RETHINKING THE ALLOCATION OF MEDICAL MALPRACTICE RISKS BETWEEN PATIENTS AND PROVIDERS

Glen O. Robinson*

I

INTRODUCTION

A distinctive feature of medical malpractice law is its acceptance of medical custom as the dispositive legal standard of care. Outside the field of medical malpractice, long established legal doctrine accords limited weight to custom in defining the general standard of care in negligence cases. In medical practice, by contrast, the legal standard of due care is virtually defined by the customary skills and practices of the profession. A showing of subconformity to custom here ordinarily yields liability, while conformity ordinarily forecloses it; as a corollary, the metes and bounds of custom are required to be established by medical experts qualified to speak to the particular skills or practice in question. As there are always exceptions to the general rule, so also there are exceptions to the general exception: liability may be established without expert testimony or independently of professional custom in cases where negligence is obvious to the lay observer or where medical judgment is not implicated.

The justification for treating malpractice cases differently from other accident cases is by no means obvious or simple. Posner has suggested that a distinction might rest on the notion that custom in malpractice cases was used essentially as evidence of the parties' contractual expectations. There is.

---

Copyright © 1986 by Law and Contemporary Problems

* John C. Stennis Professor of Law, University of Virginia. I wish to acknowledge extraordinary research and editorial assistance from Stephen Bainbridge and Anne Wilkinson. I would like also to thank Clark Havighurst, Ron Cass, Clay Gillette, Kenneth Abraham, and participants in workshops at Boston University, the University of Virginia, and Yale University for critiques of drafts of this article.

1. The classic modern statement of the rule is that of Judge Hand in The T.J. Hooper, 60 F.2d 737, 740 (2d Cir. 1932). The rule is much older, however. See, e.g., Hiltier v. McCartney, 31 Ala. 501 (1858); Mayhew v. Sullivan Mining Co., 76 Me. 100 (1884); Maynard v. Buck, 100 Mass. 40 (1868).


3. Id. at 621-31. Perceived liberalization of this exception provoked a spate of legislative enactments on the use of the doctrine of res ipsa loquitur and on the requirement for expert testimony, but it does not appear that the traditional common law rules have been much changed. See Robinson, The Medical Malpractice Crisis of the 1970's: A Retrospective, LAW & CONTEMP. PROBS., Spring 1986, at 5, 27-30.

support for this view in early medical cases. However, this explanation does not account for nonmedical cases involving contractual relationships where custom might be but is not accorded the same deference. In any case, custom-as-contract cannot adequately explain the modern judicial treatment which has crystallized professional norms into a fixed rule of public obligation unalterable by private agreement.

II
JUSTIFICATIONS FOR THE CUSTOMARY PRACTICE RULE

A. Technical Complexity and Fairness

Morris, whose treatment of the role of custom is deservedly considered a classic, thought the distinction between medical malpractice and other cases was compelled by the fact that judges or juries were not competent to determine whether a doctor had acted reasonably: "The conformity test is probably the only workable test available." According to Morris, lay people are not sufficiently knowledgeable about the complexities of medicine to be able to distinguish reasonable and unreasonable levels of skill or care, nor are they competent to evaluate relevant nonmedical circumstances.

The rationale of lay incompetence accords with modern judicial treatment, as is underscored by the cases in which customary norms are not dispositive because lay observers can detect medical negligence without expert guidance. This rationale, however, proves too much. Courts and juries are asked daily to judge disputed matters beyond their personal ken. Invariably, in complex, nonmedical cases they are informed by expert testimony and other evidence, but the ultimate evaluation of that evidence is not left to the professionals. Of course, in any such dispute, the plaintiff must offer a prima facie case, which means, as a practical matter, that on complex issues of fact and judgment there must be at least enough expert evidence to make a credible complaint. In the case of "ordinary" complexity, however, experts are free to roam among possible alternatives—and the judge or jury is free to accept their assessments—without regard to customary norms. In malpractice cases the expert evidence must be directed at a single question: Did the defendant conform to customary norms?

In some of these cases it might be better to take such complex matters out of the hands of courts, or at least out of the hands of juries; this has been partially done in some jurisdictions to the extent of permitting claims to be submitted to binding arbitration or requiring their review by screening panels before trial. But forum choice is another issue. The question here is whether

5. See, e.g., Bowman v. Woods, 1 Greene 441 (Iowa 1848); Leighton v. Sargent, 27 N.H. 460 (1853); McCandless v. McWha, 22 Pa. 261 (1853); see also F. Hilliard, Law of Torts 252-54 (2d ed. Boston 1861).
6. See, e.g., The T.J. Hooper, 60 F.2d 737 (2d Cir. 1932). R. Posner, supra note 4, at 125-26, concedes the point.
7. Morris, Custom and Negligence, 42 Colum. L. Rev. 1147, 1164 (1942).
there is any basis for separating one distinct set of complex issues from another for purposes of defining the ambit of judicial discretion.

It might be more fruitful to seek an explanation not in the complexity of the activity as such, but in the special difficulty of identifying a compensable event in a context in which many plaintiffs are predisposed to harm by the conditions that cause them to seek treatment. Moreover, quite apart from any predisposing condition, it is the nature of medical treatment that it often entails a risk of harm that could not be avoided with due care. A properly applied rule of negligence should cabin liability within acceptable bounds. Morris believes, however, that the fear of judicial erosion of the ordinary negligence standard requires special protection against "undeserved liability." 9

If "undeserved liability" is a problem, it is surely not unique to medical practice. The threat of ill-considered liability, often coupled with excessive recoveries, is a pervasive one in other areas of tort law, such as products liability, where compliance with custom is not an automatic defense. It may be that the demarcation of causal responsibility is a vexing problem for malpractice cases, but it, too, is not unique to the malpractice area. A similar problem is encountered in letting courts and juries set product design standards. 10

Morris's concern over "undeserved liability" is expressed essentially in fairness terms, although he nowhere attempts to explain why it would be "unfair" to hold physicians to the standard of care required of other tort defendants. To say that liability is "undeserved" is entirely question-begging until we have established a normative benchmark for measuring moral desert in this context.

B. Incentives: Promoting Deterrence Without Encouraging Defensive Medicine

Whether or not holding physicians to a more rigorous standard would be unfair, it might be inefficient if it had inappropriate incentive effects. Consider, for example, the rationale offered by McCoid for the custom-immunity rule:

[The medical practitioner's] judgment should be free to operate in the best interests of the patient. If the "judge" is himself to be judged by some outsider who relies on after-acquired knowledge of unsatisfactory results or unfortunate consequences in reaching a decision as to liability, the medical judgment may be hampered and the doctor may become hesitant to rely upon his developed instinct in diagnosis and treatment. If, on the other hand, the doctor knows that his conduct is to be evaluated in terms of what other highly trained medical practitioners would have done or would accept as competent medical practice, he is more likely to pursue his own judgment

when he is confident of the diagnosis and line of treatment, and is more likely to provide good medical service for his patient.\textsuperscript{11}

McCoid's rationale seems distinctly odd in supposing that the effect of the rule is to foster independent medical judgment. A rule compelling conformity to custom would scarcely seem calculated to induce the doctor to "pursue his own judgment." On the contrary, the rule creates precisely the opposite incentive, for it is only the individual's judgment, as distinct from that of the medical community, that exposes him to the risk of liability.

McCoid's defense of custom also expresses the more general concern that second-guessing medical practitioners might force doctors to practice "defensive medicine." This is an ageless complaint against malpractice liability generally\textsuperscript{12} and requires some examination.

It is noteworthy that both Morris and McCoid assume liability has a deterrent effect on physicians' behavior. Some observers, however, doubt that liability rules have an important effect on the behavior of prospective tort defendants. In particular it is argued that liability insurance usually insulates defendants from the direct costs of liability and thus destroys the economic incentives that are supposed to constrain risk-creating activity.\textsuperscript{13} Of course, insurance is not costless, and some measure of deterrence is presumably achieved by setting insurance premiums that reflect the risk-creating propensities of the insured. As is commonly observed, however, the mechanism for adjusting insurance costs to each insured's behavior is crude; much of the deterrent effect is dissipated through "community rating" that assigns accident costs to broad classes of individuals or activities, thereby externalizing the costs.\textsuperscript{14} Plainly, this observation conflicts with the widespread perception of "defensive medicine": if insurance practices effectively screen out the impact of liability, what would be the incentive to avoid it by practicing "defensively"? A physician's practice of defensive medicine would at best confer a small benefit upon the insured's risk class.

Although community rating insurance does tend to insulate insureds from direct financial costs, they may bear other costs that affect their behavior. Some carriers assess special surcharges on physicians against whom claims have been filed in prior policy years.\textsuperscript{15} Moreover, companies may refuse to renew policies for individuals deemed to be poor risks. Finally, legal liability imposes some costs that are uninsurable, such as harm to reputation, disruption of the physician's practice, and emotional stress caused by

\textsuperscript{11} McCoid, \textit{supra} note 2, at 608.

\textsuperscript{12} See Bernzweig, \textit{Defensive Medicine}, in U.S. DEP'T OF HEALTH, EDUC. AND WELFARE, REPORT OF THE SECRETARY'S COMMISSION ON MEDICAL MALPRACTICE & APPENDIX 38, 39 (1973) (surveys of physicians indicate that "between 50 and 70 percent of all physicians claim they practice defensive medicine of one sort or another").


\textsuperscript{14} Id. at 1299-1300. On class rating of health care providers in particular, see Brook, Brutoco & Williams, \textit{The Relationship Between Medical Malpractice and Quality of Care}, 1975 Duke L.J. 1197, 1206.

litigation. These are real costs which a physician presumably will seek to avoid even if they are not accompanied by direct economic penalties. Since these costs are a function of liability, which is presumed to be a function of the physician's risk-creating behavior, there are several clear mechanisms of deterrence independent of individualization of premiums. It remains to define the boundary between appropriate deterrence and inappropriate "defensive medicine."

That legal liability induces costly and unproductive avoidance strategies—"defensive medicine"—by physicians has long been an article of faith of the health care profession, and one of that profession's foremost indictments of malpractice law. As malpractice claims have increased in recent years, so too have the complaints about the costs of defensive medicine. In 1985 the American Medical Association estimated that the annual cost of defensive medicine now exceeds $15 billion.\(^\text{16}\) Even in the context of an annual health care budget that exceeds $375 billion,\(^\text{17}\) a cost of $15 billion for unproductive medical practices would be reason for grave concern. It would be, that is, if there were good reason to believe that costs of such magnitude are being incurred.

In fact, evidence of defensive medicine is notoriously unreliable. The AMA's estimate was based on the most casual empirical technique—soliciting physicians' subjective opinions through an AMA survey. The basis for those opinions in this particular survey is not revealed; however, if earlier reports of defensive medicine by health care providers are representative of the rationale underlying those opinions, one must be deeply skeptical.

Part of the problem is methodological; the absence of systematic, rigorous survey techniques and unbiased observers has been frequently noted.\(^\text{18}\) A deeper problem is the absence of an adequate definition of what constitutes "defensive medicine." Customary usage applies this term to practices that are not warranted or appropriate according to accepted standards of medical practice. However, because the canons of accepted practice allow a considerable range for different individual judgments as to what is warranted or appropriate, the definition of defensive medicine becomes a matter of individual motive that is difficult to measure. More important, the measurement parameters are skewed insofar as they are based on what the reporting physician deems warranted or appropriate, and not on whether the practices are what the legal system contemplates as appropriate. Given that the social control of medical practice is a major purpose of legal intervention,

---

\(^{16}\) Special Task Force on Prof. Liab. and Ins., Am. Medical Ass'n, Professional Liability in the '80s, Report 1, at 6 (1984). The estimate is based on surveyed physicians' reports as to the percentage of diagnostic testing and treatment procedures conducted "as a response to the increased risk of professional liability." Cost estimates were derived from these data and from estimated annual billings.


\(^{18}\) For critiques of the evidence, see Institute of Medicine, National Academy of Sciences, Beyond Malpractice: Compensation for Medical Injuries 16-17 (1978); Brook, Brutoco & Williams, supra note 14, at 1218-20.
“defensive medicine” cannot be measured by reference to medical opinion alone, but must reflect the broader social standards that justify liability. Indeed, recent evidence suggests that what some providers report as “defensive medicine” includes practices that are quite consistent with reducing iatrogenic risk, whether or not that is their intended purpose.19

The foregoing is not to suggest that health care providers do not seek to reduce the risk of liability through practices that are medically or economically unjustified, and therefore socially dysfunctional. A far more discriminating analysis of the nature of these avoidance strategies is required, however, in order to measure confidently the magnitude of the problem. In any event, assuming that health care providers’ defensive strategies are of such character and of such magnitude as to prompt concern, it by no means follows that the appropriate response is to restrict liability exposure as the AMA and others have argued. The risk of liability is a necessary but not sufficient condition of defensive medicine. The insufficiency of liability becomes clear upon examination of why the problem of dysfunctional liability avoidance (expenditures in excess of safety benefits) is thought to be special to health care rather than a generic problem of all tort liability.

Some commentators have observed that all liability rules are capable of inducing either overinvestment or underinvestment in care under certain conditions, most notably when there is uncertainty as to application of the rule.20 Outside the health care field, however, the possibility of excessive expenditures to avoid liability has not generally been perceived as a large practical problem. The reason for the difference between health care and other activities in this respect is quite simple. In the latter, the market constrains excessive expenditures. For example, product manufacturers contemplating excessive investments in care or other dysfunctional strategies to avoid liability are restrained by their limited ability to pass on their costs in the form of higher prices. To be sure, in markets characterized by inadequate information and inadequate competition, the manufacturer will have some freedom to pass on such inefficiently incurred costs. But few commercial markets permit the freedom to pass on inefficiently incurred costs to an extent permitted by the health care market. Third-party payments for health care costs significantly lessen market constraints on health care expenditures.21 To the extent that the health care market is unable to monitor and control health care expenditures generally, it is unable to distinguish expenditures incurred for purposes of avoiding liability or promoting health.

19. See Zuckerman, Medical Malpractice: Claims, Legal Costs, and the Practice of Defensive Medicine, HEALTH AFF., Fall 1984, at 128, 151, for reports of surveys of physician responses to increased premium costs. The responses include such practices as increased referrals, additional treatments, and refusal to take certain cases.


21. See infra note 53.
Recent cost containment measures, such as the diagnosis-related group reimbursement system (DRG's), or alternative health care delivery systems designed to introduce cost efficiency, such as health maintenance organizations (HMO's) or preferred provider organizations (PPO's), might reduce inefficient expenditures. Though these systems are not specifically directed against defensive medical expenditures, to the extent that they succeed in curbing inefficient expenditures generally, they presumably will curb defensive medicine as well. However, it remains to be seen how well the general objective of cost efficiency will be achieved.

It also remains to be seen how the malpractice liability rules will accommodate this objective in evaluating cost-constrained health care decisions. If cost containment measures can curb excessive expenditures incurred to avoid liability, they also might impede legitimate expenditures that the liability rule demands insofar as the rule does not take the new cost constraints into account.

If the common law reliably followed Judge Learned Hand's celebrated economic calculus, there would be little problem accommodating these new economic constraints (assuming the constraints themselves reflect a reasonable benefit-cost analysis). Unfortunately, Judge Hand's benefit-cost test does not appear to be an important element in malpractice jurisprudence.

The word "appear" is used advisedly; here, as elsewhere in

22. The DRG system, adopted by Medicare and some state Medicaid programs, fixes third-party payments for hospital costs on the basis of predetermined prices for different diagnostically related categories of treatment. Cost reimbursement is fixed according to the DRG classification given each patient on discharge, based on a grouping of factors: the patient's principal diagnosis or surgical procedure, the presence of an important secondary diagnosis, age, and whether the patient was alive at discharge. Payments for each DRG are based on average costs incurred for treatment of patients falling into that general category. See, e.g., Stern & Epstein, Institutional Responses to Prospective Payment Based on Diagnosis-Related Groups, 312 New Eng. J. Med. 621 (1984).


24. It is also possible, however, that some cost containment mechanisms will themselves produce dysfunctional responses no less inefficient than defensive medicine. For example, it has been widely noted that DRG's may skew incentives to select certain diagnoses and treatments in order to seek reimbursement under high-payment, high-profit DRG's. See, e.g., Omenn & Conrad, Implications of DRGs for Clinicians, 311 New Eng. J. Med. 1314 (1984); Stern & Epstein, supra note 22.

25. The problem of placing providers in a "double bind" by requiring them to respond to the conflicting demands of cost containment and traditional professional care standards is noted in Havighurst, Decentralizing Decision Making: Private Contract versus Professional Norms, in Market Reforms in Health Care 22 (J. Meyer ed. 1983). Havighurst argues for contractual adjustment of the liability rule in order to align these competing demands, an argument expanded on below (though not specially as a means of cost control).


27. Helling v. Carey, 83 Wash. 2d 514, 519 P.2d 981 (1974), is sometimes cited as an illustrative application of the economic approach to medical malpractice. See, e.g., Schwartz & Komesar, Doctors, Damages and Deterrence: An Economic View of Medical Malpractice, 298 New Eng. J. Med. 1282, 1283 (1978). What is most noteworthy about Helling, however, is that it did not apply the accepted professional custom standard of malpractice, and is thus out of step with prevailing legal doctrine. Of course, it is possible that professional norms may align with economic efficiency in a particular case. Interestingly, in Helling it appears that contrary to the defendant's assertion, accepted by plaintiff and the court, there was a routine practice of giving the tests which the court thought
the realm of negligence, the "reasonableness" standard is susceptible to varying interpretations about its economic content.\textsuperscript{28} To the extent medical negligence embraces the traditional norms of the medical profession it is almost forced to reject Judge Hand's test. Relatively undisciplined by market constraints on expenditures, medicine has evolved a culture that traditionally has been hostile to economic efficiency, particularly as a desideratum of "sound" medical practice.\textsuperscript{29} The emergence of for-profit medicine and concern about the level of health care costs are now changing health care's traditional insulation from economic criteria.\textsuperscript{30} The law has recognized this transformation to the extent of legislation and regulations facilitating cost containment measures and the economically oriented alternative delivery systems just mentioned. However, the common law of malpractice has given no obvious recognition to this transformation in medicine, although the absence of visible signs may simply reflect the common law's affinity for ambiguity in both rule and principle. In any event, the future test of realism for malpractice law will be how well it accommodates economic changes in health care.

Precisely what these economic changes imply in terms of liability and related legal rules governing responsibility for medical accidents can be debated. At a minimum, they should imply the necessity of taking a hard look at the role of liability rules in encouraging efficient health care practices. Among other things, they also suggest rethinking accepted legal conventions about the allocation of medical accident risks.

C. Allocating Accident Risks Between Provider and Patient

If the foregoing analysis is sound, one must question legislative efforts since the mid-1970's to give medical professionals greater protection against liability for medical accidents. Granted, prior liberalization of liability standards was partly responsible for the marked increase in claims frequency and insurance costs. Nevertheless, a plausible rationale can be offered for such a movement in the direction of stricter liability for providers of services as dangerous as medical care.

To the extent the tort system measures the social costs of medical maloccurrences with reasonable accuracy, increased liability may achieve the


\textsuperscript{29} See Havighurst, \textit{supra} note 25.

economic objective of allocating costs to those who, in general, appear best able to minimize their occurrence or to distribute them efficiently. Even if the present tort system does not accurately measure the social costs of medical maloccurrences, it is doubtful that the system as a whole overestimates the social costs of medical malpractice. The occasional excessive damage award and meritless claim are probably more than offset by underenforcement of meritorious claims. If so, one might infer that liability rules ought not be constricted but instead should be liberalized.

Of course, the inference that providers of medical care should bear more, not fewer, risks does not follow necessarily from the efficiency goal of cost internalization, which is met by patient cost bearing, as well as by physician liability. If there is an economic preference for the latter, it must rest on the intuition that physicians are better able than patients either to reduce accident costs or to distribute them efficiently—that they are, in Calabresi’s argot, cheaper cost avoiders. The term “intuition” is used to denote the fact that it is not easily demonstrated that physicians are necessarily superior risk bearers for all medical risks or for any particular subclass of such risks. As far as the loss distribution (risk spreading) function is concerned, it is not clear that providers have better access to insurance markets than do patients; indeed, most patients are already well protected by so-called collateral sources against many of the risks of medical injury. It also appears that the third-party insurance to which providers have access is both less comprehensive and more costly (due to the complex liability rules that govern whether a particular loss will be indemnified) than the first-party insurance which covers patients. As for the accident reduction function, provider liability presupposes that accident costs “feed back” into medical practice decisions. It may be presumed that this happens to a degree, even though, as noted earlier, present insurance rating practices blunt the effect of this feedback and other nonfinancial incentives must be relied upon to inspire due care.

The fact that the present system of provider liability is not optimally efficient in distributing losses and reducing the creation of risk does not, of course, establish that any other allocation of risks would be better. In particular, it is widely believed that requiring patients ultimately to bear medical risks would be even less efficient because of their relative ignorance of medical risks and their more limited control over the risk-creating events. These conditions of patient incompetence are intensified to the extent that the underlying demand for health care is perceived to be a medical necessity, implying that foregoing medical treatment is not a realistic option.

31. Studies indicating a large disparity between the number of compensable injuries and the number of claims filed point to significant underenforcement. See Robinson, supra note 3, at 14.

The search for the best risk bearer or cost avoider is made more difficult because, however risks are allocated, the insurance system passes costs through to patients and the community at large. Indeed, because of this pass-through of costs, it is questionable whether it makes much difference whether physicians are liable or the loss is left to lie with patients; either way, the public at large will ultimately bear the costs of such risks. Possibly the initial placement of liability might make a difference in how (and how well) the risk is administered. The channels of risk distribution will differ depending on whether the costs are distributed initially through physicians' liability insurance or through patients' disability and health insurance. Also, the risk avoidance strategies might be different for physician and patient, depending on their ability to perceive the costs as they are distributed through the different channels. These are, primarily, empirical questions; there is no basis for judging, a priori, that one risk distribution channel is superior to the other in all cases.

Those readers familiar with the Coase theorem will recognize that uncertainty about efficient loss allocation is essentially irrelevant if the parties can negotiate without significant impediments. In many tort contexts, transaction costs preclude any simple bargain solution; for example, in the paradigm accident involving strangers (such as an auto accident), transaction costs pose formidable, even insurmountable, obstacles, requiring the law to prescribe liability rules. No such obstacles appear to stand in the way of patient/physician bargaining, however. The parties to a medical transaction are not only not strangers but they have already entered into a contractual arrangement. That contract, in addition to specifying what the parties will do to benefit each other, might also prescribe how risks are to be allocated. If it does, the question arises whether the tort system's rules for allocating risks are thereby displaced or whether the court will ignore the contract's terms and apply tort doctrine.

The contractual character of the physician/patient relationship has become obscured in modern times. In the nineteenth century, the source of the physician's duty of care was derived as much from principles of contract as from those of tort. In the twentieth century, contractually shaped duties have been largely supplanted by conceptions of obligations derived from a general public policy independent of the parties' contractual intent. The shift in orientation from contract to tort has been reinforced, if not caused, by professional, social, and regulatory efforts to distinguish the provision of health care services from ordinary commercial transactions and to create an aura of professionalism untainted by crass commercialism. This general

34. See sources cited supra note 5. The English case usually cited as the seminal formulation of the custom standard for malpractice, Slater v. Baker, 2 Wils. K.B. 359, 95 Eng. Rep. 860 (1767), was brought as a contract action.
35. See K. Arrow, Essais in the Theory of Risk-Bearing 186-89 (1974). Arrow views this noncommercial aspect as inherent in the nature of medical services (particularly the problems of measuring quality, consumer dependency, etc.) and consumer expectations. This view seems
view of the special character of health care was carried over into traditional common law attitudes toward the physician/patient relationship. As noted earlier,³⁶ the traditional antipathy of health care professionals toward commercial practices and organization is now rapidly changing with the emergence of new, economically oriented delivery systems and cost control measures, but recognition of this fact does not yet appear in common law doctrine or attitudes toward health care services.

Nevertheless, whatever the legal view of health care in general, the provider/patient relationship is grounded in contract and most of the economic terms are set by the usual elements of contract formation. What needs to be considered, then, is whether there is any reason to treat risk of injury differently from all other terms of the bargain. This article argues that there is not; I am obviously not the first to so argue. Epstein and Havighurst, among others, have both argued for contractual risk allocation.³⁷ The argument presented in this article is parallel to theirs, but examines the negative case in more detail. In addition, as Epstein’s writings have done, it defends even the most extreme form of risk allocation contract—the exculpatory clause, which shifts all risks of injury to the patient. It follows a fortiori from the arguments presented here that less sweeping reallocations of rights and responsibilities, of the kind suggested by Havighurst and elsewhere in this symposium, should also receive a respectful judicial hearing.

III
EXAMINING THE CASE AGAINST CONTRACT

The affirmative case for contract is simple and powerful. In terms of utilitarian efficiency, contractual arrangements allow parties to achieve the most efficient combination of efforts to manage risk in accordance with their respective comparative advantages and their respective risk preferences. The moral argument proceeds along similar lines but emphasizes the fact that contractual allocation promotes individual freedom of choice, constrained only by the need to accommodate the divergent interests of the contracting parties. To justify private ordering one need not suppose that it always yields “good” or “fair” results. It is enough that, in general, private parties are somewhat naive in ignoring the role of the profession itself in promoting the conditions on which such a conception rests. See P. Starr, supra note 30, at 198-232. The profession’s efforts to remove itself and its services from commercialism were related to its historical opposition to the intrusion of corporate enterprise into medical services. The overriding objective, as Starr explains, was to avoid “the kind of hierarchial controls that typically prevail in industrial capitalism.” Id. at 216. To this end, the AMA in the 1930’s declared it was “unprofessional” for a physician to allow his services to be used by profit-seeking investors (other than himself or his professional colleagues), explaining that "where physicians become employees and permit their services to be peddled as commodities, the medical services usually deteriorate, and the public which purchases such services is injured." Id. at 217, quoting Bureau of Medical Econ., AM. MEDICAL ASS’N, Economics and the Ethics of Medicine 49-50 (1935).

³⁶ See supra notes 22-24 and accompanying text.

likely to achieve results that are at least as good and fair for themselves as would be achieved by paternalistic intervention.\textsuperscript{38}

Despite the general acceptance of contractual risk allocation in the world of ordinary commercial transactions, courts have been extremely hostile to contractual allocation of tortiously created risk. In the field of medical risks, courts have generally rejected out-of-hand attempts by physicians and hospitals to shift the risk of negligence to patients.\textsuperscript{39} Tunkl v. Regents of the University of California,\textsuperscript{40} the leading case, is reasonably representative of the judicial attitude. In Tunkl, plaintiff, as a condition of his admission to the U.C.L.A. Medical Center, executed a release form relieving the Center of "any and all liability for the negligent or wrongful acts or omissions of its employees, if the hospital has used due care in selecting its employees."\textsuperscript{41} The California Supreme Court held that the release was void as against public policy.\textsuperscript{42} The court sought to draw a line between exculpatory contracts that involved mere "private interests" and those affecting the "public interest."\textsuperscript{43}

The court considered several general criteria for defining public interest: (1) the party seeking immunity is engaged in performing a service thought suitable for public regulation or one of great public "importance" and "practical necessity"; (2) the party seeking immunity holds himself out as willing to perform this service to any member of the public (or any member meeting "certain established standards"); (3) as a result of the "essential nature of the service" and in the "economic setting of the transaction" there is a "decisive advantage of bargaining strength" favoring the party seeking immunity; (4) the exculpatory contract is a "standardized contract of adhesion" and makes no provision whereby the other party may pay "additional reasonable fees" and obtain protection; (5) the person or property of the purchaser (patient) is under the control of the party seeking immunity and thereby subject to the risk of that party's carelessness.\textsuperscript{44}

Of the criteria listed above, the first two seem either irrelevant or question-begging. Whether health care is a matter suitable for public regulation only begs the question of what kinds of public regulation are suitable—in particular whether judicial interference with contractual risk allocation is

\textsuperscript{38} Epstein, Medical Malpractice: Its Cause and Cure, in The Economics of Medical Malpractice 245, 257 (S. Rottenberg ed. 1978).
\textsuperscript{39} See, e.g., Tunkl v. Regents of Univ. of Cal., 60 Cal. 2d 92, 383 P.2d 441, 32 Cal. Rptr. 33 (1963); Emory Univ. v. Porubiansky, 248 Ga. 391, 282 S.E.2d 903 (1981); Meiman v. Rehabilitation Center, Inc., 444 S.W.2d 78 (Ky. 1969); Olson v. Molzen, 558 S.W.2d 429 (Tenn. 1977).
\textsuperscript{40} 60 Cal. 2d 92, 383 P.2d 441, 32 Cal. Rptr 33 (1963). The court's opinion cites numerous earlier decisions invalidating exculpatory contracts where the "public interest" was affected. The "public interest" in those cases appears generally, though not invariably, to have involved regulated businesses ("affected with a public interest"). Rather curiously, the court does not cite the then recent decision in Henningsen v. Bloomfield Motors, Inc., 32 N.J. 358, 161 A.2d 69 (1960), which has become the leading precedent for invalidating liability waivers in the closely related area of product liability cases.
\textsuperscript{41} Tunkl, 60 Cal. 2d at 94, 383 P.2d at 442, 32 Cal. Rptr. at 34.
\textsuperscript{42} Id. at 94, 383 P.2d at 441-42, 32 Cal. Rptr. at 33-34.
\textsuperscript{43} Id. at 96-98, 383 P.2d at 443-46, 32 Cal. Rptr. at 35-39.
\textsuperscript{44} The court's opinion does not catalogue its reasons in exactly the same order, or with the same precision, as the list above suggests.
appropriate. The criterion of holding out to the public is simply the mindless adaptation of an ancient rule for common carriers. It begs the question of what is the proper scope of the holding out; in Tunkl the extent of the holding out was presumably limited by the contractual conditions of admission, which included a waiver. The third and fourth criteria fit conventional doctrine supporting judicial scrutiny of contractual bargains generally, and they warrant some attention. The fifth criterion is a bit nebulous, but it may be interpreted—rather freely, perhaps—to denote a concern over the loss of deterrence and the creation of moral hazard.

A. Adhesion Contracts, Choice, and Bargaining Power

Disparity of bargaining strength and use of standardized contracts of adhesion are time-honored rationales for judicial supervision of private contracts. They form the core of the unconscionability doctrine in consumer transactions. Judicial surveillance of contracts to protect individuals against the superior bargaining power of others, against contracts of adhesion, or against unconscionable bargains is a subject more than adequately covered in existing legal literature. It is enough here to show that unconscionability is not a generic problem of medical risk “bargains.” Given the assumption that individual medical service contracts are not less subject to judicial monitoring than ordinary commercial transactions, there is still no warrant for a blanket prohibition against contractual arrangements dealing with medical risks.

Standardized contracts are a characteristic of virtually all routine commercial transactions, most of which do not remotely involve the kind of bargaining problem that would warrant, or attract, judicial concern. A purchase of oranges at a grocery involves a standardized contract. Standardization in such cases is merely an efficient means of minimizing the contract costs of multiple transactions and, in itself, is not evidence of unfairness. Parties to such arrangements benefit from cost-efficient transactions, however, and there is no basis for presuming that the benefits

45. On the holding-out question, the court relies on Wilmington Gen. Hosp. v. Manlove, 54 Del. 15, 174 A.2d 135 (1961), in which a private hospital was held to a duty to serve because it maintained an emergency ward. However, the traditional rule has been that health care providers are not necessarily engaged in a public calling merely by virtue of their providing medical services. While there has been some erosion of the traditional rule in the case of health care institutions, the scope of an individual physician’s “calling” generally remains a matter of private contract. E.g., Lyons v. Grether, 218 Va. 630, 239 S.E.2d 103 (1977).


48. On the costs of individualizing, see Schwartz supra note 46, at 1064-71. Standardization of terms not only avoids the costs of individualizing any given contract, but may also reduce market search costs by facilitating comparison of terms offered by different competitors. See Schwartz & Wilde, Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis, 127 U. PA. L. REV. 630, 672 (1979).
are distributed between them so inequitably as to invite judicial interference. The problem with standardized contracts arises when they are associated with one party's market power and with a concomitant limitation on options available to the other party, or when they are associated with information asymmetries, which will be discussed below. The Tunkl court's concern over the patient's lack of choice is not carefully articulated, but it appears to mix two quite distinct factors influencing choice. One is the "essential nature of the service," which is thought to compel patients to seek medical attention. The other is the hospital's "decisive" economic advantage that enables it to dictate terms.

The first factor may be important in many cases. When a patient seeks emergency, or perhaps even nonemergency care under circumstances not conducive to reasonable deliberation about the risks of medical treatment—as may have been the case in Tunkl—courts are justified in disallowing contractual arrangements "imposed" by the health care provider. Nevertheless, protecting vulnerable patients in such cases does not warrant a broad, paternalistic rule in cases involving elective therapy that is, or should be made to be, conducive to the patient's reasoned evaluation. The Tunkl court's implied generalization that medical services are intrinsically necessary ignores the obvious fact that a large percentage of medical procedures are elective—not essential to maintain life or the patient's ability to function.

Indeed, to describe some of the medical treatments as elective understates the point. There is substantial evidence that many of the most common treatments have little or no therapeutic value. Significantly, some of these

49. Rakoff, supra note 47, at 1248-55, argues to the contrary. Regardless of bargaining power, he would have courts declare all form contract terms which are "not visible" parts of the bargain or which would not be expected by a "customary shopper" to be presumptively unenforceable (subject to affirmative defenses that are somewhat vaguely formulated). In general, Rakoff's elaborate argument for judicial intervention is not persuasive. To the extent his case for judicial control turns on simple inadvertence to contract terms, however, that concern could be accommodated by requiring that liability disclaimers be brought to patients' attention.

50. There has been a growing public controversy over the extensive rise of medical treatments having little or no proven therapeutic value. Much of the controversy has been generated by the costs of such procedures and the burden they place on health care insurance costs. See, e.g., Staff of Subcomm. on Oversight and Investigations, House Comm. on Interstate and Foreign Commerce, 94th Cong., 2d Sess., Report on Cost and Quality of Health Care: Unnecessary Surgery 31 (Subcomm. Print 1976) (unnecessary surgical procedures in 1974 cost some $3.92 billion) [hereinafter cited as Unnecessary Surgery]; Califano, U.S. Must Discipline Health-Care Market, N.Y. Times, May 6, 1984, at E23, col. 3 (citing Chrysler Corp. Study indicating that 25% of hospital costs for Chrysler employees "may be due to waste and inefficiency"). The increase in unjustified risk of iatrogenic injury, however, is significant in its own right. It has, for example, been estimated that unnecessary surgery produced nearly 12,000 deaths in 1975, a figure derived by applying mortality rates for all forms of discretionary surgery to surgery deemed to be "unnecessary"—that is, lacking in therapeutic value. Unnecessary Surgery, supra, at 34. Surgical procedures that have come under challenge by medical experts include such common procedures as tonsilectomies, hysterectomies, and cholecystectomies (gall bladder removal). See id.; see also Bolande, Ritualistic Surgery—Circumcision and Tonsillectomy, 280 New Eng. J. Med. 591 (1969). The increasingly frequent resort to coronary-artery bypass surgery as treatment for chronic angina has been recently questioned. See Braunwald, Effects of Coronary-Artery Bypass Grafting on Survival, 309 New Eng. J. Med. 1181 (1983) (commenting on a 1977 V.A. study and expressing concern over the frequency with which the procedure is used—about 160,000 procedures in 1981—given the high cost—about $3.25
procedures are attributed to patient insistence more than to physicians' advice.\textsuperscript{51}

Even if particular care or treatment is essential for the patient, it need not imply compulsion to accept the terms and conditions offered by any single provider. Thus, the choice issue is reduced to a question of the availability of competitive alternatives. It is plausible to suppose that the choice of consumer options is limited; it does not follow that in every case the choice is so circumscribed as to confer monopoly power on the health care provider. In this regard, one must challenge the \textit{Tunkl} court's heavy emphasis on the hospital's exercise of superior bargaining power by its failure to provide an alternative arrangement whereby persons could buy protection against negligently caused injury for "additional reasonable fees."\textsuperscript{52} The pertinent question is not whether U.C.L.A. provided such an option, but whether such an option was available at other hospitals in the market.

Concededly, the availability of meaningful choices in the market turns critically on whether patients would search for alternatives if they were made available. Some observers argue that most patients do not actively search for competitive alternatives in medical services generally; others more modestly contend that they would not search for alternative risk arrangements.\textsuperscript{53} Having no empirical evidence on either point, this author's intuition is that searching for competitive arrangements is fairly limited, in part because for most patients there is relatively little gain from doing so.\textsuperscript{54} It may also be assumed, however, that there is sufficient market searching by marginal patients to justify some reliance on market arrangements in this area just as in

\textsuperscript{51} See Bolande, supra note 50, at 594.

\textsuperscript{52} \textit{Tunkl}, 60 Cal. 2d at 100, 389 P.2d at 446, 32 Cal. Rptr. at 38. Apart from the availability of the option to purchase physician responsibility, the court expressed concern that it would not be certain that the patient received "adequate consideration for the [risk] transfer." \textit{Id.} at 101, 383 P.2d at 446, 32 Cal. Rptr. at 38. In \textit{Emory Univ. v. Porubiansky}, 248 Ga. 391, 282 S.E.2d 905 (1981), however, a waiver contract between a patient and a university clinic was invalidated even though the court noted that the defendant's fees were generally lower than fees for like services performed by private dentists.

\textsuperscript{53} One of the bases for this skepticism is that because of third-party payments, most patients do not have a financial incentive to search out health care alternatives on the ground of cost. This is an aspect of the general moral hazard problem created by insurance. \textit{See, e.g.,} K. Arrow, supra note 35, at 142. If third-party payments eliminate any financial incentive to search out alternative arrangements in general, patients presumably would, according to this argument, have no financial incentive to accept any arrangements that did not offer provider liability (insurance) for negligent care. In other words, given a choice, patients will always opt for the more expensive full-coverage arrangement rather than accept any risk themselves. The argument proceeds from a plausible premise—the existence of some moral hazard—but carries it too far. For one thing, not all health care services are fully covered by third-party payments. In fact, patients directly pay slightly more than 30% of the nation's aggregate health care expenses. \textit{See Bureau of the Census, U.S. Dep't of Commerce, Statistical Abstract of the United States} 104 (104th ed. 1984). Thus, patients do retain some incentive to constrain health care consumption. Moreover, if one assumes that patients would have the incentive to demand full responsibility for negligence risk at any cost, why is it that they do not now demand full responsibility—and all-risk insurance—for all iatrogenic risk?

\textsuperscript{54} It must be supposed that this would not change appreciably under a regime of risk contracting. The vast number of ailments for which individuals seek out professional health care do not involve the kinds of medical hazards that would generate concern by either providers or patients.
any other area. To assume the contrary would imply that there should be not only no choice of risk arrangements but no choice of other terms (such as price) governing health care services.

The latter point bears emphasis, for it makes clear that whether a reasonable health care contract must include physician responsibility for negligent error is only part of a more general issue of the degree to which courts will supervise the terms of the contract. If there is reason to think the bargain is not fair and reasonable in regard to risk allocation, is there not also reason to think that other terms may be similarly unfair or unreasonable? If the economic bargaining power is troubling, it is odd that concern is focused exclusively on one of its less significant manifestations. Why do courts not examine the price of the services being supplied? If the answer is that legislatures decide whether to institute such regulation, why are legislatures not also the proper judges of whether marketplace bargains over the allocation of risks should be generally enforceable?

B. The Information Problem

Probably the most common rationale for judicial intervention in contractual dealings is the absence of information about relevant features of the contract, most especially the risks being assumed. The "information problem" is not the absence of information relevant to a contract between two equally ignorant persons, even one that full information would disclose to be a poor bargain for both. Instead, the problem arises when one party has relevant information the other lacks and when withholding enforcement in favor of the disadvantaged party would create an incentive for fair and efficient information sharing. Such asymmetry of information is widely perceived to be a critical problem in the health care context. Tunkl did not articulate this concern except insofar as it is embodied in the court's more diffuse concern about bargaining power and contracts of adhesion. Other observers, however, have considered this to be virtually a dispositive consideration. It is not.

First, it is not a basis for concern that providers possess relevant information not possessed by patients. Asymmetry in the possession of information about the character and magnitude of medical risks can be solved by requiring providers to convey that information to patients as a condition of allowing the latter to assume responsibility for those risks. The requirement of full disclosure could be made as exacting as the nature of the risks demands, and a court could reasonably insist upon a more complete disclosure of information than informed consent rules now require.

The difficulty arises from the common perception that relevant information about risks cannot be meaningfully conveyed to patients. As is often supposed of ordinary consumers in general, patients are presumed to be

55. See, e.g., Reder, Medical Malpractice: An Economist's View, 1976 AM. B. FOUND. RESEARCH J. 511, 536-37 (observing that the assumption is widely accepted but "not beyond dispute").
not only ignorant but inherently incapable of appreciating the significance of risks. The assumed ignorance of accident victims in general, and malpractice victims in particular, is sometimes buttressed by references to studies by cognitive psychologists demonstrating consumers' ignorance about and inability to evaluate risks.

A number of studies in cognitive psychology show various biases (or "heuristics") in risk evaluation by individuals. One of the most common biases is what Tversky and Kahneman label the "availability heuristic." This heuristic causes people to weigh disproportionately information with high salience, memorability, or currency. Suppose, for example, a physician tells a patient that the statistical risk of injury to the eye from cataract removal is one in one thousand but the patient's sister tells her that a friend recently suffered blindness from such an operation. The "availability heuristic" causes the patient to give undue weight to the latter information relative to the former.

Another bias relevant to health care is that of "representativeness," which is a common misperception of the relationship between sample events and the population from which they are derived. Representativeness bias is illustrated by the "gambler's fallacy" of believing that specific outcomes of a series (for example, a run of "heads" in tossing a coin) must be followed by outcomes of another sort ("tails")—an implied belief in a law of averages which views probabilities as necessarily inherent in (or represented by) small samples. Suppose the eye surgery patient is told the general injury statistics for cataract surgery but is also told that in the hospital where he is scheduled for surgery there has been no injury in over 600 cases. One plausible interpretation of this statistic is that the particular hospital is safer (or that its surgeons are more skilled) than average and that therefore the odds of suffering injury are less than one in one thousand. However, the representativeness bias may lead the patient to the opposite and erroneous inference that, after 600 cases without incident, the probability of injury during his operation is actually greater than one in one thousand.


57. For a selection of some of the principal studies on risk cognition, including most of the studies cited below, see D. Kahneman, P. Slovic & A. Tversky, Judgment Under Uncertainty: Heuristics and Bias (1982).


60. Id. at 435. Other examples of this heuristic include such errors as assigning equal probabilities to sample statistics based on their similarity (their equal resemblance to the population), disregarding large differences in sample size. For example, the probability of finding more than 600 boys in a sample of 1000 babies is incorrectly assigned the same likelihood as the probability of finding more than 60 boys in a sample of 100, whereas, in fact, the latter is much more probable. Id. at 437.
These perception problems raise serious questions about the rationality of patient decisions concerning risk. Zeckhauser and Nichols infer a rule against contractual risk allocation from the studies of risk perception:

"Thus a physician attempting to discuss alternative courses of action with a patient will have difficulty estimating the probabilities of various adverse outcomes, and the patient will have difficulty interpreting even the most accurate estimates. Given this limited information and individuals' difficulties in assessing it, the conditions required to ensure the efficiency of free contracting will not be met."\(^6\)

Despite the problems Zeckhauser and Nichols observed, the cognitive psychology studies do not support the conclusion that providers must bear the risks and that patients must be barred by law from assuming them. First, the fact that people systematically err in evaluating probabilities does not mean the probabilities of injury are underestimated. Although some commentators casually interpret these studies as showing insufficient appreciation of risk,\(^6\) there is no basis for such a conclusion. In the hypotheticals above, for example, the probability of adverse outcomes was exaggerated, not underestimated. The exaggeration would still interfere with "efficiency of free contracting" as Zeckhauser and Nichols state, but it would not produce an unwarranted assumption of risk by the patient, as they fear. Indeed, the availability heuristic is more likely to lead the patient to avoid risk than to assume it, because bad outcomes will be more vividly reported and remembered than favorable outcomes ("bad news travels fast," that is, has greater salience or availability). The direction of error produced by the representativeness bias is more ambiguous; however, the example suggests one plausible scenario in which the error is in the direction of exaggerating risk.

Second, even when biases in risk perception are apt to be particularly disadvantageous to patients, this finding cannot be translated into a rule irreversibly imposing liability on health care providers. Such a rule presupposes that the provider not only possesses all relevant risk information but can better interpret it. As a matter of mere statistical interpretation, however, the cited studies do not support this assumption. Indeed, the same biases are observed in experiments involving scientists with training in probability theory and statistics.\(^6\) In this connection it ought to be emphatically noted that physicians' ability to evaluate probabilistic information has been specifically tested by cognitive psychologists and found wanting.\(^6\)

---

61. Zeckhauser & Nichols, supra note 56, at 22.
62. See sources cited supra note 56.
64. See, e.g., McNeil, Pauker, Sox & Tversky, On the Elicitation of Preferences for Alternative Therapies, 306 New Eng. J. Med. 1259 (1982) (physicians unable to see the equivalence of probabilistic data presented in two different ways; preference for different medical strategies varied according to whether probable outcome described in terms of mortality or survival, even though each description contained the same mortality/survival ratio); Saks & Kidd, Human Information Processing and Adjudication: Trial by Heuristics, 15 Law & Soc'y Rev. 123, 139 (1980) (discussing studies of "availability" bias in which physicians make same errors as lay persons).
Bias, then, cuts across professional/lay boundaries. So, too, do other impediments to understanding risk, such as "cognitive dissonance." Cognitive dissonance is "a state of tension that occurs whenever an individual simultaneously holds two cognitions (ideas, attitudes, beliefs, opinions) that are psychologically inconsistent."\(^{65}\) Characteristically, dissonance arises after an individual makes an important choice and is then confronted with information at odds with the understanding that inspired that decision. A variety of behavioral responses have been observed by which individuals subconsciously seek to avoid facing the reality of error or conflict.\(^{66}\) One common response is the rejection or distortion of post-decision information which would undermine the beliefs, attitudes, or understanding that led to the prior decision.

Akerlof and Dickens have noted the implications of cognitive dissonance for health and safety risks.\(^{67}\) For example, a worker who has voluntarily accepted employment in a hazardous workplace may refuse to recognize the risk insofar as it implies that he was foolish to accept the job. In the immediate context of health care one can imagine a similar problem. Suppose, after years of smoking, a person exhibits cancer symptoms. Seeking diagnosis and therapy may be psychologically difficult insofar as it implies prior neglect of health. The individual experiencing dissonance will be disposed to minimize the risk of neglecting to seek treatment in order to protect his self-image. Once the condition is diagnosed, the individual, again seeking to spare himself from recognizing the full consequences of his previous mistakes, would be disposed to minimize the risk of therapy.

The problem of a patient minimizing therapy risk is compounded if the provider has a similarly distorted risk perception. Suppose, for example, that the doctor believes a particular therapy is optimal for the patient and has used the therapy repeatedly on other patients. Cognitive dissonance would impair his acknowledgment of risk—for a "responsible" physician would not have subjected patients to therapy that was not wholly in their best interests.\(^{68}\)

The essential point of the argument thus far is that cognitive bias is a human characteristic, as likely to affect risk perception by physicians as by patients. Of course, since health care institutions are not human, we might suppose them to be unbiased. This supposition assumes either an institutional rationality independent of human judgments or decisionmaking processes capable of correcting human bias. The first assumption represents too simplistic a reification of organizational function, but the second

---


\(^{67}\) Akerlof & Dickens, The Economic Consequences of Cognitive Dissonance, 72 Am. Econ. Rev. 307 (1982).

\(^{68}\) Such dissonance may partly explain providers' unusually sharp reaction to legal second-guessing of their own "expert" judgments concerning risk. The irony is that this very dissonance supports the case for legal liability in order to correct biased risk perspectives. This assertion does not contradict the argument for contractual risk allocation, for similar corrective effects can be achieved through contract, as discussed below.
assumption is plausible. The process of routinized (bureaucratized) decisionmaking within institutions may screen out the biases or errors of individual human judgment. If depersonalized decisionmaking by individuals within an organization permits greater rationality about risk and can surmount biases distorting personal decisions, then perhaps the burden of risk should be imposed on institutional health care providers. Indeed, if the problem were only one of risk perception, a fixed and invariant legal rule to that effect might be acceptable. But another element of risk assessment must be considered: risk preference.

Risk evaluations involve both cognitions and attitudes toward risk. A provider may indeed be better able rationally to understand risks but may be incapable of appreciating what those risks mean to any given individual. Yet, the typical discussion of risk allocation ignores individual risk preference, or rather ignores its individual variation. The usual assumption is that people uniformly wish to avoid all risks, and therefore the search is only for the “cheapest” way of doing so. Cheap risk avoidance becomes a kind of technological question. Calabresi implies this when he observes (regarding the importance of rules facilitating private adjustments of behavior through what he calls “bribes”) that contractual risk allocations show “we do not know collectively who the cheapest cost avoider is, since if we did we would have allocated the costs to him in the beginning.”

Calabresi’s observation is incomplete because it fails to recognize the positive value in private ordering based on personal preferences without regard to social optimality. Indeed, if the private bargain captures all the significant social costs, it is not useful to postulate a social optimum apart from what the parties choose for themselves (assuming fair and equal bargaining, as previously discussed).

Some would say that the present legal bar to patient assumption of risk reflects the preference of the vast majority of individuals. On this account, the legal rule expresses the result that would generally be reached anyway by contracting parties. This argument embraces two distinct points. One is that patients, acting naturally, do not or will not seek alternative risk arrangements. The second is that most patients do not wish to know or even to think about the risk of medical maloccurrences, but rather rely implicitly on health care providers to make the appropriate judgments. Essential to this reliance is provider accountability for making reasonable judgments. This paternalism by delegation is supposedly satisfactory to providers as long as their legal accountability is not enforced too rigorously.

It is unlikely that the present arrangements are precisely what most patients and providers would make for themselves if guided by their own informed preferences rather than by the preferences of judges. There is no strong empirical evidence one way or the other, although some informed consent studies indicate a general preference of patients for more risk

69. G. Calabresi, supra note 32, at 151.
Too, consumers’ voluntary insurance purchases reflect choices somewhat different from what liability rules select for victims, such as compensation for pain and suffering (though moral hazard problems may also explain the absence of first-party coverage). But it is difficult to ascertain the parties’ true preferences under a legal regime that fails to give those preferences legal significance.

The patients’ ignorance argument against letting patients voluntarily assume risks is thus poorly supported by the factual assumptions underlying it. Also, the argument is inconsistent with accepted rules governing the allocation of the risk of nonnegligent injuries. If patients are really incapable of dealing with medical risks, they should not be allowed to accept the risk of nonnegligent treatment. Present liability rules, however, assign to patients all risks that they cannot demonstrate were produced by negligence. Only a strict liability system, in which the health care provider bears all of the risks, would be truly congruent with the assumption of patient ignorance. Despite common acceptance of the consumer ignorance assumption, there is not a broad accord among legal scholars, courts, or others of the assumption’s full implications.

Finally, the assumption of patient ignorance is at odds with the trend toward greater physician disclosure as a prerequisite of informed consent, which is premised on a new appreciation of the value of patient autonomy and responsibility in making choices about health care. If a patient’s choice of medical risks is not to be subordinated to his physician’s judgment, why should his choice of legal risks be subordinated to the judgment of his lawyer or a judge? Is it possible that the different weight accorded to patient autonomy in the two situations is only a reflection of the legal profession’s hubristic assumption that, while physicians adhering to medical custom cannot be trusted to decide what is in their patients’ interests, the legal system knows what is best?

C. Deterrence and Moral Hazard

The *Tunkl* court’s fifth criterion implies a fear that contracted-for immunity will impair the deterrent function of liability. The apprehended effect is analogous to that of the “moral hazard” associated with liability insurance. Some moral hazard is already present to the extent the health care provider is

---


71. The strict liability implications of patient risk misperception have not, however, gone unnoticed. See, e.g., P. Danzon, supra note 15 (which in turn draws on the analysis of products risks in Spence, Consumer Misperceptions, Product Failure and Product Liability, 44 REV. ECON. STUD. 561 (1977)).

72. “Moral hazard” is conventionally used to describe the effect of insurance in distorting insureds’ incentives in regard to the insured event, for example, the overconsumption of health care by persons with health care insurance. See, e.g., K. Arrow, supra note 35, at 142; Marshall, Moral Hazard, 66 AM. ECON. REV. 880 (1976); Pauly, The Economics of Moral Hazard: Comment, 58 AM. ECON. REV. 531 (1968). The term is used here in a broader sense of describing all changes in incentives in response to contractual indemnity or immunity. The essential problem remains unchanged in this broader context.
insured against liability and the costs of that insurance are not precisely matched to his actions. Contractual immunity would supposedly expand the scope of moral hazard by removing any possible penalty for malfeasance.

This concern about lost deterrence is not well founded. If the bargain is fair and reasonable, the problem is not different from any other contractual arrangements involving risk. The central problem of moral hazard is one not merely of incentives but of asymmetry of information about the future risk behavior of the party who does not bear the risk. When the moral hazard problem is obvious—as it seems to be here—one must suppose that it will affect the terms of the contractual arrangement and that patients will not lightly relieve providers of accountability for the harms they cause. Patients will insist on other means of quality assurance or will demand comparable compensation. The effect of either demand on the provider is to penalize performance below the level that patients expect. In short, tort incentives thereby become contract incentives. The transformation of incentives by contract might produce less efficient social allocations than those externally imposed by tort law. The parties are presumed to understand that risk and yet still regard the overall bargain as mutually advantageous; otherwise they would not enter into the arrangement.

If the two parties regard the contract as a sound bargain, the only question remaining is whether their private agreement fully reflects (internalizes) relevant social costs. The Tunkl court's attempted distinction between private and public interests suggests a negative answer, but apart from the concerns addressed above it is not apparent what justifies this response. It could be argued that allocating risks of medical injuries implicates social responsibilities that should not be determined by the immediately affected parties. Given this premise, however, it would be hard to account for the fact that society relies almost entirely on private parties to enforce those responsibilities. If private contractual relationships cannot fairly represent the social interests at stake, presumably private tort actions cannot fairly represent them either. It would follow that there should be a mechanism for a public representative to supervise every out-of-court claim settlement between plaintiffs and defendants in order to see that broad social ("public policy") concerns are met. It would also presumably be necessary to have public representatives intervene in litigation for the same purpose and to make sure that all meritorious claims are brought to the courts' attention.

The foregoing critique of arguments for invalidating contractual risk allocation may miss the real objection to permitting waiver of negligence liability. Reading Tunkl and other discussions one senses that the real source of hostility is an unarticulated moral objection to permitting persons to avoid responsibility for "wrongdoing." It is difficult to address this argument in the absence of a clear identification of the nature and source of the moral judgment. The moral argument for allowing contractual risk assumption—

respecting individual choice and autonomy—was mentioned earlier. Even if one does not attach great moral value to choices involving risk, the hard fact of life is that some such choices cannot be avoided. The question then is what warrants singling out these particular risk choices from the rest of the moral universe. Granted, practical reasons such as those just examined might so undermine meaningful choice as to support legal interference with private ordering and external, social costs might be created which would justify interference. If, as argued above, these circumstances are not present in the particular case, is there still some residual, overarching ethical or moral (fairness) objection to private ordering? Is it possible that there is some perceived "injustice" or "unfairness" in enforcing a contractual waiver of responsibility for tortious conduct?

The answer, it seems to me, is no. The objection from morality—if such it can be called—misconceives a contractual waiver as excusing or granting an indulgence for wrongdoing. The imputation of moral significance to most medical accidents stretches the limits of moral judgment rather far, but in any case the waiver is not the same thing as an excuse or an indulgence in the ordinary (moral) sense. The purpose of the contract is not to excuse some egregious and obvious error that has already occurred, but to avoid the risk that a contemplated act will at some future date be judged by a somewhat unreliable third party (a court or jury) to be an error under a somewhat artificial legal (hardly a moral) standard. Even if providers' attempts to obtain forgiveness for errors already committed were thought to be morally offensive, efforts to escape accountability under the imperfect legal system cannot be thought to be morally repugnant. The flaws of the tort system as a moral arbiter would seem clear enough that private parties should not be bound to submit to it.

D. No-Fault Risk Allocation

A few words are appropriate about no-fault insurance alternatives to liability insofar as they may be tied to contractual arrangements. Inasmuch as no-fault insurance might be one arrangement for which the parties could contract as a means of allocating medical risk, the case for contract can be melded with the case for substituting consensual no-fault compensation for conventional tort remedies, as O'Connell and others have proposed.74 On the premise that both parties are capable of making an informed decision about their own risk-bearing preferences, there is no justification for compelling one party to accept the other party's preferences. However, there is also no basis for objecting to a mutual agreement to implement the no-fault plan.

Of course, if the parties are free to make their own risk allocation arrangements, one might question whether there is any point to developing an elaborate statutory system of no-fault insurance. The answer is that there might be advantages in having a state-approved plan which the parties could adopt (or reject) by private agreement. One advantage is the opportunity to mold a plan that is officially scrutinized for balance and fairness, ameliorating judicial concerns about contracts of adhesion: if the no-fault allocation were accepted in the contract, legislative approval of the particular terms would presumably deem the contract "officially fair." Neither party would be forced to accept it, but the very existence of such an option could force the parties to confront the question of risk allocation and not hide behind the illusion of risklessness.

A second advantage of an officially approved "risk plan" is one common to all form contracts ("adhesive" or not): it economizes on transaction costs. The term "transaction costs" is used broadly here, to include any impediments to arriving at an informed agreement. One of the strongest obstacles to contractual risk allocation arrangements is simply the perception that such arrangements are out of character with the physician/patient relationship. As noted earlier, the image of medical professionalism has been promoted in part by attempting to remove health care from the commercial marketplace. One consequence has been to create an environment in which contractual bargaining is awkward for both physician and patient insofar as it appears to undermine the implicit element of trust and confidence that is fostered by the image of the physician as a professional and fiduciary. In this regard, formalism may be most important by reducing the friction which people might otherwise perceive to be inherent in contractual dealings. Particularly when the formalism embodies officially scrutinized terms, it seems likely to gain greater acceptance by parties who are accustomed to a high degree of paternalism in health care.

There is, of course, no reason why the officially sanctioned terms must take the form of no-fault insurance. Indeed, there is no reason why they must take any single form at all. One can imagine a variety of risk allocation options that might receive an official seal of approval. Legislation providing an array of preformed arrangements would reduce the economic costs and ease the social resistance to contractual dealings, while still providing a degree of flexibility that could not be obtained from a single form contract. If a legislature should accept this suggestion and specify the terms of some "officially fair" arrangements, it should also explicitly deny the implication that other arrangements, voluntarily arrived at, are presumptively unfair and unenforceable.

75. A similar suggestion for preformed contracts to minimize transactions costs is made by Reder, An Economic Analysis of Medical Malpractice, 5 J. LEGAL STUD. 267, 277 (1976). The advantages and disadvantages of preformed contract terms, in a commercial context, are examined at length in Goetz & Scott, The Limits of Expanded Choice: An Analysis of the Interactions Between Express and Implied Terms, 73 CALIF. L. REV. 261 (1985).
E. Effects of Contractual Risk Assumption

Widespread patient assumption of risk would have effects on insurance arrangements for providers and patients that warrant brief notice even though the details of this subject are beyond the scope of this article. Implicit in much of the foregoing discussion is the assumption that a contractual shift of risk from providers to patients would also shift insurance costs from third-party malpractice coverage to other insurers such as first-party health and disability insurers. The likely magnitude of such a shift is difficult to predict. Insofar as malpractice insurance is community rated, limited use of such contracts would probably have no immediate effect on insurance costs for individual providers. Individual providers who can contract out of risk might, however, choose to self-insure ("go bare") or to reduce their coverage. Moreover, the withdrawal of a significant number of risks from the insurance pool would in time presumably affect the rates for the entire class of providers affected.

The effect on first-party insurance is more ambiguous. At first glance it might seem that a reduction of third-party insurance premiums would produce a symmetrical increase in first-party premiums covering the same risks. Such a symmetrical increase presupposes, however, that the payment obligation is merely transferred unchanged from third-party to first-party payers. In fact, first-party insurers already pay a significant portion of the costs of medical accidents, and to the extent of this double payment by first- and third-party payers, the patient's assumption of risk would not increase first-party insurance premiums.

To the extent that premium costs would be increased, it is desirable to have them borne only by those contracting with the providers rather than to spread them among the entire pool of insureds. Otherwise, the addition of new high risks could pollute the pool by adverse selection. At least in principle it should be possible to avoid the adverse selection problem; insofar as the high-risk insureds identify themselves by contracting to assume risks heretofore borne by providers, the insurer should be able to make individual premium adjustments to the extent that the reallocated risks are significant.

76. The transfer could also be from malpractice insurers to other third-party insurers when other tortfeasors are jointly liable for the medical accident costs. This article is concerned only with the transfer from third-party to first-party payers.

77. Double payment assumes, of course, subrogation of the first-party payer to the victim's claim. Subrogation in this context varies. For health care insurance, subrogation is generally dependent on express provision in policies, and the use of such provisions, while increasing, is not universal. To make matters more complex, some courts have invalidated express subrogation clauses in this context. See R. KEETON, BASIC TEXT ON INSURANCE LAW 150-53 (1971).


79. Some such adjustment is impliedly assumed in the discussion of shifting incentives, supra notes 53, 72 and accompanying text.
In marked contrast to the usual abstract context in which Coase's "solution" to accident cost allocation is discussed, the medical situation provides at least a plausible occasion for considering the contract option. There are no insurmountable economic impediments to bargaining; the only significant transaction costs are the impediments imposed by a social (or social *cum* legal) culture that views the relationship between patient and health care provider as one of status more than contract—a culture in which the allocation of risk is, in Walzer's phrase, a "blocked exchange." While this paternalistic attitude agrees with the image that the medical profession has promoted and defended aggressively since the turn of the century, it is out of touch with the reality of modern health care services. It is ironic that physicians are blocked from negotiating different risk allocation arrangements with their patients largely by the world view that they themselves so assiduously cultivated.

Social culture aside, there are more positive, tangible objections to freedom of contract. It is said that the particular vulnerability of patients demands special legal solicitude and judicial restrictions on their freedom to assume the risks of medical error. Although I am skeptical about the need for aggressive judicial intervention in so-called "unconscionable bargain" cases generally, I concede that there may be appropriate instances for judicial protection—even for "paternalism," if you will. Certain types of medical treatment cases suggest themselves as presumptively within that category. Emergency treatment would certainly qualify for special solicitude; so too might certain nonemergency cases when treatment is urgent and patient options are restricted by inherent limitations of the market, or when the patient's situation suggests a special vulnerability to overreaching. The traditional legal posture that medical risks must be judicially assigned is not convincing, however. That posture reflects several dubious notions about health care in general and about medical risks in particular. As to health care, the central misconception is that it is so inherently essential and the course of treatment so obviously prescribed by medical "science" that the patient has no real choice. In fact, medical science has not overcome all uncertainty, and most health care services are not compelled in the sense that there are no practical options with respect to the nature of the treatment or the manner in which and conditions on which they are delivered.

The assumption of full disclosure is, of course, central to the argument. It is widely thought that patients' inability to evaluate probabilistic information precludes rational risk assessment. But the inability of individuals to evaluate risk is largely irrelevant in the context of contractual transactions between

---

80. See M. WALZER, SPHERES OF JUSTICE 100-03 (1983). Walzer's examples of "blocked exchanges" are too general to be applied to liability waivers, although certain cases could fall within his category of "disparate exchanges" which are prohibited.
individuals who share the same disability. Nor is there anything peculiar about medical risks that sets them apart from other accident risks. If inability to evaluate risk is a basis for restricting assumption of risk in the medical context, it should also be a basis for superintending all individual behavior involving judgments about risks.

It is appropriate to close with a disclaimer. If there was a medical malpractice “crisis” in the 1970’s, I doubt it could have been resolved simply by allowing doctors to contract out of liability for risks. Nor do I think such a change in legal policy can itself correct the alleged present crisis. Most health care providers would probably not want to use exculpatory clauses because they would not want to risk arousing patient suspicion or patient demands for some consideration in the form of price concessions. (If individual patients did not demand concessions, organized purchasers, such as employer groups, probably would.) Nevertheless, some health care providers might well seek to shift liability for some types of risks. For example, obstetricians or birthing centers might seek to avoid liability for deformed infants (liability might be predicated on failure to detect and disclose the abnormality) and might indeed be willing to meet demands for a contractual quid pro quo. That the number of such arrangements is likely to be too limited to transform tort law into contract law does not distinguish this small reform from dozens of other sensible measures that could improve, at the margin, the functioning of the legal system.

As to improving the health care system, the disclaimer must be made even more forcefully. The “political pathology of health care policy,” to borrow Wildavsky’s phrase, is a matter largely beyond the ken of the common law. The most that can be said of changes in legal liability rules along the lines suggested here is that they are consistent with the direction of larger reforms aimed at diversifying health care practice and containing the hemorrhaging of health care expenditures. No more can be expected; even if one were dramatically to reduce and not merely reallocate the costs of medical malpractice, the impact on overall health care costs would be very small. Again, however, we ought not be paralyzed by the thought that behind one problem solved there remain larger ones yet unsolved.


82. Taking medical malpractice premium costs as a rough measure of the costs of malpractice risks—and ignoring the costs of defensive medicine—the ratio of such costs to total health care expenditures is less than one percent. See MILLIMAN & ROBERTSON, INC., ACTUARIAL ANALYSIS OF AMERICAN MEDICAL ASSOCIATION TORT REFORM PROPOSALS 2 (1985) (1984 premium costs were $2.7 billion); Levit, supra note 17 (1984 health care expenditures totalled $387.4 billion).