Physicians' Predictions of Elderly Outpatients' Preferences for Life-Sustaining Treatment

Jennifer A. Druley, MS; Peter H. Ditto, PhD; Kathleen A. Moore, MA; Joseph H. Danks, PhD; Aloen Townsend, PhD; and William D. Smucker, MD
Kent and Akron, Ohio

Background. Research has shown that physicians are poor predictors of patients' life-sustaining treatment preferences. Our study examined the association between three aspects of physician experience and their ability to accurately predict patients' preferences for two different life-sustaining treatments in the event of two serious medical conditions.

Methods. Seventeen physicians predicted the treatment preferences of 57 patients and then interviewed patients regarding their actual treatment preferences. Physicians' professional experience, length of their relationship with the patient, and experience with direct feedback were measured to determine the association of these factors with the accuracy of the physicians' predictions.

Results. Physicians became more accurate predictors as they interviewed more patients and received direct feedback regarding the accuracy of their predictions ($P < .001$). Residents were more accurate than faculty in predicting patients' preferences ($P < .05$).

Conclusions. Increased experience with life-sustaining treatment discussions improved the physicians' abilities to accurately predict patient preferences. Although possibly resulting from small sample size, neither greater professional experience nor longer relationship with a patient improved the accuracy of physicians' predictions. Future research should examine whether discussing end-of-life issues with patients more often makes physicians more sensitive predictors of patients' life-sustaining treatment preferences.

Key words. Judgment; life-support care; aged; advance directives; physician-patient relations. (*J Fam Pract 1993; 37:469-475*)

The majority of decisions about the use of life-sustaining medical treatments are made by physicians and family members after patients become incapacitated.[1–3] When a patient is no longer able to voice his or her treatment preferences, surrogate decision-makers attempt to make a "substituted judgment" based on their knowledge about the patient's values, goals, and prior treatment preferences. This task is simplified if the patient's treatment preferences are documented in an advance medical directive or through discussion with a physician. Less than 20% of the general public has an advance directive,[4] and even fewer have had discussions with a physician concerning advance directives and related issues.[5,6]

Although family members generally play the dominant role in surrogate life-support decisions,[7] physicians are influential in this decision-making process.[8–10] Because of the influential role played by physicians in life-sustaining therapy decisions, it is important that their ability to accurately predict their patients' treatment preferences be assessed and improved. Most patients believe that their physicians understand and would honor their life-support wishes.[11] Research suggests, however, that physicians' predictions of patient preferences are usually

Submitted, revised, July 15, 1993.

Portions of this study were presented at the annual meeting of the Society of Behavioral Medicine, New York, New York, March 28, 1992; and at the annual meeting of the Midwestern Psychological Association, Chicago, Illinois, May 2, 1992.

From the Department of Psychology, Kent State University, Kent (J.A.D., P.H.D., K.A.M., J.H.D., A.T.), and the Family Practice Center of Akron (W.D.S.), Ohio. Ms Druley and Dr Ditto were principal co-authors. Requests for reprints should be addressed to Peter H. Ditto, PhD, Department of Psychology, Kent State University, Kent, OH 44242.

ISSN 0094-3509

no more accurate than the accuracy expected due to chance alone.[11–14]

Little or no research has gone beyond the documentation of physicians' predictive inaccuracies and attempted to identify factors that may improve substituted judgments. The current study examined the agreement between elderly outpatients' preferences for two types of life-sustaining therapies in two hypothetical medical situations and physicians' predictions of those preferences. In addition to examining overall levels of predictive accuracy, information obtained about patients and physicians from interviews, questionnaires, and patients' medical records were analyzed to determine their relation to physicians' predictive accuracy.

Of particular interest in this study was the relation between physician experience and predictive accuracy. Three aspects of physician experience were measured: (1) length of professional experience (ie, resident vs faculty status), (2) length of relationship with the patient (eg, the number of months the physician had been seeing the patient), and (3) amount of experience with direct feedback regarding the accuracy of life-sustaining treatment predictions. To obtain this final measure, the current study used a novel design: physicians made multiple predictions across the course of the study, receiving accuracy feedback after each prediction. It was predicted that direct and ongoing feedback on the accuracy of life-sustaining treatment decisions would be more beneficial than other more general types of feedback in improving physicians' predictive accuracy.

Methods

Participants

The participants were 57 outpatients and their respective physicians (n = 17) from the Family Practice Center of Akron. The center is a university-affiliated, community-based family practice residency training program. For a 6-month period, all patients aged 65 years and over who had made at least one prior visit to their physician and who either were scheduled for a routine visit at the center or had just completed a routine visit at the center had their medical records screened for exclusion criteria. Exclusion criteria were a spouse's participation in the study and a charted diagnosis of dementia, depression, schizophrenia, severe communication disorder, or terminal illness. The data for this study were collected as part of a larger study examining elderly outpatients' emotional reactions to physician-initiated discussions of advance directives.[15]

Patient Data

Patients eligible for participation were contacted by telephone and invited to participate in a study involving patients' treatment preferences. Participants were told that they would be interviewed by a research assistant before their scheduled physician visit and would spend an extra 15 minutes at the end of their visit discussing health care treatment preferences.

The preexamination interview began with the Short Portable Mental Status Questionnaire (SPMSQ)[16] to assess cognitive impairment (all participants' SPMSQ scores were less than 5) and the Center for Epidemiologic Studies–Depression Scale (CES-D)[17] to assess depressive symptomatology (two potential subjects scored greater than 16 on the CES-D and therefore were excluded). Patients then completed the Medical Outcomes Study (MOS) Short Form,[18] a 20-item index of functional status and perceived health and well-being, and the Multidimensional Health Locus of Control (MHLC) scale,[19] an 18-item scale measuring control-related beliefs specific to health outcomes. Finally, participants completed two instruments developed specifically for this study: a Medical Care Satisfaction Index on which participants indicated their agreement with four statements concerning their level of satisfaction with their medical care (eg, "I feel that my medical needs are in good hands"), and a Life Support Attitude Index on which participants indicated their agreement with four statements about issues of life support (eg, "I believe my physician understands my wishes with respect to life-support measures").

The following information was collected from patients during the preexamination interview or from patients' medical records: age, sex, race or ethnicity, religion, education, number of hospitalizations in the last 5 years, number of medications, number of months since first seen by current physician, and number of visits with current physician. Patients also indicated whether they had a will, a living will, a durable power of attorney, or any other advance care document.

After completing the preexamination interview, patients had their regularly scheduled office visit. Next, the physician interviewed the patient regarding his or her life-sustaining treatment preferences. To assure basic uniformity among participating physicians, all interviews began with the provision of scripted information about the purpose of advance directives, cardiopulmonary resuscitation (CPR), and artificial nutrition and hydration (ANH). Physicians then read two scenarios excerpted from Emanuel and Emanuel's Medical Directive document[20] describing (1) a coma with a low probability of recovery, and (2) an advanced stage of a progressive

dementing illness. Patients stated their preferences regarding the provision of CPR and ANH in each scenario on a 5-point scale ("definitely do not want," "probably do not want," "unsure," "probably want," and "definitely want").[13]

Physician Data

One week prior to each visit with a participating patient, physicians were asked to predict the patient's treatment choices using materials and response scales identical to those used by the patients in the study. The age, sex, and professional status (resident or faculty) of each physician was recorded.

Statistical Methods

Cronbach alpha coefficients were calculated to assess the internal consistency of all scales administered in the pre-examination interview. Alpha coefficients for all scales and subscales were >.6 with the exception of the life-support attitude items (α = .3). Accordingly, the life support attitude items were examined individually in all analyses.

Following the analysis strategy used by Uhlmann et al,[13,14] patient and physician responses to the 5-point treatment preference scale were dichotomized such that "definitely want," "probably want," and "unsure" responses were grouped together as a "want treatment" response, and "definitely do not want" and "probably do not want" responses were grouped together as a "do not want treatment" response. An "unsure" response was treated as a "want treatment" response because, under conditions of uncertainty, the recommended protocol is to treat.[13,14,21]

These dichotomized versions of physician predictions and patient preferences were analyzed in three ways. First, chi-square tests for categoric data were conducted to examine whether, at a group level, patients and physicians made similar treatment decisions. Second, the level of agreement between the treatment decisions made by each patient-physician pair was examined by comparing the observed level of agreement with the level of agreement expected by chance alone using the kappa statistic.[22] Finally, incorrect physician predictions were broken down into "overtreatment" errors (the physician predicts that the patient wants the treatment when the patient actually does not want it) and "undertreatment" errors (the physician predicts that the patient does not want the treatment when the patient actually does want it). McNemar's chi-square test was conducted on the judgments for each of the four scenario-treatment combinations (ie, coma and dementia scenarios and CPR and ANH treatments) to examine whether physicians showed a significant tendency to make one type of error more than another.

To examine the relation between specific patient, physician, and patient-physician relationship characteristics and overall predictive accuracy, a series of ordinary least-squares multiple regression analyses was conducted. The criterion variable was an *overall predictive accuracy index* created for each physician-patient pair by dividing the number of correct dichotomized physician predictions by the total number of predictions made (four). An additional predictor variable was also created to represent the amount of direct feedback physicians had previously received regarding the accuracy of their predictions. Physicians interviewed an average of three patients during the study (range from 1 to 10 patients). Data from each interview were coded with a number from 1 to 10 representing the number of study participants (including the current patient) that the physician had previously interviewed.

Because of the large number of potential predictors relative to sample size, the regression analyses were performed in two stages. For the initial set of analyses, variables were grouped into six categories: (1) patient demographics (eg, patient age), (2) patient psychological health (eg, patient CES-D score), (3) patient physical health (eg, patient MOS subscale scores), (4) patient attitudes regarding life support, (5) physician-patient relationship characteristics (eg, months since the patient's first visit to the physician), and (6) physician characteristics (eg, resident vs faculty status).

A maximum of six variables were included in each category in order to maintain an approximate 10:1 ratio of subjects to variables in any one regression equation.[23] The significant predictors from each category (based on an inclusion criterion of $P < .15$) were then entered into a final equation to determine their relative predictive power. The final regression equation used the more stringent inclusion criterion of $P < .05$.

Results

Demographic Characteristics of Participants

The mean age of the patient sample was 72.0 years. The sample was 70% female, 84% white , and 79% Protestant. Of the 57 participants, 40 (70%) had recorded a will, but only 5 had executed a living will and only 4 had assigned a durable power of attorney.

Four faculty physicians and 13 residents participated in the study. The mean age of the physicians was 31.4

Table 1. Preferences of 57 Patients Regarding Cardiopulmonary Resuscitation (CPR) and Artificial Nutrition and Hydration (ANH) to Sustain Life in Hypothetical Coma and Severe Dementia Scenarios

Condition and Preferred Treatment	Patient Responses				
	Definitely Want	Probably Want	Unsure*	Probably Do Not Want	Definitely Do Not Want
Coma scenario					
CPR	2	5	5	5	40
ANH	5	6	4	8	34
Dementia scenario					
CPR	0	9	1	10	37
ANH	2	5	8	9	33

**"Unsure" was interpreted as "want treatment," in keeping with established protocol when there is uncertainty about the patient's preference.*

years. Of the 17 physicians, 12 were male and all but 1 were white.

Most participants had a well-established relationship with their physician, evidenced by both the mean number of months since their initial visit (29.6 months) and the total number of visits (13.3 visits) to their physician.

Patient Preferences and Physician Predictions

The number of patients and number of physicians giving each response for the two scenarios and two life-sustaining treatments are shown in Tables 1 and 2. In both scenarios, the majority of patients opted to forgo life-sustaining treatment. In response to the question of CPR use, only 12 (21%) and 10 (18%) of the 57 patients gave "want" responses to the coma and dementia scenarios, respectively. In response to the question of administering ANH, 15 (26%) gave "want" responses in each scenario.

In aggregate, physicians were sensitive to patients' desire to forgo life-sustaining therapies. As shown in Table 2, in all four scenario-treatment combinations, physicians' predictions regarding the percentage of patients who would "want" treatment are very similar to the percentage of "want" responses actually given by patients (all four chi-square values $<.84$, all P values $>.35$).

Accuracy of Physician Predictions

A better indicator of physicians' predictive accuracy is the level of agreement between treatment decisions made by each patient-physician pair. The percentage of pairs making concordant decisions in each of the four scenario-treatment combinations and the associated kappa statistic are shown in the left-hand portion of Table 3. Physicians were somewhat more accurate predictors of patients' preferences regarding CPR (72% and 75% accurate predictions for the coma and dementia scenarios, respectively) than ANH (61% accurate for both health states). In none of the four scenario-treatment combinations, however, did physicians' predictions exceed the level of agreement that could be expected due to chance alone ($P > .21$ for all kappas).

A breakdown of patient-physician disagreements into overtreatment and undertreatment errors is shown in the right-hand portion of Table 3. In all four scenario-treatment combinations, disagreements were relatively evenly split between physicians predicting that the pa-

Table 2. Physician's Predictions of Preferences of 57 Patients Regarding Cardiopulmonary Resuscitation (CPR) and Artificial Nutrition and Hydration (ANH) to Sustain Life in Hypothetical Coma and Severe Dementia Scenarios

Condition and Preferred Treatment	Physician Predictions				
	Definitely Want	Probably Want	Unsure*	Probably Do Not Want	Definitely Do Not Want
Coma scenario					
CPR	0	7	3	36	11
ANH	0	9	8	30	10
Dementia scenario					
CPR	0	3	5	34	15
ANH	0	9	12	23	13

**"Unsure" was interpreted as "want treatment," in keeping with established protocol when there is uncertainty about the patient's preference.*

Table 3. Percentage of Physician-Patient Pairs (N = 57) Making Concordant and Discordant Life-Sustaining Treatment Decisions for Cardiopulmonary Resuscitation and Artificial Nutrition and Hydration in Hypothetical Coma and Severe Dementia Scenarios

Condition	Concordant, %	Kappa (±SD)	*P*	Discordant, %	
				Overtreatment Error	Undertreatment Error
Coma scenario					
CPR	72	.10 (±.20)	.30	12	16
ANH	61	.05 (±.16)	.39	21	18
Dementia scenario					
CPR	75	.08 (±.22)	.36	10	14
ANH	61	.12 (±.15)	.21	25	14

SD denotes standard deviation; CPR, cardiopulmonary resuscitation; ANH, artificial nutrition and hydration.

tient would want the treatment when the patient did not and physicians predicting that the patient would not want the treatment when the patient did. McNemar's tests confirmed that in none of the categories were physicians' errors significantly biased in either direction (all *P* values $> .28$).

Factors Associated with Predictive Accuracy

Two variables were statistically significant predictors of physicians' overall accuracy in the final regression equation. These variables accounted for 21% of the total variance.

First, physicians' professional status was significantly related to overall predictive accuracy, with residents showing greater overall predictive accuracy than faculty ($R^2 = .10$, $\beta = -.34$, $t = -2.60$, $P < .05$). Residents accurately predicted patients' treatment preferences in an average of 71% of all cases, whereas faculty physicians predicted accurately in only 61% of all cases. If the predictive accuracy of residents and faculty is examined in each scenario-treatment combination individually, the only significant difference between the predictions of residents and those of faculty occurred in the scenario involving the decision to use CPR if the patient had dementia (mean predictive accuracy = .87 for residents vs .55 for faculty, $\chi^2 = 6.95$, $P < .05$).

Second, the physicians' experience in discussing advance directives was significantly related to overall predictive accuracy ($R^2 = .11$, $\beta = .47$, $t = 3.54$, $P < .001$). Physicians' ability to accurately predict their patients' treatment preferences improved as physicians interviewed more patients and received feedback regarding the accuracy of their predictions. For example, in the first interview, the average physician was able to accurately predict patients' preferences on fewer than half (.45) of the four scenario-treatment combinations. By the third interview, average predictive accuracy jumped to .80 and remained near that level. Experience had the greatest effect on physicians' ability to predict patients' preferences regarding ANH, for which average predictive accuracy increased from .29 in the first interview to over .50 for both the coma and dementia scenarios in later interviews.

To obtain a better understanding of why experience in making treatment predictions improved predictive accuracy, zero-order correlations were calculated between the experience variable and (1) patient responses and (2) physician predictions in each of the four scenario-treatment combinations. In each case, the correlation was calculated using the original undichotomized 5-point scale. Experience was not related to patient preferences in any of the four scenario-treatment combinations. Experience was significantly related to physician predictions in two of the four scenario-treatment combinations. Regarding the decision to administer ANH, for both the coma ($r = .34$, $P < .01$) and dementia scenarios ($r = .38$, $P < .01$), the more patients a physician had interviewed, the greater was his or her tendency to predict that patients would want to forgo artificial nutrition and hydration.

Discussion

The current study concurs with the findings of previous studies, showing that potential surrogate decision-makers have serious deficiencies in their ability to predict patients' life-sustaining treatment preferences.[11–14] Patients in this study showed relatively little interest in receiving either CPR or ANH in either the coma or severe dementia scenario. Although the physicians as a group were sensitive to this preference, their individual predictions for specific patients were no more accurate than the accuracy expected due to chance alone. In the current study, physicians were equally likely to make overtreatment and undertreatment errors in all four hypothetical situations, although they showed a slight ten-

dency to incorrectly predict that patients would want ANH in the dementia scenario. This finding might have been statistically significant if a larger sample size had been used. A crucial issue for future research is to clarify the differential patterns of surrogate prediction errors that have been reported in the literature.[11–13]

This study extends past research by examining whether characteristics of the patient, physician, or physician-patient relationship influence the accuracy of surrogate decisions. Two aspects of physician experience were significant predictors of a physician's overall accuracy as identified in the multiple regression analyses. First, resident status was associated with enhanced predictive accuracy. One possible explanation for this finding is that, although residents have less long-term professional experience than faculty physicians, they may have greater day-to-day experience with patient care and do-not-resuscitate orders than faculty. Care must be taken not to overinterpret this finding, however, given the relatively small number of faculty physicians participating in the study (four) and that the effect applied to only one of the four scenario-treatment combinations.

The other physician factor clearly associated with improved ability to predict patients' treatment preferences was the cumulative experience physicians gained during the course of the study. Unlike previous research on surrogate decision-making, the current study used physicians as "interviewers" of patients and followed physicians as they received direct feedback regarding the accuracy of their predictions. As physicians received this feedback, they seemed to adjust their predictions accordingly. Thus, as feedback was received suggesting that they were overestimating patients' desire for ANH, physicians altered their predictions to correct this overestimation.

The current study had two possible limitations. First, because of the relatively small sample size, it is possible that additional correlates of predictive accuracy were not identified as a result of the lack of statistical power. Second, questions can be raised regarding the validity and generalizability of our conclusion that experience with discussions of advance directives improves physicians' predictive accuracy. An alternative explanation for the improvement observed in physician accuracy with interview experience is that the questions or response options as framed by the physicians became more leading as the study progressed. If this were true, however, patient preferences should have changed with greater physician experience. In fact, only the change in physicians' predictions as experience increased was statistically significant. This suggests that physicians were not altering patients' preferences, but were adjusting their own predictions based on the feedback they were receiving.

It is also unclear whether physicians were learning about general perceptions among elderly patients or whether they were learning only about individual patient responses to the specific study scenarios. The scenarios used represented extreme clinical conditions with poor prognoses, and physicians may have simply learned that in these situations, most patients would not want to have CPR or ANH administered. Future research should include a measure of how often in the past physicians have discussed life-sustaining treatment issues with patients. If this more naturally occurring feedback also led to improved physician predictions, considerable confidence in the effect seen in our study would be gained.

Although patients believe that physicians and family members understand their treatment preferences, research shows that this belief is often unfounded. Physicians play a key role in surrogate decision-making and thus must understand the limitations of their predictive accuracy when advising family members about the use of life-sustaining treatments. The findings from this study suggest that as physicians gain experience discussing life-sustaining treatments, their ability to predict patients' preferences improves. Future research should focus on developing methods to enhance the ability of surrogates to accurately choose the life-sustaining treatments that patients would choose for themselves.

Acknowledgments

This project was supported by grant No. HS07660 from the Agency for Health Care Policy and Research and by a research grant from the Summa Health System Foundation. Additional financial support was provided by the Family Practice Center of Akron's Clinical Research Center, the Kent State University Applied Psychology Center, and the Department of Family Medicine at the Northeastern Ohio Universities College of Medicine (grant D15 PE55048–01 from the Department of Health and Human Services).

Barbara Bailey, MD, Margaret Abernathy, and Michelle Swain assisted in data collection and manuscript preparation.

References

1. Bedell SE, Pelle D, Maher PL, Cleary PD. Do-not-resuscitate orders for critically ill patients in the hospital: how are they used and what is their impact? JAMA 1986; 256:233–7.
2. Bedell SE, Delbanco TL. Choices about cardiopulmonary resuscitation in the hospital: when do physicians talk with patients? N Engl J Med 1984; 310:1089–93.
3. Stolman CJ, Gregory JJ, Dunn D, Levine JL. Evaluation of patient, physician, nurse, and family attitudes toward do not resuscitate orders. Arch Intern Med 1990; 150:653–8.
4. Emanuel LL, Barry MJ, Stoeckle JD, Ettelson LM, Emanuel EJ. Advance directives for medical care: a case for greater use. N Engl J Med 1991; 324:889–95.
5. Lo B, McLeod GA, Saika G. Patient attitudes to discussing life-sustaining treatment. Arch Intern Med 1986; 146:1613–5.

6. Shmerling RH, Bedell SE, Lilienfeld A, Delbanco TL. Discussing cardiopulmonary resuscitation: a study of elderly outpatients. J Gen Intern Med 1988; 3:317–21.
7. Danis M, Southerland LI, Garrett JM, Smith JL, Hielema F, Pickard CG, et al. A prospective study of advance directives for life-sustaining care. N Engl J Med 1991; 324;882–8.
8. Orentlicher D. The illusion of patient choice in end-of-life decisions. JAMA 1992; 267:2101–4.
9. Charlson ME, Sax FL, MacKenzie CR, Fields SD, Braham RL, Douglas G. Resuscitation: how do we decide? A prospective study of physicians' preferences and the clinical course of hospitalized patients. JAMA 1986; 255:1316–22.
10. Lo B, Jonsen AR. Clinical decisions to limit treatment. Ann Intern Med 1980; 93:764–8.
11. Seckler AB, Meier DE, Mulvihill M, Cammer Paris BE. Substituted judgment: how accurate are proxy predictions? Ann Intern Med 1991; 115:92–8.
12. Hare J, Pratt C, Nelson C. Agreement between patients and their self-selected surrogates on difficult medical decisions. Arch Intern Med 1992; 152:1049–54.
13. Uhlmann RF, Pearlman RA, Cain KC. Physicians' and spouses' predictions of elderly patients' resuscitation preferences. J Gerontol 1988; 43:M115–21.
14. Uhlmann RF, Pearlman RA, Cain KC. Understanding of elderly patients' resuscitation preferences by physicians and nurses. West J Med 1989; 150:705–7.
15. Smucker WD, Ditto PH, Moore KA, Druley JA, Danks JH, Townsend A. Elderly outpatients respond favorably to a physician-initiated Advance Directive discussion. J Am Board Fam Pract 1993; 6:473–82.
16. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. J Am Geriatr Soc 1975; 23:433–41.
17. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977; 1:385–401.
18. Stewart AL, Hays RD, Ware JE. The MOS short-form general health survey: reliability and validity in a patient population. Med Care 1988; 26:724–35.
19. Wallston KA, Wallston BS, DeVellis R. Development of the multidimensional health locus of control (MHLC) scales. Health Educ Monogr 1978; 6:161–70.
20. Emanuel LL, Emanuel EJ. The Medical Directive: a new comprehensive advance care document. JAMA 1989; 261:3288–93.
21. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. Deciding to forego life-sustaining treatment: ethical, medical, and legal issues in treatment decisions. Washington, DC: US Government Printing Office: 1983:132–6.
22. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960; 20:37–46.
23. Tabachnick B, Fidell L. Using multivariate statistics. New York: Harper Collins, 1989.