The Accuracy of Surrogate Decision Makers

A Systematic Review

David I. Shalowitz, AB; Elizabeth Garrett-Mayer, PhD; David Wendler, PhD

Background: Clinicians currently rely on patient-designated and next-of-kin surrogates to make end-of-life treatment decisions for incapacitated patients. Surrogates are instructed to use the substituted judgment standard, which directs them to make the treatment decision that the patient would have made if he or she were capacitated. However, commentators have questioned the accuracy with which surrogates predict patients’ treatment preferences.

Methods: A systematic literature search was conducted using PubMed, the Cochrane Library, and manuscript references, to identify published studies that provide empirical data on how accurately surrogates predict patients’ treatment preferences and on the efficacy of commonly proposed methods to improve surrogate accuracy. Two of us (D.I.S. and D.W.) reviewed all articles and extracted data on the hypothetical scenarios used to assess surrogate accuracy and the percentage of agreement between patients and surrogates.

Results: The search identified 16 eligible studies, involving 151 hypothetical scenarios and 2595 surrogate-patient pairs, which collectively analyzed 19,526 patient-surrogate paired responses. Overall, surrogates predicted patients’ treatment preferences with 68% accuracy. Neither patient designation of surrogates nor prior discussion of patients’ treatment preferences improved surrogates’ predictive accuracy.

Conclusions: Patient-designated and next-of-kin surrogates incorrectly predict patients’ end-of-life treatment preferences in one third of cases. These data undermine the claim that reliance on surrogates is justified by their ability to predict incapacitated patients’ treatment preferences. Future studies should assess whether other mechanisms might predict patients’ end-of-life treatment preferences more accurately. Also, they should assess whether reliance on patient-designated and next-of-kin surrogates offers patients and/or their families benefits that are independent of the accuracy of surrogates’ decisions.

Arch Intern Med. 2006;166:493-497

Clinical Practice emphasizes the importance of allowing patients to make their own medical decisions. This approach allows individuals to determine the course of their medical care, thereby respecting patient autonomy. However, this approach also raises concern about how clinicians should make treatment decisions for patients who lack the functional capacity to make their own decisions.

Clinicians currently rely on patient-designated and next-of-kin surrogates to make treatment decisions for incapacitated patients. The Patient Self-Determination Act guarantees patients the right to formally designate a surrogate to make treatment decisions for them if they become unable to make their own decisions. When patients lose the capacity to make their own decisions and have not designated a surrogate, most states have statutes to identify a next-of-kin surrogate for them.

Patient-designated and next-of-kin surrogates are instructed to make decisions based on the substituted judgment standard, which involves making the treatment decision that the patient would have made if he or she were capacitated. Surrogates should use the best interests standard, which directs them to make decisions based on what is in the patient’s best interests, but only when they lack sufficient evidence to determine what decision the patient would have made.

See also page 560

Use of the substituted judgment standard is typically defended on the grounds that it extends patient autonomy, allowing the preferences and values of the patients to guide their medical care even after they lose the ability to make their own treatment decisions. In practice, reliance on surrogates offers an effective way to implement the substituted judgment standard only if patient-designated or next-of-kin surrogates can accurately predict what decisions patients would have made if they were capacitated. Yet, commentators argue that surrogates are “frequently inaccurate,” “disagree at a striking rate” with patient preferences, and are “not better than chance” at predicting the decisions patients
Assessment of Surrogates’ Predictive Accuracy

A comprehensive literature search was conducted using PubMed, the Cochrane Library, and manuscript references for studies published in English between 1966 and 2003 that report quantitative data on how accurately surrogates predict patients’ treatment preferences. We also assessed the impact of the 2 most commonly proposed methods for improving surrogate accuracy.

Methods

Assessment of Surrogates’ Predictive Accuracy

Using a Bayesian approach, a Markov Chain Monte Carlo estimation procedure was employed. The estimates provided are means from the posterior distributions of parameters and 95% credible intervals (CIs) are the 2.5th and 97.5th quantiles of the posterior distributions. Uninformative priors were used. We also considered main effects for both health state and intervention in the regression model, but because of sparseness, these results are not included or discussed. Sensitivity analyses were conducted to determine the influence of individual studies on parameter estimates. Four reanalyses, each of which excluded 1 study from the analysis, were performed. The 4 studies chosen were selected because they had large sample sizes or a large number of scenarios and would therefore be most likely to influence results. A parameter estimate was considered to have been sensitive to a study if the parameter estimate when the study was excluded was not within the 95% CI of the estimate when the study was included.

Results

The 16 studies that were included varied widely in both the number of surrogate-patient pairs sampled (range, 225-122615) and the number of scenarios (range, 115-3025). We systematically analyzed the existing empirical literature on surrogate accuracy to determine how well surrogates predict patients’ treatment preferences. We also assessed the impact of the 2 most commonly proposed methods for improving surrogate accuracy.

The hypothetical scenarios offered patients and their surrogates the option of accepting or refusing the proposed intervention. Nine studies9-16,19,21,23,25 assessed respondents’ confidence in their choices using a Likert scale, which the study authors then collapsed into either “accept” or “refuse” the intervention. Seven studies9-17,19,21,23,25 included uncertainty as a response option, which the study authors categorized as acceptance of the intervention based on recommendations that physicians treat patients under conditions of uncertainty. While this assumption may not be appropriate in all cases, the individual studies did not provide the data necessary to assess surrogates’ accuracy with the “uncertain” responses excluded from the analysis.

Assessment of Methods to Improve Surrogates’ Predictive Accuracy

When patients do not designate a surrogate while they are incapacitated, most states appoint a next-of-kin surrogate for them.2 To assess the accuracy of patient-designated vs legally assigned surrogates, we compared the accuracy data in the 11 studies that asked patients to designate their own surrogates9-13,15-19,21,22,25 with the accuracy data in the 5 studies that assigned patients’ surrogates using the relevant state’s legal hierarchy.9,10,20,21,23,24

To improve surrogate accuracy, many authors recommend that patients discuss their values and treatment preferences with family members or other potential surrogate decision makers.13,19,21 Two of the eligible studies12,25 were designed to assess the effect of such discussions on surrogate accuracy. One of these studies12 was excluded from the original analysis because it reported data from an included sample population but was used in this comparison because it explicitly assessed the impact of prior discussions on surrogate accuracy. No data were considered twice.

Statistical Analysis

Meta-analytic techniques were used to combine results across studies. The β-binomial model was used to estimate the overall percentage of agreement and the agreement within each study. For models assessing differences across health states and interventions, a random-effects grouped logit model was used. The model is essentially a generalized linear model from the binomial family with a logit link, and a random effect is included for each study.

A Bayesian approach was used for both the random-effects and the β-binomial models. The software used in the study was WinBUGS, version 2.0.1 (Medical Research Council Biostatistics Unit, Cambridge, England), which implements a Markov Chain Monte Carlo estimation procedure. The estimates provided are means from the posterior distributions of parameters and 95% credible intervals (CIs) are the 2.5th and 97.5th quantiles of the posterior distributions. Uninformative priors were used. We also considered main effects for both health state and intervention in the regression model, but because of sparseness, these results are not included or discussed.

Sensitivity analyses were conducted to determine the influence of individual studies on parameter estimates. Four reanalyses, each of which excluded 1 study from the analysis, were performed. The 4 studies chosen were selected because they had large sample sizes or a large number of scenarios and would therefore be most likely to influence results. A parameter estimate was considered to have been sensitive to a study if the parameter estimate when the study was excluded was not within the 95% CI of the estimate when the study was included.

Conclusion

The 16 studies that were included varied widely in both the number of surrogate-patient pairs sampled (range, 225-122615) and the number of scenarios (range, 115-3025). We systematically analyzed the existing empirical literature on surrogate accuracy to determine how well surrogates predict patients’ treatment preferences. We also assessed the impact of the 2 most commonly proposed methods for improving surrogate accuracy.
Also, the studies sampled different populations, including terminally ill patients, outpatients from hospital practices, a convenience sample of patients with chronic disease, and women older than 69 years. Fifteen of the 16 studies focused on standard clinical care; the remaining study involved enrollment in clinical research. Description of health states in hypothetical scenarios also varied across studies. Studies did not, for example, use standardized descriptions of coma or dementia when describing scenarios to participants.

Overall, surrogates predicted patients’ treatment preferences with 68% accuracy (95% CI, 63-72). Figure 2 shows the distribution of surrogate accuracy percentages for the individual scenarios. We also assessed surrogate accuracy as a function of the patient’s health state in the individual scenarios (Table 1) and as a function of the proposed intervention (Table 2). Surrogates appear to be most accurate in scenarios involving the patient’s current health (79%; 95% CI, 74-83) and in scenarios involving antibiotics (72%; 95% CI, 66-77). Surrogates appear to be least accurate in scenarios involving dementia (58%; 95% CI, 52-64) and in scenarios involving stroke (38%; 95% CI, 52-64).

Twelve studies assessed the type of error surrogates make when they misjudge patients’ treatment preferences: Three studies found that surrogates tend to err by providing interventions that the patient does not want; 1 study found that surrogates tend to err by withholding interventions that the patient does want; and 8 studies found mixed results or no consistent trend in surrogates’ mistakes.

It has been suggested that surrogates’ projection of their own values onto patients may affect their ability to predict patients’ choices. However, the only study to address this concern concluded that in the absence of explicit prior instructions, projection may assist surrogates in predicting patients’ preferences.

In 11 studies, reporting a total of 108 scenarios, patients designated their own surrogates. In 5 studies, reporting a total of 43 scenarios, investigators assigned the patients’ surrogates using the relevant state’s

Table 1. Surrogate Accuracy by Health State*

<table>
<thead>
<tr>
<th>Health State (No. of Scenarios)</th>
<th>Accuracy, %</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coma (52)</td>
<td>70</td>
<td>63-75</td>
</tr>
<tr>
<td>Dementia (36)</td>
<td>58</td>
<td>52-64</td>
</tr>
<tr>
<td>Current health of patient (33)</td>
<td>79</td>
<td>74-83</td>
</tr>
<tr>
<td>Cancer (15)</td>
<td>62</td>
<td>56-67</td>
</tr>
<tr>
<td>PVS (12)</td>
<td>68</td>
<td>61-74</td>
</tr>
<tr>
<td>Stroke (7)</td>
<td>58</td>
<td>52-64</td>
</tr>
</tbody>
</table>

Abbreviations: CI, credible interval; PVS, persistent vegetative state.
*Health states are ordered by frequency of appearance in scenarios. Only health states described in more than 3 scenarios were considered.

Table 2. Surrogate Accuracy by Intervention*

<table>
<thead>
<tr>
<th>Intervention (No. of Scenarios)</th>
<th>Accuracy, %</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPR (33)</td>
<td>69</td>
<td>64-74</td>
</tr>
<tr>
<td>Intubation (30)</td>
<td>70</td>
<td>64-75</td>
</tr>
<tr>
<td>ANH (30)</td>
<td>69</td>
<td>64-74</td>
</tr>
<tr>
<td>Antibiotics (16)</td>
<td>72</td>
<td>66-77</td>
</tr>
<tr>
<td>Amputation (9)</td>
<td>61</td>
<td>53-70</td>
</tr>
<tr>
<td>Chemotherapy (9)</td>
<td>62</td>
<td>53-70</td>
</tr>
<tr>
<td>Dialysis (9)</td>
<td>67</td>
<td>59-74</td>
</tr>
<tr>
<td>Gallbladder surgery (8)</td>
<td>70</td>
<td>63-76</td>
</tr>
<tr>
<td>Blood transfusion (6)</td>
<td>70</td>
<td>61-78</td>
</tr>
<tr>
<td>Surgery† (5)</td>
<td>62</td>
<td>53-71</td>
</tr>
</tbody>
</table>

Abbreviations: ANH, artificial nutrition and hydration; CI, credible interval; CPR, cardiopulmonary resuscitation.
*Interventions are ordered by frequency of appearance in scenarios. Only interventions offered in more than 4 scenarios were considered.
†The included studies did not specify the type of surgery being offered.
Table 3. Surrogate Accuracy: Effect of Method of Surrogate Selection

<table>
<thead>
<tr>
<th>Method of Surrogate Selection</th>
<th>Accuracy, %</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient designated*</td>
<td>69</td>
<td>63-74</td>
</tr>
<tr>
<td>Legally assigned†</td>
<td>68</td>
<td>59-75</td>
</tr>
</tbody>
</table>

*Surrogates selected by the patient; 108 scenarios, 2068 surrogate-patient pairs.
†Surrogates assigned using legal hierarchy of appropriate state; 43 scenarios, 527 surrogate-patient pairs.

Table 4. Surrogate Accuracy: Effect of Prior Discussion of Patient’s Treatment Preferences and Values

<table>
<thead>
<tr>
<th></th>
<th>Accuracy, %</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ditto et al*</td>
<td>71</td>
<td>69-74</td>
</tr>
<tr>
<td>Without discussion</td>
<td>74</td>
<td>72-77</td>
</tr>
<tr>
<td>Mathies-Kraft and Roberto†</td>
<td>58</td>
<td>55-62</td>
</tr>
<tr>
<td>With discussion</td>
<td>64</td>
<td>60-67</td>
</tr>
</tbody>
</table>

*Scenarios summarized for 9 health states, 315 surrogate-patient pairs.
†Scenarios, 60 surrogate-patient pairs.

Abbreviation: CI, credible interval.

Making end-of-life treatment decisions for patients who have lost the capacity to make their own decisions poses one of the most difficult ethical challenges in clinical medicine. In an attempt to extend patient autonomy, current practice is to rely on surrogates and to instruct them to attempt to make the decision that the patient would have made if he or she were capacitated. Despite widespread acceptance of this practice, the present analysis reveals that patient-designated and next-of-kin surrogates fail to predict patients’ end-of-life treatment preferences accurately in one third of all cases.

The present findings also reveal that the 2 most widely endorsed methods for improving surrogate accuracy are ineffective. Specifically, patient designation of surrogates does not appear to improve surrogate accuracy.29,30 Also, the 2 controlled studies that were designed to assess the impact of prior discussions of patients’ treatment preferences found that these discussions do not improve surrogate accuracy. These findings are consistent with 3 other studies that found unclear impact of prior discussions on surrogate accuracy.20,22 However, none of these 5 studies was controlled, and all of them relied on patient reports of whether a prior discussion had taken place. Therefore, taken together, available data suggest prior discussions of patient preferences do not improve surrogate accuracy.

The present findings also reveal that the ability of surrogates to predict patients’ treatment preferences. However, they also reveal that surrogates are more accurate than physicians in predicting patients’ treatment preferences. Therefore, in the absence of alternative methods, current reliance on surrogates may be defended as the best available method for implementing the substituted judgment standard. Future studies should consider whether there are other ways to improve surrogate accuracy. They should also investigate alternative methods to make treatment decisions for incapacitated patients and evaluate whether these methods more accurately predict patients’ preferences. Finally, future studies should assess the impact of relying on surrogates vs alternative methods and determine whether patients or their families prefer one method over the other.
Our analysis has 5 limitations. First, assessing agreement using the κ statistic was not possible given the included studies’ presentation of data. However, we share skepticism expressed elsewhere regarding the appropriateness of the κ statistic for measuring surrogates’ predictive accuracy.21 Second, some studies classified “uncertain” responses from patients and surrogates as acceptance of the intervention in question, which may have influenced the results. Third, many of the scenarios did not provide possibly relevant data, such as the patient’s chances of reaching the described postintervention health state. While these abbreviated descriptions may mimic clinical uncertainty, they may have led patients and surrogates to interpret the same scenarios in different ways. Fourth, the existing literature focuses primarily on surrogates’ ability to predict patients’ preferences for life-saving interventions. The results may not reflect surrogates’ ability to predict patients’ preferences for nonlife-saving interventions. Fifth, hypothetical scenarios were used to assess surrogate accuracy. While surrogates may perform differently in actual cases compared with hypothetical scenarios, it is impossible to measure surrogate accuracy in actual cases because it is not possible to know the preferences of patients when they are incapacitated.

Accepted for Publication: August 28, 2005.

Correspondence: David Wendler, PhD, Department of Clinical Bioethics, National Institutes of Health, Bldg 10, Room 1C118, Bethesda, MD 20892 (dwendler@nih.gov).

Financial Disclosure: None.

Disclaimer: The opinions expressed are the authors’ own. They do not represent any position or policy of the National Institutes of Health, Public Health Service, or Department of Health and Human Services.

Acknowledgment: We thank Marion Danis, MD, Neal Dickert, BA, Ezekiel Emanuel, MD, PhD, Lindsay Hampson, BA, and Franklin Miller, PhD, for their helpful comments.

CONCLUSIONS

On average, patient-designated and next-of-kin surrogates incorrectly predict patients’ end-of-life treatment preferences in one third of cases. Also, it appears that the 2 most commonly endorsed methods for improving surrogate accuracy—patient designation of a surrogate and prior discussion of treatment preferences with surrogates—are not effective. Assuming one goal of surrogate decision making is to predict what decision the patient would have made, future studies should attempt to identify methods to improve surrogate accuracy. They also should consider novel methods to predict incapacitated patients’ end-of-life treatment preferences. Alternatively, our data could imply that it is time to place less emphasis on predicting patients’ treatment preferences accurately and that we should begin to assess whether patients and their families prefer to rely on surrogates, even when surrogates fail to predict patients’ treatment preferences accurately. Finally, we should try to evaluate the impact that various methods of making treatment decisions has on surrogates, families, and loved ones.

REFERENCES


