A MULTIATTRIBUTE UTILITY ANALYSIS OF LEGAL SYSTEM RESPONSES TO MEDICAL INJURIES

Michael J. Saks
Daniel Strouse
Nicholas Schweitzer
A MULTIATTRIBUTE UTILITY ANALYSIS OF LEGAL SYSTEM RESPONSES TO MEDICAL INJURIES

Michael J. Saks,* Daniel Strouse,** and Nicholas Schweitzer***

INTRODUCTION

The tort system is only one of a number of possible legal responses to the problem of accidental injuries. If an alternative system can better accomplish the desirable goals of the tort system while reducing undesirable side effects, then that alternative should be preferred. Any such choice depends, of course, on an accurate assessment of the relative benefits and harms of the tort system and of the alternatives to it.

Thanks to empirical research conducted over the past several decades, we can be reasonably sure that quite a lot of discussion about medical malpractice and corresponding law proceeds on erroneous assumptions, speculations, and anecdotal impressions about the operation and effects of the existing system. The tort system's most serious problem in the medical malpractice area is not that plaintiffs bring claims too readily and too often, or that the system affords relief too easily, or that the relief granted is overly generous. Rather, the best empirical evidence indicates that medical treatment is a widespread cause of avoidable serious injuries and deaths, that very few victims with an actionable injury bring claims, and that in response to those valid claims the system typically fails to compensate losses fully or, most commonly, fails to provide any compensation at all.1 Thus,
health care system providers and consumers enjoy a windfall of lower costs, paid by others that is far out of proportion to the actual risk of being sued for causing negligent injury or, further, being found liable for such harm.4 Ironically, then, the shortcomings of the medical malpractice system are largely the opposite of the imagery held by most of the press, the public, and the legislative community: The problem of medical injury is large and serious while the tort system touches relatively few cases of actionable injury.5

Alternatives to the tort system are often advanced on the claim (or the hope) that they will bring about various improvements on some dimensions.6 The difficulty is that, so far, most of those alternatives remain either under-studied or untested entirely.7 If nothing else, deliberate and limited experimentation with alternatives, including careful empirical evaluation of the performance and effects of the tested alternatives, should be encouraged. With such evidence in hand, the performance of various alternatives on multiple dimensions could be compared to each other and to the tort system, and some conclusions could be drawn regarding which system would serve society best.

The study reported in this Article is an exploratory attempt to make those comparisons, notwithstanding the limited data available. As a proxy for actual data about the performance of the different systems, we turned to people who have spent important parts of their careers generating and studying the empirical data on medical malpractice and corresponding law, and asked them to use their informed judgments so we could tentatively fill in the knowledge gaps. Then, we aggregated those judgments.8 We thereby acquired data on the positive questions—those exploring the likely impact of the tort system, and its alternatives, on various criteria of effectiveness.

In order to obtain estimates of the relative importance of those system-performance criteria—the normative questions—we also collected data from various groups: ordinary citizens, judges, tort scholars, physicians in high-risk specialties, physicians in low-risk specialties, tort plaintiffs’ attorneys, and tort defense attorneys.9 The responses of these groups were not aggregated, but were examined separately by group for possible insights into the similarities and differences between what different groups believe are the more and less important features of any legal system response to medical injuries. The study we present in this Article is by no means a definitive answer to the questions at hand, but rather it is a model of how one might approach the comparisons that must be made if the alternatives to the tort system are to be helpfully appraised and if the legal policy addressing these alternatives is to be better informed.

II. Multiattribute Utility Analysis

The approach we have taken is known as multiattribute utility analysis (MAUA). MAUA was developed for problems involving multiple outcomes which must be evaluated simultaneously.10 MAUA facil-


2. Externalities are costs that are created by an actor’s conduct and imposed on others. The goal of much legal policy is to shift those externalities back to the actor, thereby compelling the actor to take those into account as part of the cost of the conduct—that is, “internalizing” costs back onto the party that creates them.


5. Ironically, most legislative reforms of medical malpractice have been aimed at further insulating tortfeasors from liability and further reducing compensation.

6. E.g., that they will reduce transaction costs, compensate more victims of injury, or reduce the incidence of injury.

7. See generally Barry Furrow et al., Health Law (2d ed. 2000).

8. Aggregation is supported by the theory that the average of many thermometers gives the most accurate temperature. Here, aggregation will cancel out each person’s random errors and dilute the nonrandom biases of some of their responses.

9. See infra Part IV.B.

tates the rational comparison of an array of different choices (concerning for example, materials, technologies, organizational practices, legal policies) which lead to an array of different outcomes.\textsuperscript{11} MAUA accommodates the use of empirical data when available, or the judgments of experts when empirical data are not available.\textsuperscript{12} MAUA also makes explicit the contribution of known or estimated empirical relationships on the one hand, and normative values on the other hand: which outcomes are valued more and which are valued less?\textsuperscript{13} By weighting the effects by the values associated with each attribute of an option, MAUA results in the calculation of “utilities” which reflect the aggregate desirability of each of the several options compared.\textsuperscript{14}

Consider this illustration: How would you determine the “best” material from which to build a rocket to carry your rover to Mars? Using MAUA you would: (1) define the attributes of interest for the material—strength, weight, cost, retention of properties under changing conditions (e.g., temperature, gravity, radiation), and so on; (2) identify candidate materials and obtain measures of their performance on each of the attributes of interest; (3) obtain measures of the relative importance of each of the attributes; and (4) finally, by weighting the performance of each material on each attribute (step 2) by the importance of each attribute (step 3), obtain a measure of the utility of each material for the task at hand.\textsuperscript{15} The material with the highest summed utility would then be the best choice.\textsuperscript{16}

Similar analyses can be, and have been, carried out in a variety of fields, from materials science to business planning to policy analysis.\textsuperscript{17} In the present study, we used MAUA to compare different possible legal responses relating to the problem of medical injuries, with the goal of identifying those that yielded the highest utilities.

As the Mars rover-rocket example shows, utility is a function not only of the empirical characteristics of a material (or technology or policy)—the attributes or criteria—but also of the values associated with each attribute or criterion. These values, even if empirically informed, are ultimately normative in nature; they represent a judgment that a particular attribute or criterion is more or less important, desirable, or meritorious than another. Because MAUA separates facts from values, the results are better insulated from any preconceived preferences of decisionmakers.\textsuperscript{18} This virtue of MAUA is, of course, likely to be most present when actual empirical relationships are obtained and then weighted by normative ratings which are themselves dictated by relatively objective concerns (e.g., the desire for a material that is strong and light). In a study such as ours, the “facts” are the informed opinions and speculations of experts, and the normative dimensions are not compelled by physical reality, but rather by policy preferences (e.g., the extent of demand for a system that emphasizes deterrence over compensation, or vice versa). Even in such circumstances, however, by separating the “facts” from the values—by collecting data in separate phases of the study from different sets of respondents being asked questions framed to elicit only one or the other kind of information—all respondents are, to some extent, placed behind a veil of ignorance.\textsuperscript{19}

Particularly in regard to social and legal policy, different people or groups might hold different values. Those differences could affect the conclusions reached as to which systems have the greatest utility, and therefore such differences need to be studied. This illustrates one kind of sensitivity analysis to which MAUA should be subjected in order to find robust conclusions.

\section*{III. Medical Injury Response System Options and Evaluation Criteria}

Our MAUA was concerned with the evaluation of different systems of legal responses to medical injury by using a set of criteria widely regarded as relevant to such evaluation.\textsuperscript{20}

\subsection*{A. Systems}

We identified and described six different legal systems through which redress for medical injuries might be approached. Some of the systems currently exist, in varying or partial form, among the states or within particular health care delivery sectors;\textsuperscript{21} others are policy proposals not yet implemented.\textsuperscript{22} We recognize that views differ (albeit not greatly) as to which of the proposed alternatives to traditional tort

\textsuperscript{11} See generally Edwards & Newman, supra note 10; Guttentag et al., supra note 10.
\textsuperscript{12} Edwards & Newman, supra note 10, at 8.
\textsuperscript{13} Id. at 52-53.
\textsuperscript{14} Id. at 65.
\textsuperscript{15} In this instance, that is selecting a material for a rocket.
\textsuperscript{16} Assuming, of course, that everything that went into the calculation is correct.
\textsuperscript{17} See sources cited supra note 10.
\textsuperscript{18} See generally Guttentag et al., supra note 10.
\textsuperscript{19} Those who give empirical ratings do not know what the normative ratings will be and vice versa.
\textsuperscript{20} See infra Part III.B.
\textsuperscript{21} E.g., traditional tort and binding arbitration.
\textsuperscript{22} E.g., no-fault and enterprise liability.
law merit independent consideration, and how best to characterize them. Accordingly, we sought to present a fairly broad range of options. To ensure common understanding among respondents, we provided system descriptions closely reflecting those that follow.

First is the "traditional" American tort system, in which an injured patient seeking compensation must file a judicial claim against the person(s) and institution(s) believed to be responsible. The great majority of claims are resolved prior to trial, but if the claim goes forward, the claimant must ordinarily prove the following: a patient-provider relationship, usually contractual, existed with the physician or other defendant; the provider owed a duty to the claimant; the provider breached the duty by failing to meet the applicable professional standard of care; and the provider's breach caused the claimant's harm and damages. Notwithstanding that complexities and exceptions (and "reforms") exist in connection with many of these key elements, the "traditional" fault-based tort system has long served as the baseline against which proposals for all alternative systems are compared.

The second category is the "modified" or "reformed" tort system. Of course, traditional tort law has never been static. But many state legislatures (and, to a lesser extent, courts), in response to perceived crises in tort litigation generally and in malpractice litigation in particular, have recently modified substantive and procedural rules in hopes of achieving one or more reform goals. Reforms commonly enacted in this period include caps on general damages, shortened statutes of limitations, limits on collateral source recovery, limits on attorneys' contingency fees, changes in the applicable standard of care, periodic (rather than lump-sum) payment of damages awards, elimination of the ad damnum clause in tort complaints, and the like. We recognize that this "system" is a general aggregation—perhaps a "grab bag"—of reforms, and that perhaps not all of these reforms even move in the same legal direction. Nonetheless, we asked respondents to do their best to evaluate "reformed tort" as a general category, identifying (and, in some cases, briefly explaining) the examples noted above.

The third approach, enterprise liability, retains the traditional tort system and its core doctrines, but differs in that only the health care organization—hospital, hospital system, or managed care organization—with which the physician is affiliated can be held liable. Physicians cannot be held personally liable (unless acting without any 23. For example, compare our typology with that in David M. Studdert et al., Health Policy Report: Medical Malpractice, 350 NEW ENG. J. MED. 283, 287–90 (2004). The authors identify three main categories of systemic-reform proposals: (1) alternative dispute resolution (which embraces our sixth system: arbitration); (2) "no-fault" or other alternatives to prove negligence (our fourth and fifth systems); and (3) enterprise liability (our third system). Id. Whereas we treat "reformed tort" as an alternative system (our second), they simply describe key features of tort reform—modifications to the current regime that limit access to court, modify liability rules, and limit the size of awards—all of which are, in substance, elements of our reformed tort "system." Conversely, Studdert and colleagues include (as one reform within their alternative dispute resolution category) the fault-based administrative system of compensation that originated with the American Medical Association (AMA), a proposed reform that we omit. See Studdert et al., supra, at 289 n.90. See also Paul C. Weiler, Medical Malpractice on Trial, 114–22 (1991); Kirk B. Johnson et al., A Fault-Based Administrative Alternative for Resolving Medical Malpractice Claims, 42 VAND. L. REV. 1365 (1989). We chose to omit this alternative because, due to its low cost and accessibility, the AMA has come to recognize that it has the potential to attract a vastly increased number of meritorious, currently unfilled claims, which has led the organization itself to stop promoting it. Mark A. Hall et al., Health Care Law & Ethics 359 (6th ed. 2003).
24. See Weiler et al., supra note 1, at 136.
25. Id. at 14, 136; Studdert et al., supra note 23, at 283.
27. See Weiler et al., supra note 1, at 136.
28. See generally Furrow et al., supra note 7.
organizational nexus). Under this system, institutional liability for practitioner negligence would be vicarious or strict.

Under the fourth approach, a no-fault administrative system, no-fault liability replaces the conventional tort lawsuit with a state administrative proceeding operating under a set of modified rules. Where traditional tort doctrine requires proof of the defendant's negligence (fault)—often the most difficult, uncertain, and contentious part of tort litigation—in no-fault, the claimant generally must instead demonstrate that the harm constitutes a "compensable event" under the scheme, with no need to prove provider negligence. Sometimes this is characterized as a shift to "strict" liability. In many versions, damages, economic and/or non-economic, would be awarded according to a schedule.

The fifth system is an elective no-fault proposal articulated by authors of the Harvard Medical Malpractice Study. Under this approach, health care organizations would be authorized by statute to offer patients, in advance of treatment, no-fault compensation for all provider-inflicted harms in exchange for the patient's relinquishment of any tort claims. No-fault benefits would fully cover medical expenses unpaid by health insurance, eighty percent of net lost earnings (up to two hundred percent of the state's average earnings level), and specified payments for hedonic losses associated with particular physical impairments. Patients, after being fully informed of these terms, could accept them or decline and obtain medical services elsewhere. Effective quality assurance would be required of health care institutions.

The final system we presented is binding arbitration. A form of alternative dispute resolution, binding arbitration generally involves a pre-treatment (in the case of managed care, often a pre-enrollment) agreement to submit injury-based claims to an arbitrator, to be bound by the results of the arbitration, and to relinquish recourse to the courts. Arbitration thus shifts the forum of dispute resolution out of the courtroom and replaces judge and jury with arbitrator(s) as decisionmaker(s). Arbitration is frequently more brief and less formal, but generally retains and applies basic principles of substantive tort law (including liability and damages) and, often, many of its procedures. The approach has been implemented by contract in some American health care settings (e.g., parts of the Kaiser system).

B. Criteria

We identified eight criteria by which to measure the performance of any system for handling medical injuries. All of our criteria reflect widely accepted goals for injury-management regimes: compensation, deterrence, corrective justice, and efficiency. For the effects-of-systems approach would be voluntary, and it would thereby avoiding the hazards of a more or less "blind leap" to a radically different, mandatory system.

---


39. See generally Weiler et al., supra note 1.

40. See id. at 148-52; Weiler, supra note 23, at 151-52 (noting, inter alia, that physicians would be immune from tort liability under this system). As explained by its proponents, a key distinction between this approach and wholesale mandatory no-fault is that the elective no-fault approach would be voluntary, and it would initially "test" no-fault on a relatively small scale, thereby avoiding the hazards of a more or less "blind leap" to a radically different, mandatory system.


43. There is widespread agreement that the three basic goals of tort litigation relating to medical injuries are: compensation of injured patients, deterrence or prevention of harmful conduct and practices, and corrective justice (simply stated, that the one responsible for the harm be the source of compensation for the one injured). See Studdert et al., supra note 23, at 283. See also Weiler, supra note 23, at 44-47; Weiler et al., supra note 1, at 16-19, 20-25. But see id. at 15, 78 (discounting the importance of the "corrective justice" goal because of the widespread existence of liability insurance). Commentators also often include as an important criterion efficiency, i.e., the measure of the costs of the process that detract from the payment of compensation. See, e.g., Furrow et al., supra note 7, at 359, 361. Finally, the patient safety movement, with its emphasis on "non-punitive, systems-oriented, cooperative strategies," is in some tension with the existing malpractice regime (Studdert et al., supra note 23, at 286), but enables consideration of new perspectives on some of the existing criteria.

Accordingly, we included criteria, or measures, of all of the foregoing. The first three of our criteria are measures of "compensation." The fourth through sixth are measures of "deterrence" or "prevention," and also incorporate novel dimensions of the patient safety movement. The seventh is a measure of "justice," and the eighth a measure of "efficiency."
lems phase of data collection, we asked our malpractice researcher-scholar respondents to evaluate how well each of the six systems was likely to perform with respect to each criterion. By contrast, for the importance-of-criteria phase of data collection, we asked all of the respondents to rate the importance of the criteria as measures of the overall performance of any medical injury management regime. These criteria were the following:

1. **Most Nearly Full and Complete Compensation of Those Injured Patients Who Bring Claims:**
   - This criterion is the (most nearly) complete compensation of those who bring valid claims. Compensation, as used here, includes both special losses (e.g., lost earnings, medical expenses) and general losses (e.g., pain, suffering, disability, disfigurement).

2. **Essential Compensation for the Largest Number of Injured Patients:**
   - This criterion is the provision of "economic" compensation (special losses: lost earnings, medical expenses, and so on) to the largest number of injured patients.

3. **Predictable Compensation:**
   - This criterion is the predictability of compensation.

4. **Incentive for Individual Professionals to Avoid the Occurrence of Adverse Events:**
   - This criterion is the incentive for individual professionals to take care not to inflict avoidable injury.

5. **Incentive for Institutions or Industry to Develop Means to Avoid Occurrence of Adverse Events:**
   - This criterion is the incentive for organizational units within the health care industry, or some industry-wide organization, to take whatever steps they can to reduce the incidence of avoidable injuries.

6. **Facilitation of Communication Among Members of Health Care Organizations, So That Adverse Event Prevention Systems Can Be Developed:**
   - This criterion is the facilitation of (rather than interference with) safety regimes designed to prevent adverse medical events, through information flow, organizational restructuring, or other steps.

7. **Corrective Justice:**
   - This criterion is the compensation of an injured person by the individual or institution responsible for inflicting his harm, rather than by someone else (such as by a collective fund or by the injured patient).

8. **Minimization of Transaction Costs:**
   - This criterion is the transfer of compensation dollars from their source to an injured patient with the least loss of dollars to inefficiencies of collection, third party profits, determination of eligibility for the compensation, and so on.

---

44. See infra Part IV.

---

IV. Method

Two separate data collections were necessary to conduct the MAUA. The first involved ratings by experts of the systems concerning their effectiveness on the defined criteria. The second involved judgments by a number of groups regarding the importance of each of those criteria. Two separate web-based surveys were constructed, one for each component of the data collection.

A. Ratings of System Effects

Twenty experts participated in this portion of the study. These experts were chosen based on prior scholarly work that showed that they were highly knowledgeable about the empirical phenomena of medical malpractice systems (e.g., having conducted original empirical research or written reviews of that literature). Accordingly, they were drawn from a national (indeed, an international) domain of such scholars.

The participating experts were directed to a website, where they were asked to rate each of the six systems by their likely effectiveness in achieving the goals defined by eight criteria. Respondents could see definitions of each system and each criterion simply by placing their computer's cursor on the name of the system or criterion they wanted more information about. Respondents were instructed to indicate how well each criterion was likely to be achieved within each system by entering a value on a scale from 0–100, where a score of 100 would indicate that a particular system fully achieved a particular criterion. Thus, expert respondents were asked to make forty-eight systems-by-criteria ratings.

B. Ratings of Importance of Criteria

Members of a number of different groups were recruited to supply normative judgments about the criteria. These groups consisted of attorneys, physicians, tort professors, judges, and laypersons. In addition, lawyers were differentiated into those who primarily represent plaintiffs and defendants, and physicians were differentiated into those in "high-risk" and "low-risk" specialties. Other than the tort professors, who were drawn nationally using the directory of the Association of American Law Schools as the sampling frame, all respondents for this phase of data collection were from Arizona. Attorneys
were sampled from the state bar directory and an invitation to participate was extended to them by email. They were given the choice of completing an enclosed paper questionnaire or going to a web-based questionnaire. Judges were obtained by having hard copy questionnaires handed to them at an annual state judicial conference or in their chambers by the chief trial judges of the two largest counties. They, too, were offered the option of filling out the paper version of the questionnaire or visiting the project website.

Laypersons were non-law faculty and staff from our home university, using the campus directory as the sampling frame. The obtained sample sizes were as follows: eleven plaintiffs’ attorneys, five defense attorneys, twelve physicians in low-risk specialties, eighteen physicians in high-risk specialties, fourteen judges, thirteen tort professors, and twelve laypersons.

Except as noted above, respondents were directed to a website where they were presented with the eight criteria (along with additional information defining and explaining each criterion) and asked to rate the importance of each criterion. More specifically, participants were asked to indicate their "judgment" of each criterion in light of "what is best for the overall benefit of society." Respondents were asked to rate how important they felt each criterion was on a one to ten scale, where ten represented the highest importance.

A MULTIATTRIBUTE UTILITY ANALYSIS

V. Results

The medians of each group's ratings of the respective criteria for evaluating legal system responses to medical injury were used to represent the criteria weights for their group. The medians of the experts' ratings of probable impact on the forty-eight system-by-criterion ratings were used to compute the utilities presented below.

We relied on the medians, rather than means, as the more stable measure of central tendency.

First, we will discuss the differences in values placed on the various criteria by the different groups. Then we will present the results of the utility analysis.

A. Values Placed on the Different Criteria

Figure A presents the median ratings of the different criteria, aggregated across all of the groups. Communication within the medical profession (in order to find ways to reduce the incidence of iatrogenic injuries) and deterrence of industry were rated significantly higher than the other criteria, while corrective justice was viewed as having the lowest value among the various criteria. It is interesting to note that compensation of any kind was not the most highly valued criterion. Instead, mechanisms for reducing the incidence of harm to future patients, such as communication and deterrence, were valued highest.

Also noteworthy is the variability in judgments of the different criteria. The group medians varied the least (showing the greatest consensus) for the importance of facilitating communication and

48. Attorneys were asked an additional question: Approximately what percentage of their cases involve the representation of tort plaintiffs or defendants? We used those responses to divide attorneys into those who principally represented plaintiffs (being those who indicated that at least ninety percent of their cases involved tort plaintiff representation) or attorneys with "defense/balanced" caseloads (those whose caseloads consisted of zero to fifty percent plaintiff representation). For simplicity, we refer to the former group as "plaintiffs' attorneys" and the latter group as "defense" attorneys.

49. The physicians solicited as respondents consisted, in approximately equal numbers, of board-certified specialists in internal medicine and family practice (both of which we characterize as "low-risk" in terms of malpractice liability, and consider together for purposes of data collection and analysis), and neurosurgeons and obstetricians/gynecologists (both of which we characterize as "high-risk," and consider together for purposes of data collection and analysis). To the extent that malpractice premiums across specialties are a measure of exposure, these choices reflect the relatively low or high malpractice premiums paid by Arizona physicians in these several specialties.

50. This language was used to try to orient respondents to concerns broader than their own, or their group's, self-interest.

51. A copy of the web page instructions and answer form are provided in the appendices.

52. Means are the arithmetic average of a set of data. Accordingly, they are sensitive to every datum in the set, including aberrations, errors, and outliers. Medians are the middle score in a distribution. In order for a median to change, data need to move from below to above or vice versa.

53. An 8 (criterion) x 7 (respondent group) mixed-factor analysis of variance (ANOVA) was conducted for this analysis. The different criteria were rated significantly differently from each other. F(7, 525) = 5.60, p < .01.

54. Iatrogenic injuries, injuries caused by the healer, are the subject of this Symposium.

55. Deviation contrasts showed that the mean corrective justice rating is significantly lower than the mean of the other criteria ratings, F(1, 81) = 13.57, p < .001. The communication criterion rated significantly higher than the mean of the other criteria ratings, F(1, 81) = 17.29, p < .001, as did the deterrence of industry criterion, F(1, 81) = 15.23, p < .01. Note that our statistical tests employ raw scores and compare means, while the descriptive statistics we report in the tables below report the group medians. In an exploratory study such as this one, we thought it best to present the more stable medians in the tables of findings, while at the same time conducting more sensitive significance tests in a search for differences.

56. See infra tbl. 1.
providing essential compensation. The greatest variability (least consensus) was observed for the value of providing full and complete compensation to the largest number of injured patients and for corrective justice.

As shown in Table 1, the data can be disaggregated by respondent group, so that we are able to see whether the various groups differentially valued the criteria. Overall, there was no significant difference among groups—that is, some groups did not on average and overall give higher or lower ratings than the others. However, on four of the criteria the respondent groups differed in their value ratings: full compensation, predictable compensation, individual deterrence, and corrective justice.

When valuing full compensation, physicians in both high-risk and low-risk specialties rated full compensation of victims as a less important goal than did plaintiffs' attorneys, defense attorneys, judges, and torts professors. In valuing predictable compensation, plaintiffs' attorneys regarded the goal of predictable compensation as having less value than did laypersons, torts professors, and physicians in both high-risk and low-risk specialties. In addition, judges placed less value on predictability of compensation than did laypersons or physicians in high-risk specialties.

When valuing individual deterrence, plaintiffs' attorneys saw the deterrence of individuals as being significantly more important than did torts professors or physicians in low or high-risk specialties. Physicians in low-risk specialties saw individual deterrence as being less important than did defense attorneys or laypersons.

When valuing corrective justice, plaintiffs' attorneys saw significantly greater value in corrective justice than did physicians in both low-risk and high-risk specialties, judges, or torts professors. Further, both defense attorneys and laypersons valued corrective justice more than physicians did.

We should not lose sight of the variables that the groups did not rate significantly differently: essential compensation, deterrence of error by motivating the medical industry, communication, and efficiency or transaction costs. The lack of differences suggests something of a consensus among the groups on the relative importance of these goals.

Certain other differences among the patterns displayed by the different groups might be worth noting. Laypersons placed relatively high value on all of the attributes of medical accident response systems. Laypersons, as the people with the most unvarnished stake in an effective system of legal response, placed a high value on just about all of the attributes.

57. Medians for facilitating communication ranged from 7.5 to 10 and for essential compensation from 6.5 to 9.
58. Medians for full and complete compensation ranged from 5 to 10 and for corrective justice from 4 to 10.
59. See infra tbl. 1.
60. The groups of respondents did not differ significantly on their overall ratings of the criteria, $F(6, 73) = 1.80, p = .11$.
61. See infra tbl. 1.
62. See infra tbl. 1. A one-way univariate ANOVA, comparing across groups, yields $F(6, 77) = 6.995, p < .001$.
63. All of the comparisons in the preceding two sentences were significant by the least significant difference post-hoc test (LSD) at significance levels ranging from $p < .001$ to $p < .028$.
64. See infra tbl. 1. $F(6, 74) = 3.792, p = .002$.
65. All of the comparisons in the preceding two sentences were significant at levels ranging from $p < .001$ to $p < .032$.
66. See infra tbl. 1. $F(6, 71) = 2.538, p = .025$.
67. All of the comparisons in the preceding two sentences were significant at levels ranging from $p < .002$ to $p < .048$.
68. See infra tbl. 1. $F(6, 78) = 4.996, p = .001$.
69. All of the comparisons in the preceding two sentences were significant at levels ranging from $p < .001$ to $p < .023$.
70. See infra tbl. 1.
71. Note that none of their median ratings was lower than 8.0. See infra tbl. 1.
TABLE 1: MEDIANs OF CRITERIA RATINGS (DISAGGREGATED BY RESPONDENT GROUPS)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DeF/DefLs</td>
<td>8.5</td>
<td>7.5</td>
<td>6.5</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>8.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Pi Ls</td>
<td>10.0</td>
<td>7.0</td>
<td>5.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Lo-Risk Docs</td>
<td>6.0</td>
<td>6.5</td>
<td>6.0</td>
<td>6.0</td>
<td>7.0</td>
<td>8.5</td>
<td>4.0</td>
<td>8.0</td>
</tr>
<tr>
<td>Hi-Risk Docs</td>
<td>5.0</td>
<td>6.5</td>
<td>9.0</td>
<td>7.5</td>
<td>6.5</td>
<td>8.0</td>
<td>8.0</td>
<td>9.0</td>
</tr>
<tr>
<td>Judges</td>
<td>10.0</td>
<td>7.0</td>
<td>5.0</td>
<td>8.0</td>
<td>8.5</td>
<td>7.5</td>
<td>5.0</td>
<td>6.5</td>
</tr>
<tr>
<td>Tott Profs</td>
<td>8.0</td>
<td>9.0</td>
<td>7.0</td>
<td>8.0</td>
<td>9.0</td>
<td>9.0</td>
<td>5.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Laypersons</td>
<td>9.0</td>
<td>8.0</td>
<td>8.5</td>
<td>8.5</td>
<td>9.5</td>
<td>8.0</td>
<td>8.0</td>
<td>8.5</td>
</tr>
</tbody>
</table>

everything, which is perhaps understandable, though "having it all" is likely impossible.72 Professional groups, on the other hand, placed value on the criteria in ways that seem to reflect their more specialized concerns or interests.73

Both plaintiffs' and defense attorneys placed higher value on full compensation and on individual deterrence than did a number of other groups.74 Plaintiffs' attorneys placed lower value on predictable compensation than did several of the other groups.75 Corrective justice was viewed as more important by plaintiffs' attorneys than by most other groups, and more important by defense attorneys than by either physician groups.76 No significant differences emerged between plaintiffs' and defense attorneys in the values they assigned to any of the eight goals.77

Physicians stand out for placing a significantly lower value on full compensation, higher value on predictable compensation, a lower value on deterrence, and lower value on corrective justice78 than most of the other groups. In regard to other criteria, physicians placed their highest value on minimizing transaction costs,79 and their next highest

72. The reluctance to trade away some benefits in order to gain others could be dealt with in future research by requiring respondents to divide a fixed number of points among the criteria. See supra tbl. 1. Perhaps the data empirically confirm George Bernard Shaw's observation about professions vis-a-vis the laity. See generally GEORGE BERNARD SHAW, THE DOCTOR'S DILEMMA act I (Penguin Books 1948) (1911).
73. See supra tbl. 1.
74. See supra tbl. 1.
75. See supra tbl. 1.
76. See supra tbl. 1.
77. The difference between plaintiffs' and defense attorneys is not significant for any single criterion; even for corrective justice the significance level was only p = .45.
78. In addition, a one-way repeated measures ANOVA with deviation contrasts shows that physicians (combined into one large group) rated corrective justice far lower than they rated any other goal. F (1, 27) = 20.10, p < .001. See supra tbl. 1.
79. F (1, 27) = 7.13, p < .05.
80. F (1, 27) = 7.13, p < .05.
81. See supra tbl. 1.
82. See supra tbl. 1.
83. See supra tbl. 1.
84. A multivariate ANOVA confirmed that different groups differed in the pattern of ratings that they gave the different criteria. The sum of the Ω roots derived from criteria ratings was significant. F (48, 438) = 1.77, p < .01, and Pillai's Trace = .975. Subsequent univariate tests indicated that groups differed on their ratings of the four criteria. See supra text accompanying note 61-69.
85. See infra tbl. 2. These correlations are based on the values presented in Table 1. See supra tbl. 1. Thus, they are correlations of group medians, and the sample size of the group is only eight. This should be taken into consideration when interpreting the significance of the correlations. Those correlations with absolute values between .887 and .658 are significant at p < .10; those between .607 and .517 are significant at p < .20. All correlations reported in the text are significant at p < .05, which we suggest is still quite conservative given the sample size on which the correlations are based.
employing different sets of criteria was to see the extent to which the utilities associated with different systems are robust and whether the order of preferences remains stable as various criteria are used in calculating utilities.

1. All Criteria

Overall—that is, combining the values of all groups—Administrative No-Fault and the Harvard No-Fault proposal accumulated the highest total utilities. At the other end of the spectrum, with the lowest utility, was Modified-Reformed Tort. This system was found to have the lowest utility by the values of each and every group—though it may be that different groups disdain it for different reasons.

### Table 3: Utilities Derived from All Criteria

<table>
<thead>
<tr>
<th>System</th>
<th>Def Docs</th>
<th>Pl Ls</th>
<th>Hi-Risk Docs</th>
<th>Judges</th>
<th>Tort Docs</th>
<th>Laypersons</th>
<th>Torts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tort</td>
<td>3645</td>
<td>3835</td>
<td>2695</td>
<td>3008</td>
<td>3125</td>
<td>3005</td>
<td>3655</td>
</tr>
<tr>
<td>Mod/Ref Tort</td>
<td>2860</td>
<td>2960</td>
<td>2253</td>
<td>2525</td>
<td>2475</td>
<td>2475</td>
<td>2925</td>
</tr>
<tr>
<td>Enterprise Liab</td>
<td>3345</td>
<td>3460</td>
<td>2660</td>
<td>2935</td>
<td>2960</td>
<td>2900</td>
<td>3460</td>
</tr>
<tr>
<td>Modified Ref Tort</td>
<td>4005</td>
<td>4010</td>
<td>3560</td>
<td>3802</td>
<td>3643</td>
<td>3800</td>
<td>4330</td>
</tr>
<tr>
<td>No-Fault Admin</td>
<td>3705</td>
<td>3715</td>
<td>3260</td>
<td>3492</td>
<td>3358</td>
<td>3485</td>
<td>3988</td>
</tr>
<tr>
<td>Harvard No-Fault</td>
<td>3288</td>
<td>3420</td>
<td>2663</td>
<td>2943</td>
<td>2908</td>
<td>2870</td>
<td>3440</td>
</tr>
<tr>
<td>Binding Arb</td>
<td>2525</td>
<td>2475</td>
<td>2475</td>
<td>2925</td>
<td>18.476</td>
<td>21.552</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>20,848</td>
<td>21,400</td>
<td>17,121</td>
<td>18,005</td>
<td>18,469</td>
<td>18,545</td>
<td>21,801</td>
</tr>
</tbody>
</table>

Next, we disaggregate by respondent groups, through whose eyes the utilities of the different systems are seen to differ. Examination of Table 3 suggests the following pattern of differences. Based upon the values of defense attorneys, physicians in high-risk specialties, judges, tort professors, and laypersons, the systems with the greatest

### B. Utilities

Recall that each system-by-criterion rating of efficacy was weighted by the importance of that criterion, and then the utilities were summed across criteria to produce a utility for each system. Utilities were calculated for each system not only by the lights of each group providing normative ratings, but also by using different sets of criteria. These variations are reflected in the tables below. The purpose of

---

86. See supra tbl. 2. The correlation with defense attorneys is significant at \( p < .10 \), while that with plaintiffs’ attorneys is significant at \( p < .20 \).
87. See supra tbl. 2.
88. See supra tbl. 2.
89. See supra tbl. 2.
90. See infra tbl. 3–7. In the significance tests reported below, the significance of repeated-measures effects are calculated using the Greenhouse-Geisser Correction, since the sphericity assumption has been violated. All between-subjects pairwise comparisons are made using the Tukey HSD test. All within-subjects pairwise comparisons are made using the Bonferroni Correction.
91. See infra tbl. 3 (this analysis included all eight individual criteria).
92. See infra tbl. 3.
93. See infra tbl. 3.
94. A one-way repeated measure ANOVA with a deviation contrast was conducted on respondents’ individual utilities. Both Administrative No-Fault, \( F (1, 83) = 634.14, p < .001 \), and Harvard No-Fault, \( F (1, 83) = 240.73, p < .001 \), were of significantly more utility than the remaining systems. Modified-Reformed Tort was found to be of significantly less utility than the other systems, \( F (1, 83) = 2158.85, p < .001 \).
95. A 6 (system) x 7 (respondent group) mixed-factor ANOVA was conducted on the calculated utilities. A significant interaction was found, meaning that each system’s calculated utility (and its relative position among the different systems) differed between the respondent groups, \( F (30, 385) = 3.64, p < .001 \).
utility, in descending order, were: Administrative No-Fault, Harvard No-Fault, and Conventional Tort. The values of plaintiffs’ attorneys led to the same three systems having the greatest utility, except the order was slightly changed: Administrative No-Fault had the greatest utility, followed by Conventional Tort, and then Harvard No-Fault. Based upon the values of physicians in low-risk specialties, Administrative No-Fault had the greatest utility and Harvard No-Fault the second greatest; the utility of other systems did not come anywhere near these two systems.

2. Major Criteria (Indexed)

For this analysis, we created a single index of compensation (by averaging the ratings given to the three compensation criteria), a single index of deterrence (by averaging the ratings given to the two deterrence criteria), and used only those two indices to reflect the two major criteria: compensation and deterrence. The third “major criterion” employed for this analysis was transaction costs.

Table 4: Utilities Derived from Major Criteria: Compensation, Deterrence, Efficiency

<table>
<thead>
<tr>
<th>Def La</th>
<th>Pl. Ls</th>
<th>Lo-Risk Does</th>
<th>Hi-Risk Does</th>
<th>Judges</th>
<th>Tort Prof</th>
<th>Laypersons</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tort</td>
<td>940</td>
<td>926</td>
<td>877</td>
<td>906</td>
<td>923</td>
<td>928</td>
<td>1064</td>
</tr>
<tr>
<td>Mod/Ref Tort</td>
<td>804</td>
<td>773</td>
<td>799</td>
<td>816</td>
<td>813</td>
<td>824</td>
<td>945</td>
</tr>
<tr>
<td>Enterprise Liab</td>
<td>1096</td>
<td>991</td>
<td>1099</td>
<td>1058</td>
<td>1021</td>
<td>1037</td>
<td>1158</td>
</tr>
<tr>
<td>No-Fault Admin</td>
<td>1533</td>
<td>1513</td>
<td>1614</td>
<td>1705</td>
<td>1594</td>
<td>1628</td>
<td>1866</td>
</tr>
<tr>
<td>Harvard No-Fault</td>
<td>1396</td>
<td>1377</td>
<td>1458</td>
<td>1539</td>
<td>1446</td>
<td>1476</td>
<td>1691</td>
</tr>
<tr>
<td>Binding Arb</td>
<td>1025</td>
<td>1010</td>
<td>1045</td>
<td>1099</td>
<td>1049</td>
<td>1307</td>
<td>1224</td>
</tr>
<tr>
<td>Totals</td>
<td>6704</td>
<td>6610</td>
<td>6802</td>
<td>7143</td>
<td>6846</td>
<td>6960</td>
<td>7978</td>
</tr>
</tbody>
</table>

Using those three criteria, over all respondent groups, the utility derived for each system was found to be significantly different from the utility for every other system. As is evident from Table 4, Administrative No-Fault and Harvard No-Fault had the highest utilities, considerably outdistancing the other systems. Modified-Reformed Tort had the lowest utility of all the systems examined and Conventional Tort had the next to lowest utility. The pattern described above and shown in Table 4 held for every one of the respondent groups separately.

3. Major Criteria (Indexed) Plus Communication

This analysis used the same criteria as the analysis in the immediately preceding subsection, except that ratings of communication were included in the utility calculations.

Table 5: Utilities Derived from Major Criteria Plus Facilitation of Communication

<table>
<thead>
<tr>
<th>Def La</th>
<th>Pl. Ls</th>
<th>Lo-Risk Does</th>
<th>Hi-Risk Does</th>
<th>Judges</th>
<th>Tort Prof</th>
<th>Laypersons</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tort</td>
<td>1340</td>
<td>1356</td>
<td>1217</td>
<td>1246</td>
<td>1223</td>
<td>1289</td>
<td>1444</td>
</tr>
<tr>
<td>Mod/Ref Tort</td>
<td>1254</td>
<td>1243</td>
<td>1182</td>
<td>1219</td>
<td>1150</td>
<td>1229</td>
<td>1372</td>
</tr>
<tr>
<td>Enterprise Liab</td>
<td>1506</td>
<td>1491</td>
<td>1434</td>
<td>1485</td>
<td>1396</td>
<td>1487</td>
<td>1563</td>
</tr>
<tr>
<td>No-Fault Admin</td>
<td>2233</td>
<td>2213</td>
<td>2209</td>
<td>2300</td>
<td>2119</td>
<td>2258</td>
<td>2331</td>
</tr>
<tr>
<td>Binding Arb</td>
<td>1525</td>
<td>1510</td>
<td>1470</td>
<td>1524</td>
<td>1424</td>
<td>1517</td>
<td>1699</td>
</tr>
<tr>
<td>Totals</td>
<td>9,854</td>
<td>9,760</td>
<td>9,480</td>
<td>9,821</td>
<td>9,208</td>
<td>9,796</td>
<td>10,970</td>
</tr>
</tbody>
</table>

Once again, over all respondent groups, Administrative No-Fault and Harvard No-Fault showed the highest utility, while Modified-Reformed Tort had the lowest, and Conventional Tort had the second lowest utility. The disaggregated results for each respondent group reproduce the overall pattern described above and shown in Table 5.

100. See supra tbl. 4. A 6 (system) x 7 (respondent group) mixed-factor ANOVA was conducted on the individual utilities of the criteria included for this analysis. A main effect of system was found: different systems had significantly different utilities, F (5, 390) = 13.324, p < .001. Pairwise comparisons with the Bonferroni Correction revealed that each system had significantly different utilities from every other system.

101. See supra tbl. 4. No main effect of respondent group was found, F (6, 78) = 0.74, p = ns. No system x group interaction was found, F (30, 390) = 1.14, p = ns, meaning that by each group's value judgements, the patterns of utilities were the same across groups.

102. See infra tbl. 4.

103. See supra tbl. 5. A 6 (system) x 7 (respondent group) mixed-factor ANOVA was conducted on the individual utilities of the criteria included for this analysis. A main effect of system was found: different systems had significantly different utilities, F (5, 390) = 1.44, p < .001. Pairwise comparisons with the Bonferroni Correction revealed each system had significantly different utilities from every other system.

104. See supra tbl. 5. No main effect of respondent group was found, F (6, 78) = 1.94, p = ns. No system x group interaction was found, F (30, 390) = 1.45, p = ns, meaning that by each group's value judgements, the patterns of utilities were the same across groups.
4. Major Criteria (Indexed) Plus Corrective Justice\textsuperscript{106}

In this analysis, corrective justice replaced communication as the fourth criterion in the model.

**Table 6: Utilities Derived from Major Criteria Plus Corrective Justice**

<table>
<thead>
<tr>
<th></th>
<th>Def La</th>
<th>Pl La</th>
<th>Lo-Risk Docs</th>
<th>Hi-Risk Docs</th>
<th>Judges</th>
<th>Tort Pros</th>
<th>Laypersons</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tort</td>
<td>1580</td>
<td>1726</td>
<td>1197</td>
<td>1546</td>
<td>1523</td>
<td>1329</td>
<td>1704</td>
<td>10,405</td>
</tr>
<tr>
<td>Mod/Ref Tort</td>
<td>1204</td>
<td>1293</td>
<td>999</td>
<td>1236</td>
<td>1063</td>
<td>1074</td>
<td>1345</td>
<td>8,224</td>
</tr>
<tr>
<td>Enterprise Lab</td>
<td>1406</td>
<td>1491</td>
<td>1209</td>
<td>1458</td>
<td>1271</td>
<td>1287</td>
<td>1588</td>
<td>9,710</td>
</tr>
<tr>
<td>No-Fault Adm</td>
<td>1773</td>
<td>1813</td>
<td>1774</td>
<td>1945</td>
<td>1744</td>
<td>1774</td>
<td>2106</td>
<td>12,843</td>
</tr>
<tr>
<td>Harvard No-Fault</td>
<td>1636</td>
<td>1677</td>
<td>1578</td>
<td>1797</td>
<td>1596</td>
<td>1626</td>
<td>1931</td>
<td>11,823</td>
</tr>
<tr>
<td>Binding Arb</td>
<td>1505</td>
<td>1610</td>
<td>1285</td>
<td>1579</td>
<td>1349</td>
<td>1367</td>
<td>1704</td>
<td>10,399</td>
</tr>
</tbody>
</table>

Overall, the now-familiar pattern holds—with Administrative No-Fault and Harvard No-Fault having the greatest utility and Modified-Reformed Tort the lowest—though Conventional Tort and Binding Arbitration rose to the tier of highest utility after the No-Fault systems.\textsuperscript{107} Examination of the disaggregated patterns across groups indicated more variation from respondent group to respondent group, but no clear pattern of differences emerged.\textsuperscript{108}

5. Selected Individual (Highest Valued) Criteria\textsuperscript{109}

This analysis employed the three individual criteria that received the highest ratings from all of the respondent groups combined (with the limitation that only one criterion was included reflecting each gen-

\textsuperscript{106} See infra tbl. 6.

\textsuperscript{107} See supra tbl. 6. A 6 (system) x 7 (respondent group) mixed-factor ANOVA was conducted on the individual utilities of the criteria included for this analysis. A main effect of system was found: different systems had significantly different utilities, F(5, 390) = 428.60, p < .001. Pairwise comparisons with the Bonferroni Correction revealed that each system had significantly different utilities from every other system.

\textsuperscript{108} See supra tbl. 6. A main effect of respondent group was found, F(6, 78) = 2.72, p < .05. A system x group interaction was found, F(30, 390) = 3.18, p < .01, meaning that by the values of different groups, the pattern of systems' utilities differed. Pairwise comparisons using the Tukey HSD test showed that, although a significant main effect of respondent group exists, no individual group is significantly different from any other groups. Although the system x group interaction is significant, the Administrative No-Fault system had the highest utility across all groups (all ps < .05). The Modified-Reformed Tort system had the lowest utility across all groups (all ps < .05).

\textsuperscript{109} See infra tbl. 7.

\textsuperscript{110} This is another way to avoid double counting where we had a "family" of criteria—which we did with compensation and deterrence—as an alternative to the "indices" of these measures used above. See supra note 109 and infra notes 111-114 and accompanying text.

\textsuperscript{111} The forced inclusion of transaction costs is based on our assumption that any serious reform of the medical injury compensation system will be sensitive to costs and aim to gain the most benefits with the least wasted resources.

\textsuperscript{112} See infra tbl. 7. See also infra fig. A.

\textsuperscript{113} See infra tbl. 7. A 6 (system) x 7 (respondent group) mixed-factor ANOVA was conducted on the individual utilities of the criteria included for this analysis. A main effect of system was found: different systems had significantly different utilities, F(5, 390) = 848.52, p < .001. Pairwise comparisons with the Bonferroni Correction revealed that each system was significantly different from every other system.

\textsuperscript{114} See supra tbl. 7. A main effect of respondent group was found, F(6, 78) = 2.45, p < .05. A system x group interaction was found, F(30, 390) = 2.06, p < .05, meaning that by the values of different groups, the pattern of systems' utilities differed. Pairwise comparisons using the Tukey HSD test showed that, although a significant main effect of respondent group exists, no individual group's values produced system utilities different from any other group's values. Although the system x group interaction was significant, the Administrative No-Fault system had the highest utility across all groups (all ps < .05). The Modified-Reformed Tort system had the lowest utility across all groups (all ps < .05).
VI. Conclusion

With respect to judgments of the importance of the various criteria, respondents were concerned primarily with the prevention of injuries. The single most important criterion for respondents was facilitating communication within the medical community. This criterion enables the development of reforms aimed at reducing the incidence of iatrogenic injury. Close behind in importance was deterrence to motivate injury prevention. This presents a challenge to designers of reforms, because while it is thought that elimination of a fault-based system would facilitate communication, such systems also are likely to diffuse responsibility in ways likely to vitiate deterrence. In any event, it is clear that respondents generally placed high value on steps that would prevent iatrogenic injuries—with the exception of high-risk physicians (for whom the most important goal is the minimization of transaction costs and the predictability of compensation) and judges (for whom the principal goal of the law’s response to medical injury should be the full compensation of victims).

Through most of the above analyses, Administrative No-Fault and Harvard No-Fault emerged as the systems with the highest utilities, suggesting that they are so robust as to withstand variations in choice of criteria and among the respondent groups. Conventional Tort and Binding Arbitration approach the utility of the No-Fault alternatives only when corrective justice is included among the criteria used to compute utilities. Enterprise Liability approaches the No-Fault alternatives only when the highest valued criteria are used to compute the utilities.

As noted at the outset, these findings should not be regarded as providing any sort of definitive answer to the questions addressed. Rather, this project is a demonstration of how one could usefully proceed to fashion the most beneficial kind of system for responding to the problem of medical injuries, in light of the criteria that are most important for assessing such systems.
APPENDIX B

QUESTIONNAIRE FOR VARIOUS GROUPS RATING THE IMPORTANCE OF THE SPECIFIED CRITERIA

[Image of questionnaire form]

When you have finished, underline the box below if you think the form is complete.

[Signature]