Medical Malpractice: Some Implications of Contract and Arbitration in HMOs

CARL M. STEVENS

Department of Economics, Reed College

Most malpractice claims are processed by the tort-law system, which includes, in addition to the decision rules embodied in the law and the process by which these rules are applied, the settlement-negotiation process, which disposes of 90 percent or so of claims without resort to trial and 95 percent or so of claims without a finally adjudicated outcome. In recent years, in the context of the medical malpractice "crisis," the tort system has been the subject of considerable study, evaluation, and adverse criticism, a good bit of which has advocated abandoning the whole fault-finding approach in favor of some form of no-fault approach (American Bar Association, 1976; Institute of Medicine, 1978; State of New York, 1976; Schwartz, 1976; Havighurst and Tancredi, 1973; Havighurst, 1975; O'Connell, 1975).

One response to alleged failures of the tort system has been to urge the superiority of contract in this domain. In what follows, I suggest that health maintenance organizations (HMOs) afford a peculiarly appropriate institutional setting for developing a contract-based approach to medical malpractice. Indeed, it may be argued that only by explicitly contracting on provider liability and related performance standards can the parties to HMOs fully exploit the inherent advantages of this kind of delivery system.

Although most medical malpractice claims have been and are dis-
posed of by the tort-law system of dispute management, for many years a few such claims have been disposed of by arbitration systems. Recently there has been some move toward more arbitration of such claims, although the volume of arbitration in this domain is still modest. In what follows, I will suggest that arbitration (rather than trial) is the appropriate adjudication process for disposition of contract-based malpractice claims against HMOs. More generally, in the design of dispute-management systems, evaluation of the relative merits of different theories on the basis of which claims might be asserted requires that these be considered in the context of the adjudication modes to be associated with them. And, in turn, evaluation of the relative merits of different adjudication modes requires that these be considered in the context of the theories that provide the basis for asserting the claims to be adjudicated.

HMOs and the Contract Approach: Initial Considerations

Usually, claims against providers charge negligent conduct, the plaintiff claims that the defendant provider has committed a tort. At the

---

1 According to the Research Institute of the American Arbitration Association (AAA), fourteen states between 1975 and 1979 enacted statutes specifically authorizing medical malpractice arbitration, viz: Alabama, Alaska, California, Georgia, Illinois, Louisiana, Maine, Michigan, North Dakota, Ohio, South Dakota, Vermont, Virginia, and Puerto Rico. Courts in all states enforce some agreements to arbitrate medical malpractice disputes, for the most part under modern general arbitration laws. The AAA further reports that there are fifteen state-wide or local private programs for malpractice arbitration, four in California, two each in New York and Washington, and one each in Illinois, Massachusetts, Minnesota, New Mexico, Pennsylvania, Virginia, and Wyoming.

As of July, 1979, the AAA’s Medical Malpractice Research Data Base included 205 closed cases that entered arbitration after 1970 in fourteen states under ten different formal arbitration plans and various ad hoc arrangements. Seventy-five percent of these cases were processed under two state-wide programs—the private California Hospital and Medical Associations’ program, and the statutory Michigan program (both administered by AAA). Although the Kaiser-Permanente (K-P) health plans in California (and elsewhere) have had arbitration systems for several years, no K-P cases are included in the AAA data set. In addition, the AAA has identified individual provider plans in nine states (American Arbitration Association, 1979). Heintz (1975, 1979) has reported extensively on the Southern California Arbitration Project.
same time, the provider-patient relationship is ordinarily considered
to be based on contract, expressed or implied, such that, theoretically,
if not frequently in practice, many malpractice claims could be cast
in terms of breach by the provider of a provider-patient contract
(Waltz and Inbau, 1971:40). It has been urged that contract law
might afford a superior basic theory of medical malpractice. In a
perceptive discussion, Epstein makes the case for contract, pointing
out, in part:

The typical malpractice case raises issues of both tort and contract
law. The physical injury suffered by the patient is quite sufficient
to place the tort element in sharp relief. The contractual element
in medical practice is borne out by the simple fact that the physician
does not conscript unwilling persons to be his patients. When
malpractice cases are treated as though they raise only tort issues,
there is the unmistakable tendency to treat the judicial rules as the
inflexible commands of positive law. It becomes, therefore, a natural
if unfortunate tendency for courts to overlook the possibility, indeed
the desirability, of having the rules that they have laid down varied
by the agreements between the parties. Where the situation is
looked upon as contractual, the basic rules governing the relation­
ship between physician and patient are best understood as approx­
imations of the rules which the parties themselves would choose
to govern their own relationship.... And within the contractual
orientation, we encourage the parties by private means to develop
a set of individuated responses of the sort precluded by the rigid
form of the tort law.... There are of course problems with private
agreements and there are imperfections in the marketplace.... Yet
it is not possible to dismiss contract solutions and a market ori­
tentation simply by pointing to these problems. For while it is easy
to say that contract rules shall be disregarded, it is very difficult
to fashion public standards, be they judicial or legislative, that
function better than the contract rules they replace. (Epstein,
1976:93–95)

In making the case for contract, Epstein remarks upon arbitration
only in passing, expressing his view that arbitration is consistent with
his theme that contract solutions to medical malpractice problems are
in general superior to administrative solutions.

With the decentralized, fee-for-service delivery system, and al­
though there is a contract implied by the parties' relationship, there
will usually be no formal, explicit contract between provider and
patient. The HMO-type delivery-financing system, on the other hand,
does feature an explicit contract between the parties. This kind of delivery system features a very distinctive provider-patient relationship in which the HMO contracts with members to arrange the provision of a stipulated bundle of health care services in exchange for periodic dues payments by the members. (Under conventional health insurance, on the other hand, the insurer contracts only to pick up some or all of the tab for services utilized by the insured, leaving it up to the insured to find a provider of the services as best he can.)

As matters now stand, HMO membership contracts do not spell out the HMO's duty to the member with respect to proper performance of the substantive terms of the membership contract (i.e., in terms analogous, say, to definition of the provider's duty afforded by legal negligence rules). However, as Curran and Moseley have pointed out, a contract approach to malpractice claims seems natural in the HMO context:

Contractual liability, however, seems particularly appropriate in the HMO context. Although there is an express contract between the member patient and the HMO which may not contain specific assurances of high quality care, these terms may usually be implied. The HMO, after all, has agreed to meet the member's every health need up to well-defined limitations and to furnish an acceptable physician for these purposes, and whether that physician is considered an agent of the HMO or the HMO an agent of the physician it would not be unreasonable for the member to infer a guarantee that high standards of quality will be met. (Curran and Moseley, 1975:75)

There is one major problem in the contract approach to performance standards in the medical services sector generally to which the HMO form of delivery system may be regarded as responsive. The general case for freedom of contract (namely, that generally the parties are wiser about their own affairs than are others, including government regulating through its legal institutions) may seriously be questioned if the parties are in some essential way unequal, such that one may opportunistically take advantage of the other. For the usual run of economic transactions, competition in the marketplace, providing the marketeers with alternatives, is supposed to take care of this. In the medical marketplace, however, competition operates only weakly. Moreover, as has frequently been emphasized, there is a large in-
equality in the amount of information possessed by the patient and by the provider, such that the latter may have a substantial advantage. How might the contracting patient cope? In other domains the “collective contract” has been regarded as affording an answer. For example, Wellington, after considering a number of noncontract alternatives to the admittedly frail individual contract of employment, remarks:

> The drawbacks to these alternatives to contract ... and the values which support the freedom of contract doctrine make the case for collective bargaining an appealing one. If the union bargains for the worker, perhaps the contract between union and employer can be treated according to the usual freedom of contract dogma.... The collective contract simply is less likely to be unfair to one of the parties than an individual contract, and it is more likely to reinforce important societal values than its alternatives. (Wellington, 1968:37)

Similarly, the members of an HMO, although not in any usual sense “organized” qua members, do comprise a collectivity in their dealings with the HMO. Indeed, their substantive contracts (those spelling out the services to be delivered in exchange for periodic dues payments) are in a real sense “collective agreements.” Most members of HMOs are group members in consequence of employment-related health plans—the same terms apply to all members of the group, and the employer (and union, if any) are parties to the contract.² Contract terms spelling out the HMO’s duty to the members with respect to standards of care (proper performance of the substantive terms of the membership contract) would, like the substantive terms themselves, be part of a collective agreement. Bovbjerg (1975:n.63, 1395) has alluded to the possibility of HMO members bargaining about care standards, commenting that “an HMO’s subscribers, at least as a

²That these circumstances may warrant distinguishing HMO provider-client contracts has been recognized: “The primary feature of *Wheeler* that distinguishes it from *Madden* is that *Madden* involved a prepaid health plan (Kaiser), and *Wheeler* involved St. Joseph’s Hospital, a nonprepaid facility. The *Wheeler* court assumed that the plaintiff in *Madden* was represented by her employer in bargaining with Kaiser concerning the inclusion of the arbitration agreement in the contract for group medical services. Therefore, both parties had equal bargaining power. The patient in *Wheeler*, on the other hand, was ‘negotiating’ with the hospital by himself” (Bukata, 1978:n.35, 406).
group, may well be able to bargain over the general style of their medical care, including, for example, what facilities are to be provided and whether physician assistants are acceptable.” He also pointed out that “the extent to which agreements on risk reduction between provider and patient or enrollee ought to influence or supersede malpractice standards is an important, difficult, and seldom considered question” (Bovbjerg, 1975:n.119, 1412).

Contract Terms: General Considerations and Barring Claims Based on Tort Theories

Generally speaking, a major advantage in moving from the tort system to a contract system *cum* arbitration is that the latter would facilitate development of definitions of the provider’s duty, and facilitate development of ways to manage disputes about compliance with that duty that would prove superior for both provider and consumer to those featured by the tort system. Thus both stand to gain from their bargain on this score and this is the spirit that should inform the devising of contract terms. (Although both stand to gain from their cooperation, there may also be a competitive element, i.e., there may be room for difference of opinion about the terms on which they shall cooperate.) Pursuant to this, the parties to HMO performance-standards contracts would best serve their mutual interests by explicitly barring claims based on liability for negligence as defined in tort law. In addition to advantages in principle, such a bar has a practical aspect. From the point of view of the HMO, development of performance-standards contracts to spell out the provider’s duty to the members might entail an unacceptable increase in exposure, unless such contracts could also bar claims that in effect contend that the provider has committed a tort as conventionally defined. That is, without such a bar, the provider would confront all of the exposure historically confronted on tort account, plus additional exposure in consequence of other terms of the contract.³

³These considerations raise the question of whether provider-client performance-standards agreements *cum* arbitration legally may preclude claims based on public rights. Parties to such agreements legally may give up their rights to trial by jury in a court of law. But giving up the right of access to a particular forum is not the same thing as giving up a cause of action. More
Social Functions of the Law of Negligence

The law of negligence has been supposed to serve not only the private interests of plaintiffs who may sue in tort but also important social functions, e.g., as by promoting efficient rates of resource allocation to accident prevention. Suppose that the parties (as suggested foregoing), deeming it to their mutual advantage, agree to bar claims based on negligence so defined. They might in this way serve their own private interests. But, might they not at the same time deny service to the public interest in the social functions of the law of negligence? The answer to this question depends partly upon how effective the conventional negligence rule may be expected to be in securing the social benefits attributed to it. If it is not very effective, then, in any event (i.e., whatever the contract terms the parties may substitute for it), not much will be lost on public-interest account by abandoning it. This, it may be argued, is indeed the case. From the point of view of service to its supposed social functions, the law of negligence suffers from a number of technical problems that have not been adequately remarked. Moreover, as will be pointed out, parties to HMO performance standards contracts are in a good position to respond to some of these problems.

The general common law rule mandates that a physician (or other provider) has the obligation to the patient to possess and employ such reasonable skill and care as are commonly had and exercised by reputable, average physicians in the same general system or school of practice in the same or similar localities. Owing in part to the vague-

particularly, is the concept of negligence as defined in tort law a "vested common law cause of action" such that a statute authorizing voluntary agreements to give up the right to assert claims based on this cause of action would be unconstitutional? Fortunately, it would appear that statutes can constitutionally authorize the execution of such agreements (Amicus Curiae Brief on Behalf of the California Hospital Association in Support of Petitioners, 2nd Civil No. 51239, 31 et seq.). We may note that under workers' compensation and no-fault auto laws, plaintiffs give up the right to sue in tort and this appears to pass constitutional muster (O'Connell, 1975: Appendix 5).

Various dysfunctions (private and social) of fault-finding litigation under tort law have effectively been exposed (see, e.g., Havighurst, 1975; O'Connell, 1975). I restrict my attention here mainly to a few technical problems that seem to me to warrant additional attention.
ness of this negligence concept (as applied operationally in particular cases), many students and practitioners believe that the Learned Hand formulation of the negligence standard affords a more useful approach. Thus, according to Posner:

It is time to take a fresh look at the social function of liability for negligent acts. The essential clue, I believe, is provided by Judge Learned Hand's famous formulation of the negligence standard—one of the few attempts to give content to the deceptively simple concept of ordinary care.... In a negligence case, Hand said, the judge (or jury) should attempt to measure three things: the magnitude of the loss if an accident occurs; the probability of the accident's occurring; and the burden of taking precautions that would avert it. If the product of the first two terms exceeds the burden of precautions, the failure to take those precautions is negligence. (Posner, 1972:32)

Thus, under the Hand rule, the expected costs of accidents are to be weighed against the costs of avoiding them. If an injurer has failed to take accident-avoidance steps that would have entailed costs less than the expected costs of the accident, he has been negligent and bears the cost of the accident. The victim, on the other hand, would bear the costs of those accidents "not worth avoiding"—i.e., such that the costs of avoidance are greater than the expected cost of the accident. A social-function advantage claimed for this rule is that it establishes incentives that should, in principle at least, contribute to minimizing the total costs of accidents and accident prevention. In this sense, proponents contend, the rule helps to achieve efficient rates of resource allocation to the various economic activities.

Schwartz and Komesar (1978) have urged the appropriateness of the Hand formulation for analysis of the function of the law of negligence in the context of medical malpractice. They suggest that, since in practice not all untoward events can be prevented, the Hand rule be modified to define negligent behavior as the failure to invest resources in accident prevention up to a level that equals the expected saving in accident cost.

The Hand rule, per se, seems to be silent on the matter of what role, if any, is to be assigned to contributory negligence. Calabresi and Hirsch (1972) have pointed out that the addition of a modified contributory negligence rule would improve the theoretical efficiency of the Hand rule.
Critical in the evaluation of any negligence rule, or, more generally, any liability rule, is not only the theoretical efficiency of the rule, if properly implemented, but also the probability that, in practice, the rule can properly be implemented. The Hand rule leaves a good bit to be desired from this point of view, particularly in the context of existing medical malpractice institutions.

However, the Hand rule (at least as modified by Schwartz and Komesar) in the domain of medical malpractice confronts a number of conceptual problems that run deeper than what properly might be characterized as implementation problems, per se. One such problem is how to operationalize the concept of “failure to invest resources in accident prevention.” Some mishaps, e.g., those owing to incomplete diagnostic workups, might seem to fit this paradigm in a fairly natural way—as, say, failures to invest enough provider time. However, the paradigm would not seem naturally to comprehend some categories of claims that, most would agree, might appropriately be asserted pursuant to a negligence rule. For example, to characterize mishaps owing to lack of provider knowledge or lack of provider skill or expertise as instances of that provider’s “failure to invest resources” would surely strain the meaning of that concept in many such cases. Such mishaps might be comprehended as failure of the medical services system to invest enough resources, e.g., in the selection and training of providers. Negligence law and the dispute-management procedures associated with it address incentives to the parties and prospective parties to negligence actions, namely, the consumers of services and the individual providers of those services (as well as counsel for these parties and insurers). The decision-making process that determines a medical services-system parameter, such as the rate of resource investment in the selection and training of providers, is a complex, multiparty process. Included among these parties are the medical schools and their associations (e.g., the Association of American Medical Colleges), the physicians and their associations (e.g., the

---

6 A point emphasized by Calabresi and Hirschoff (1972), who propose a formulation they regard as more likely than the Hand formulation, in practice, to accomplish accident-prevention cost minimization.

7 Schwartz and Komesar (1978) afford a discussion of some of these implementation problems. They compare what they regard as the “ideal” negligence signal called for by the Hand formulation and the negligence signal generated in the “real world.”
American Medical Association), the hospitals and their associations (e.g., the American Hospital Association), the specialty boards, the state and federal legislators who provide funding for medical education, and various additional parties. These parties are not, as such, prospective parties to negligence actions claiming provider malpractice. That is, most of the decision makers who, collectively, are important for determining the rate of resource allocation to selection and training of physicians and, hence, to this aspect of accident prevention, are insulated from the incentives afforded by administration of the law of negligence in accord with the Hand Rule (or, indeed, any other rule). These considerations mean that, at least in the medical services sector, to contend, as do Schwartz and Komesar (1978:3), that "the Learned Hand Rule serves to assure that resources are being efficiently allocated ... by establishing procedures that minimize the total cost incurred by accidents and accident prevention" is to make a claim that is dubious at best, even at the level of the theoretical efficiency of the rule.

How might a remedy for the foregoing problem be found within the context of the law of negligence? In rather abstract principle, one approach that would be responsive to this problem would be for plaintiffs in malpractice actions to bring not only professional-services suits against physicians but also what might be thought of as product-liability suits against, say, the medical schools and residency programs responsible for selecting, training, and professionally motivating the physicians who turn out to be defendants. To characterize this approach as responsive, in principle, to the problem helps to elucidate the nature of the problem. But this approach can scarcely be regarded as a practicable solution. In what sense can medical schools and residency programs, say, be expected to guarantee the knowledge, skill, and professional responsibility of each of their "products"? It is true that such training programs must be accredited. The nature of and rationale for the accreditation process would seem to imply that these training programs ought to be able to guarantee that each physician possess some minimum level of knowledge, skill, and professional responsibility at the time of exit from the program. But, for events and circumstances beyond that exit point, these programs can assume no responsibility.

There probably is no solution to this problem strictly within the context of administering the law of negligence. A more hopeful ap-
Medical Malpractice and HMOs

approach would seek some institutional bridges such that the outcomes of individual malpractice cases would provide information to other agencies, which would then take appropriate action on the basis of the information. Some such arrangements do exist (although not, so far as I know, addressed to problems in training programs).

For example, a big factor affecting the management of malpractice suits in California is the Board of Medical Quality Assurance. All recoveries against providers in excess of $3,000 are reported to the board, which may then elect to investigate the circumstances of the case. Investigation by the board is mandated when the total of recoveries against a provider is in excess of $30,000. If the board finds something amiss, the sanctions available to it include restriction or limitation of a physician's practice to certain types of procedures or, in an extreme case, revocation or suspension of a provider's license. Pursuant to the problem here being considered, the purview of such a board could be extended. The board could review all cases in which there were recoveries against physicians, to determine whether there was any tendency for the physicians involved in misadventures to be the product of certain training programs or certain kinds of training programs. If the record revealed such tendencies, the board could investigate these programs. In principle, a possible outcome of such investigation could be a recommendation by the board that an increased allocation of resources to these training programs would be an investment in accident prevention that would be worth the cost in terms of accidents averted. An investigation or trial would be required to determine whether, in practice, one might anticipate any useful yield from adopting such procedures. What can be said is that, in principle, such procedures would provide a more general link between the outcomes of malpractice actions and decisions to invest in (some aspects of) accident prevention than can now be provided by the law of negligence as it is administered.

An assessment of the social significance of following a Hand-type negligence rule suggests an additional, important technical issue. This turns on how the "cost" of accidents is to be measured. One approach

---

8 Established by the Medical Injury Compensation Reform Act of 1975 (MICRA) to replace the former Board of Medical Examiners. The board's Division of Medical Quality is directed to take action against any certificate holders guilty of "unprofessional" conduct, with "incompetence" identified as a form of unprofessional conduct.
would be to appeal to the performance of the malpractice system and reckon as the cost of any given kind of accident plaintiffs' actual average recovery. This measure would be directly relevant for the management of, say, a self-insured delivery system making decisions about how much to invest in accident prevention. It would not, however, necessarily be relevant from the point of view of the social function of the negligence rule. In selecting an appropriate measure of accident costs, it is important to keep in mind that consumers in the aggregate pick up the tab for accident costs and the costs of accident prevention—by incurring the costs (monetary and psychological) of accidents, by paying health insurance premiums and taxes, and by making out-of-pocket payments for services. (Damages assessed, premiums to insure liability, and prevention costs are, from the provider's point of view, costs of doing business that will be reflected in the price of the product.) The question to which the negligence rule must generate the correct answer is whether any given investment in accident prevention is worth the cost to consumers. That is, the relevant evaluation standard for investment in accident prevention is whatever it is worth to consumers, \textit{ex ante} the occurrence of any of various mishaps to reduce the risk that they will experience the mishaps.\textsuperscript{9} It would appear that only by chance would the \textit{ex ante}-accident evaluation correctly be measured by \textit{ex post}-accident recoveries generated by the medical malpractice dispute-management system. The costs represented by plaintiffs' recoveries do, of course, have some relevance for decisions about investment in prevention. The malpractice system will generate some rate of recovery by victims, and consumers in the aggregate will pay this tab.

If an additional dollar in prevention will save more than a dollar in accident costs, rational consumers will want to invest the dollar

\textsuperscript{9}This same standard is the relevant one for evaluating efficient rates of resource allocation, not only to accident prevention, per se, but also to any life-saving, morbidity-reducing activity, e.g., the medical services sector as a whole. For discussion of this see Schelling (1966). As Schelling points out, it may be difficult for consumers to establish their own preferences with confidence on this score. The problem is that, even given good technical data relating investment in life-saving and morbidity-reducing activities to risk reduction, the consumer would still typically confront the problem of evaluating the worth of small reductions in very small probabilities of the occurrence of untoward events—the very prospect of which may evoke high levels of anxiety.
in prevention. But this investment decision is based upon only a subset of the factors the consumer will want to take into account, and this subset of factors may well not be decisive for the preferred rate of investment in prevention. Suppose that any given consumer wants to invest at a given rate to reduce the risk that he, or members of his family, or his friends (or others in his utility function) will, say, suffer brain damage owing to anesthesia accidents. That preference is based on the consumer's evaluation of the expected disutility of the untoward event, an evaluation that is made independently of whatever rate of recovery the extant malpractice system happens to be generating for victims of these particular accidents—and that rate of recovery will only by chance motivate the rate of investment in prevention the consumer would elect on the basis of his more fundamental risk-aversion preferences.

It seems likely that consumers' *ex ante* accident risk-aversion preferences would call for a higher rate of investment in prevention than would be warranted on the basis of *ex post* accident recoveries generated by the extant malpractice system. This is so because for various reasons the extant system probably tends to undercompensate victims in the aggregate. For one thing, a significant number of meritorious claims are never asserted. Also, according to data published by the National Association of Insurance Commissioners (NAIC), indemnities paid for incidents with economic loss (past and anticipated medical expense, past and anticipated wage loss) of $100,000 and over are less on average than the economic loss itself, i.e., there is no compensation for the real costs of pain and suffering. These data show, however, that indemnities paid for incidents with small economic loss are in excess of that loss—a result one might expect for very small claims where the parties are likely to settle for an amount that is largely determined by what it is worth to avoid the cost of adjudicating the claims (National Association of Insurance Commissioners, 1977).

If the extant malpractice system were to generate recoveries that would warrant a higher rate of investment in accident prevention than would be motivated by the consumers' more fundamental, *ex ante* preferences, the former will presumably determine the investment rate.

Much more could be said about the social efficiency of the law of negligence. The foregoing analysis has been intended only to direct attention to a couple of technical points that mean that, even at the
level of theoretical efficiency, the law of negligence is unlikely to achieve the resource allocation benefits sometimes attributed to it. Moreover, the analysis points to a fundamental advantage for consumers in contracting with providers on performance standards. Rather than relying on administration of the law of negligence to determine efficient rates of investment in accident prevention, consumers would be wiser to negotiate such rates with providers. Consumers could, in this way, directly map their *ex ante*-accident risk-aversion preferences into the decision. In the decentralized, fee-for-service delivery setting, there may be no very feasible way to accomplish such negotiations. It is a peculiar advantage of organized delivery systems such as HMOs that it would appear feasible to take more or less systematic account of consumers’ risk-aversion preferences in making decisions about investment in prevention. Thus, for example, HMO members might, through suitably selected representatives or agents, negotiate with HMO management the decision about prevention programs and strategies to be adopted, weighing the cost of such programs (reflected in the dues the members pay) against what the reduction in risk is worth (as best this might be estimated). In practice, such negotiations would probably be addressed not just to accident prevention, per se, but more generally to the overall quality of delivery system performance as this might be affected by such factors as the supply of physicians and other health manpower and facilities to be afforded per member.10

Contract Terms: Some Further Considerations

Some students appear to contend that there is one appropriate social function of liability rules, namely, to promote economic efficiency. The fact is, however, that malpractice institutions based on such rules,

10 Although my discussion has focused on medical malpractice, the points on the law of negligence likewise may be relevant to the case of product liability. Markets for some products may permit the consumer to bargain with producers by choosing among a number of models, each featuring a different rate of investment in accident prevention and each a different price reflecting the investment rate. In medical services markets, however, to bargain on this score with providers, consumers will probably have to negotiate, i.e., rather than play take-it-or-leave-it in the marketplace.
and the dispute-management systems associated with them, in practice discharge a number of different social functions. Among these is the compensation of victims. In consequence, they have a distributional impact, transferring income from consumers in the aggregate to the victims. For this function, the malpractice system should be responsive to canons of equity and justice (not just to canons of economic efficiency).

The problem of achieving distributional equity is a peculiarly vexing one in the context of the extant malpractice system, in part because of the way the system structures the decision-making process. Eligible victims recover for economic loss and pain and suffering in amounts determined by adjudicators (judges, juries, arbitrators) or by settlement negotiations. Perhaps there would be general agreement that equity requires that victims be made whole for economic loss. But what is an appropriate recovery for pain and suffering (a rubric that is supposed to comprehend all of the disutility suffered by the victim other than economic loss)? One answer is to accept as the appropriate recovery on this score whatever recovery is yielded by the proper decision-making process (e.g., trial by jury, or arbitration, or—as in most instances—settlement negotiations). Given the recovery rate, whatever "tax," levied on consumers in the aggregate, was necessary to finance it would likewise be regarded as appropriate. That is, the tax loading (on prices of services) would simply follow, given independent determinations of appropriate recoveries.

This approach, it may be argued, leaves a good bit to be desired, because it does not come directly to grips with the fact that gains for persons in their roles as victims mean losses for persons in their roles as nonvictims, such that an appropriate decision-making process should simultaneously take account of these distributional effects. What kind of decision-making process can achieve this?11 Trial by jury (or

---

11 In a recent study sponsored by the California Medical Association (CMA) and the California Hospital Association (CHA), a systematic attempt was made to define and measure the prevalence of "potentially compensable events" in a sample of California general hospitals. These events were defined in a way suitable to inform the design of various models of patient-disability (more or less no-fault) compensation systems (Mills et al., 1977). An actuary's analysis of the data developed by this study concluded that the cost in California of a no-fault system would run between $700 million and $1.5 billion per year. But, if we assume an efficient dispute-management system, as the proponents of no-fault contend such a system would be, most of the
by judge, or by arbitrator) is not well suited to accomplish this. Even if one makes the assumption that these adjudicators strive to perform as effective, responsible agents representing the interests of both victims and consumers in the aggregate, it would remain to explain how, in fact, they are able validly to represent these competing interests. It is the peculiar advantage of an HMO delivery setting, on the other hand, that it does afford the institutional context for a properly structured decision-making process addressed to this distributional question. The members of the HMO are, at once, prospective victims who will benefit from whatever rates of recovery might be agreed upon and those who will, in the aggregate, finance the recoveries. Thus each member can simultaneously take account of the competitive utility implications for these two roles in arriving at preferences about appropriate rates of recovery—and, given that the members can reach some agreement, they can negotiate the result with the HMO.

Whatever mix of principles (e.g., economic efficiency, equity) the parties resort to to inform the drafting of their performance-standards contracts, the process of selecting particular contract terms may be assisted by considering various suggestions that have been made about the definition and administration of liability concepts in this domain. For example, they might adopt a straightforward no-fault approach, following the lead of the California Medical Insurance Feasibility Study, which took the view that, to the extent possible, compensation in a no-fault system should not be predicated upon whether or not the disability was preventable. With this approach, they might make no attempt to build into the liability system, per se, incentives to reduce the rate of mishaps.

Alternatively, they might try an approach along the lines suggested by Havighurst's "medical adversity insurance" (MAI) concept. According to Havighurst (1975:1249), "the central tenet of the MAI scheme is that financial incentives supplied by liability rules can be

---

total premium expense should be simply an income transfer, i.e., from nonvictims to victims. It is not at all clear how large an income transfer of this kind would be regarded as appropriate by consumers in the aggregate. It should be kept in mind, however, that even a very generous income transfer to victims would entail only a relatively small loading on each individual's cost of health care.
a major guarantor of good-quality medical care.” Pursuant to this, he suggests that “if experience rating could not be expected as a spontaneous development, a possible means of creating quality-assurance incentives in an MAI program ... would be simply to impose a share of the cost of each claim directly on the responsible providers through some kind of deductible or coinsurance requirement” (Havighurst, 1975:1251). Adapting this approach to the HMO context would mean that each physician member of the medical group would be at some direct financial risk. MAI is essentially a no-fault scheme, but it embodies the concept of “liability without fault,” a notion that, as Havighurst observes, physicians and others have some trouble getting used to.

O'Connell (1975) has put forward the concept of “elective” no-fault liability. Under this scheme, providers would be authorized unilaterally to elect no-fault for themselves and their patients, i.e., if a provider so elected, patients would be barred from suits in tort. He anticipates constitutional challenges to this approach, remarking, “Perhaps to some the most disturbing constitutional feature ... is that private persons ... are allowed unilaterally to alter the common law rights of those they insure” (O'Connell, 1975:206). The answer to this problem is for both parties to elect no-fault, i.e., by explicitly contracting on no-fault if they deem it to their mutual advantage. If, as O'Connell holds, patients (and product customers) are really better served by no-fault, then contract is the natural way to get it.

The parties might abjure the no-fault route, staying within the framework that defines liability by reference to customary standards of care. There would still be room for innovation, e.g., along the lines suggested by Bovbjerg. As matters stand, negligence in HMO practice is defined by reference to customary standards of care developed on the basis of experience in the health services sector generally, which is dominated by the fee-for-service delivery mode. Bovbjerg argues that the application of this standard may tend to inhibit efforts by HMOs to develop innovative, more efficient delivery systems. He urges that HMOs be given legal authorization to substitute “HMO custom” as the standard of care against which to measure liability (Bovbjerg, 1975:1408–1409). He points out that his suggestion would in effect allow the subgroup of medical practitioners in HMOs to set their own malpractice standards. Statutory authorization would
be required for HMOs to impose such standards unilaterally. Such authorization presumably would not be required for HMOs and their members explicitly to contract on such standards.

In addition to the foregoing, the parties to performance-standards contracts in HMOs could consider any of various other possibilities. The general point, exemplified by the foregoing suggestions, is that there is a rather rich mix of contract provisions among which the parties might choose in negotiating a package that best accommodates their mutual interests.

Arbitration of Malpractice Claims Against HMOs

I now develop the proposition that arbitration (rather than conventional litigation) is the superior adjudication mode for managing those disputes that arise pursuant to performance-standards contracts in HMOs—particularly if the full potential of the contract approach is to be realized.\(^{12}\)

Arbitration has been a feature of various HMO (e.g., the Kaiser-Permanente Health Plans) malpractice-dispute management systems. For example, the Amendment to Group Medical and Hospital Services Agreement of the Kaiser Foundation Health Plan of Oregon provides, in part:

**ARBITRATION OF CLAIMS**

A. *Claims Subject to Arbitration.* Any claim arising from alleged violation of a legal duty incident to this Agreement shall be submitted to binding arbitration if the claim is asserted: (1)....
(2) On account of death, bodily injury, disease or ailment allegedly arising out of the rendition of, or failure to render, services under this Agreement, irrespective of the legal theory upon which the claim is asserted.

\(^{12}\) I should make it clear that I do not undertake in this essay a general assessment of the relative merits of arbitration and the trial in the domain of medical malpractice. To do so would not only require a very large amount of space, but it would also divert attention from the particular points I want to make—namely, those concerned with the relations between contract and arbitration. There has been much controversy over medical malpractice arbitration, both in the literature and in the field. I have discussed this controversy at some length elsewhere (Stevens, 1979).
Medical Malpractice and HMOs

The membership contracts of the Kaiser Foundation Health Plans in California contain similar arbitration provisions.

A common objection to voluntary agreements to arbitrate medical malpractice claims is that, in many cases, they may not be really voluntary on the part of the patient. When the patient presents to the provider seeking care, it is argued, he may well be somewhat confused by the unfamiliar situation and he may, moreover, be anxiety-ridden, preoccupied with his medical condition, such that he is unable to reflect in a self-serving way on the relative merits of adjudicating any disputes that might subsequently arise in one way rather than another way. Under these circumstances, it is argued, the patient may unwittingly become party to an arbitration agreement. It may be remarked that precisely these same circumstances might result in the patient’s unwittingly failing to become party to an arbitration agreement. Generally speaking, of course, individuals are made better off, not worse off, by an increase in the number of alternatives freely available to them. It requires some special argument to reach the conclusion that the law should not permit the choice of arbitration. The problem is to devise suitable hedges against the possibility of unwitting agreement to arbitrate or unwitting agreement not to arbitrate.

Thus, for example, California’s Code of Civil Procedure Section 1295 provides in part:

(b) Immediately before the signature line provided for the individual contracting for the medical services must appear the following in at least 10-point red type:

NOTICE: BY SIGNING THIS CONTRACT YOU ARE AGREEING TO HAVE ANY ISSUE OF MEDICAL MALPRACTICE DECIDED BY NEUTRAL ARBITRATION AND YOU ARE GIVING UP YOUR RIGHT TO A JURY OR COURT TRIAL, SEE ARTICLE 1 OF THIS CONTRACT.

In addition, the patient may rescind the agreement to arbitrate by written notice within thirty days of signature.

The HMO form of delivery system affords an inherent advantage from the point of view of this problem. The clients agree to arbitrate when they become members of the organizations, as part of their overall membership contract. In the usual case, at the decision point
they are well and going about their ordinary business, not preoccupied with the anxieties of illness. Those who have argued that arbitration agreements in HMOs should not be allowed because, since they are a condition of membership in the health plan, they smack of adhesion, have distracted attention from the real significance of such arbitration agreements. The clients are not, after all, forced to choose membership in an HMO. They have a choice, the HMO delivery system cum arbitration or some other type of delivery system cum actions in tort. They can choose the "package" that seems to them, on balance, the best. The "package" approach is predicated on the notion that the procedure for managing disputes about performance standards is and ought to be regarded as an integral part (along with the delivery system) of the "health care services systems" to which individuals may attach themselves. There is a fundamental reason why this way of looking at the matter is sound. As emphasized earlier in this essay, the choice of contract rather than tort as the legal-theory basis for asserting malpractice claims may have important consequences for the performance of the health services system qua delivery system (e.g., appropriate rates of investment in accident prevention and, more generally, quality assurance). Similarly, the procedures adopted for managing malpractice disputes may have an important bearing on the quality of the overall provider-patient relationship. It would be an error to suppose that the choice of dispute-management procedure is neutral to the performance of the delivery system component of the health care services system of which both are a part.

The association of arbitration and contract-based claims is natural. The parties to a contract establish by mutual assent the substantive rules to govern their relationship: these rules are not imposed by the outside authority of public law. Likewise, their agreement to arbitrate disputes over interpretation and application of their contract has a consensual basis, agreement by the parties upon their own dispute-management process. In short, the contract creates private rights that may appropriately be adjudicated by resort to private tribunals.

For the parties to performance-standards contracts in HMOs, the evolution of a formal grievance procedure that would set up steps before arbitration would greatly facilitate the administration of their contracts.13 Under such a procedure, a grievance would be a claim

---

13 I have discussed this matter elsewhere (Stevens, 1974). My discussion in the text draws on my earlier treatment. For a general discussion of grievance
by the member that there had been a violation of the performance-standards contract and, if the claim were disputed, it would be processed through one or more steps at each of which the parties and their representatives would attempt to reach a resolution of the problem. Failing resolution, the claim would go to arbitration. Thus, under a formal grievance procedure, arbitration of a malpractice claim would be the final step in a multistep dispute-management process.

It should be pointed out that grievance procedures as thus far developed in HMOs are for the most part not intended to process malpractice claims; these are processed through more conventional procedures. Rather, extant HMO grievance procedures are in the main addressed to administrative matters such as dues payments, the services to which members are entitled, and so on. The American Arbitration Association (AAA) has developed a grievance procedure for processing malpractice claims in organized delivery settings, which includes arbitration as the final step (Ladimer, Solomon, and House, 1979). Invoking the AAA procedure, however, does not depend upon the existence of explicit contracts on performance standards such as those advocated in this essay.

A special feature of the relation of arbitration to contract to which I wish to direct attention is that the arbitration of contract disputes is frequently more than just a matter of contract administration; it can also be a matter of contract-making. Indeed, for some of the potentially most important terms of performance-standards contracts, there will be no way other than by administering the contract under arbitration and accompanying grievance procedure to develop an acceptable contract. The most parsimonious way to make this point is by resort to a collective bargaining analogue (Stevens, 1974).

Most collective bargaining agreements provide (in the so-called management-rights clause) that the management may discipline or discharge employees “for due cause”—with this concept being nowhere explicitly defined in the agreement. If an employee feels that management has violated this contract clause, he may grieve. The outcome of many grievance arbitrations, in many contexts, over many years has clothed the concept of “due cause” with operational meaning—a meaning that might be said to reflect the “common law” of the procedures including arbitration in HMOs see Ladimer, Solomon, and House (1979). In general, these authors adopt a systems perspective in which they evaluate arbitration as part of a larger system of dispute management, of which it is but one component.
workplace, those customs of the workplace that generally are regarded as equitable and viable. There really is no other way in which satisfactory meaning can be ascribed to the "due cause" concept in this context. Administering the collective contract under the grievance procedure, for this kind of subject matter, is very much a matter of contract-making.

Turning to the medical care sector for an analogue, we may remark that among the potentially most important terms of explicit contracts on performance standards will be language addressed to the provider's duty of "full disclosure" in the therapeutic relationship. ("Full disclosure," which imposes a more demanding duty of communication on the provider than that imposed by the traditional "informed consent" doctrine in most jurisdictions, is a controversial issue in the provider community. The general statements about it in the text, sufficient for present purposes, do not engage this controversy.) The general idea comprehended by the duty of full disclosure can readily be set out, viz: the provider shares with the patient information about the (differential) diagnosis of his condition, about the treatment alternatives available (including no treatment), and the probabilities (as best these can be estimated) of risks and benefits associated with each. The therapeutic decision becomes a genuine two-party decision. The informed consumer elects the preferred regimen in light of his own preferences over the possible outcomes and in light of his own risk-aversion propensities. It may be argued that a properly complied-with duty of full disclosure can go a long way toward improving the quality of medical care, and it may also diminish the number of malpractice claims. Brittain, one of a group of physicians and physician-attorneys who examined in depth a consecutive series of more than 1,000 malpractice claims, has reported:

As strange as it may seem to many physicians, only a few malpractice suits are initiated specifically because of the damages which the patient will later claim. To the contrary, a majority of malpractice suits are brought because of patient or patient-family anger over something totally peripheral to the event leading to the claimed damages. This may be an excessive bill, or, more commonly, a misunderstood bill, hostility, inattentiveness, abruptness, or any one of many other human characteristics which would cause any of us to turn hostile.... The second most frequent reason why patients consult attorneys about potential medical liability is real
or alleged "surprise."... The law is clear that at least for elective procedures, it is the patient who has the right to decide on whether to be treated or not. Truly "informed" patients are rarely surprised. (Brittain, 1978:19)

Albeit the general idea comprehended by full disclosure can readily be set out, realistically the parties must recognize that in the actual administration of the standard there will be problems and legitimate exceptions to literal compliance with the standard. For one thing, some patients in some circumstances may not want to be as fully informed as literal compliance might urge. Also, in some situations, full disclosure might, in the professional judgment of the provider, have a negative medical impact. Also, there are problems with communication in this domain, both because of the sometimes technical nature of what must be communicated and because the patient's anxiety and fear may impede his comprehension. What is the answer? One answer would be simply to abandon any effort to administer such a standard. But the potential importance of full disclosure may be regarded as too great to accept this solution. Alternatively, the parties might attempt to draft contract language in sufficient detail explicitly to take account of all contingencies. But this is not really a practical solution. The best solution is to leave the contract language setting the duty of full disclosure rather general in nature, and to permit consumers who feel that the duty has not been complied with to grieve. We might anticipate that (as with the "due cause" analogue) arbitration awards would clothe the concept of "full disclosure" with operational meaning—a meaning that would be sensitive to and would reflect the realities and the equities of the provider-patient relationship. Parties to continuing relationships agree to arbitrate, rather than litigate, their contract disputes, in part because they seek a forum with this kind of capacity to contribute constructively to the making of their contract.

In the domain of labor relations, the parties, with their collective agreements, grievance procedures, and arbitration, have constituted a pervasive system of "industrial jurisprudence." Rather than resorting to the public law and its institutions for the enforcement of contracts, the parties have created their own system of private law for these purposes—and, I again emphasize, it is a system with various complementary parts. Similarly, in the domain of organized, medical care
delivery systems, the parties with their performance-standards contracts, grievance procedures, and arbitration could constitute an analogous system of "private medical care jurisprudence." And this private system, as I have elsewhere remarked, could well develop an expertise, sensitivity, and dispatch in the handling of malpractice and related matters scarcely to be anticipated under formal litigation at public law (Stevens, 1974). The development of such a system should be recognized as one of the major potential advantages of organized delivery settings such as HMOs.

As I hope the discussion in this essay has suggested, the implications of contract and arbitration in organized delivery systems such as HMOs are very far-reaching. A medical services system delivers medical services and achieves medical outcomes in the context of a somehow structured provider-patient relationship. From the point of view of the utility experienced by consumers in consequence of participating in the system, it is the whole package that counts. In the decentralized, fee-for-service practice setting, the provider-patient relationship is in important part structured by negligence (and other liability) rules, conventional settlement negotiations, and the prospect of trial. On the other hand, in organized delivery system settings, the provider-patient relationship might, as has been suggested herein, in important part be structured by explicit contracts on performance standards, formal grievance procedures, and arbitration. This system is more likely than is the conventional system to serve the interests of the parties.14

Private Adjudication of Public Rights

I have suggested in this essay that arbitration is peculiarly appropriate for the adjudication of contract-based malpractice claims—peculiarly appropriate, that is, for the adjudication of private rights. Does this imply the other side of the coin—namely, that arbitration is not appropriate for the adjudication of public rights (e.g., tort negligence rules)? This is an important question for the design of optimal medical malpractice dispute-management systems. And, given the increasing

14 See Ladimer, Solomon, and House (1979) for a discussion that urges this same conclusion.
resort to arbitration in this domain and the character of various arbitration provisions, it is far from an academic question.

Some arbitration provisions are very broad, in the sense that they will accommodate claims irrespective of the legal theory upon which the claim is asserted, e.g., the Kaiser Health Plan provision cited earlier in this essay, and the State of Michigan arbitration statute. For example, Sec. 5140 (1), Act No. 140 (State of Michigan Public Acts of 1975) provides:

The provisions of this chapter shall be applicable to the arbitration of a dispute, controversy, or issue arising out of or resulting from injury to, or the death of, a person caused by an error, omission, or negligence in the performance of professional services by a health care provider, hospital, or their agent, or based on claimed performance of such services without consent, in breach of warranty, or in violation of contract.

The Michigan statute, however, is restrictive in a special way. Section 5043 (1) (b) provides: "The prevailing standard of duty, practice, or care applicable in civil action shall be the standard applied in the arbitration." This provision would appear to mandate the private adjudication of public rights. More generally, the argument in this essay would suggest that such a restriction is unfortunate in that it ties the hands of parties to performance-standards contracts in HMOs, such that they may be precluded from realizing the advantages of designing their own standards to reflect their own peculiar preferences.

Proponents of arbitration contend that broad scope is necessary if arbitration is to be a really effective alternative to litigation such that the maximum benefits of arbitration are to be realized. That is, proponents of arbitration generally see no reason why the alleged benefits of arbitration (e.g., lower cost, more expeditious disposition of claims) should be peculiar to claims asserted under some theories but not other theories. Nevertheless, it may be argued that the arbitration of public rights, those established by public law, does raise questions of propriety in a way that the arbitration of contract disputes does not. In the malpractice domain, the proliferation of arbitration schemes represents a kind of invasion by private tribunals of legal turf historically presided over by public tribunals. This development, over the longer term, might lead to some displacement of public law and public legal theories by private law and private legal theories—a result
that, this essay has argued, would be of benefit to the parties. This
development might also lead, however, to private construction and
application of public law—a result that may have untoward impli-
cations for the development of public law. Consideration of some
recent developments in labor law may help to inform judgment about
this matter.

Recent developments in labor law have seen an invasion, by public
law and public tribunals, of turf long presided over by private law
and private tribunals (the reverse of the situation with malpractice
disputes, where private tribunals have been the invaders).

Historically in this country, the “web of rules” to govern in the
workplace has largely been fashioned by the collective bargaining
system, relatively few terms and conditions of employment being
determined by external public law. In recent years, however, there
has been an increasing tendency to substantive federal regulation of
the terms and conditions of employment, including Title VII of the
Civil Rights Act, the statutory provision that has resulted in most
of the private-law/public-law jurisdictional conflict in this domain.
Most collective bargaining agreements include antidiscrimination pro-
visions. In some instances, these provisions are virtually identical with
or incorporate Title VII by reference. In this case, the grievance
arbitrator in a discrimination case, although interpreting and applying
the collective agreement as usually instructed by that agreement, will
also find himself or herself in effect or explicitly interpreting and
applying federal law. Is this appropriate? There is opinion in the
labor-relations community that this is not appropriate (Feller, 1976a,
1976b; Edwards, 1977). In Feller’s view, the labor arbitrator should
stick to his last, should confine his attention to the contract; otherwise,
the whole system of private industrial jurisprudence and arbitration
as part of it will be threatened:

> Deference to arbitral competence was and is difficult to achieve.
And I suggest it will be impossible to maintain if arbitrators extend
themselves and regard arbitration as the tribunal in which broader
policies than those contained in the agreements themselves are to
be enforced.... My view is that the profession and the process are
best protected to the extent that the process is regarded as a spec-
ialized one rather than a generalized one." (Feller, 1976a:110–111)

Edwards is in agreement with Feller’s view that arbitrators should
not take on public-law issues, but contends:
At issue here is not whether arbitration will suffer if arbitrators go beyond collective bargaining agreements in settling disputes.... At issue is the private development of public law. Where arbitrators decide issues of public law, two major problems arise. The first is that they may be wrong. The second is that their errors, if honored by a public tribunal out of deference to arbitration, may distort the development of precedent. (Edwards, 1977:90)

Similar concerns have been voiced in other arbitration contexts, viz:

Arbitration is power, and courts are forbidden to look behind it. The protection of awards against judicial interference and, under that umbrella, of the development of organized arbitration as a rule maker have established “judicial powers” other than those provided by federal and state constitutions. It is not possible to maintain any legally established policy or order ... if courts abdicate their power in favor of private tribunals serving private interests. (Hessen, 1965:64)

Obviously, labor relations and medical malpractice relations represent very different kinds of institutional situations—e.g., disputes with entirely different topical content, and a very different legal context as this engages the relation of private tribunals to public law. Nevertheless, the issues raised are relevant to the arbitration of medical malpractice disputes. Some medical malpractice arbitration schemes, such as the Michigan statute cited earlier, put the arbitrator in the business of adjudicating public rights established by public law. Are concerns such as those expressed by Feller, that the arbitration system will lose viability if it is extended to the adjudication of public rights, relevant in the domain of medical malpractice? Or, are concerns such as those expressed by Edwards and Hessen, namely, that the private development of public law may have untoward consequences, relevant in this domain? In the extensive debate over the merits of arbitrating malpractice disputes, these questions get very little attention. In my view, these concerns are relevant to an evaluation of the merits of medical malpractice arbitration and do cast doubt on the propriety of arbitrating public rights. By themselves, however, they cannot be regarded as decisive for the choice between arbitration and litigation. To fully inform this choice, a number of additional factors, namely, all of those upon which the quality of justice yielded by these dispute-management systems depend, must be taken into account.
References


Amicus Curiae Brief on Behalf of the California Hospital Association. 1977 (September). In the Court of Appeal, Second Appellate District, State of California, Division One. 2nd Civil No. 51239.


This work has been supported by the National Science Foundation Program in Law and Social Science. The author expresses his thanks to Prof. Clark C. Havighurst for helpful comments on an earlier draft.

*Address correspondence to:* Prof. Carl M. Stevens, Department of Economics, Reed College, Portland, OR 97202.