The problem of medical malpractice is difficult only if one believes that liability rules are important tools for achieving what I have elsewhere called primary accident cost reduction - that is, the minimization of the sum of medical accident costs and medical safety costs. In the medical context this 'economic efficiency' motive is no different from the knotty problem of 'achieving the highest quality of medical care' where (if we are not to be silly or fatuous) highest quality implies 'considering the price.'

I have not said that the desire to reduce the sum of medical accident and medical accident avoidance costs by itself makes medical malpractice a difficult problem. If one believes that collective, regulatory approaches suffice to limit medical maloccurrences to those, and only those, which would be too costly to avoid, then medical malpractice as we know it today, or generally discuss it in reform proposals, becomes unimportant. If governmental rules or peer group controls, roughly akin to building codes - I believe they are called professional standards review organizations (PSROS) in the United States - can determine what medical care is worthwhile - and to be mandated - and what is noxious or simply too costly - and to be proscribed (I exaggerate the ideological purity of the approach, of course) - then efficient medical care can be achieved without the dubious benefits of torts law, and one need not discuss the issue further.

Similarly, though conversely, medical malpractice has no role to play if one believes that negotiations as to appropriate treatment occur between

---

* John Thomas Smith Professor of Law, Yale University, New Haven, Connecticut
† This paper is a slightly revised version of the C.A. Wright Memorial Lecture delivered at the University of Toronto on 9 February 1977. A preliminary version of the paper was presented for discussion at the American Enterprise Institute Conference on the Economics of Medical Malpractice held in Washington, DC on 15-16 December 1976. The American Enterprise Institute will publish the paper as part of a collection of papers there presented, together with the discussions which followed.
2 PSROS were authorized by Congress in 1972 to monitor the appropriateness of health services financed by Medicare, Medicaid, and other health programs. U.S. Congress, House of Representatives, Social Security Amendments of 1972, Pub. Law 92-603, 92d Congress, 2d Session, 1972, H.R. 1. For a recent review of quality assessment programs in medical care, including PSROS, see, Institute of Medicine, Assessing Quality in Health Care: An Evaluation, Final Report (1976).
doctors and patients with sufficient parity of knowledge of risks. In such a
case also, torts law is not needed because patients will choose those and just
those treatments whose risks and costs they desire, and just those doctors
whose abilities, for the price, they find optimal.3

The first non-tort law approach to achieving optimal quality of medical
care that I described is a parody of a collective ideologue's search for
utilitarian efficiency in medical care. That this regulatory or criminal law
model is not infrequently suggested will frighten some and amuse others,
but, from my point of view, what is important is that most thoughtful people
are likely, soon, to recognize that it is not very promising as an effective way
of assuring optimal medical care. It can work to control extreme abuses - the
butcher can be deterred, the obviously unnecessary treatment can be for­
bidden. But this approach, which is so rigid that it fails even in relatively
stable technological areas (like the building trades), is apt to be a very poor
guide indeed when the issue (of what care is worth risking) changes as
rapidly as modern medicine.

The second non-tort law approach that I mentioned is a parody of the
market ideologue's dream of utilitarian efficiency in medical care. This
contractual model, at first glance, seems particularly appealing because here
those transaction costs, which frequently make the pure market a poor road
to an utilitarian's heaven, seem absent. The victim does not, as in auto­
pedestrian cases, have to seek out every possible injurer and bribe him to
employ better brakes - he is staring at his potential injurer across the
examining room. Unfortunately, parity of knowledge of risks is needed for
contractual arrangements to lead to Professor Coase's optimum,4 and the
cost of giving that knowledge may come very high indeed. It is not only costly
to enforce rules designed to give parity of knowledge, that is, to make patient
consent be 'informed,' but, more important, parity of knowledge and in­
formed consent frequently entail significant costs of making people know
things they do not wish to know. This fact makes the medical malpractice
problem far more complicated than the analogous problem of knowledge in
other contractual situations. Extreme cases of bad treatment and bad doc­
tors may well be controlled ultimately by the unfettered market - but those
cases could also be controlled by regulation and are not the key to our
problem. As a result, as many despair of the contractual approach to optimal
medical care as do of the regulatory-criminal law approach.

All that I have so far said, however, is in no way unique to medical

3 Cf, Epstein, Medical malpractice: The case for contract, (1976), 1 American Bar Foundation
Research J., 96-107. For a very different contractual approach, one that assumes medical
malpractice law as a starting point but allows persons to contract out of that, fault-based,
liability by electing a system of no-fault liability and compensation, see, O'Connell, An

4 For the classic description of how contractual approaches can, in the absence of transaction
costs, lead to efficiency, see, Coase, The problem of social cost, (1960), 3 J. Law and
Economics, 1-44.
malpractice. It applies as well to most areas where safety costs must be traded off against accident costs. In those areas too, regulation or criminal law are occasionally deemed effective and employed while at other times contractual negotiations are allowed to establish the desired level of safety. The major difference between those areas and medical malpractice is the fact that in those areas the tort law approach — whatever its flaws, and be it fault or non-fault based — seems to be a reasonably effective supplement when regulation and contract do not suffice to minimize the sum of accident and safety costs. What is so troublesome about medical malpractice, instead, is that neither the fault-based approach, which so far has been dominant, nor any proposed non-fault substitutes has seemed to be a satisfactory means of medical care assurance in that — apparently — broad set of situations in which neither regulation nor contract seems likely to work.5

It may, of course, be the case that our society is not interested in reducing the sum of medical maloccurrence costs and the costs of avoiding them, or — what is the same thing — that the cost of any system which would do so beyond what can be done by regulation and contract is so high as not to be worth undertaking. It may even be the case that the fuss about malpractice has little to do with the quality of medical care at all. It could, instead, have to do with a desire to compel spreading of the costs of certain catastrophic medical maloccurrences. Or, it could have to do with a desire to compel spreading of the costs of certain catastrophic medical maloccurrences. Or, it could have to do with a desire to remove the cost of such maloccurrences (whether spread or not) from victims and doctors and have them be borne — for purely wealth distributional reasons — by taxpayers at large. If either of these goals is what we are seeking (and in part they surely are), then, once again, the malpractice problem becomes relatively simple. Those losses which we would have spread, we will require people to insure against; the premiums can be borne by patients and by doctors — if, like good monopolists, doctors are already charging the most the traffic will bear⁶; or they can be removed and paid out of the general fisc — that is, assessed to whomever, for wealth distributional reasons, the society wishes to assess. If they are removed, of course, all financial incentives for optimal quality of medical care are also removed but, by the hypothesis that I made, the society

5 Cf, Keeton, Compensation for medical accidents, (1973), 121 U. Pennal L.R., 590–617. The matter is more complicated than the text suggests. All approaches, regulatory, contractual, and mixed (like torts law), get overburdened. Regulation may work well in any given context, and yet fail if that context is the last of a long list of areas which we are trying to regulate. The knowledge needed for a contractual approach to work well may be made available at reasonable costs in a limited number of contexts and yet be impossible, or prohibitively expensive, if one tries to make it available in all contexts. Similarly, the tort law approach may be able to handle adequately some problems (perhaps even malpractice), if these were the only problems it had to handle, and fail miserably if the same problems were part of a much larger set assigned to it. In such situations, decisions based on the comparative advantage of using one approach as against another become crucial. Thus the fact that torts might do a slightly better job than 'pure contract,' for instance, would not suffice to justify the use of a torts approach if such use made torts less effective in other areas where its advantage over contractual methods was greater.

would not care or could achieve what it sought by other means, like regulation.

Unfortunately, however, I am not convinced that we can so readily ignore the possibility that financial incentives established through allocation of liability for medical maloccurrence are important to the achievement of that level of medical care our society desires. It may be that after considering the miserable job that current malpractice law does in this regard, and after examining the difficulties with possible reforms, we are reduced to saying: let us rely on the internal pressures for good care that the profession develops, let us deal with extreme cases through regulation or criminal law, or through the kind of pressures which the contractual approach (that is, patient-doctor negotiations) will give despite limited and unequal knowledge, let us spread losses or redistribute burdens to our (ideological) heart's desire and let us accept that, since we can do no better, what we have (without more complicated financial incentives) is efficient, indeed optimal, even if pretty dreadful. 7 It may be, yet we cannot begin with that gloomy conclusion.

II

I cannot in this paper spend much time on the problems, real and fancied, with the existing system of malpractice law. The theory of that branch of law is simple enough and not even totally unsound; the reasons it does not work adequately are real and have been written about quite sufficiently. Like most fault-based approaches to liability, medical malpractice begins with the assumption that the costs of any medical maloccurrences which cannot be judged, by a jury, court, administrative agency, or legislature, to be worth avoiding, should lie on the victim. If, instead, a governmental agency of the kind just listed is prepared to say that the cost of avoiding the maloccurrence is less than the cost of the untoward event discounted by its risk, and that the doctor should have known this, then the doctor is at fault and must pay for the maloccurrence. In theory the same fault approach would charge doctors with unnecessary avoidance costs: costs the doctors should have known were not worth the maloccurrences they avoided. Such unnecessary avoidance costs are just as much at fault, in a technical sense, and ought, from an efficiency point of view though not from a spreading point of view, to be deemed malpractice.

That in the malpractice area the governmental-utilitarian, or cost-benefit, analysis just described has been influenced (to an extent unknown in other areas) by the custom of the trade (medical practice as shown by expert testimony) is not crucial, since if that aspect is considered undesirable it could with relative ease be cured. Similarly, one need not worry unduly about the fact that this traditional approach leaves on the hapless victim all

7 Cf, Epstein, supra note 3.
the medical maloccurrence losses which are not deemed worth avoiding. One need not worry about that since, if we wish, insurance against such, or some particularly catastrophic set of such, losses could be required, and the burden of the losses could thereby be spread among all users, or even removed to distributionally preferred payors. All this could be done without substantially altering the incentive element on doctors to avoid those and just those medical maloccurrences the governmental agency deems, after the fact, to have been worth avoiding before the fact. (Spreading or removing of losses from victims might, in theory, reduce victim care. But the notion that these losses – by hypothesis deemed not worth avoidance by doctors – are subject to significant reduction through victim care, seems, to me at least, most unlikely.)

Similarly, the basic problem with fault-based malpractice as a vehicle for assuring optimal quality of medical care does not lie in the very real inadequacies of the insurance market. It may well be true that (since the law in this area is changing rapidly [and here I am speaking mainly of the United States, of course], though perhaps not more rapidly than tort law generally – and since malpractice cases do not get decided until many years after the alleged tort) insurance companies are incapable of approximating the actuarial risk they undertake in insuring doctors against distant future losses arising out of today’s events. Because insurance companies cannot be sure that they can charge future generations of doctors high enough premiums to cover old liabilities which are just coming to judgment (new entrants in the insurance market could undersell old companies in insuring such future doctors since these new insurers would not bear the liability backlog), present insurance rates tend to be ‘gambling’ rather than actuarial rates. They are designed to cover an unknown future risk. Insurance companies, being more averse to risk than almost anyone – they are by specialization actuaries not entrepreneurs – are most unwilling to take on a gamble. This leads to few insurers and rates, among those few, which may, in retrospect, turn out to be exorbitant but which cannot ex ante be shown to be wrong. The rate of return demanded for bearing an entrepreneur’s risk – true profit, F.H. Knight called it – has always been high, even when sought by those less averse to risk than insurance companies. Still, this problem could be alleviated by systems designed to assure insurers that they will be able to compel future generations of doctors to bear a share of the past risk backlog. Indeed, that is what reforms like ‘medical association’ malpractice insurance often come down to (if the medical society believes it will be able to compel future doctors to insure with it, it need not worry if current rates turn out to be inadequate to cover future suits and can compute those rates without a ‘gambling’ premium). Or the problem can be alleviated by government intervention. The government can ‘insure’ against the entrepreneurial risk,

much as it 'insured' long-term low-deposit mortgages after World War II in
the United States. If the entrepreneurial risk turns out to have been low (as
in the post-war mortgage market), the government will have simply been an
insurer. If instead the risk turns out to have been higher than the premiums
charged could cover, then, in retrospect, the government will have sub-
sidized medical care by its intervention. Either way, the problem would have
been dealt with.

No, the basic problems in fault based malpractice law do not lie in these
real but on the whole correctible problems. Nor do they lie in that much
maligned, but nevertheless admirable, lawyer's fee mutual-insurance
scheme that we call the contingent-fee system. They stem, rather, from the
fact that we have little or no faith in courts, juries, administrative agencies,
and even peer-group tribunals when they purport to tell us which medical
maloccurrences are worth avoiding, which are not, and, perhaps more
important, which maloccurrence-avoidance costs are too expensive and
should be avoided. I again except the easy cases of butchery, because those,
as I have said, probably can be dealt with adequately under any approach.
Short of those cases, however, we have virtually no assurance that the
administratively highly expensive, stigma spewing, approach we use creates
those incentives to good medical care which, ultimately, can be its only
reason for being. And if it does not, then we would be justified, it would
seem, in abandoning it, and substituting for it whatever systems would give
us the spreading or wealth distribution allocation of burdens we desire, and
forget about incentives.

Why does not the fault-based system give us the incentives which are
needed to justify it? In theory, as I said, there is no reason why it should not.
As a practical matter, however, it is clear that the governmental or peer-
group decision maker will consistently find only certain categories of mal-
occurrences to be avoidable. In almost every instance, for every treatment
which, if it goes wrong as sometimes it will, gives rise to a compensable loss,
there exists a substitute treatment or non-treatment whose harm in practice
would not be recognized as a compensable loss. This problem in the
decision-makers is compounded by the fact that the victim is equally 'biased'
in what he or she can recognize as a cost which can give rise to compensation.
In other words, almost inevitably, the fault-malpractice system penalizes
some medical maloccurrences and some avoidance costs while systematically
failing to penalize the medical maloccurrence and avoidance costs of substi-
tute approaches. One need not be an economist to realize that any approach
which so biases incentives is unlikely to lead to an efficient result. Defensive
medicine is an apt term for what unfortunately is only the tip of this iceberg.
Unless we can round out the circle and create roughly equivalent incentives
for avoiding medical maloccurrence costs and safety costs in substitute forms
of treatment and non-treatment, we will not do what alone can justify the
existence of this cumbersome and expensive area of the law. But if the
unsolved problem of creating unbiased incentives is the source of our fundamental dissatisfaction with existing malpractice law, it equally bedevils proposals for its reform, to which I must now turn.

III

The most promising attempt at reform, I think, is that tried by Havighurst and Tancredi. They recognize the need to establish incentives to good medical care for treatment and non-treatment alike, as well as for substitute forms of treatment. Their solution is to select areas of medical care in which all significant maloccurrence costs of substitute treatments either can be compensated for on a non-fault basis, or can be prevented by regulation. That is, if non-treatment in a particular circumstance is so grotesque as to be readily proscribed (or even without proscription, will be avoided by doctors regardless of complex financial incentives) — say, a failure to operate or to use anesthesia in the operation — then that non-treatment need not be worried about in establishing an incentive scheme for medical care. It is not a meaningful substitute. This fact permits us to charge — on a non-fault basis — for any harm arising out of the inevitable operation and anesthesia, thereby creating appropriate incentives for choosing among all meaningful substitutes.

Unfortunately, the number of situations in which such a circle of possible substitute treatments is either complete or can satisfactorily be completed by regulation is quite limited. Much remains out of Havighurst’s and Tancredi’s control. More important, their approach tends to create a series of border situations where victim compensation depends on whether the injury occurred within the scheme or without. Whenever that is the case administrative costs multiply, because the victims have a strong incentive to describe the facts so that they fall within the scheme, and the injurers have the opposite incentive. One major source of the high cost of workmen’s compensation plans in the United States is the litigation expense of fighting out the often crucial issue of whether an injury arose ‘out of and in the course of employment,’ and is covered by the scheme, or did not, and is covered by the riskier (but occasionally more liberal) general law of torts. Unless we are prepared to compensate victims to the same extent — whatever the source of their injury: treatment, or non-treatment; if in one area of medicine, then in all others (much as the Woodhouse approach did for all accidents in New Zealand) — we will have created precisely the borderline costs which have so bedevilled other reform schemes.

What would be a complete system of financial incentives to good medical care? What would round out the circle so that on a non-fault basis the

appropriate incentives would operate and we would be free to abandon both fault and regulatory approaches? It is not difficult to state such a complete circle – in one context. But to state it is to suggest immediately the problems – the costs, if you will – of rounding out the circle. Because of this and because there are many other, equally flawed, ways of rounding out the circle, I do not describe my complete circle as a proposed reform, but rather as a 'joke,' a way of raising the question of whether it is worth pursuing any reform which proposes to take seriously the goal of achieving optimal quality of medical care through financial incentives.

Let us posit a system of health maintenance organizations (HMOS) with sufficient competition among HMOS so that a substantial choice among them exists for any consumer. Let us posit also that the choice of HMOS will be made in the employment context through union-management negotiations, so that costs of treatment and non-treatment become available as understandable statistics to those who are to opt for more avoidance (at more expense) or the converse. (The obvious costs of all this in terms of individual choice, even if we posit representative unions, need not be dwelt upon.)

Such an approach would immediately internalize to each HMOS some cost of 'bad' medical care, and cause these to be reflected in the rates each HMOS would charge in contradistinction to 'safer' competitors. Without more, however, the circle would not be rounded out. Specifically, the costs of subsequent medical care as a result of either inadequate or too much treatment would be internalized to the HMOS, and an incentive to treat initially so as to reduce such subsequent costs appropriately would exist. But there would be no incentive linked to the HMOS to avoid treatment which, if it failed, would produce an untreatable, but perhaps unemployed and unemployable, victim.

Again, if individual consumers could know this risk, adequately, without insurance, and therefore make it a part of the HMOS negotiation, there would be no problem. But this is no more than saying that, in theory, pure contract law could cope with the problem of adequate medical care – and were that so, we would have ended this discussion long ago. No, to convert this un-

10 'HMOS' distinguishing characteristic is that they undertake to provide all the medical care their enrollees need in exchange for fixed, advance capitation payments.' Bovbjerg, The medical malpractice standard of care: HMOS and customary practice, (1975) Duke L.J. 1976. The same article, at 1975–414, contains a careful description of HMOS and of how malpractice law would affect and be affected by them.

11 See supra text and note 2. To the extent that HMOS, by working in an employment context, generate greater knowledge of such risks than would otherwise be available to individual patients, then, as Frech (in a discussion of this article to be published by the American Enterprise Institute as part of the collection of papers on the Economics of Medical Malpractice mentioned in note 1, supra) points out, some of the impediments to a pure contractual solution would be removed. One may well doubt, however, that the union-employer-HMOS nexus will, without specific incentives, do an adequate job of internalizing the costs of those injuries which would make a patient, unemployed and unemployable; that is, injuries which make the patient not only no longer a charge on the HMOS, but possibly no longer the responsibility of either the employer or the union.
employment risk into a statistical figure which could be meaningfully assimilated in the HMO-employment contract some sort of wage maintenance insurance would be required. And, since it would be hard to define when wage maintenance was needed because of treatment and non-treatment errors and when it was needed for non-medical reasons, the most likely form such insurance would take would be universal wage maintenance. (The 'moral hazard' costs involved in such an approach should be readily apparent.) With such insurance, any given employer-union combination, in any given industry, could readily compare the cost of competing HMOs not only in terms of their present and future treatment costs, but also in terms of their 'absenteeism' costs.

The circle would still not be complete, however. The HMO which either cured or killed could underbid the HMO which preserved life at a higher present or future medical and absenteeism cost. The same would be true for the HMO which reduced safety and accident costs by increasing chances of pain which could not be alleviated by treatment and which did not bring on absenteeism. The second is perhaps too fanciful to be worried about. The first may well not be. In order to deal with the first, we would have to require that life insurance for a substantial amount be made part of the union-employer package, along with wage maintenance and insurance against future medical care needs. Then the circle would be complete, the buyer would know that virtually all treatment and non-treatment costs were reflected in the price of any HMO, and that, therefore, incentives to reduce such costs, appropriately, would lie on the HMO and its employees, the doctors.

Reduction in HMO price by reducing amenities of care would, of course, remain possible, as would increased comfort, at increased price. But as to these, I readily admit, I become a pure contractarian. I think the consumer has adequate knowledge to decide whether standing in line is worth what it saves, and having an HMO which permits relatives, comfortably and easily, to visit the hospitalized patient is worth what it costs. As a result, such comparison among competing HMOs would not, in my view, require any intervention to maximize satisfaction. I hope so, for, if I am right, the long and rather tedious discussion between doctors and sociologists of what constitutes adequate medical care, and of whether it includes comfortable waiting rooms and bedside manners, can be shortcircuited. And, more important, rounding out the circle, it should be clear, would be exceedingly costly even without further intervention designed to deal with these problems.

It may be worthwhile spending a moment considering why it is so complex and costly. That may best be done by thinking about what it is that we have done when we completed the circle. We have, in effect, made any death which occurs sooner than the appropriate life expectancy for individuals in that industry a part of the medical care cost. We have also made any absenteeism, whether treatable or not, part of the same cost. And, of course,
we began by making subsequent medical care part of the cost as well. And finally, we made the costs of avoiding all of these costs part of medical care costs. To the extent that an HMO does an average job on any of these counts (as presumably would be the case if death or absenteeism were genuinely unrelated to treatment or non-treatment), the HMO and its competitors would be affected equally. A superior or inferior job, however seemingly unrelated to medicine, would instead be reflected in the competing HMOs' costs. Financial incentives would now be present because any medical maloccurrence and any costs of avoiding that maloccurrence would now be part of the scheme and compensable. But, to create the scheme, all sorts of non-medical costs were included in it and made subject to compulsory insurance. Since insurance is, obviously, not cost free – administratively, in terms of incentives to avoid harms, and in terms of permitting individuals the option to choose risks rather than coverage and spend the money saved elsewhere – the well-rounded circle comes anything but cheap.

If, however, one believes that, except for extreme cases of butchery, anything short of a complete circle of incentives is worse than no incentives, what are we to do? The fault approach, by trying to concentrate on those treatments and non-treatments which, ex post, we could say were worth avoiding ex ante, was one solution. It is plausible; unfortunately, if its object was efficiency in medical care, it has failed. We could try a Havighurst-Tancredi (Medical Adversity Insurance) approach, but it doesn't, ultimately, promise enough of an answer. We could go to regulation, and in my judgment either create a system which does nothing (but may give people at large the feeling that something is being done, until they wake up and realize they have been fooled), or, worse, create a system which applies real standards that will quickly go out of date and lead to poor medical care. Finally, we could give up on any incentives to good medical care, except in the extreme cases, and concentrate solely on spreading and wealth-distribution goals. This would entail the abolition of malpractice law, the establishment of medical catastrophe insurance, and the creation of peer or criminal law bodies to punish the occasional scoundrel.

I have been told that this last approach would prove unacceptable to the mass of people, especially as medical care plans become more universal. In such circumstances people demand the appearance of controls, even if ineffective and costly. And certainly, our tendency to look for scapegoats or for what appear to be solutions, even when they don't exist, in other areas of law supports this cynical view. Perhaps the old, old, rarely successful malpractice suit, ineffective both as a source of incentives and mis-incentives, as it was, served this very function of creating a useful illusion of control where no control was feasible. In the United States, it is, nonetheless, dead and cannot be resurrected. Old subterfuges, once exposed, can almost never regain credibility. One may wonder whether, had non-fault compensation for medical catastrophes been established and paid for out of the general
fisc, we would have felt pressure to enlarge, and ultimately to expose as a sham incentive system, the liability of doctors for malpractice. But whatever one may think about that question, the option it represented is probably no longer open to us in the United States. Whether it is still open to Canada, I do not know, but given the United States' proximity, it seems hardly likely.12

As a result, we are faced with an uncomfortable problem. We can try to round out the circle of incentives, all the while being deeply uncertain as to whether such schemes will be worth their costs. We can openly abandon incentives in this area (except for those that either pure contract or regulation can give at the extremes) and concentrate on spreading and redistributing some medical burdens, knowing all the while that perhaps a more sophisticated incentive scheme might have been worth it. We can try to control by regulation what in easier areas has proved very hard to regulate intelligently. Or, finally, we can create new subterfuges, less adequate than the old precisely because chosen rather than found, which only give the semblance of assuring 'quality' medical care while freeing us to spend or redistribute to the extent we wish. I, it should be obvious, would opt for either of the first two. I strongly suspect, however, that both will be rejected in favour of regulation or subterfuge, or most likely a mixture of regulation and subterfuge like the American PSROS.

12 One may usefully compare England where, in a legal environment which early promoted compensation, malpractice law has seemed quite capable of withstanding pressures to expand liability.