TORT REFORM: THE CASE OF MEDICAL MALPRACTICE

PATRICIA M. DANZON
The Wharton School, University of Pennsylvania

I. INTRODUCTION

In most countries the prevailing rule of liability for medical injuries is some form of negligence rule (Schwartz, 1992). Many countries, including the UK, the US, and Canada, are increasingly dissatisfied with this traditional system. In theory, the tort system is designed to deter medical negligence and to compensate patients injured as a result of negligent care. The evidence suggests that it performs these functions imperfectly, at best, and at high cost, including high overhead costs. Although rough estimates suggest that the frequency and cost of malpractice claims is several-fold higher in the US than in other countries (Danzón, 1990), there is a common concern over the frequency of medical injuries and claims, and the costs of compensation and of malpractice insurance premiums.

Over the last two decades most states in the US have enacted some tort reforms for medical malpractice, including caps on awards, offset of benefits from other collateral sources, shorter statutes of limitations, screening and mediation panels, etc. Some of these reforms have moderated the growth in number of claims and size of awards (Danzón, 1984a, 1986; Zuckerman et al., 1990), but fundamental criticisms of the tort system remain. More radical alternatives, including enterprise liability and no-fault systems of compensation, have been proposed, drawing partly on the no-fault schemes that have been in operation in Sweden and New Zealand for two decades.

In the UK the rapid increase in number and cost of medical claims in the 1980s led to the introduction of the National Health Service (NHS) indemnity, a form of fault-based enterprise liability whereby the Health Authorities assume responsibility for defending all claims arising from NHS treatment by employed doctors; similarly, self-governing NHS hospital trusts assume liability for the negligence of their employees. Other alternatives, including several no-fault schemes, have been proposed by the British Medical Association, the Royal College of Physicians, and the NHS (Fenn, 1993b).
Since concern over medical malpractice has been triggered by its budget cost to payers and apparent inequity in compensation, reform proposals tend to focus on these features. However, a broader view of the real social cost of medical injuries indicates that the primary function of a medical liability system should be quality control ('deterrence'). Compensation can be provided at lower cost and more equitably through other public and private insurance systems. Thus the tort system and tort reform must be evaluated in the context of the full network of systems of quality control and injury compensation that exist in all countries. In this paper, section II outlines the economic theory of professional liability. Section III summarizes the shortfalls between this theory and the actual operation of malpractice systems. Sections IV, V, and VI evaluate proposed reforms, including traditional tort reforms, the Swedish and New Zealand no-fault compensation schemes, and proposals for an administrative fault-based system (AMA, 1988) and enterprise liability. Section VII concludes.

II. THE THEORY OF TORT LIABILITY

Physicians and other learned professionals—including architects, attorneys, and accountants—have been singled out from other occupations in their professional liability to clients. The traditional basis for professional liability is negligence. Under a negligence rule, the plaintiff must show that the defendant owed the duty of care, that he failed to conform to the required standard of care, and that this failure was the proximate cause of the plaintiff's injury. Traditional rules of tort damages provide for full compensation of pecuniary and non-pecuniary damages. Thus, in principle, the law of medical malpractice holds health-care providers liable only for medically-caused (iatrogenic) injuries that are caused by negligence; adverse outcomes that are consistent with the normal risks of customary medical care are the burden of the patient. Nevertheless, most professionals consider liability insurance to be a prerequisite of professional practice. Tort liability performs two primary functions. First, by providing compensation it acts as a source of insurance. Second, by imposing sanctions on persons found negligent, it deters future negligent behaviour. However, if the tort system is evaluated on grounds of economic efficiency, then it can be justified, if at all, only by its performance in deterring negligence. Compensation and risk spreading can be accomplished at lower cost—and arguably more equitably—through either public or private first-party insurance. In the US roughly 40 cents of the malpractice-insurance-premium dollar reach the patient as compensation, compared to over 90 cents for large first-party health-insurance programmes. Much of the difference—about 40 cents of the liability-insurance dollar—is spent on litigation, equally divided between plaintiff and defence. Other real but hidden costs of tort liability include time and anxiety costs borne by the litigants, and liability-induced distortions in medical practice—'defensive medicine'. These additional costs of tort liability are worth incurring only if there are offsetting deterrence benefits, in terms of future injuries averted.

Economic models have examined the efficiency of alternative liability rules in performing this deterrence function.\(^1\) Optimal or efficient investment in injury prevention minimizes the total societal cost associated with injuries, including costs related to injuries, prevention, litigation, and overheads.\(^2\) This requires that, at the margin, a dollar spent on prevention saves a dollar of expected injury-related costs, including overheads.

A fundamental principle of liability rules is that, if all parties are fully informed about risks and contracting is costless, then the allocation of resources to loss prevention will be the same, regardless of whether the liability rule is *caveat emptor* (all losses reside with the victim) or strict liability (all losses shifted to the injurer) (Coase, 1963). But if consumers misperceive risks or contracting is costly, then *caveat emptor* leads to non-optimal accident rates and non-optimal insurance (Spence, 1977; Shavell, 1980). Such asymmetric information pro-

---

1. See, for example, Brown (1973), Shavell (1980).
2. "Optimal" and "efficient" are used here in this technical sense. They have normative content only to the extent that efficiency is a major goal of social policy, without implying that it is or should be the only policy objective.
vides a rationale for the professional liability of learned professionals. In the case of medical care, if patients are less well informed than providers about the benefits and risks of alternative treatments and cannot readily monitor the quality of care delivered, then the rate of risky procedures and care per procedure may be non-optimal. 3

In principle, a negligence rule of liability can correct these distortions and create incentives for efficient care and risk-taking, under certain conditions. These conditions include that courts set the standard of due care at the efficient level, that damages be optimally set, that providers be liable for failure to obtain informed consent, and that suits be brought and compensation awarded if and only if negligence occurs. 4 Efficient deterrence incentives can, in theory, also be achieved by a rule of strict liability, whereby providers are liable for all injuries caused by medical care, regardless of negligence.

However, negligence and strict liability differ in their allocation of risk, number of claims, and overhead costs. Under strict liability all iatrogenic injuries would be compensable through tort. This is inefficient if tort compensation is more costly to administer than first-party compensation. By contrast, under a perfectly functioning negligence rule there should be no negligence and no claims, since by definition it is cheaper to prevent injuries that would be deemed negligent than to pay for the resulting damages (Shavell, 1982). Injuries that optimally are not prevented could be covered through private first-party or social insurance.

The negligence system operates in practice very differently from this theoretical ideal, primarily because the decision-makers—courts, doctors, patients, liability insurers—lack the perfect information that is assumed by the models (Danzon, 1991a). Because courts lack perfect information about appropriate care, the standards applied in practice are unpredictable and possibly systematically biased. With uncertain legal standards, a negligence rule may create non-optimal deterrence incentives (Craswell and Calfee, 1986), including incentives for 'defensive medicine'; many valid claims are not filed and many invalid claims are filed. Uncertain legal standards lead to a demand for liability insurance. Perfectly experience-rated liability insurance would not interfere with deterrence. But in practice experience-rating is very crude, both because insurers lack the necessary information and possibly because of political pressures. A rule of strict liability in theory eliminates the need for courts to define due care. But determining whether an injury was caused by medical care, rather than by the underlying disease, would require a similar inquiry, as would the no-fault rules that exclude 'normal risks' of medical care. Moreover, a strict liability rule is more vulnerable to court errors in setting damages and requires administration of many more cases. Thus, once the assumptions of perfect information are abandoned, the choice between liability rules becomes ambiguous a priori and we must turn to empirical evidence on costs and benefits.

Unfortunately, however, accurate empirical evidence on key components of the costs and benefits of alternative liability regimes is unavailable. Most problematic is the measurement of deterrence benefits—the injuries averted because liability makes providers more careful. 2 It has proved impossible to distinguish empirically between efficient liability-induced changes in medical practice (deterrence) and wasteful defensive medicine.

III. CONCERNS WITH THE TORT SYSTEM

(i) Rising Claim Costs

Although medical malpractice liability has existed for centuries, such actions were rare until the late 1960s. In the US from the early 1970s to the mid-

---

3 This prediction holds if providers act as self-interested income-maximizers. Altruism, professional or ethical concerns, or other quality-monitoring mechanisms may modify the result (Danzon, 1991b).

4 Liability for lack of informed consent is necessary to control the rate of risky procedures. Strictly, these conditions are sufficient but not necessary. Efficiency could in theory be obtained if deviations are offsetting, e.g. a shortfall in claims is offset by higher damage awards.

5 It is often argued that the lack of experience-rating of liability-insurance premiums undermines the deterrence potential of tort liability and that reforms should therefore focus on compensation. But even with flat-rated premiums or under the NHS indemnity in the UK, the psychological and reputation costs of liability may still deter carelessness, albeit bluntly.
1980s malpractice-claim frequency increased at more than 10 per cent a year, and claim severity (average payment per paid claim) increased at twice the rate of general inflation. Claim frequency is now about 13 claims per 100 physicians per year, down from a peak of 16 in 1986. This unexpected surge in claim costs precipitated 'crises' in liability insurance markets in the mid-1970s and mid-1980s, which in turn led many states to adopt tort reforms designed to reduce claim costs, including caps on awards, collateral source offset, and shorter statutes of limitations. Some of these reforms—in particular, caps on awards—have slowed the rate of growth in costs (Danzon, 1984b; 1986; Zuckerman et al., 1990). Others appear to have negligible or unexpected effects. For example, if arbitration reduces the costs of dispute resolution it may increase the claim filings and number of patients compensated; however, this increases budget costs.

During the 1980s the rate of increase in number of claims and size of payments was at least as rapid in the UK and Canada as in the US. But in 1987 physicians in the US were still five to six times more likely to be sued than physicians in Canada and the UK, and awards for comparable injuries were several times larger in the US (Danzon, 1990). However, this overstates the difference in real compensation to victims, because the attorney's contingent fee (typically one-third of the award) is subtracted in the US and because medical costs are shifted to public health-care systems in the UK and Canada. The increase in malpractice premiums outpaced the increase in claims costs in the UK and Canada, particularly for surgeons, as the medical defence unions introduced rate differentials across specialties and attempted to shift from pay-as-you-go to partial funding of incurred liabilities. The squeeze of sharply rising premiums but constrained reimbursement under public health systems generated intense pressure for reforms.

(ii) Mismatch between Claims and Injuries

The high cost of malpractice claims is not by itself evidence of system malfunction. The number of malpractice claims in the US appears to fall far short of the number of negligent injuries; for other countries there is no evidence, but the shortfall is likely to be even larger. Two detailed studies of hospital records in California (Mills et al., 1977) and New York (Weiler et al., 1993) have concluded that the incidence of negligently caused injury is just under one per 100 hospital admissions. However, both studies used a broad definition of injury and make no attempt to define negligence by weighing marginal costs and benefits of additional precautions. The New York study defined as an iatrogenic injury 'any disability caused by medical management that prolonged the hospital stay by at least one day or persisted beyond the patient's release from hospital'. This begs the question of the appropriate level of care, the implied duration of stay, and expected outcome. Given these broad definitions, it is perhaps not surprising that almost 60 per cent of the injuries were minor, or that the number of claims filed was less than one-tenth of the number of negligently caused injuries as defined by the study.

However, this apparent mismatch between claims and injuries does not necessarily imply that deterrence incentives and compensation are too low. Compensating small claims through the tort system is probably not cost-effective, given other lower-cost compensation mechanisms. Overall incentives for care depend on the pecuniary and non-pecuniary penalties from the claims that are filed, on providers' risk aversion, and on other quality-assurance mechanisms. The ratio of claims to negligent injuries was much higher for serious injuries: roughly one claim is filed for every three such injuries and one in six is paid (Weiler et al., 1993). Given an iatrogenic injury, the probability of suit is substantially greater if there is a valid basis for a claim. Although the Harvard study concluded that many of the claims filed lacked evidence of a medically caused injury, this could reflect the limited information available to the reviewers. In other studies using more complete information, independent reviewers have concluded that negligence was certainly present in roughly 31 percent of cases and not present in 44 percent, with the remainder uncertain.

4 More recent estimates for the UK show the rate of new claims per 100 hospital doctors at 10.5 (Fenn, 1993b). This is not directly comparable to the US rate of 13 claims per 100 doctors, since the US figure applies to all medical specialties, including primary care doctors who are much less likely to be sued than surgical specialists.

7 With costly litigation, it may be optimal to have a low probability of suit but high penalties. The optimal tort award for deterrence purposes is lower if market forces or other quality-assurance mechanisms are partially effective (see Spence, 1977).
(Farber and White, 1991). For claims with negligence, the probability of payment was 0.64, and the average payment was $258,000; for claims without negligence the probability of payment was 0.24 and the average payment was $65,900. This suggests that the most extreme criticisms of the tort system as a random lottery are exaggerated.

(iii) Unequal Compensation

Another common criticism is that tort awards often provide very unequal compensation for similar injuries. However, although equal compensation for similar injuries might be appropriate if compensation were the sole purpose of the tort system, deterrence may require unequal payment for similar injuries. Theory and empirical evidence suggest that the settlement process adjusts payments for the degree of negligence, even though comparative fault is not the typical rule, and this is consistent with efficient deterrence (and with some definitions of fairness). Several studies confirm that the disposition of claims conforms to some degree to legal rules (Danzon and Lillard, 1983; Farber and White, 1991). Nevertheless, considerable unpredictability remains and this undermines deterrence, creates incentives for defensive medicine, and contributes to volatility in liability-insurance markets.

(iv) Inappropriate Compensation

Another valid concern is that awards for pain and suffering, which account for a large and probably growing fraction of malpractice pay-out, may exceed levels necessary for optimal compensation.\(^8\) Theory cannot determine optimal compensation for irreplaceable loss, but there is a strong presumption that payments for pain and suffering are too high in the US (Cook and Graham, 1977; Danzon, 1984b). The unpredictability of these awards undermines their value for patient insurance and for deterrence; it also contributes to volatility in liability-insurance markets. Scheduled limits on awards for non-economic loss, related to the plaintiff’s age and injury severity, are used explicitly in countries such as Sweden. This type of reform is likely to reduce litigation, by reducing uncertainty and the parties’ influence over the outcome, with little if any loss in efficiency of deterrence and compensation.

(v) High Overhead Costs

A final area of concern is the high cost of litigation and implied high overhead rate on patient compensation, relative to other compensation mechanisms. The fact-finding undertaken in liability systems is worth incurring only if there are offsetting deterrence benefits. While this is unproven it is plausible (see Danzon, 1985) and there is some empirical evidence of significant deterrence effects. Weiler et al. (1993) find evidence that the proportion of injuries attributable to negligence was lower in hospitals facing a higher probability of being sued, given a negligent injury. Extrapolating, they conclude that the proportion of negligent injuries per hospital admission would be 80 per cent higher if tort liability were eliminated. This is probably a lower bound on the deterrent effect of tort liability, assuming that the average deterrent effect exceeds that marginal effect observed from cross-sectional analysis. Moreover, elimination of tort liability would probably result in some reduction in other quality assurance and professional monitoring systems, that have been strengthened in response to liability. Thus, in practice, liability and other quality-control efforts may be complements, not substitutes.

IV. TRADITIONAL TORT REFORMS

All of the perceived defects of the status quo—imperfect deterrence, imperfect experience-rating of liability insurance, imperfect compensation, and high litigation costs—are ultimately attributable to imperfect information on the part of courts, plaintiffs, and providers. Although imperfect information under caveat emptor is a rationale for provider liability, changing the liability rule does not of itself create better information. The practical choice is thus between imperfect alternatives. In evaluating proposed reforms and alternatives, the practical question is whether they are likely to improve

---

\(^8\) Five per cent of claims account for 50 per cent of dollars paid in compensation (Danzon and Lillard, 1983).

\(^9\) The Harvard study (Weiler et al., 1993) did not find statistically significant evidence that a higher risk of suit reduces the absolute number of negligent injuries. However, there are statistical reasons why such an effect may be hard to detect.
efficiency in deterrence, compensation, and administration, or at least improve one dimension without loss along others.

(i) Traditional Tort Reforms

Some modifications of traditional rules would plausibly offer net benefits, at least in the US (Danzon, 1985). In particular, annuitized payments (but with amounts fixed at claim disposition) and scheduled limits on awards for non-pecuniary loss, based on injury severity and the plaintiff's life expectancy, are consistent with optimal insurance and would reduce litigation with minimal if any loss in deterrence. Collateral source offset significantly reduces cost internalization and is therefore, in theory, less desirable than subrogation; however, since subrogation may entail higher transactions costs, determining the optimal mechanism for eliminating double compensation is an empirical question, the answer to which may differ across countries. Proposals for screening and mediation panels, intended to streamline claim disposition, may simply increase delay and costs, unless significant penalties are imposed for appeal from their decisions to the courts. Adopting the English rule for allocation of court costs could reduce frivolous suits. However, in order to protect risk-averse plaintiffs, defence costs should be applied against the plaintiff's attorney, if paid on a contingent basis, rather than against the individual plaintiff.

These reforms are less relevant to the UK and other European countries to the extent that damages are already constrained by schedules or implicit rules, payments from public health-care systems and other social insurance are netted out of the tort award (collateral source offset), rules of discovery and procedure are less prone to exploitation, and rules for cost allocation discourage frivolous suits.

V. NO-FAULT ALTERNATIVES

Even if the most extreme criticisms of the tort system are exaggerated, the question remains whether alternatives would be more cost-effective. The Swedish model has been adopted in Norway and Finland and has been suggested in the UK (see Fenn, 1993a) and Canada. The Swedish and New Zealand models are often cited as illustrating the potential savings from a no-fault, i.e. causation-only, test of compensability (for example, Weiler, 1991), by analogy with workers' compensation systems. However, analogies between these models and workers' compensation or proposals for strict enterprise liability are misplaced. A causation-only rule of liability is neither necessary nor sufficient for low litigation and overheads as a percentage of premiums in Sweden and New Zealand. Nevertheless, the experience of these two systems is instructive.11

(ii) The Swedish Patient Compensation Insurance

The Swedish Patient Compensation Insurance (PCI) was established in 1975 by voluntary contract between medical providers and a consortium of insurers to pre-empt the threat of statutory expansion of tort liability. Although patients retain the right to sue in tort under traditional negligence rules, tort claims have been extremely rare until recently. A key feature of the Swedish model is decoupling of compensation and deterrence. Patient compensation is provided by the PCI, while the discipline of medical providers is handled by the Medical Responsibility Board (MRB). There is no transmission of information between them, which is said to be necessary to elicit the doctors' cooperation with the PCI.

The superficial appeal of the Swedish model is its relatively low budget cost and administrative overhead rate, and its widespread acceptance by medical providers. After an intended initial increase in claims, claim frequency has stabilized at about 21 per 100 physicians per year, compared to 13–16 claims per 100 physicians in the US; roughly 40 percent of these claims receive compensation in both countries. But the PCI costs roughly $2.38 per capita, or 0.16 percent of health-care costs in Sweden, whereas medical malpractice insurance

10 Schwartz (1992) describes differences in procedural rules between the US, Japan, and several European countries.

11 For a more detailed description and evaluation of the Swedish and New Zealand systems, see Danzon (1993, 1994).

12 Proposals for statutory expansion of liability grew out of concern that very few patients (roughly ten a year) received compensation under traditional tort liability.
premiums are about 1 per cent of (higher) health expenditures in the US—thus more than a tenfold difference. Administrative overheads are 14–18 per cent of total PCI premiums, compared to roughly 60 per cent in the US. This low overhead rate is often cited as evidence of the potential savings from switching from a negligence rule to a no-fault (causation-only) rule of compensability for medical injuries (Weiler, 1991), analogous to the strict liability of employers for workplace injuries under workers’ compensation.

However, these inferences are based on a misunderstanding of the PCI. The low budget cost of the PCI, despite the higher claim frequency, reflects primarily two factors. First, the collateral offset rule shifts most of the wage loss and medical expense to other social insurance programmes, thereby undermining cost internalization and general deterrence. This is cost-shifting, not real cost reduction. Payments through the PCI vastly underestimate the true cost of compensating iatrogenic injuries in Sweden.

Second, awards for non-economic loss are below those in most other European countries and roughly one tenth of those in the US. Underlying this difference is the much less pro-plaintiff tort regime in Sweden. Since the PCI is a voluntary alternative, it must offer plaintiffs an expected pay-off, net of costs, that at least matches their expected tort recovery, in order to deflect tort claims. Thus other countries that have more generous tort systems could not adopt the Swedish model or other voluntary contractual alternatives and expect to realize comparably low expenditures. Indeed, out-of-court settlements (which are one contractual alternative) already offer some of the gains that might be expected under a voluntary Swedish model.

The PCI’s low overhead percentage is not the result of using a causation-only test for compensability. Although the PCI is often called no-fault, this is misleading. From the patient’s perspective, the criteria of compensability are quite similar to a custom-based negligence rule. An injury is compensable if (i) it occurred with ‘substantial probability’ as a direct consequence of medical intervention, and (ii) either the treatment was not medically justified or the injury could have been avoided by performing the treatment differently. Normal and even most abnormal risks of standard medical care are explicitly excluded. But from the provider’s perspective, the PCI is truly no-fault and no-liability. The PCI eliminates all reference or inquiry into fault, requires no proof of negligence by an individual provider, and entails neither financial nor reputational consequences for the provider. This ‘no-fault’ scheme bears no resemblance to strict liability, either in theory or as applied to workers’ compensation and product liability.

The low expenditure on litigation reflects the fact that neither party has strong incentives to oppose or appeal the insurer’s decision. Physicians have no personal stake in the outcome, so generally cooperate rather than oppose compensation. Patients face low expected net benefits from appealing to the review panel or to arbitration, and are probably uninformed about the appeal process, which is closed to the press and public and has ruled in favour of the insurers in 90 per cent of cases. Thus the primary factors contributing to the low overhead percentage are the elimination of all links between compensation and deterrence, and the modest level of patient rights, compared to a US tort plaintiff (although not necessarily compared to a tort plaintiff in Sweden). Other contributing factors are the simple claim-filing process; administration

---

13 ‘General deterrence’ refers to the internalization of injury costs to the responsible activity or industry; it operates via effects of prices and demand elasticity. ‘Specific deterrence’ refers to internalization to the individual responsible; it operates by changing individual incentives for prevention.
14 The mean payment for non-economic loss was $3,800 in 1987, with a maximum of $117,070. Nevertheless, payments for non-economic loss account for roughly 74 per cent of total PCI payments, because economic loss is covered through collateral sources.
15 Among other obstacles, plaintiffs allegedly have difficulty obtaining the expert testimony required under the custom-based negligence rule.
16 Of course if the Swedish model were adopted as the mandatory alternative, eliminating the right to sue in tort, then the need to match tort compensation would not be a binding constraint and benefit levels could be set at any level.
17 This is discussed in detail in Danzon (1994, forthcoming).
18 Since 1992, major panel decisions and all arbitration decisions are published.
by a monopoly consortium of insurers, which eliminates insurers' incentives to compete by vigorously opposing claims or experience-rating premiums;\(^{19}\) and lack of competition and provider-specific accountability for costs in the health-care system, which makes providers more willing to tolerate flat-rated premiums,\(^{20}\) despite significant geographic differences in claims experience. It remains to be seen whether the PCI will survive the 1992 reforms of the Swedish health-care system, which have introduced more competition and provider accountability.

Although the PCI database on iatrogenic injuries could, in principle, be used for risk-prevention purposes, in practice the information collected is insufficient. Moreover, although clinics and hospitals are informed about their claims experience, the responsible individuals and sometimes even the nature of the injury are not identified.

Patients can file a claim with the MRB if they feel that their treatment was negligent or contrary to the code of medical practice. They bear their own filing costs and receive no compensation. Providers may be sanctioned by a reprimand or warning, but this has no financial consequence and probably at most a minor reputation effect. There are roughly six MRB claims per 100 physicians per year, of which one in six receives some sanction. Thus the ratio of MRB sanctions to paid PCI claims is less than one in ten—a rough measure of the loss in potential deterrence that results from decoupling compensation from medical discipline.

The main lesson from the Swedish PCI experience is that a sufficient and, possibly, a necessary condition for low overhead costs and provider co-operation in patient compensation is to forgo all links between compensation and injury prevention. Whether or not the loss in deterrence outweighs the reduction in litigation costs is an empirical question, the answer to which may differ across countries, depending on their tort systems and on the costs and effectiveness of other systems of quality control. Although in principle tort liability and other systems of quality control should be substi-

\(^{19}\) This monopoly structure would probably have to be changed if Sweden joins the EC.

\(^{20}\) The PCI is financed by premiums paid by county councils, that are responsible for financing and provision of the public health-care system in Sweden, and by private physicians, dentists, and other paraprofessionals. Premiums are assessed on a flat per capita basis, regardless of claims experience.
pensable. This is similar to the criterion of ‘unintended and unexpected’ adverse consequences proposed by Weiler (1991), which was rejected by the founders of the PCI as unworkable (Oldertz, 1988).

Defining compensability in terms of an event that is unexpected or of unexpected severity suppresses but does not eliminate the need to determine whether the care was appropriate. If ‘expected’ is defined as a statistical probability, this depends on the level of care delivered, relative to the condition of the particular patient. If a subjective measure of ‘expected’ is used, this presupposes some notion of informed consent and is surely impossible to determine ex post. Rulings and commentators have sometimes used objective criteria, sometimes subjective.21

Difficulties in implementing this definition led to numerous proposals for change, including relying on ICD-9 (International Classification of Diseases) definitions of injuries. Others urged extending the system to all incapacity, arguing logically that a no-fault scheme cannot inequitably distinguish among victims with similar conditions, on grounds of the cause of their injuries. This objection applies to any system that provides compensation selectively to victims of medical injury, but with no deterrence rationale for the discrimination.

The 1992 reforms adopted a far more restrictive definition that goes a long way towards restoring a negligence standard of compensability. ‘Medical misadventure’ is now defined as ‘personal injury resulting from medical error or medical mishap’. ‘Medical error’ is ‘the failure . . . to observe a standard of care and skill reasonably to be expected in the circumstances’. ‘Medical mishap’ is determined largely on the basis of ‘rarity and severity’ of the outcome, specifically, less than a 1 per cent probability of occurring, provided that the injury severity exceeds a threshold. This category specifically excludes abnormal reactions and complications of procedures, and injuries related to lack of informed consent, misdiagnosis, or treatment omissions, unless resulting from negligence.22

The 1992 reforms eliminated the shifting of costs to the public health-care system and internalized to the ACC all medical costs incurred by beneficiaries. Previously, medical expenses incurred by ACC beneficiaries were largely borne by the public health system, except that the ACC paid directly for services in private hospitals, co-payments, and services not covered by the public system. This cost-shifting undermined the ACC’s incentives to monitor claim duration.

Overhead costs are less than 10 per cent of total expenditures and payment is prompt. However, far from indicating efficiency, this simply reflects the ACC’s practice of accepting over 80 per cent of claims as filed, relying largely on physicians as gatekeepers to certify that a claim is a ‘personal injury by accident’ and, in cases of permanent disability, that continued benefits are necessary. But physicians have no incentive to oppose claims; indeed, until recently physicians could benefit from certifying a claim, since the ACC paid higher fees than did the NHS and ACC compensation acted as a bar to a tort suit. Thus this mechanism of claims adjudication may have saved overhead costs but has contributed to the rapid escalation of total claims costs. The ACC databases did not identify iatrogenic injuries. Thus premiums could not be levied on medical providers and the frequency and causes of medical injuries could not be monitored for risk-management purposes. Again, such economizing on overheads may be ‘penny wise but pound foolish’, skimping on budget costs but with higher real social costs.

Prior to 1992 the costs of iatrogenic injuries were hidden in the payroll and general taxes that financed all injury compensation to workers and non-workers, respectively. The 1992 reforms authorize the

---

21 Venell (1992, p. 4) notes that in one case appealed to the High Court, ‘Bisson J. appeared to move away from the previous objective approach, that if the risk was one that was known to the medical profession then it was not medical misadventure. He adopted a subjective approach which involved looking at things from the point of view of the victim (and her medical advisers).’ Duncan (1984) refers to an injury ‘which is unexpected and undesigned by the person injured’.

22 ‘It has been apparent that difficult questions of causation manifest themselves when the task is to establish a causal link rather than, as in negligence, having to prove that a potential tortfeasor has failed to attain an appropriate standard of care, to which the subsequent damage was causally linked’ (Venell, 1992). The suggestion is thus that causation is more, not less difficult to establish, once the element of negligence is removed.
ACC to establish a medical misadventure account, funded by premiums paid by registered health-care professionals, with experience-rating and no-claims bonuses. This is consistent with the shift towards greater autonomy and accountability of health-care providers as part of the reform of the health-care sector. Whether the ACC will in fact exercise its authority to assess medical providers remains to be seen. In the past it has compressed employer premiums rather than exploit its full statutory authority to use experience-rating. As in Sweden, lack of competition—the ACC is a public monopoly insurer—is a necessary condition for the survival of flat-rated premiums. The 1992 reforms also authorized the ACC to report potentially negligent medical misadventures to the appropriate disciplinary body. However, since reporting is apparently discretionary, the effect of this clause also remains to be seen.

The New Zealand experience under the original ACC structure illustrates pitfalls to be avoided rather than providing a useful prototype that other countries might copy. The original definition of a compensable event raised practical and philosophical issues that led almost inevitably to proposals to expand the system to cover incapacity. However, the huge budget costs of such a system and the difficulty of defining incapacity led to the restoration of a quasi-negligence criterion of patient compensation, but without provider-specific liability, that resolves in some ways the PCI criterion.

As in the Swedish PCI, the low administrative costs should not be interpreted as a measure of efficiency. Rather, low overheads reflect the elimination of all links between compensation and deterrence. The causes of medical injuries are not investigated and there is no feedback to the individual providers that are responsible for the injuries. The proposed introduction of experience-rated premiums for physicians is likely to raise providers’ opposition to patient compensation and hence raise litigation and overhead costs. As in Sweden, the elimination of all provider liability, explicit or implicit, is crucial to the non-adversarial adjudication of claims. In addition, in New Zealand the very low overhead percentage reflects the rapid increase in claims payments (the denominator) owing to minimal claims investigation. The true overheads of an insurance or accident compensation scheme include not only the measured overheads, but also the deadweight loss from unnecessary injuries and inappropriately compensated claims (Danzon, 1992). Unfortunately this is not observable, but in the ACC it is likely to be very high.

VI. OTHER PROPOSED ALTERNATIVES TO TORT

Reform of several key dimensions of the malpractice system have been proposed: the rule defining a compensable injury (causation only, low probability/high severity, medical error); provider responsibility (individual, enterprise, none); measure of compensable damages (schedules, deductibles, collateral source offset); forum of adjudication (administrative agency rather than traditional tort trial); and financing (premiums paid by individual providers or health-care organization, broader taxes). Some of these changes could be applied to the traditional tort system, while retaining its other features—for example, the NHS indemnity scheme replaces liability of individual doctors with a form of enterprise liability, but otherwise retains traditional tort rules; some, such as scheduled damages for pain and suffering and collateral source offset, are already in place in several countries. In evaluating comprehensive changes, such as the NHS indemnity scheme or proposed alternatives, it is important to identify the problems that the scheme is intended to address and evaluate the scheme in the context of its overall effect on the social costs of injuries, including deterrence and true overhead costs, not merely the more visible budget costs. Here I comment briefly on two comprehensive US proposals and on the NHS indemnity scheme in the UK.

(i) The AMA Administrative Fault-based System

The AMA (1988) has proposed an administrative fault-based system (AFBS) that would remove medical malpractice claims from the courts to a new specialized administrative ‘Medical Malpractice Review Board’ in each state. The aim is to extend

33 Like the NHS reforms, the goal of these reforms was to separate financing and provision of hospital care. Hospitals have been reorganized as 'Crown health enterprises', with autonomous boards, and must compete to provide services to regional health authorities.
compensation to more patients; streamline administration, including screening out invalid claims more promptly; strengthen deterrence and spread the costs of compensation more broadly.

The AFBS retains a fault-based rule of liability, modified to define the standard of care in terms of a `range of reasonableness', based on the standards of a prudent and competent practitioner in the same or similar circumstances. The traditional contributory negligence rule is replaced by comparative negligence. Whereas, traditionally, compensation is in principle denied if the relative contributions of the patient's underlying condition and the provider's actions were 55 and 45 per cent, respectively, under the proposed standard the provider would pay 45 per cent of the damages. Apportioning damages would extend lower compensation to more patients, and could be considered more equitable to patients and providers; it is also consistent with efficient deterrence of injuries.\(^\text{24}\) Damage rules are reformed to include an indexed schedule of payments for non-economic loss, depending on the patient's life expectancy, and collateral source offset. Future damages would be annuitized, but the amount is determined at time of claim disposition to maximize incentives for rehabilitation.

The administrative procedures for claims adjudication are intended to be less costly, permitting greater access to those with valid claims, while screening out non-meritorious claims. Nevertheless, the process remains adversarial, with attorney representation, as a necessary condition of protecting rights of both patients and defendants. Patients whose claims are deemed valid on initial screening are offered free attorney assistance. In order to encourage settlement, both the plaintiff and defendant(s) would be required to make blind settlement offers prior to the hearing, and would be subject to sanctions if they rejected an offer that is not significantly bettered at the hearing. The hearing resembles a traditional trial, except that it is adjudicated by an examiner experienced in medical malpractice claims, rather than judge or jury; appeal is to the appellate courts, but on rules of law only. Thus the court cannot review the facts or the finding of liability in a particular case or set medical standards; the Board has ultimate authority over these functions.

To strengthen professional discipline, the Board would operate a clearing-house for reports from several sources, including settlements and awards in malpractice cases, hospital reviews, reports from other physicians (who are required to report suspected incompetence, impairment, and drug or alcohol dependence of their colleagues), and other state disciplinary actions. The Board can also investigate reports of substandard performance from several sources, including members of the public and, following a full due process proceeding, may impose sanctions including fines and licence revocation.

The proposal does not specify in detail how the system would be financed. The optimal system of financing would depend on the extent to which the system implicitly includes pure social insurance components, for example, as a result of offering free legal aid to patients whose claims pass the initial screen. Although the incremental deterrence value of provider-specific premiums might be small, given the direct feedback from claims adjudication to the disciplinary process, the incremental cost may also be small, in terms of increased incentives for providers to oppose claims.

This administrative approach resembles in some ways the Swedish PCI, but with important differences. The AMA proposal retains and, in some respects, reinforces links between compensation and deterrence. Unlike the PCI, the AMA retains the notion of individual provider fault. Both limit compensation to injuries caused by medical error (`avoidable' injury in the case of the PCI), defined relative to customary medical practice. Both use written clarification of the criteria of compensability, although the PCI rules are much more detailed. By adopting a comparative negligence standard, the AMA adopts a more expansive definition of causation and hence expands the number of potentially compensable injuries.

Because the AMA retains provider liability, provider incentives to oppose claims remain. In part

\(^{24}\) Haddock and Curran (1985) show that a comparative negligence standard is potentially efficient if applied conditionally on a violation of the standard of care. This is implicit in the AMA proposal which requires fault in addition to the comparative measure of causation.
this reflects the intent to use the adversarial process to eliminate frivolous claims. The AMA offers free legal representation to claims that pass an initial screen, and permits patients to represent themselves at that screening, whereas the patient's submission to the PCI (and subsequent appeals) must be in writing only, unless special permission is received for oral representation. The AMA process is public and would certainly be subject to continued public scrutiny for fairness to patients, whereas the PCI faces no public scrutiny. For all these reasons, the AMA approach is unlikely to yield low per capita budget cost and overhead cost comparable to the PCI. However, it offers greater deterrence and more generous rights to plaintiffs, partly reflecting the political requirement that any serious reform proposal must offer gains to both sides, relative to the status quo in that country.

The PCI experience suggests that this administrative alternative would be held to some accountability because of the implicit or explicit threat that, if it operated unfairly, the tort system would be re-established. The Swedish experience indicates that, if providers prefer the administrative alternative, they will design it such that patients are at least as well off under the administrative alternative as they would be under the tort system. In that case the administrative alternative is clearly a Pareto-improvement: it survives only as long as both sides are better off.

(ii) Elective Strict (No-fault) Liability (ESL)

This proposal, modelled on the workers' compensation system (Weiler, 1991; Weiler et al., 1993), would empower hospitals and other health-care organizations to offer an administrative alternative, in return for a waiver from common law tort liability. The proposal is for an elective system initially, in order to gain experience before moving to wider implementation.

Victims of iatrogenic injury would be compensated without regard to provider negligence or fault. The intent is a system of strict enterprise liability, analogous to the liability of employers for workplace injuries. However, medical causation is a necessary but not sufficient condition. Excluded are injuries resulting from the normal risks of medical care and imperfect cures of the underlying condition. Thus in contrast to some earlier no-fault models, ESL recognizes that attributing an adverse outcome to medical care is complex. Because the patient enters the health-care system in less than perfect health, the issue is to identify the incremental harm caused by the medical system, recognizing that some adverse outcomes reflect the inevitable natural course of the underlying disease, and that appropriate medical care entails some positive risks of traumatic injury, in order to reduce the risk of deterioration from the underlying condition.

A simple causation-only strict liability system would create incentives for providers to avoid risky treatments, even though these treatments might on balance be optimal, because providers would be penalized for bad outcomes but not rewarded for good outcomes. To counteract this, the fault principle must be retained for errors of omission.

Weiler argues that the causal inquiry is far less difficult than is the additional fault judgement, citing evidence from the Harvard study and from Sweden and New Zealand: 'in Sweden and New Zealand, the two countries that have provided no-fault compensation for medical injuries ... it has been possible to draw a causal dividing line without any pronounced administrative burden for the no-fault programs as a whole.' As argued above, the lower administrative cost percentage in Sweden and New Zealand results from decoupling of deterrence from compensation, not from use of a causation-only standard of compensation (see also Danzon, 1993, 1994). Since ESL retains full cost internalization, litigation and administrative costs would certainly be higher than under the PCI or ACC.

Any savings that ESL realizes in administrative costs would result primarily from the use of an administrative disposition process and scheduled damage payments. These changes can be made without switching from fault to a rule of strict liability, as illustrated by the AMA proposal. Similarly, if channelling liability from the individual
doctor to the health-care enterprise offers real efficiency savings, this could be done within the context of a fault-based rule, and could be done by voluntary contract. Consistent with this, the Kaiser chain of health maintenance organizations (HMOs) assumes liability for all practitioners within its organization. Not surprisingly, such contractual election of enterprise liability is confined to fully integrated, exclusive HMOs, whose providers treat only HMO patients and whose patients receive all their treatment from HMO providers. In looser, non-exclusive networks, which are far more common because they offer greater choice and flexibility to providers and patients, enterprise liability could add administrative costs and reduce accountability, contrary to the intent of the proposal.

(iii) The NHS Indemnity

In 1990 the NHS introduced an indemnity scheme for its employees, whereby the Department of Health assumes liability for all treatments provided by its employees.26 GPs’ medical malpractice subscriptions (premiums) have always been reimbursed as a practice expense, but hospital doctors paid their own subscriptions. This system broke down in the 1980s with the rising cost of claims. The Medical Defence associations were forced not only to raise subscriptions across the board but also to introduce specialty differentials, in order to meet the threat of competitive entry by commercial insurers, and this exacerbated the cost increase for the high-risk specialties. Canadian specialists similarly experienced a treble shock in the 1980s: an overall rate increase owing to rising claims costs was exacerbated by a move to specialty-specific rates and pre-funding rather than pay-as-you-go financing, precipitated by the threat of competitive entry into the liability insurance market. This contrasts with the Swedish experience, where flat rating and partially pay-as-you-go financing have persisted because the PCI is operated by a monopoly insurer consortium. As noted earlier, Swedish entry into the EC may disrupt this tranquil monopoly. Although switching to specialty rating—a rough proxy for claims experience—is temporarily disruptive, in the long run it is generally consistent with efficient internalization of costs to activities that generate high injury costs. However, this presumes that the claims process is accurate or at least unbiased. If in practice the higher claims rate for surgical specialties reflects the fact that surgical errors are more obvious than medical errors, some inter-specialty cross-subsidies may be justified.

Although the NHS indemnity may be viewed as a form of fault-based enterprise liability, efficiency gains are likely to be realized only as liability is transferred to self-governing trust hospitals and fundholding GP practices. Enterprise liability is intended to increase deterrence and reduce overhead costs by placing liability on the single party that has the information and the authority necessary to make decisions with respect to risk management. Employer liability for workplace injuries fits this model, as do fully integrated HMOs that are exclusive for providers and patients. If the NHS internal market is effective, the role of district health authorities is as purchasers on behalf of patients. Hospitals and, in particular, the self-governing trust hospitals, are intended to be autonomous entities that assume responsibility for the cost and quality of the services that they deliver. Thus placing liability on district health authorities is at odds with the separation of purchasers and providers, whereas transferring liability to self-governing hospitals and fundholding GP practices could realize efficiency gains. Similarly, proposals in the US to transfer liability from individual physicians and hospitals to health plans, would probably reduce deterrence and increase administrative cost if applied to traditional fee-for-service plans and loose networks HMOs (independent practice associations). Enterprise liability is likely to be efficient only for fully integrated health-care systems such as staff model HMOs, where the health plan has the information and authority to manage the delivery of care.

VII. CONCLUDING COMMENTS

There is no simple solution to the problems of professional liability. The rationale for professional liability arises from asymmetric information between patients and providers. But changing the liability rule transfers decision-making to courts and liability insurers which also lack good information. Unpredictable and sometimes erroneous deci-

---

26 This is discussed in detail in Fenn and Dingwall (1990).
sion-making by these parties creates incentives for filing invalid claims, defensive medicine, and investments in litigation to influence the outcome. The ideal reforms would improve the accuracy of the decision-making process, structure benefits according to sound insurance principles, and impose sanctions for abuse of the system. This in turn should assure efficient deterrence and compensation.

My personal judgement is that an administrative fault-based system, with scheduled payments for non-economic loss, written clarification of the rules for determining economic loss (e.g. inflation and discounting), and written criteria of compensability, is the most promising alternative. This could be combined with an elective enterprise liability option and, for the US, the English rule for allocating costs applied to the plaintiff’s attorney if cases are taken on a contingent-fee basis. This maintains cost internalization to parties responsible for injuries and hence preserves deterrence incentives, while reducing some of the uncertainties of the traditional common-law rules that encourage wasteful litigation. No-fault schemes such as the Swedish model, that reduce litigation expense by eliminating all attempt at deterrence, reduce to systems of social insurance that single out victims of medical injuries for special compensation. Such reforms have little basis in equity. They may shift costs from healthcare budgets but ultimately may lead to higher real social costs of iatrogenic injuries.

REFERENCES

—— (1990), 'The 'Crisis' in Medical Malpractice: A Comparison of Trends in the United States, Canada, the United Kingdom and Australia', Law, Medicine and Health Care, 18(1–2).
Duncan, V. M. (1984), 'The New Zealand Experience and Possibilities for the Future', in J. L. R. Davis and D. F. Partlett (eds), Personal Compensation for Injury, Proceedings of a Seminar sponsored by the Faculty of Law, Australian National University, August.


Harvard Medical Practice Study (1990), Patients, Doctors, and Lawyers: Medical Injury, Malpractice Litigation, and Patient Compensation in New York, report of residents and fellows, Cambridge, MA, Harvard University Medical School.


