviewed by hospital-based physicians providing care to acutely ill patients.

Accepted for publication June 6, 2000.

Financial support for this research was provided in part by a grant HS08180 from the Agency for Healthcare Research and Quality (formerly the Agency for Health Care Policy and Research), Rockville, Md, the Applied Psychology Center at Kent State University, Kent, Ohio, and Sigma Xi.

The research was conducted as a part of Dr Coppola’s doctoral dissertation at Kent State University.
The authors gratefully acknowledge Janet Long and Margaret Abernathy for their support in data collection and manuscript preparation and each of the physicians for their participation.

Corresponding author and reprints: Kristen M. Coppola, PhD, Department of Psychology, Monmouth University, West Long Branch, NJ 07764 (e-mail: kcoppola@monmouth.edu) or Peter H. Ditto, PhD, Psychology and Social Behavior, 3340 Social Ecology II, University of California, Irvine, Irvine, CA 92697-7085 (e-mail: phditto@uci.edu).

REFERENCES