DOCTORS, DAMAGES AND DETERRENCE: AN ECONOMIC VIEW OF MEDICAL MALPRACTICE

WILLIAM B. SCHWARTZ, NEIL K. KOMESAR

R-2340-NIH/RC
JUNE 1978

Rand
SANTA MONICA, CA 90406
DOCTORS, DAMAGES AND DETERRENCE: AN ECONOMIC VIEW OF MEDICAL MALPRACTICE

WILLIAM B. SCHWARTZ, NEIL K. KOMESAR

R-2340-NIH/RC
JUNE 1978

Rand
SANTA MONICA, CA. 90406
The research described in this report was supported in part by a grant to Dr. Schwartz from the Robert Wood Johnson Foundation, Princeton, New Jersey; in part by a general research support grant (5 S07 RR05710-07) from the National Institutes of Health; and in part by The Rand Corporation from its own research funds. The opinions, conclusions, and proposals in the text are those of the authors and do not necessarily represent the views of the Robert Wood Johnson Foundation or the National Institutes of Health.
The growth during the 1970s in the number of malpractice suits and in the size of damage awards has led to widespread attacks on the medical malpractice system. It is the thesis of the present study—being published simultaneously as a report by The Rand Corporation and as an article in the June 8, 1978, issue of The New England Journal of Medicine—that damages awarded in a medical malpractice suit should be viewed not only as compensation for victims but also as a means of deterring health-care providers from negligent behavior. Economic analysis of the malpractice system indicates that such awards can send a "signal" to providers which informs them how much to invest in avoiding mishaps.

The work reported here was sponsored in part by a grant to Dr. Schwartz from the Robert Wood Johnson Foundation, in part by a general research support grant (5-S07-RR05710-07) from the National Institutes of Health to The Rand Corporation, and in part by Rand from its own research funds.

Dr. Schwartz is Vannevar Bush University Professor and Professor of Medicine, Tufts University, and Principal Program Adviser to the Rand Health Sciences Program. Dr. Komesar is Professor of Law, University of Wisconsin Law School.

*The two pieces are identical except for a few minor editorial changes and the addition to the report of a preface and two footnotes.
ABSTRACT

Damages awarded in a malpractice suit must be viewed not only as compensating the victim but also as deterring health-care providers from negligent behavior. Economic analysis of the malpractice system indicates that awards can send a signal to providers that informs them how much to invest in avoiding mishaps.

The malpractice system is beset by difficulties, but not the ones commonly incriminated. The signal to the physician, as determined by the number of claims and the size of awards ("expected damages"), appears to be insufficient for ideal deterrence. Moreover, the deterrence signal is attenuated because malpractice premiums are set for groups of physicians, not for individuals according to their record of previous malpractice incidents.

Replacing the present tort system with a no-fault insurance scheme would not necessarily be cheaper, and might well abolish the deterrent signal or distort clinical decisionmaking.
ACKNOWLEDGMENTS

The authors wish to acknowledge the valuable criticisms and comments of Drs. Patricia Munch, Joseph P. Newhouse, Charles E. Phelps, Rodney T. Smith, and Albert P. Williams of The Rand Corporation, Dr. William Bennett of Harvard University, Professor Paul I. Joskow of the Massachusetts Institute of Technology, Professor Robert C. Ellickson of Stanford Law School, and Professor Richard A. Posner, The University of Chicago Law School.
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PREFACE</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>ECONOMIC ANALYSIS OF NEGLIGENCE</td>
<td>1</td>
</tr>
<tr>
<td>MEDICAL MALPRACTICE AND PROFESSIONAL CUSTOM</td>
<td>4</td>
</tr>
<tr>
<td>RECEIVING THE SIGNAL: THE PHYSICIAN'S RESPONSE</td>
<td>5</td>
</tr>
<tr>
<td>THE IDEAL SIGNAL</td>
<td>6</td>
</tr>
<tr>
<td>THE SIGNAL IN THE REAL WORLD</td>
<td>9</td>
</tr>
<tr>
<td>Size of Awards</td>
<td>11</td>
</tr>
<tr>
<td>Number of Claims</td>
<td>11</td>
</tr>
<tr>
<td>Administrative Costs</td>
<td>12</td>
</tr>
<tr>
<td>Erroneous Findings</td>
<td>13</td>
</tr>
<tr>
<td>MALPRACTICE INSURANCE AND ATTENUATION OF THE SIGNAL</td>
<td>14</td>
</tr>
<tr>
<td>CONTINGENCY FEES AND SIGNAL QUALITY</td>
<td>16</td>
</tr>
<tr>
<td>ALTERNATIVES TO MALPRACTICE</td>
<td>18</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>21</td>
</tr>
</tbody>
</table>
DOCTORS, DAMAGES, AND DETERRENCE: AN ECONOMIC VIEW OF MEDICAL MALPRACTICE

Attacks on the malpractice system are widespread and intense. According to some critics, too many claims are brought, and excessive damages are awarded. Others decry the lawyers' contingency fee as an inducement to unwarranted suits, or assert that the system penalizes physicians randomly. These, and other criticisms, have led to demands for major reform—even to proposals that the tort system be replaced wholly or in part by other methods for dealing with medical injury. Curiously enough, the perceived failings and proposed correctives are rarely subjected to critical scrutiny; rather, they are taken as self-evident.

These critiques of the existing system rest largely on an assumption that the primary or sole purpose of malpractice law is to compensate injured patients. But it can be argued, on the basis of modern economic analysis, that this rationale is unduly narrow: findings of negligence are seen not only as redressing past wrongs but also as giving providers an incentive to avoid future careless injuries. Viewed in this way, the malpractice system and its problems dramatically change character.

ECONOMIC ANALYSIS OF NEGLIGENCE

As early as 1881, Oliver Wendell Holmes, Jr., in a classic treatise on the subject, observed that compensation should not be viewed as the rationale of negligence law. Because individuals could, if they wished, insure themselves against accidents, he could see little reason for a government-operated system of compensation. More recently, critics of the compensation rationale have pointed out that patients could be indemnified against medical accidents through insurance programs, private or social, at a much lower administrative cost than that exacted by tort proceedings. Indeed, the legal apparatus required to determine negligence and place blame is so expensive that only about 35 cents of every dollar in malpractice premiums is paid to
successful claimants; \(^{12}\) by contrast, 80 cents of every premium dollar reaches patients with health insurance.\(^ {13}\)

Even if negligence litigation were not expensive, as a means primarily to compensate the innocent injured, it would have puzzling features: two different individuals may sustain an identical injury and be equally innocent of its cause, and yet their prospects for receiving damages may be entirely different. Eligibility for compensation is determined by the behavior of the person responsible for the injury; so if negligence is present in one case and not in the other, one person will be compensated and the other will not.

The negligence system makes a great deal more sense if it is understood primarily as a means to deter careless behavior rather than to compensate its victims. By finding fault and assessing damages against the negligent provider, the system sends all providers a signal that discourages future carelessness and reduces future damages.

The concept of negligence is fundamental to malpractice law, but the usual definitions of negligence, even in law books, suffer from vagueness. Legal formulations, such as "conduct which involves an unreasonably great risk of causing damage" or "conduct which falls below the standard established by law for the protection of others against unreasonably great risk of harm"\(^ {14}\) are intuitively appealing. But such definitions only substitute, by implication, a term like "unreasonable conduct" for "negligence"; they do not specify what "unreasonable" means.

Some 30 years ago, the distinguished jurist Learned Hand formulated an explicit definition of negligence, \(^ {15}\) which has now become a textbook definition.\(^ {14}\) Negligence occurs, he stated, whenever it would cost less to prevent a mishap than to pay for the damages predicted to result from it. More explicitly, the cost of preventing a mishap (C) must be less than the probability that it will occur (P) multiplied by the loss suffered when it does occur (L). Restating the rule in symbols: negligence occurs whenever \( C < P \times L \). The product \( P \times L \) is usually called "expected loss" or "expected damages."

In many situations it is impossible to prevent all injuries, and the Learned Hand Rule must, therefore, be modified as follows: negligent
behavior is the failure to invest resources up to a level that equals the anticipated saving in damages.

There is an intuitive appeal to the notion that "unreasonable" or "careless" behavior can be equated with underinvesting in accident avoidance. The harmony between common sense and this interpretation of the law is best illustrated by an example. A customer walking down the aisle of a supermarket knocks a jar of baby food to the floor, where it shatters. A few seconds later, another customer turns into the aisle, fails to see the puddle on the floor, slips, and breaks a leg. It would not have been "reasonable" to expect that the owner of the market would station personnel in each aisle to offer continuous warnings about wet spots and to provide for instant cleanup. In this case, the cost of preventing the injury would far exceed the expected loss ($P \times L$). Common sense would almost certainly lead a jury to conclude that negligence had not occurred. But if the wet spot had been left for thirty to forty minutes, the verdict would be likely to go the other way, because the cost of policing the aisles at such an interval is relatively low, and the likelihood of injury occurring during the longer interval is considerably greater. As the cost of preventing a mishap rises and the likelihood becomes more remote, failing to avoid it becomes increasingly reasonable.

More than just an explicit formulation of what is "fair" and "reasonable," the Learned Hand Rule serves to assure that resources are being efficiently allocated. It does so by establishing procedures that minimize the total cost incurred by accidents and accident prevention.

As the supermarket cases illustrate, someone who is injured is not automatically entitled to compensation, even though everything about his injury is identical to that of another person who is compensated. What matters is that someone was negligent. Thus, litigation, beyond providing a means to redress the loss and suffering caused by carelessness, signals potentially negligent individuals that it will cost them more to be careless than to invest in an appropriate level of prevention. Damages awarded to a victim induce potentially negligent individuals to compare the cost of avoiding an injury with the cost of
paying for it.* The importance of compensation in this view of the system is that it provides the victim with an incentive to bring suit; only thus can the signal be initiated. Compensation, in this sense, is an indispensable ingredient of the deterrence mechanisms, even though compensating the victim is not its main purpose.

The Learned Hand Rule has a rather obvious corollary, which some people find difficult to accept: there are accidents which, from society's point of view, are not worth avoiding. The cost of prevention far exceeds the expected loss; so the accident should be allowed to occur.

MEDICAL MALPRACTICE AND PROFESSIONAL CUSTOM

Because judges and juries may find it exceedingly difficult to evaluate a medical dispute, the courts have largely accepted "customary standards of medical practice" as a measure of reasonable investment in mishap avoidance.¹⁴ The physician can plead that he has followed customary practice in his care of the patient, and the courts will usually substitute expert testimony about professional custom for a more specific accounting by the negligence standard.

This substitution reflects the faith of the legal system that physicians as a group, or at least those who set standards, are correctly investing in mishap avoidance and that medical custom can be used as the benchmark of adequate performance. And it conforms to the view, held both by providers of health care and by observers of the system, that most physicians tend to guard the welfare of their patients without need for external regulation. The malpractice system exists to discipline the occasional physician who does not (or cannot) protect his patients.

* A simple example will illustrate this point: a power lawn mower causes a severe injury to the leg of an operator. This particular type of accident could be eliminated if a foot guard were added to each lawn mower at an ultimate cost to the manufacturer of $10 per machine. Without the protective device, the probability of the injury is one for every 100 lawn mowers sold, and the damages entailed by each injury are assumed to be $2000. The expected loss ($P \times L$) is $0.01 \times 2000$, or $20$ per lawn mower. Therefore, the cost of prevention, $10$, is less than the expected loss, and the manufacturer is found to be negligent. If, on the other hand, the ultimate cost of adding the foot guard were $30$ per lawn mower, negligence could not be established.
In the face of conflicting testimony as to what constitutes customary and acceptable practice, the courts may undertake more explicit (if informal) application of the Learned Hand Rule. Even if a defendant can show that his performance has met the accepted standard of medical practice, the courts retain ultimate responsibility for determining whether that standard also constitutes "reasonable care." And they may find that it does not.

For example, custom has been rejected as the defense in a suit brought by a 32-year-old woman who became blind from glaucoma because tonometry was not performed. The appellate court weighed the cost of the test ("inexpensive" and "harmless") against the probability that a young person would develop glaucoma ("one in 25,000") and the magnitude of the loss (blindness). The court concluded that "reasonable prudence" required the test and faulted the profession for not routinely carrying out tonometry in patients under the age of 40. Here, custom did not meet the standard of efficiently allocating resources; the court measured custom against the Hand standard and set custom aside. This decision, when expressed quantitatively, yields a result consistent with the Learned Hand Rule. If we take cost of tonometry as not exceeding $5 and the average jury verdict for total or legal blindness at $678,000 (for period 1973-1977), we arrive at the following formulation:

\[ \$5 < 1/25,000 \times \$678,000 , \]
\[ \$5 < \$27. \]

No reasonable adjustment of the figure for the cost of avoidance would alter this conclusion. Cases like this one indicate that professional custom serves only as a provisional substitute for the negligence rule.

**RECEIVING THE SIGNAL: THE PHYSICIAN'S RESPONSE**

As with negligence law in general, the malpractice system may be understood as a mechanism for signaling the potentially negligent—in this case, a subpopulation of physicians. The signal, when properly received, informs the doctor how much to invest in avoiding mishaps; the "correct" response, then, is for him to invest his resources up to the level of expected damages.
To the extent that the usual fee for a service accurately reflects the training and the investment of time needed to meet customary standards of care, non-negligent physicians are appropriately remunerated. By contrast, the negligent physician, who fails to provide the full service but accepts the full price, is shortchanging his patient. With an effective malpractice signal, the potentially negligent physician would be stimulated to invest more time for no increase in pay because he probably could not set his fees higher than those of his more competent colleagues.

In practice, the negligent physician may modify his behavior in one of several ways. The doctor who tends to skimp on history or physical examination or to rush through procedures must take the time needed for more careful work. But increasing his investment of time on each case may be insufficient. An inadequately trained physician is notified by the damages award that he should invest in further training. The cost of training must then be amortized over future cases. Alternatively, a physician may abandon procedures that he is not competent to perform, even though these procedures are relatively more remunerative than others in his practice.

THE IDEAL SIGNAL

The ideal negligence signal is achieved only when every significant incident of malpractice leads to a claim and every valid claim to a full award. Then (in an ideal or, so to speak, "frictionless," system) the physician is stimulated to invest appropriately in mishap reduction.

Let us consider a theoretical world in which there is a doctor with 100 patients in his practice. He works only to maximize his income and is completely aware of these facts: the incidence of a given mishap, the cost to him of reducing that incidence, the loss to him if the mishap occurs, and the proportion of such mishaps that lead to a suit and award.

To reduce the occurrence of mishaps he must invest his own resources, specifically his time, to prevent each occurrence. As with any attempt to approach a zero incidence of accidents, each further
reduction in probability is more expensive than the previous one. So, as shown in Figure 1, reducing the frequency of a mishap from .05 to .04 costs more than reducing it from .06 to .05.

If we look at it the other way around, this is the familiar phenomenon of diminishing returns: every additional $5 spent to reduce the frequency of a mishap buys a smaller reduction than the previous $5 did. In real life, for example, repeatedly checking a cross-match of bloods or rereading a label before administering a drug would cost the same each time it was done, but would yield progressively smaller increases in accuracy as perfection was approached.

Against this background, let us examine the hypothetical situation illustrated in Figure 1. In this example, the physician knows that a particular injury will result in damages of $5000. Thus, every time he reduces the probability of mishap by one percentage point (an average of one patient per 100) he saves $50 in expected loss per patient (.01 × $5000). He also knows that reducing the frequency of a $5000 mishap from .10 to .09 costs $5, as shown by the vertical bar, but it yields an expected saving of $50, as indicated by the dashed horizontal line. For each successive decrement of .01, the costs are higher, but the saving remains constant at $50 for each .01 reduction. When the probability is reduced from .06 to .05 the cost is $50, but this expense exactly equals the saving in damages. At this point of indifference, further investment will yield no further saving. To stop here and invest no more in protection would not, therefore, be negligent. The purpose of the malpractice system is to bring the physician to this point; further investment would actually increase the net costs to society.

If only some mishaps result in suits, the physician can be expected to invest less than the optimum. Suppose, for example, that only one-fifth of his incidents lead to a negligence suit. As shown in Figure 1, the savings to the physician would then amount to only $5 when he reduces the frequency of a mishap from .10 to .09. That is, he is now saving only the difference between a $10 reduction in expected damages per patient (.20 × $50) and $5 spent to achieve the reduction. Thereafter the cost of averting a mishap exceeds the
Each vertical column represents the cost to produce a reduction of one percentage point in the probability of a mishap; the top dashed line represents the corresponding reduction in expected loss. For each reduction, the cost is higher than the previous one. At a mishap probability of .06, the cost of another .01 reduction exactly equals $50, the amount by which the expected loss would be diminished if all valid claims were brought. At this so-called point of indifference, savings are balanced by the costs, and the ideal negligence standard has, in theory, been reached.

If the incidence of suits is less than 100 percent, the incentive to invest in mishaps avoidance is reduced. For example, at a 20 percent incidence (represented by the lowest of the dashed lines), the expected loss is only one-fifth of the actual. The physician who maximizes income now has no financial incentive to invest more than $10 in mishap reduction, and thus will be satisfied with a rate of .09.

Figure 1. Costs of Progressively Reducing the Probability of a Mishap and the Effect of Reduced Rates of Claims.
savings. Should the rate of suits be increased to 60 percent, it would become worthwhile to reduce the frequency of mishaps to .07, but not lower. Pushing it down to .06 would cost $37 but save only $30. Thus, in this "ideal" system, the optimum response is achieved only when the maximum number of valid suits is brought. Below 100 percent, the signal is too weak.

Although the signal itself is suboptimal at a low frequency of suits, the gains are greatest when physicians are stimulated to move from the lowest levels of investment to a somewhat higher level, as shown by Figure 2, which diagrams the cumulative rather than the incremental gains achieved. Because costs of avoidance at low levels of protection are small relative to the reduction in damages, the benefits are large. But as the negligence standard--the point of indifference--is approached, additional savings become small and then vanish.

The doctor in our example is far more explicit in his quantitative analysis than we would expect him to be in real life, even if a real-life physician did seek only to maximize his income. The example does, however, correspond to the intuitive analysis that an income-maximizing doctor would have to carry out.

It is unlikely that more than a handful of physicians, if any, are concerned solely with financial return. A substantial number, it can fairly be assumed, incorporate the welfare of their patients into their calculations and, therefore, invest appropriately in avoiding mishaps. Such physicians would continue to behave appropriately even if the malpractice system did not exist and there were no other sanctions against negligence.

THE SIGNAL IN THE REAL WORLD

For the malpractice system to work as an efficient deterrent to negligent behavior, physicians must encounter an expected loss ($P \times L$) sufficient to make them fully aware of their deficiencies, but not significantly above that level. If awards do, in fact, exceed the losses suffered by claimants and if there are too many successful claims, we have a "malpractice problem" much as the critics represent it.
The costs of mishap reduction accumulate more rapidly than the reduction in expected losses. As a result, the cumulative savings are greatest when the negligent physician is stimulated to move from the lowest level of investment to a somewhat higher level. As the negligence standard is approached—in this instance a probability of mishap of .06—additional savings become small and then vanish. Further reduction of mishaps eventually depletes the initial savings. (Note that last two cost bars are not drawn to full height.)

Figure 2. Effect on Cumulative Savings of Expenditures on Mishap Reduction.
Size of Awards

The theoretical aim of the tort system is to set the value of an award equal to the loss suffered. Such an award has three components in a malpractice judgment: medical expenses, both past and future; lost earnings, past and future; and compensation for pain and suffering.

It is a widespread belief that damages being awarded by the American courts are excessive when judged by these criteria. Criticism of high damages appears to be based, in part, on the attitude that pain and suffering are not legitimate losses for which compensation is warranted. But someone whose injury affects his life style—say the ability to participate in recreation or the ability to have children—is paying a price that is properly included in a damages award.19

The highest, and therefore the most conspicuous, awards—those in excess of $50,000—have the largest element of compensation for pain and suffering.20 Such awards occur almost exclusively when a catastrophic injury, such as paralysis or severe brain damage, has resulted from negligence. In such circumstances, the effect on life style is enormous, and a jury properly assesses the loss as a real injury that can only be compensated monetarily.19 In these terms, even the largest awards do not seem excessive.

The average payment in the award range of less than $50,000 was substantially below medical expenses and lost earnings—typically in the range of one-third to two-thirds of such losses. Indeed, in the aggregate, awards in 1974 were slightly below medical expenses and lost earnings.20

Number of Claims

Many more incidents of malpractice occur, it appears, than result in a claim for damages. Records of patients discharged from two hospitals during 1972 revealed a large number of significant injuries resulting from malpractice; of these, only one in every 15 led to malpractice claims.21 The same study also suggested that many more incidents of malpractice had actually occurred than could be established from the records (this finding will hardly surprise physicians familiar with hospitals' record-keeping practices).
The conclusion that many incidents never lead to a claim is bolstered by the observation that 40 percent of the file entries held by insurance companies consist of mishaps reported by a physician but never pursued by the patient.\textsuperscript{22} There is, moreover, a large pool of malpractice incidents in which the decision to carry out a procedure, rather than a bad outcome, is a form of negligence. Unnecessary surgery, most often hysterectomies and tonsillectomies, probably comprises the bulk of these instances. Numerous as they are, unnecessary operations rarely lead to malpractice claims and have not been included in the analysis of incidents versus claims.\textsuperscript{21,23}

Since 1972, the number of malpractice claims has risen by nearly 50 percent.\textsuperscript{24} Even after estimates are adjusted to account for this change, at most only one out of every 6 or 7 incidents can be expected to result in a claim. A recent study in California supports this estimate: of malpractice incidents detected in hospital records, no more than one-sixth eventuated in a claim. This conclusion is based on the calculation that there were approximately 24,000 instances of malpractice in California hospitals during 1974\textsuperscript{23} but that no more than 4000 of the injured filed claims.\textsuperscript{25}

The fraction (detected from the record) of malpractice incidents manifested in claims against the provider is thus quite small, probably below 20 percent. Given the degree of uncertainty in the data, we calculate, however, that the ratio of claims to incidents could be as high as .3 or less than .1. In any case, the number of suits, even though it is burgeoning, remains far below the theoretical level required to fully signal the expected loss (P \times L) resulting from negligence. On the other hand, from society's point of view, some of the suits required to convey the full message may actually not be "worth" bringing. This seeming paradox results from the effect of administrative costs.

**Administrative Costs**

The balance sheet laid out earlier in Figure 2 is incomplete because it has ignored a significant component of costs. The time spent by lawyers, physicians, patients, witnesses, judges, and juries are all resources that society must expend to maintain the malpractice system, as are the costs of maintaining the courts and the administrative outlays of such tangential institutions as the insurance industry. These
costs reduce whatever level of savings may be achieved by the system. As illustrated in Figure 2, when the cumulative costs of preventing mishaps are subtracted from the savings in expected damages, a curve of net savings is obtained. When administrative costs are also subtracted, the net savings at each level of mishap prevention are reduced and the point of maximum saving is thus shifted to the left. This reduction in savings means that some number of mishaps is not worth preventing through the malpractice system, and it implies that some fraction of otherwise justifiable suits is not worth bringing.

Erroneous Findings

The net saving to society that can theoretically be achieved by a system of malpractice litigation is reduced when the courts make errors. These mistakes can be of two kinds—failure to detect real negligence or penalizing non-negligent behavior—and both are costly.

A court's failure to recognize negligent behavior has the effect of reducing the expected loss and thus permits physicians to set their investment in mishap avoidance below the ideal level. At the same time, the malpractice system as a whole is made more expensive because the proceeding raises administrative costs but yields no benefit.

When non-negligent physicians are penalized, the result is also costly. Physicians, realizing that such errors occur, are induced to practice inappropriately "defensive medicine," i.e., to provide medically unjustified care to reduce the probability of a malpractice suit. The availability of third-party payment permits the physician to use diagnostic tests and hospitalization without restraint from the patient or cost to himself, and thus to demonstrate a level of care so painstaking that neither a patient nor a jury would be likely to make the error of calling him negligent. But this immunity is bought at a cost to society far in excess of anticipated benefits. The excesses of

*The availability of health insurance may have another perverse effect on the malpractice signal. Physicians may respond even to a valid finding of negligence by substituting resources that are "free" to them (such as X-rays or consultations) for increased expenditure of their own time. Although the result seems desirable—an appropriate reduction in the incidence of mishaps—it may be purchased at too high a social cost.
medicine will be discouraged only if providers of health care can be encouraged to compete with each other. When a provider, such as a prepaid health plan, lowers its expenditures on defensive medicine and than passes the savings on to its members, it can attract patients away from the more wasteful, and thus more expensive, providers.

Any system is liable to error, the costs of which must be subtracted from the profits of the system as a whole. In the case of negligence procedures, there are two incentives to avoid mistaken judgments: the economic loss sustained by society and the personal injuries suffered by a plaintiff or defendant.

MALPRACTICE INSURANCE AND ATTENUATION OF THE SIGNAL

Even if the signal were perfect--every incident of malpractice leading to a proper claim, and every award equal to the loss suffered by a patient--under existing conditions it would fail to elicit the appropriate response from physicians. Malpractice insurance, as it is currently administered, virtually insulates the negligent physician from the damages award and, thus, from the malpractice signal.

The reason is simple: the malpractice premiums of individual physicians are rarely influenced by their record of claims, settlements, and verdicts. (Haldi, J.: personal communication, April 1, 1977.) Rather, premiums are usually set for an entire specialty group in a given region. Thus the physician with a record of frequent negligence bears no larger a share of the burden than his colleagues with excellent records. No individual physician has more than a slight pecuniary incentive to reduce the expected losses resulting from his own behavior.

Under the prevailing system of group rating, we can imagine a physician who could avoid an expected loss of $1000 by spending $100. And yet, if his malpractice premiums are set for a group of 100 physicians, his own premium will rise, as a result of the mishap, by only $10 (and so will that of 99 other physicians). The individual is thus assessed only one-tenth of what it would have cost him to prevent the injury. In reality, the physician may not lose even the $10; for in most instances all physicians in a given group can pass on nearly all of the cost of their rising premiums, as increased fees, to third-party
payers. Thus, the burden of higher malpractice premiums falls largely on the patient in the form of a higher health insurance premium.

As a rule, the system of group rating successfully interferes with transmission of the signal, but there is one partial exception. Part-time practitioners of a specialty must pay the full premium, even though only a fraction of their income derives from it. To the extent that a part-time practitioner—say, the otolaryngologist doing some plastic surgery—is incompetent, he is successfully weeded out by the group rating. In this situation, the malpractice system still has a deterrent effect, as was demonstrated by a recent study of California physicians who have been relinquishing part-time activity and reclassifying themselves exclusively into their primary area of competence. Of course, the effect is clumsy, in that it also eliminates part-time specialists who are competent but cannot or will not pay the full specialty premium. Group rating is, then, an exceedingly insensitive method of identifying the negligent physician.

Malpractice insurance, on the other hand, has important social value and our comments should not be taken as an argument for abolishing insurance protection. Because insurance spreads risks, it protects the physician against the financial catastrophe that could result from even a single large finding against him, and also against erroneous findings of negligence. A balance must therefore be struck between the risk-spreading value of insurance and the need for experience rating as a means of sending the deterrent signal to the physician.

The lack of individual experience rating in malpractice premiums has, in the past, been economically acceptable to both the insurance industry and the medical profession. For the industry, the cost of identifying individual bad risks has probably not been worth the potential savings. Now, as malpractice litigation comes to involve ever larger amounts of money, the interests of both the industry and the profession may shift toward a system of experience rating.

If a physician were rated by his individual experience, his premium would reflect, to at least some degree, the risk that he poses to the insurer. Moreover, given the procedures employed by health insurance companies, such physicians would probably not be able to
raise fees much above those prevailing in the community. Thus, physicians who pose substantially more risk than others would themselves have to absorb much of the extra burden of premium costs.

Experience rating of physicians is, of course, justifiable only if suits are not brought randomly or capriciously. It has been argued that "good" physicians are sued as often as the "bad." In support of this contention, the belief is expressed that doctors with board certification are sued as often as those who are less well trained. Even if this statement should prove correct, it is not evidence that negligence is inaccurately found. Many factors could lead well-trained specialists to perform below their presumptive level of competence. Some well-trained physicians may take on too many patients and invest less time per case than is appropriate. Others may assume responsibilities in areas for which they are not trained or experienced. Moreover, the thesis that suits are brought randomly is not supported by a recent study of 8000 physicians in the Los Angeles area. In a four-year period, 46 physicians (0.6 percent of the 8000) accounted for 10 percent of all claims and 30 percent of all payments made by the insurance plan. The average number of suits against the 46 doctors was 1½ per year. Analysis indicates that doctors against whom multiple suits are brought do, indeed, represent a higher-risk population than their colleagues.

Virtually the only penalty currently paid by the negligent physician is the value of his time spent in defending a suit and the costs of his embarrassment. In the absence of experience-rated premiums, these factors may be the only effective component of the signal and may serve some useful purpose. But because the costs cannot be purposely set, they are not likely to equal the damage assessment that the physician should experience if resources are to be efficiently allocated. "Time and embarrassment" costs may be either larger or smaller than the optimum.

CONTINGENCY FEES AND SIGNAL QUALITY

According to a popular line of reasoning, lawyers are encouraged
to bring malpractice suits that lack merit because they are paid a percentage, usually one-third, of the damages awarded; if they were paid a flat fee for service they would not be so indiscriminate. Even if it is true that too many frivolous suits are brought, this analysis of the contingency-fee system would be incorrect.

The lawyer who is paid a contingency fee is compensated only if he succeeds in obtaining either a settlement out of court or an award from a jury. He is not likely to invest time and several thousand dollars in out-of-pocket expenses on a case with little prospect of success. Under the system of contingency fees, lawyers thus have the incentive to filter out capricious suits, which otherwise would overload the courts, harass physicians, and produce no social benefits.

Payment by contingency fee not only encourages the lawyer to turn away baseless claims but permits him to accept clients lacking the funds to pay a fee for service. The contingency fee offers "a key to the courthouse door" to some individuals with valid claims that would otherwise be lost to the system. Without contingency fees, the deterrent signal to the physician would be reduced.

Why then, if contingency fees are such an incentive to careful selection of cases, are 80 percent of claims that reach trial ultimately resolved in the physician's favor? During periods of rapidly changing technology, like the present one, we must expect that individual doctors and lawyers cannot always be certain of correct medical practice. Litigation can serve to establish proper standards and thus to provide important information about the changing limits of malpractice. The suits providing this information may well be appropriately brought, even if the accused physician is exonerated.

Fee-for-service payments to lawyers would be likely to have perverse effects on the malpractice system because lawyers would actually have an incentive to encourage suits, regardless of their merits. Responsibility for screening suits would then devolve on the client himself, who is unlikely to be adequately informed about prospects for success. The number of suits might be reduced, but not because the frivolous ones had been eliminated. Rather, those who simply could not afford to pay for litigation would be discouraged.
ALTHERNAVES TO MALPRACTICE

Dissatisfaction with malpractice proceedings has produced various suggestions for changing the system, such as ceilings on the size of awards and a shift to fee-for-service payment of lawyers, changes that might well have undesirable effects. Also advocated is a system of arbitration whereby a panel of experts would reach a finding through less formal means than a full-fledged jury trial. With this system, the gains in efficiency might be offset by significant losses: the protection against bias and influence that a broadly based, rotating jury provides and the accuracy attained by complete and careful presentations in court.

The negligence standard might be abandoned altogether and another form of liability substituted for it. One approach, for example, would establish a system of negotiations between patient and physician who would agree, in advance, to a given investment in mishap avoidance. This "contractarian" approach, as it is called, neglects the basic problem that necessitated regulation of medical care, in the form of malpractice law, to begin with: the consumer of medical services, more than the consumer of most goods and services, lacks the information and sophistication needed for bargaining. It is precisely this market failure that created the need for regulation of health services, whether in the form of licensing or litigation.

At another extreme and currently receiving the widest attention, is a proposal to make the physician automatically responsible for many types of mishap, regardless of fault. The no-fault alternative deserves some scrutiny because it may prove to be a problematic "solution" to existing problems.

Under most of the proposed no-fault programs, the physician would be held liable for any bad outcome associated with a particular "compensable event"—any complication of a blood transfusion, for example—and would be insured accordingly. Such no-fault liability would remove from litigation a number of events in which bad outcomes are usually, but by no means exclusively, the results of negligence. In such instances, the costs and unpleasantness of malpractice proceedings could be avoided and, proponents argue, a net saving relative to current expenditures would be achieved.
It is not clear how great the real savings under a no-fault system would be, because a no-fault system could hardly eliminate disputes and litigation. As with workmen's compensation (another form of no-fault insurance), disagreements might arise as to whether a particular outcome, say hepatitis, was in fact the result of a compensable event, such as a blood transfusion, or was the untoward effect of a drug concomitantly administered and known occasionally to cause liver damage. Litigation would be inevitable. The combined cost of such litigation, and of payments to patients who were injured but not as a result of negligence, could well offset the overall saving that a no-fault system might otherwise achieve.

In the design of a no-fault system, the most serious question must be: Who will pay the damages? If the individual physician's premiums do not reflect the costs of compensable events attributed to him, he will not be deterred from negligence. To achieve deterrence, the individual physician must be rated by his experience. But the use of individual experience ratings for no-fault insurance premiums could be expected to have some distinctly perverse effects. The practice of "defensive medicine" based on third-party resources would be encouraged because the physician would be liable for any bad outcome and would thus gain maximum protection from highly redundant safeguards. The savings would go directly to the physician; the costs would be borne by the patients.

Perhaps more seriously, physicians would be deterred from undertaking risky procedures, even when such procedures were, on medical grounds, most appropriate and should, for the patient's sake, be ventured. The cost of a bad outcome from the procedure would be borne solely by the physician, and in this case, the patient would likely be deprived of an opportunity for proper care. Skilled physicians, to whom difficult and high-risk patients are referred, would be especially penalized. In principle, a complex schedule of premiums could offset this tendency by comparing liability payments with the risk profile of a given doctor's practice. In reality, such a scheme would be difficult to implement and open to abuse.

If, on the other hand, experience ratings were abandoned, the physician would have little or no economic incentive to prevent
mishaps, and a rise in the number of bad outcomes--compensable events--could reasonably be predicted. (Not only would premiums fail to reflect negligence, but time and embarrassment costs would have disappeared.) An increased frequency of bad outcomes would also raise the costs of a no-fault system above currently predicted levels.

The no-fault alternative does not a priori appear to offer significant advantages over the present system. And even if it could be implemented for some selected group of events, the malpractice system would still be needed to deal with the very large number of cases that could not feasibly be dealt with under a no-fault system. The prospect of two systems, both of them complex and fraught with problems, is not an attractive one.

Our purpose has not been either to defend the system of malpractice litigation or to gloss over its shortcomings. Rather, we have demonstrated an analytic approach that helps to clarify the role of malpractice proceedings in maintaining the quality of medical care. The same kind of analysis should be applied to proposals for change before policy decisions affecting the system are made.
REFERENCES


12. Munch P: Costs and Benefits of the Tort System if Viewed as a Compensation System. Santa Monica, California, Rand Corporation, 1977


16. The T. J. Hooper, 60 Federal Reporter 2d 737 (1932)


30. Lipson AJ: Medical Malpractice: The response of physicians to premium increases in California. Santa Monica, California, Rand Corporation, 1976

32. Phelps CE: Experience Rating in Medical Malpractice Insurance. Santa Monica, California, Rand Corporation, 1977
