

Medical Malpractice and the
Propensity to Litigate

by Jerry R. Green

Preliminary draft: November 1976

Medical Malpractice and the Propensity to Litigate

I. Introduction

The rapid increase in the number of cases and the increasingly large awards being granted is ample cause for concern about the reliability and practicality of the present malpractice system. The right of patients to sue their physicians for negligence has several complex implications. On one level, it provides a way for the patient to be compensated when he suffers an injury, just as automobile accident liability does. At the same time, the risks imposed on physicians who might be sued act as an indirect system of control over the precautions they take to avoid adverse legal findings. The difficulty of adequately monitoring the quality of medical care requires that some methods of this type be employed to increase the incentives for better services.

On the other hand the real costs of administering such a system of checks and balances is enormous. The burden of malpractice payments falls only partially on the physician. It may be passed on to the public through the mechanism of increased premiums for malpractice insurance, rising fees for medical and surgical services and finally, higher medical insurance premiums. To be sure, some of the incentives for better care rests with the physician even if he is insured for malpractice. The widespread concern about defensive, as opposed to best-practice, medicine should be sufficient evidence that physicians feel the impact of these imposed risks in a real way. If the potential burden of malpractice suits does not induce a significant increase

in the quality of care, the extra expenses incurred in litigation will have been wasted. But if medical care for everyone can be improved by this means, the net benefit to society may far exceed the resulting higher costs.

In this paper I intend to focus on the desirability of using an adversary system of justice as an indirect way of exerting influence on the quality of medical care being offered. Special attention will be paid to the costs of the legal system itself. Its efficiency therefore depends upon the frequency with which litigation is undertaken, as well as on the care level that it can produce. I will contrast this system with a type of peer review procedure. It will be argued that the latter has advantages over litigation to the extent that it is a more flexible policy. The standards for non-negligent care can be raised without inducing a rise in the cost of administering the process. On the other hand, we will argue that peer review, or investigation by any third party, has two adverse effects. It dulls the incentives for exceeding any fixed negligence standard and it results in a somewhat lower case selectivity in choosing cases for investigation. I will try to delineate the circumstances under which the balance will be tipped in either direction.

Throughout the analysis I will be neglecting the function of malpractice as a provider of a type of insurance for the patient. Presumably this kind of insurance could be arranged by other means. Coverage for adverse physical and economic circumstances need not be linked to the cause of the event.¹ It is therefore not impossible to arrange for such insurance in other ways, and it may even be preferable to do so. At any rate : this paper will concentrate entirely on the other role of malpractice, that is as a regulator of care.

There is one salient difference between litigation for medical malpractice and other forms of accident liability. It is this: The complexity of modern medicine, and its inaccessibility to the layman, prevent a patient whose outcome has been worse than expected from knowing whether his unfortunate circumstance is the result of malpractice, or whether it is a chance event not under the control of his physician. In automobile cases, for example, quite the opposite is true. The transparency of the matter usually leaves little doubt as to whether or not an accident in fact occurred. Placing the fault, however may be a highly contested issue, not easily resolvable by the parties at the time of the incident.

The implication of this rather simple observation is that while virtually all serious automobile liability cases are contested and ultimately settled, it is quite likely that some equally serious incidents of malpractice will escape the system, disguised

as bad luck. More specifically, the decision to undertake a malpractice suit on the part of the patient and his attorney will be influenced by their perceptions of the probability of winning the case. This has a very strong bearing on the efficiency of the current malpractice system. Suppose courts increase the standard used for negligence, making doctors responsible for damages in some situations in which they would previously have escaped liability. There is a large body of evidence² indicating that this has occurred during the last ten years. Such a change would increase the propensity to undertake litigation. The risks faced by physicians are thus increased in two ways: Stricter standards for care will be applied, and they are more likely to be sued than previously. A similar shift in the automobile fault standards would not induce an increase in the litigation rate, so that only the direct effect of higher standards would be operative.

There is also some indirect evidence to this effect, though it is not entirely conclusive because several alternative hypotheses would have produced the same results. If the propensity to litigate were not endogenous, the higher standards applied would lead to a higher proportion of successful suits than previously. In fact, exactly the opposite has been the case. Over half the malpractice actions initiated now result in no award. This has gone up substantially in recent years.³ Although, as we shall argue, the increase in this "zero-award rate" is not a necessary implication of higher standards for negligence, they are consistent under the hypotheses we maintain and they would be incompatible if one were to use a more standard framework for studying this type of liability.

In the following sections I will attempt to describe how the optimal negligence standard should be set under the present system and to contrast this with the workings of an optimal negligence rule for accidents in which the propensity to litigate is not endogenous. Let me preface this by analyzing the simpler case in which the occurrence of an accident can be immediately recognized by the party who bears the direct costs. In such circumstances, if we set aside complexities induced by risk-averse behavior, the optimal rule for negligence is simply to hold the injurer liable in all such circumstances, independent of the precautions he may be able to prove had been taken. Operating in this way is simple--litigation is reduced to the question of whether an accident did or did not occur--and with the full incentives on potential injurers, they will choose their care level so as to minimize the expected costs of accidents plus the cost of care, which is precisely the social objective.

Such a method would clearly not work in the medical context.⁴ If physicians had to compensate all patients whose outcomes were unfortunate, the entire health care delivery system would break down. Therefore a real standard for measuring care taken must be instituted. It must be set so as to make litigation sufficiently likely that physicians feel stronger incentives to take care, but it must also reflect the real costs of the induced litigation imposed on both defendants and plaintiffs.

II. Complexities of the Medical Context as a Setting for Adversary Proceedings.

There are many other differences between ordinary accident

law as an incentive system and the malpractice situation. It would be creating a falsely simplistic impression if these were not mentioned at the outset, although we will not be addressing ourselves directly to these issues herein. One of the most interesting features of medical injuries is the extraordinarily long lag between the actual incident and the date at which injury is perceived and litigation ensues. The American Insurance Association's recent closed claim survey reveals that although 91% of cases involved a lag of less than one month from incident to injury, over 60% had a lag of over 6 months until the date they were reported. On the basis of weighting claims by the dollar value of the resulting settlement, this figure is over 70%, indicating that the more serious cases may involve a longer lag than the average. This lag has implications for the nature and accuracy of evidence that can be gathered and brought to bear on such cases. By the time legal action is initiated, the opportunity would have been missed to perform some of the medical tests or observations which could have indicated whether prior negligence played a role in creating the patient's condition. Further, the extent of the injury suffered might be greatly magnified by the delay.

A second difference is related to the multi-faceted nature of medical care. More specifically, the least-cost method of reducing the probability of an adverse medical outcome might not be readily observable by the court. An example would be the time taken during a physical examination, and the thoroughness with which it is performed. The court's readily observable attributes

of care would typically be those procedures of which records are kept: tests ordered, x-rays, frequency of reexamination, length of hospital stay. Focusing on these aspects of care when others that would be more efficient are by-passed is defensive medicine. The problem of defensive medicine is a most serious one, but, unfortunately, one with which we cannot deal except by urging our courts or monitoring agency to redirect their attentions. It obviously will be working at cross-purposes to improve the incentive structure for physicians to take precautions that we do not really want to encourage. Throughout this paper I will be speaking of care as an entity that can be measured at least in a rough way. By care we shall mean simply whatever mixture of activities for accident reduction are encouraged by the system. If these are not the best activities for this purpose, costs considered, a restructuring of the measured attributes of care is called for. To the extent that this is impossible, the resulting inefficiencies are a reflection of the costs of imperfect information--which are real economic costs to be reckoned with.

Finally, and most to the point, the highly technical nature of medicine makes the litigation process very costly--more so relative to claims and insurance coverage than in other aspects of tort law. Using data from the American Insurance Association's recent closed claims survey, I have calculated that over one-fifth of the dollar settlements made were payments of plaintiff's legal fees. This neglects the legal fees paid in losing cases and all of the direct legal costs of the insurance carriers themselves. Overall the figure is likely to be around one-half. These, too, must be regarded as real economic costs to

the system. Our analysis below will discuss the ways in which these costs affects the propensity to undertake litigation and, through this, the overall structure and performance of the medical system.

III. A Model of the Effects of the Endogenous Propensity to Sue on the Efficiency of Malpractice Litigation

Accepting these stylized facts as an accurate reflection of the structure underlying medical malpractice, we proceed to describe the workings of the system of litigation in detail. In particular, we will try to analyze the ways in which the policies implicitly followed by the courts shape the behavior and performance of the medical sector.

The most important aspect of the system, and its distinguishing characteristic compared to applications of liability rules in other areas, is the way in which information becomes known and its accuracy as a representation of the true circumstances concerning the accident. Unlike other accident problems, the long lags surrounding medical claims often preclude knowing whether malpractice was really involved.

Nevertheless it is natural to presume that whatever the inaccuracies of the final determination may be, the grounds on which the legal action is initiated are even less precise. We shall view the patients' a priori estimate of the extent of care taken by the physician as a random variable which is positively correlated with the court's ex post estimate. Because of the difficulty of assessing the strength of the evidence that could be brought to bear by the defense, and because of the highly variable time to settlement, the plaintiff and his lawyer have to make the decision to press suit based on the limited information available to them.

Although in actuality information on a case continually accrues through the process of litigation, and in principle the litigation can be terminated by the plaintiff at any date if the evidence starts to accumulate adversely, it will be simpler if we view the process in two discrete stages. First, the patient establishes a cutoff a priori indication below which he will undertake the suit, and pay the costs of litigation. Second, the outcome is determined by the court on the basis of the observations available at the end of the litigation. The court sets a cutoff level such that if the ex post indications are below this, the case is decided in the plaintiff's favor.

It may be, of course, due to the large amount of randomness inherent in the system, that some cases which could have been pressed successfully will not be tried because the a priori indications were too unfavorable. Many cases may be tried unsuccessfully. But the vast majority of cases are probably of the type that have little or no a priori indication of a malpractice, would, if tried, uncover none, and rightly so because the level of care given was in fact entirely satisfactory.

The court system, in executing this role as a regulator of care provided, cause substantial resource costs to be incurred in the process. Some of these are the direct costs of legal fees. Assuming that there are many lawyers whose services can substitute quite closely for each other, we would not make a great error if we identified the social costs of malpractice lawyers' time with the private costs borne by plaintiffs and defendants or their insurance companies. In addition, the vast number of cases

litigated crowds court calendars and induces massive bottlenecks for other types of cases as well. In 1974 the average length of time from the reporting of a malpractice claim to its closing was over one year.⁵ Even if only part of this interval represents the actual settlement process (part of it is the waiting period due to claims previously filed) the costs of delay would be a significant fraction of total litigation costs. These time costs are difficult to measure, but they are likely to be very heavy. Consider the fact that every new case brought to trial delays every case whose position in the queue is later by the full amount of the time taken to settle it. Since the queues are very long, a short trial period will still impose costs on many others so that its total social costs will be very large.

We will be ignoring the time costs because of the difficulty of measuring them even approximately. Moreover, until the current malpractice crisis results in a change in procedures and a return to normalcy, the court system will not be able to assess the demands that will be made on it and adjust its capacity to meet the flow of cases without imposing such delays. In the long run, the real costs of excessive and frequent litigation of malpractice claims is the expansion of the judicial system required.

How will the two principal groups of decision-making individuals, potential plaintiffs and defendants, react to the system in which they are embedded? We will take the usual economist's view, probably a correct one in this case, that they will each decide in their own best interest using the available information, under the

assumption that the behavior patterns of the opposition will remain unaffected.

The individual patient will undertake a suit if the expected gain from this action exceeds his expected costs of litigation. Calculations of these two critical parameters depend on the information available relevant to his case, on his perception of the general level of medical care being offered and on the way in which his legal costs will depend on the settlement. The impact of the contingent fee system is most strongly felt in the last question. But the relationship between the legal fee and settlement received is important only if considerations of risks weigh more heavily in the decision to sue than average returns. Taking the expected net return as the relevant decision variable means that it is the expected costs of litigation that enter into the individual's calculations, so that the precise relationship between fees and awards will not affect behavior under this assumption.

The physician is assumed to know how his care level would affect the probability of losing a suit. He chooses a care level that maximizes the difference between the profitability that is associated with it, in the absence of malpractice claims and expected costs of losing malpractice litigation. These costs are compounded from two factors, the probability of being sued, and the probability of losing a suit given that it has been initiated.

Putting these patterns of behavior together, we can see that the standard set by the court gives rise to both a prevailing level of care taken by physicians and a propensity to sue by

patients with the property that each is an optimal pattern of behavior given the actions of the other. By varying the negligence standard these actions will be shifted. Our problem is to evaluate how they will respond and thus to characterize and evaluate the optimal standard for this type of liability system, and to compare it to other possible systems.

The care level of physicians is of direct and central importance to the issue at hand. Ideally, as mentioned in section I, we would seek a balance between the costs and benefits of additional care. The problem is that attaining such an optimum is precluded by the costs of litigation. In order to evaluate the litigation costs it is necessary to combine the ex ante cutoff level used by patients with the care level; the probability of litigation going up with the latter and down with the former.

In our search for the optimal negligence standard, we note therefore that as long as higher standards can induce higher care levels and lower cutoff values they will be superior. Tradeoffs will come when, in order to induce higher care levels, progressively higher cutoffs must be induced, and the latter are forced to increase so rapidly that the propensity to litigate increases.

IV. Choosing the Best Care Standard

We begin the analysis of the negligence standard by looking at the way in which each of the two types of individuals will respond to an upward shift in its level. The patients respond by choosing a higher cutoff level for a priori indications at each underlying level of care. The reason for this is clear. If it becomes easier to win malpractice cases, it is profitable to press some suits which previously would not have been worthwhile. Similarly, if the physician is facing a higher probability of loss given that the suit is undertaken, he will endeavor to protect himself against this. He does so by taking more care. This actually has two effects when the cutoff level for filing claims is viewed as fixed. It lowers both the likelihood of being sued, and of losing a case given that legal action has been undertaken.

In order to trace out the effects of these responses on the equilibrium of this system, it is necessary to know how the actions of each group depend on the other given a fixed negligence standard. More care by physicians will lower the cutoff level for filing suits. The suits previously filed would be less successful, on average, when the care level has increased, and therefore those in the marginal category would not now be pressed.

A higher a priori cutoff by patients would lead to more care by physicians. They are at greater risk of being sued and, since some of the cases added will involve adverse decisions for them, (even though they will have a lower success rate than all those currently being tried), the effective level of successful suits will increase.

Thus the situation can be depicted by the following diagram:

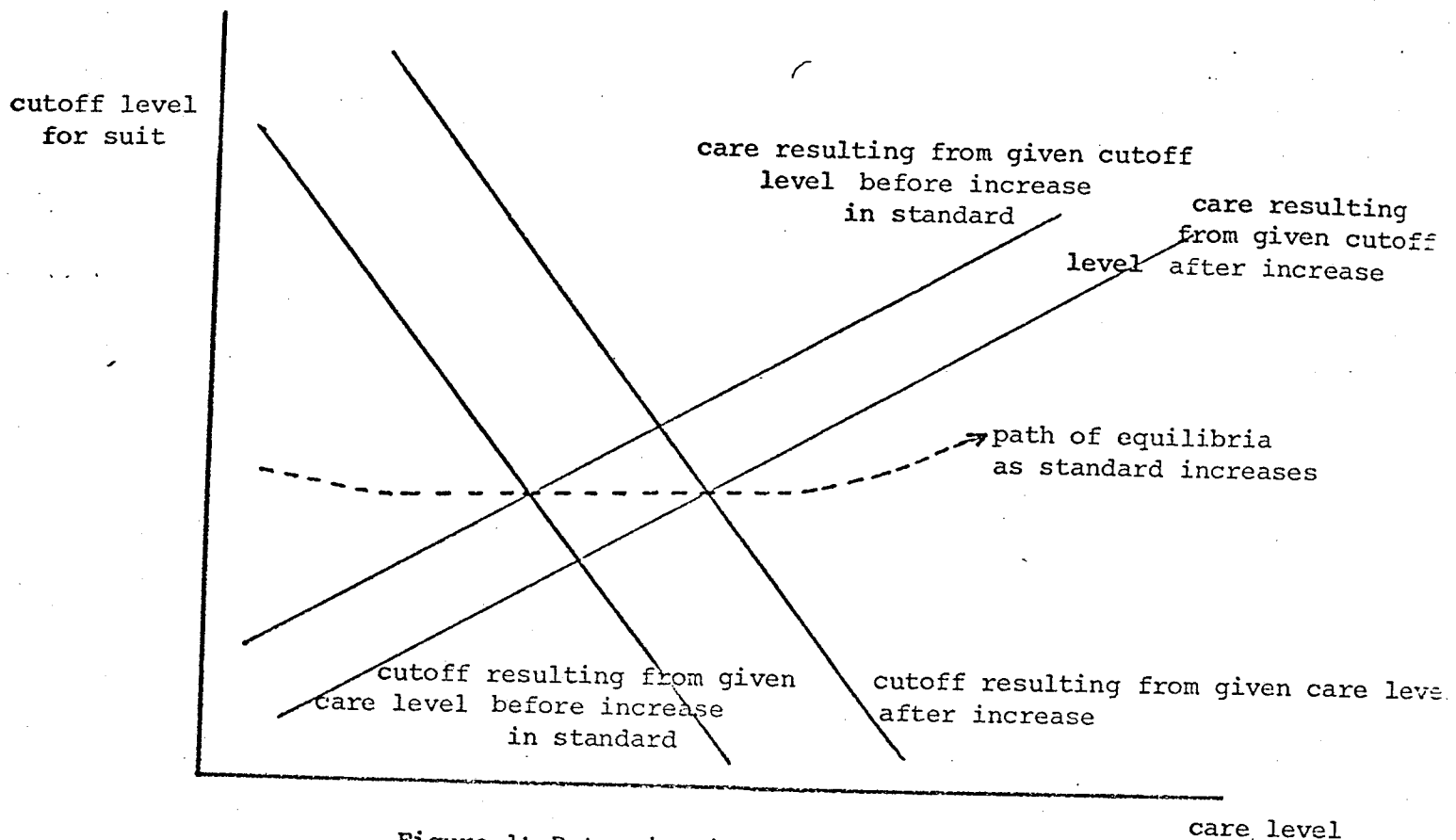


Figure 1: Determination of Equilibrium for Each Negligence Standard

The upward-sloping curves represent the care level chosen in response to each cutoff level for suits to be filed at two different levels for the negligence standard. In accordance with the discussion above, the line farther to the right corresponds to the higher negligence standard: a larger probability of loss giving rise to more care taken for each a priori cutoff level for suits to be initiated. The downward-sloping lines describe the reactions of the patients to increasing care levels. The higher curve is associated with the higher negligence standard because if it is easier to win a suit, cases that are increasingly less

clear-cut on an a priori basis will be pressed.

It is clear from this analysis that an increase in the negligence standard will raise the level of care taken and that the effect on the cutoff level of a priori indications for a suit to be pressed could go either way. The reason for this ambiguity is clear. A higher care standard induces more suits because it is more profitable to do so, at the margin. But on the other hand the induced increase in care makes suing less likely to be profitable because the conditional distribution of ex post evidence will have improved for every a priori indication that that might have been realized. Without knowing how much extra care will be taken in response to an increase in the standard, it is impossible to predict which of these effects will pre-dominate.

What is the effect of these shifts on the efficiency of the system as a whole? The care level is a central focus of this analysis but the cutoff level for suits is only indirectly relevant. What really matters is the frequency with which litigation is undertaken. Broadly speaking, it is this that determines the level of resources expended in the legal process through which medical practice is, at least partially, regulated.

This frequency is itself dependent on the care level and the cutoff for suits. If the cutoff level for suits is rising when care is rising, the resulting change in frequency will be in doubt. If the cutoff level were falling while care is rising, then the frequency of suits would be decreasing as the negligence standard is rising.

Although either case is in principle possible, the recent evidence would surely indicate that we currently are in a regime of rising negligence standards and propensity to undertake suits. This means, necessarily, that cutoff standards have been rising as well--the public as consumers of medical services has become more "litigation-minded" in the words of the Secretaries Commission on Medical Malpractice.

If we graph the path of the relevant objectives--frequency of suit and care taken--as the negligence standard increases, we are likely to see a pattern such as the following:

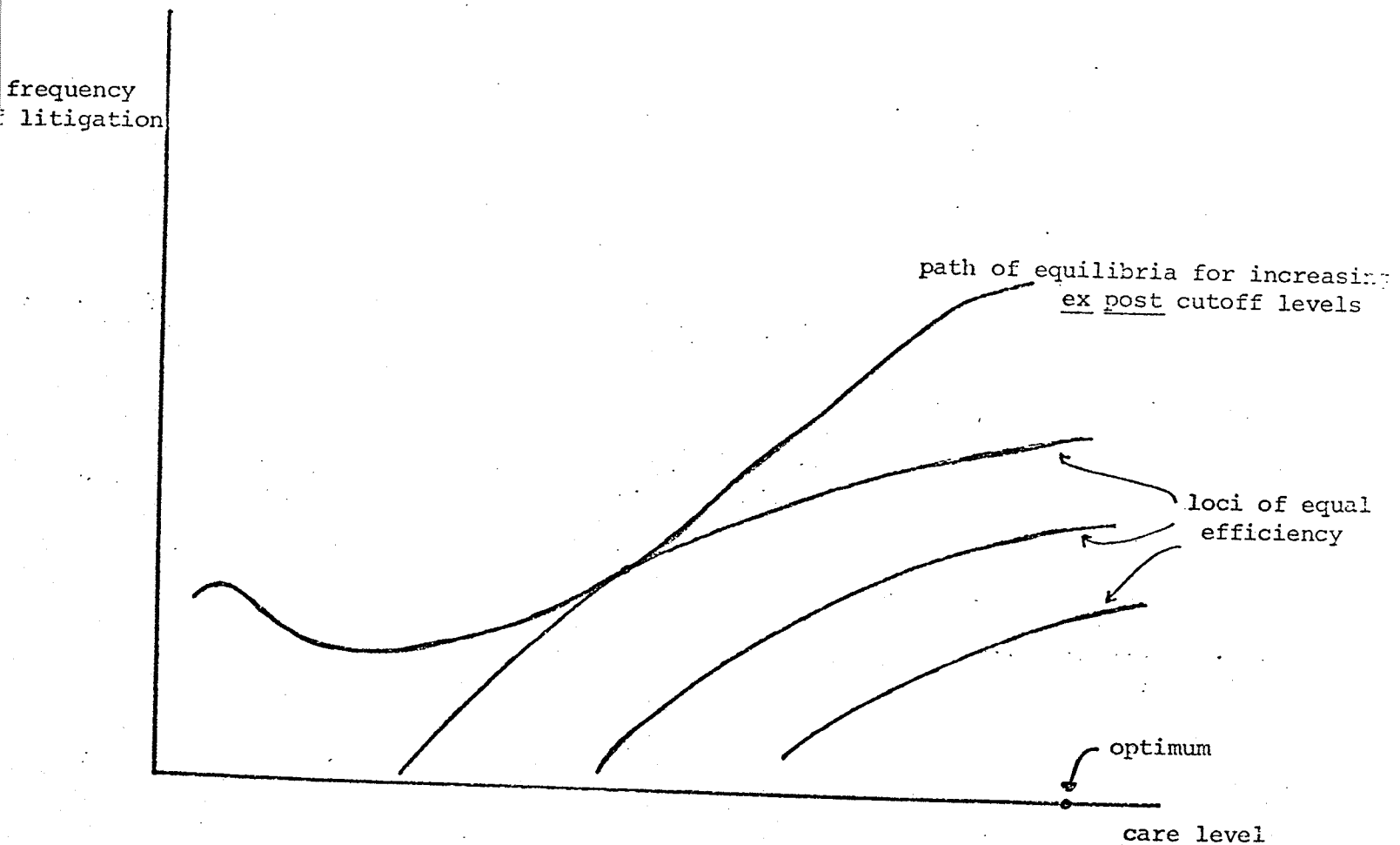


Figure 2: Determination of the Optimal Negligence Standard

How well can this system do in attaining a socially superior situation even when the optimal negligence standard is being used? It is clear that the overall optimum is not attainable. The incentives of all parties involved preclude it. It would be to set the care level so as to balance the costs and benefits of increased care and to avoid all litigation. The best attainable situation is illustrated in the diagram above. It involves some litigation and a somewhat lower level of care than the optimum. Note that in order to attain it, we must be in a situation in which higher negligence standards are associated with higher litigation frequencies and hence with increasingly less severe requirements for the a priori indications required for patients to undertake legal action.

V. Alternatives to the Negligence Standard System

The method outlined above, which describes our current practice in its broadest outlines, is not capable of attaining an optimal situation. Resources will have to be used up in the litigation process. We may therefore justifiably enquire as to whether there are other systems through which control over the care of physicians may be exercised, without adversely impinging upon the practice of medicine in the large.

One such system is where review of a certain proportion of cases is undertaken by a medical review board, or a governmental health care monitoring agency, rather than relying on the patient's own perceptions of his care to determine whether or not a review is undertaken. Such a board might function in the same way as a court, exacting penalty fees from negligent physicians as a way of establishing the appropriate incentives for better medical care. Whether these fees are rebated directly to the injured party or are used to subsidize health insurance schemes in general is really irrelevant to the matter at hand, as no one's actions would be affected by the use of the funds collected. The dual role of litigation as a form of insurance would have to be treated separately--but this facet of the problem is outside the scope of the present study.

There is one clear advantage to such a system and one clear disadvantage, in comparison to the litigation method. It is superior in that the proportion of cases investigated may be selected independently from the negligence standard employed, rather

than being set endogenously as a result of interactions between the two sides. But the system loses efficiency to the extent that it cannot hope to be as selective in its undertaking of investigations as the negligence litigation method can be. The information in the hands of the patient, his family and his attorney is much more accurate and detailed than that available to the outside agency. And unless patients have an incentive to put their own cases under review this information will never be transmitted in sufficient depth. If cases are not undertaken selectively the administering agency is forced to investigate a much higher proportion of them in order to find the same number below the negligence standard. To keep the analysis simple, we presume that the administering agency's information does not help it select the cases in which a negligence finding is at all more likely.

Let us consider how the care level of physicians would respond to the rate of investigation, holding the negligence standard fixed. The physicians would choose care to minimize the direct costs of taking care plus expected losses due to adverse investigation findings. Additional care is beneficial in avoiding the latter only in so far as it reduces the frequency with which the ex post observed value falls short of the standard. It does not decrease the probability that a given case is examined more closely, as it would under a negligence system. Therefore, for a given negligence standard, if the rate of investigation is set at the same level as would be induced by patients bringing suits under the litigation system, the corresponding level of care would be lower. However, the extra flexibility of the method currently

under consideration is that it is able to have higher rates of investigation at the same negligence standard, or alternatively higher negligence standards without incurring more costs of investigations. For each frequency of investigation it is therefore natural to ask what the optimal, or care maximizing, negligence standard is under this system. The answer depends in general on the precise nature of the statistical dependence between the care taken and the result of the investigation. In qualitative terms, the negligence standard will have its strongest impact if it is set at the level where there is the greatest concentration of the results of investigations. In this way the effort to avoid adverse findings by raising the care level taken will be strongest. If the statistical relationship between the investigations' findings and the care taken is a precise one, the care level can be made quite high by raising the standard for negligence while keeping the frequency of investigation quite low. In these cases, a sufficiently high standard could induce more care than in the negligence situation with the same rate of litigation. When the relationship between care and the findings is weaker, the potential gain to the physician from an increased care level is much lower since he will have little chance of affecting the result of the investigations by taking more care. Here, the added impetus to avoid litigation will be an important factor in sustaining care levels above those which could be attained through investigations initiated on a random basis.

frequency
of investigation
or litigation

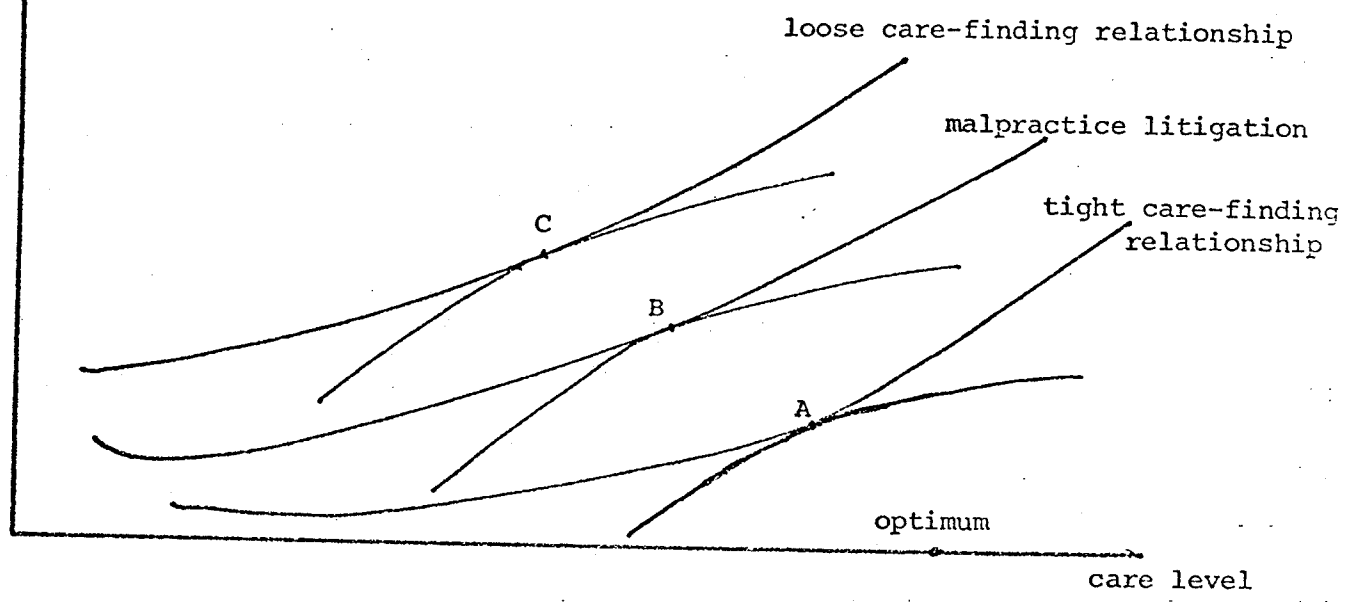


Figure 3: Investigation vs. Litigation
A, B, and C are conditional optima

VI. Conclusions

We have tried to present a model of medical malpractice litigation that is consistent with the broad qualitative aspects that have evolved markedly in recent years. Viewing the present environment in this context, it is possible to contrast the efficiency of the system of litigation with alternatives such as peer review or medical-legal panels. We find that litigation has the advantages of inducing more selectivity in that the incidence of malpractice observed among cases pressed to suit will be higher. In this way it will economize on costs of administration relative to a non-litigation alternative. Moreover it has a double-edged incentive built into it in that it causes care to be taken both to avoid litigation and to reduce the risk of losing those suits that are brought. The alternative scheme operates via only the latter effect. However litigation links the care standard used with the frequency of litigation together. It is inherently unable to treat these as separate policy tools without at the same time destroying its natural advantages as just discussed. In this way the alternative schemes will prove more flexible. Which method is best depends on the nature of the imperfection of information in the hands of plaintiffs and courts. It is an empirical matter. The current existence of a variety of such procedures in different states should serve as a natural laboratory in which their relevant attributes we have highlighted herein can be measured and evaluated.

AppendixAn Analytical Model of the Propensity to Litigate

The care level being taken by physicians is parameterized by a real number, x , that represents the cost of this care. When a level of x prevails in the medical sector, a frequency distribution of outcomes results. Outcomes are described by three variables:

$a \in \{0,1\}$ indicates the occurrence ($a = 1$) or non-occurrence ($a = 0$) of an adverse medical outcome.

$y \in R$ indicates the ex post care level that would be indicated by the investigation of a court or other outside agency.

$z \in R$ indicates the ex ante care level as perceived by the patient and his attorney.

All adverse medical outcomes cost C . Only those cases in which an adverse outcome occurs can lead to suits. The values of y and z when $a = 0$ are therefore really irrelevant.

Let $f(y,z|x)$ be the conditional density of y and z given x and given that $a = 1$. Let $p(x)$ be the probability that $a = 1$.

The cost of filing suit to a potential plaintiff is S .

Let

$\bar{y}$ = court's cutoff level for finding in plaintiff's favor, when $y \leq \bar{y}$.

$\bar{z}$ = patient's cutoff for pressing suit, when $z \leq \bar{z}$.

It is natural to assume that

$$\frac{\partial f(y, z | x)}{\partial x} < 0$$

and
$$\int_{y \leq \bar{y}} \frac{\partial f(y, z | x)}{\partial x} dy < 0$$

whenever y and z are on the lower tail of their respective distributions. Since malpractice is still a rare event in the realm of all medical treatment, these assumptions are valid throughout the range in which such a claim might be made--namely when $z \leq \bar{z}$ and/or $y \leq \bar{y}$.

The frequency of litigation, given a bad outcome, is given by:

$$L(x, \bar{y}, \bar{z}) = \Pr(a=1, z \leq \bar{z}) = p(x) \cdot \int \int_{z \leq \bar{z}} f(y, z | x) dy dz$$

L is the frequency of litigation and $p(x)$ is the frequency of adverse outcomes.

The social objective is to minimize

$$p(x) (SL(x, \bar{y}, \bar{z}) + C) + x$$

Of course, x , $\bar{y}$, and $\bar{z}$ are not all policy variables.

The Malpractice Litigation Equilibrium

The only policy variable is $\bar{y}$. The others are determined by the equilibrium of the system, as follows:

Patient's Decision

Choose $\bar{z}$ to satisfy

$$(1) \quad c \int_{y \leq \bar{y}} f(y, \bar{z} | x) dy = S$$

The result is $\bar{z}(\bar{y}, x)$.

We want to find the signs of $\frac{d\bar{z}}{dx}$ and $\frac{d\bar{z}}{d\bar{y}}$. Totally differentiating (1) with respect to x and $\bar{z}$:

$$c \int_{y \leq \bar{y}} \frac{\partial f(y, \bar{z} | x)}{\partial \bar{z}} dy d\bar{z} + c \int_{y \leq \bar{y}} \frac{\partial f(y, \bar{z} | x)}{\partial x} dy dx = 0$$

we have

$$\frac{d\bar{z}}{dx} = \frac{- \int_{y \leq \bar{y}} \frac{\partial f}{\partial x} dy}{\int_{y \leq \bar{y}} \frac{\partial f}{\partial \bar{z}} dy} < 0$$

under the assumptions that

$$\frac{\partial f}{\partial x} < 0 \quad \text{and} \quad \int_{y \leq \bar{y}} \frac{\partial f}{\partial \bar{z}} dy < 0$$

over this range of integration.

Totally differentiating (1) with respect to $\bar{y}$ and $\bar{z}$:

$$c \int_{y \leq \bar{y}} \frac{\partial f}{\partial \bar{z}} dy d\bar{z} + c f(\bar{y}, \bar{z} | x) d\bar{y} = 0$$

and thus,

$$\frac{d\bar{z}}{d\bar{y}} = \frac{-f(\bar{y}, \bar{z} | x)}{\int_{y \leq \bar{y}} \frac{\partial f}{\partial \bar{z}} dy} > 0$$

under the assumption that y and z are positively related given x .

Doctor's Decision

Choose x to minimize

$$(2) \quad x + C_p(x) \int_{z \leq \bar{z}} \int_{y \leq \bar{y}} f(y, z | x) dy dz$$

The optimal care level is denoted $x(\bar{y}, \bar{z})$. Thus, the first-order condition is

$$(3) \quad 0 = 1 + C_p'(x) \int_{z \leq \bar{z}} \int_{y \leq \bar{y}} f(y, z | x) dy dz$$

$$+ C_p(x) \int_{z \leq \bar{z}} \int_{y \leq \bar{y}} \frac{\partial f(y, z | x)}{\partial x} dy dz$$

Totally differentiating (3) with respect to x and $\bar{y}$ we find,

$$\frac{\partial x(\bar{y}, \bar{z})}{\partial \bar{y}} = - \frac{p'(x) \int_{z \leq \bar{z}} f(\bar{y}, z|x) dz + p(x) \frac{\partial f(\bar{y}, \bar{z}|x)}{\partial x}}{\Delta}$$

where Δ , the second derivative of (2) with respect to x , is positive by the second-order condition. The numerator is negative since $p'(x) < 0$ and $\frac{\partial f(\bar{y}, \bar{z}|x)}{\partial x} < 0$ in this range of integration. Thus

$$\frac{\partial x(\bar{y}, \bar{z})}{\partial \bar{y}} > 0.$$

Similarly we find that

$$\frac{\partial x(\bar{y}, \bar{z})}{\partial \bar{z}} = - \frac{p'(x) \int_{y \leq \bar{y}} f(y, \bar{z}|x) dy + p(x) \int_{y \leq \bar{y}} \frac{\partial f(\bar{y}, \bar{z}|x)}{\partial x} dy}{\Delta} > 0.$$

The properties of these response functions assumed in the text are therefore verified.

Equilibrium

Let the functions $x^*(\bar{y})$ and $\bar{z}^*(\bar{y})$ satisfy

$$x^*(\bar{y}) = x(\bar{y}, \bar{z}(\bar{y}, x^*(\bar{y})))$$

and

$$\bar{z}^*(\bar{y}) = \bar{z}(\bar{y}, x(\bar{y}, \bar{z}^*(\bar{y})))$$

These define the outcome for every fixed $\bar{y}$. The outcome is

clearly unique because $\frac{d\bar{z}(\bar{y}, x)}{dx} < 0$ and $\frac{dx(\bar{y}, \bar{z})}{d\bar{z}} > 0$.

It is straightforward to derive that $\frac{dx^*(\bar{y})}{d\bar{y}} > 0$ and that $\frac{dz^*(\bar{y})}{d\bar{y}}$ is ambiguous in sign. Clearly the frequency of litigation, L , can also respond in either direction.

The Peer Review Equilibrium

Here, $\bar{z}$ does not play any role since investigations are done at random among all cases experiencing adverse outcome. Let ρ be the proportion of such cases investigated.

Physicians set x to respond to ρ and $\bar{y}$, which are now independent controls of the regulatory agency. They minimize

$$(4) \quad x + Cp(x)\rho \int_{y \leq \bar{y}} f(y, z|x) dz dy$$

This gives the first order condition:

$$(5) \quad 0 = 1 + Cp'(x)\rho \int_{y \leq \bar{y}} f(y, z|x) dz dy \\ + Cp(x)\rho \int_{y \leq \bar{y}} \frac{\partial f(y, z|x)}{\partial x} dz dy$$

If ρ is fixed at the level of the propensity to litigate in a malpractice equilibrium with care standard $\bar{y}$, a comparison of (5) and (3) leads to the conclusion that litigation will induce a higher equilibrium care level than peer review. However, since peer review can raise $\bar{y}$ without changing ρ , it can be the case that a higher care level would be attainable at this ρ , and hence that a superior outcome would be attained.

Footnotes

1. This has been forcefully argued in S. Shavell's paper "Theoretical Issues in Medical Malpractice" in this volume.
2. Some of this is documented by R. Epstein's paper "Medical Malpractice: Its Cause and Cure" in this volume.
3. Data from the American Insurance Association's Special Malpractice Review: 1974 Closed Claim Survey indicate that somewhat under 50% of claims closed in that year received no award. Comparable data in the Secretary's Commission on Medical Malpractice indicates a 40% zero award rate in the 1965-1971 period.
4. There are two special points that should be made in this connection. The first is that if the benchmark is too unfavorable to physicians, they will respond simply by refusing to take on the riskier portion of the patients and cases presented to them. Since it would be virtually inconceivable to avoid this problem while maintaining an adversary malpractice system, the benchmark necessarily is set so that malpractice, defined as a shortfall from the standard, is a rare event. In this way the problem of a patient-selection bias is largely offset for most specialties.
 Secondly, the benchmark is actually arbitrary, but its location will affect the identity of the plaintiff in an adversary proceeding. Differential treatment of plaintiffs and defendants will then have an impact on the choice of the benchmark.
5. Data from the American Insurance Association survey cited above imply an average time from filing suit to closing of 12.2 months on all claims closed with positive awards in 1974 and, weighted by the dollar value of settlements, this lag is 14.7 months.