Medical Malpractice in Crisis

R. Lee Akazaki

This publication first appeared in the February 1999 volume of The Advocates’ Quarterly.

Introduction

In medical malpractice, perhaps more than in any other field of civil personal injury compensation, the blunt instrument of tort law is as well suited to the social needs of contemporary Canadian society as the feudal art of trial by combat was in settling civil disputes in earlier times. Plaintiffs have long complained that lawyers are reluctant to take on the cases because malpractice insurers "will not settle," and it is too expensive to try them. Defendant doctors, obstetricians, orthopedic surgeons, neurosurgeons, and cardiovascular surgeons in particular, view themselves as target defendants in a litigation process fuelled by the premiums paid to their mutual defence organization. They also harbour the not unfounded perception that the likelihood of being sued by their patients owes as much to chance as it does to their skill and conduct.

Most iatrogenic injuries and illnesses, by their nature, occur when something is already going wrong. One might exclude from this description cosmetic surgery and clinical trials but include obstetrics. Even the excluded fields share with the rest of medicine several or all of the following possible medical outcomes, no matter what the cause:

1. The treatment will succeed without unwanted adverse effects.
2. The treatment will fail.
3. The treatment will succeed but result in adverse effects.
4. Treatment or diagnosis missed or delayed will not have any consequences.
5. Treatment or diagnosis missed or delayed will cause condition to worsen or result in death.

If, as perceived by some doctors, any adverse outcome can lead to an allegation of negligence and a claim for damages, it seems inevitable that a legal system which permits such suits to be brought will also be flooded by them, and by their attendant costs, including the cost to an overburdened court system. From the plaintiff patient class' perspective, two opposing forces in the tort system come to bear on the likelihood of becoming a victim of a doctor's mistake: moral hazard and the pressure of mounting premiums. Paradoxically, they are capable of producing the same type of result. Moral hazard is the human tendency to be less careful in the presence of the comfort of insurance. The pressure of mounting premiums completes the vicious cycle by requiring doctors to see more patients in the same amount of time, in a fee-for-service regime. Whatever the allocation of blame in this equation, one side, on behalf of physicians, has thrown up the white flag and has called for reform of the tort regime. But what does this mean? Historically, the law of torts as it applies to medical cases has favoured doctors, in comparison with other target defendants.

The CMPA’s Call for Tort Reform

The logical point of departure for any analysis of the crisis is the Canadian Medical Protective Association (CMPA), which provides legal protection to more than 90% of Canadian physicians. It is not an insurer but a "medical mutual defence organization." The CMPA is the favourite whipping post of many, both within and without. The size of the CMPA’s "fully-funded" reserve has drawn considerable attention. In an independent study, former Ontario Chief Justice Charles Dubin concluded that the CMPA’s 1996 reserve of $1.29 billion was justified because that was the size of the case load, in terms of claims and legal costs. He did criticize the current method used to determine fees, adopted in 1984, which charges physicians according to the risk ratings of their specialties. Thus, it was reported, in 1996 a family physician paid $1,932 for CMPA coverage, while an obstetrician paid $23,340. The report advocated a return to a flat premium for all doctors, even though such a fee structure may be both unfair and actuarially unsound. There is an overriding danger of driving doctors away from high-risk specialties such as obstetrics and orthopedics. It also concluded that continued state subsidy of the reserve was warranted.

This independent study of the CMPA must be accepted in context: the very existence of the organization presupposes the continuation of a tort-based medical malpractice regime. The admitted compromises of insurance principles and fairness
among members of the profession emerge from the friction between doctors' individual economic interests and the collective needs of public health. The challenge of public medicine in Canada is to guard universal access while limiting the disincentives which drive medical school graduates from high-risk fields and tempt practitioners south. Reverting the fee structure to the pre-1984 scheme should, in large part, be considered a tax on the majority of members to ensure continued public access to certain minority specialties. But why should low-risk doctors bear the tax, as opposed to the general public?

Regardless of the criticisms of its size, the CMPA reserve is most useful, from a macroeconomic perspective. It is a measure, calculated by a legion of actuaries, of the cost of medical misadventure in a tort-law system. Apart from the CMPA's long overdue initiatives to exploit some untapped efficiencies of scale, such as the use of bargaining power to reduce legal costs, reforming the organization cannot, as a matter of first principles, alter either the nature or magnitude of the problem. It cannot control the number of law suits, the standard of care, or the influence of a law-obsessed America on the litigiousness of Canadians.

At its annual meeting in August, 1997, the organization unanimously adopted a resolution calling for it and the federal and provincial medical associations "to push for tort reform." One of the advantages of a non-commercial defence organization is that it is not stopped by private institutional concerns from plotting its own demise. Despite some significant developments in the law of medical informed consent, this very important area of compensation law has undergone no significant changes since the leading cases were decided in the 1950s. Recently, the postponement of the delivery of defence expert opinions until the eve of trial has been blamed by the judiciary for add to the delay and expense in the process. This has led to changes in the Ontario Rules of Civil Procedure to require all expert opinions to be delivered at least 90 days in advance of trial. "Trial by ambush" as a defence tactic has resulted in more cases proceeding to the trial-preparation stage of litigation than if the opinions were disclosed earlier. Defence counsel, on the other hand, have for years complained about the commencement of actions without the benefit of a medical opinion corroborating the allegations of negligence. The current state of medical malpractice litigation remains highly litigious.

The leading case in the Canadian law of medical negligence was actually reported a year before the leading English case. For this reason, Canadian law has charted a different course than in England, although the distinctions may be more philosophical than real. Both of these jurisdictions were at least a half century behind the development of analogous rules in the United States, where the law appears to have developed from the turn of the century in direct response to the
development of modern medicine. Before legislators heed the CMPA’s call for reform, it may be prudent for them to examine more closely the sources of discontent with the tort system for medical malpractice.

The Tort Model

The underlying corrective justice of tort law, attributed by Weinrib and others to Immanuel Kant, is that of individual responsibility. Invoking Kant’s "categorical imperative," tort theorists have isolated the source of this law in that balancing of "rightness" and "necessity" in the rules which govern our everyday conduct. Since there exists no net social benefit to shifting the effect of the loss from the victim to the perpetrator, the compelling reason for the law to step in and do so had to be moral. Medical malpractice, although not unique in this regard, is nevertheless a conspicuous paradigm where the tort law model exists as a pure contest between individual rights. The remedy is a straight-up turning of tables: to the extent that an award of damages may replace the use of a vital organ or function, an eye for an eye, &c. From the doctor’s perspective, a law suit is the worst nightmare come true. Unlike the accepted practice for other personal injury and casualty claims, there is no conventional routine for negotiation at the adjuster level prior to the commencement of formal proceedings. If it is the practice of the applicable mutual defence association not to settle, at least until proceedings are instituted, plaintiffs and their lawyers will not waste resources on pointless negotiations or exchange of polite letters. Doctors are thus likely to receive their first notice of a claim in the form of a statement of claim, delivered in person at the hospital or office. This is probably not the preferred method of engagement from either the doctor or patient’s perspective.

The burden on Canadian plaintiffs for medical malpractice is not as high as it is in England. There, the Bolam test, providing that if medical professionals can show that they "acted in accordance with the practice accepted as proper by a reasonable body of medical men skilled in that particular art," has been criticized by a generation of jurists as permitting defendant doctors "to be judged on the least demanding professional standards prevailing." Over the years, this certainly has been the prevailing view. The House of Lords has recently restated the Bolam test thus in the context of psychologists:

"[T]hey are only bound to exercise the ordinary skill of a competent psychologist and if they can show that they acted in accordance with the accepted views of any one other reputable psychologist at the relevant time they will have discharged the duty of care, even if other psychologists would have adopted a different view."
In fact, the test has undergone further refinement in the case of narrowly defined subspecialties (or "superspecialties"). In Defreitas v. O'Brien et al., the English Court of Appeal upheld the ruling of the trial judge that orthopaedic and neurosurgeons engaged mainly or wholly in spinal surgery, could embark upon exploratory spinal operations which would be considered unreasonably risky for the ordinary specialist. It has been argued that this application of Bolam to subspecialties could prove to be a mixed blessing to potential defendant doctors because it could require someone who holds himself out as belonging to such an exclusive club to be judged by a higher standard of skill and achievement. It should be remembered that professional malpractice is not exclusively a tort domain, but rather one of concurrent tort, contract and quasi-contract principles. However, in practical terms it is difficult to envision English courts second-guessing a superspecialist for a procedure that only such a consultant could perform, once his qualifications have been established. In this sense, both in Canada and in England, the law of professional negligence generally accords protection from liability in direct proportion to the risk inherent to the procedure, and extends virtually to full immunity where the outcome is, more or less, out of the professional’s hands.

The Canadian test refers to "the judgment of the generality or average of the special group of technicians to which he belongs." This statement from the reasons of Rand J. in Wilson v. Swanson is often cited as the definitive statement of law, despite the fact that it was obiter. The case was actually decided on the issue that an exercise of judgment leading to a poor result is not actionable as negligence. The test is often mentioned in the same breath as the Ontario Court of Appeal's decision in Crits v. Sylvester, [1956] O.R. 132, in which the Court held that (a) liability is imposed for negligence and not misadventure, (b) res ipsa loquitur may be invoked to prove medical negligence, and (c) the "standard" practice of medical professionals is not the final word on the standard of care. The last element from Crits is consistent with Australian cases in which Bolam has been expressly rejected. It should be noted, however, that the refusal of the Australian and Canadian courts to hand over the domain of standard of care arises from cases where the custom, however widespread, defies common sense. In the context of informed consent, as seen in negligence as opposed to battery, ordinary negligence principles apply, and medical evidence, while relevant, is not determinative.

The difference between the Canadian and English views is that in Britain, the standard of care is legally the province of the medical profession to which the courts observe absolute curial deference. Here, the standard of care is for the court to decide. In Crits, however, the failure to ground sources of static electricity in an operating theatre where inflammable agents were used to anaesthetize the
patient was not representative of the type of problems which give rise to medical malpractice actions. Many problems which give rise to risk in modern medicine require more than common sense to understand. To the extent that the problems of human error have become more technical as medicine has become more dependent on machines and "superspecialists," Bolam is often de facto the test in most common-law jurisdictions.

The Bolam test must be seen for what it is. As a principle of tort law and corrective justice, it is a pure application of the rule that the standard of care may be determined by custom and the reasonable man, unless custom defies common sense. Furthermore, matters of professional judgment, competently made, cannot result in liability. If physicians are not the ones to determine what the standard of care is, then who is? Critics of Bolam contend that doctors watch after themselves or that the principle is a throwback to the English class system. On first glance, such views had some merit. On closer inspection, the arguments are thinly veiled ad hominem attacks on members of the profession based on social and economic observations and generalizations. In the tradition of the common law as an arbiter of individual rights and—perhaps more importantly—as an unseen instrument of "Natural Law," the root principle of Bolam is good law. Doctors, as opposed to judges and juries, should determine the standard of care. Ideally, however, the rule is substantive and not evidentiary. It should not be sufficient for the defence to produce one reputable expert witness to exonerate the defendant. Rather, it is the court’s task to determine which of the experts, for the plaintiff or for the defendant, has formed a more credible opinion of the standard of care based on the facts of the case. In this regard, the Canadian and Australian versions of the test are probably preferable to the English, except that the leading cases in this important area remain conspicuously enigmatic.

Legal Sources for the Inadequacy of the Tort Model

The nature of the legal test, however, imports inherent practical ramifications, two of which have a direct inflationary effect on the costs to doctors and patients: disproportionately low access to legal remedies and cases lacking in merit.

The historical complaints on behalf of plaintiffs of the barriers to access to judicial remedies need not be detailed here. The fact that each case requires locating at least one expert in the same medical field as the defendant to review the case and express an opinion supporting a finding of negligence immediately separates this from most personal injury actions in term of the up-front expenses. According to one English study, conducted by the Oxford Centre for Socio-Legal Studies, 45% of actions were abandoned due to evidentiary difficulties in proving fault. This
problem has a ripple effect of institutionalizing delay, which in turn inflates cost and denies justice to both sides. Transferring some of the burden to lawyers by permitting contingency fees has lowered this barrier in the United States, but making the lawyer the financial stakeholder has the secondary effects of limiting access to clients with strong or catastrophic cases and of inflating claims that are brought.

The siege mentality, alleged by some critics, must exist in some measure, especially among high-risk specialties, if only because of the economic pressure of mounting premiums. Such increases in premium levies are the costs associated not only with winning but also with keeping cases difficult to bring and maintain: the greater the circumference of the wall, the greater the cost of feeding the troops to police it. From the perspective of defence associations as consumers of legal services, their mandate is to protect doctors from damage awards and to guard their professional reputations. This is not a factor in insurance companies’ duty to defend and indemnify, in respect of standard commercial third-party liability policies. There is an inflationary factor associated with any product which overlays its utilitarian features with the emotional and ritual urgency of keeping up appearances. In this sense, the cost of defending a doctor is expensive in the way that weddings and funerals are expensive. It is not the proliferation of litigation but the non-utilitarian aspects of medical defence litigation which raise the marginal demand for services, compared to more commercial areas of tort and insurance litigation, such as casualty and automobile. An allegation of negligence against a factory owner or highway commuter does not generally carry the same capacity to polarize emotion as the same allegation against a doctor. As long as the system of compensation for iatrogenic injuries remains based on fault, the legal services associated with the regime will suffer from this distortion of the market forces in play.

Various failures, from the ambiguity of the judicial expressions of the Bolam test to the participation of inexperienced plaintiffs’ lawyers, have contributed to the proliferation of lawsuits lacking in merit. To the extent that other factors raise extraordinary barriers to success in the field, it may partly be a problem of perception. We must, as a matter of principle, differentiate frivolous or vexatious suits from those in which there is no substantive merit. If the vast majority of medical tort claims are not settled, if a small minority go to trial, and if an even smaller minority are successful, what is to be done with the frivolous claims? Any statistical analysis must be misleading because the short limitation period is probably responsible for the issuance of many claims against lawyers’ better judgment, in order to preserve the plaintiff’s right to sue. Proposals such as the requirement for written medical corroboration prior to the commencement of an
action are probably indefensible in law, because medical evidence is not a sine qua non of recovery in Canada. A notice provision which provides for the extension of the limitation period, as employed in various statutes such as those which govern proceedings against the Crown or against municipalities, could and perhaps should alleviate some of this pressure to commence formal proceedings.

Unlike commercial litigation, there is no principled reason for bringing medical claims for collateral purposes. Except in psychiatric cases, where remedies such as habeas corpus may have some application, the only purpose for taking a claim to court is to recover an award of damages. Thus, by definition, no case should be considered frivolous or vexatious in which a plaintiff has suffered an adverse outcome and can raise at least a triable issue. Actions perceived as lacking in merit will be instituted for the same reasons as in the past. Either the actions are inappropriate and brought by inexperienced plaintiffs’ lawyers, or defence counsel have been inadequately advised by association experts and the merit in claims are overlooked. (The inadequacy of the advice to defence counsel is not necessarily a reflection of competence—leading experts are usually consulted—but rather of misunderstood purpose.) The fact that plaintiffs’ lawyers who used to practice in automobile litigation have had their practices diminished due to no-fault insurance may have driven many into this field. Such factors, beyond the capacity of the system for slow improvement, are driven by factors of historical inevitability.

A significant factor which has been ignored is the "informational disadvantage" which is exploited as a defence tactic and which may encourage avoidable litigation either by forcing the parties into adversarial positions or by disrupting the medical treatment of the patient. This problem is probably not inextricable from the tort system, but it is difficult to see how it can be reformed in any meaningful way in the foreseeable future.

The Collateral Costs to the Health Care System of the Tort Model: Forms of Defensive Medicine

A recurrent complaint about the tort model for medical malpractice has been the effect of law suits on the conduct of doctors. The Bolam test, including its international variations, is said to encourage conservative and defensive medicine and discourage innovation for the benefit of individual patients. The two styles so fostered are, perhaps paradoxically, antonyms.

Conservative medicine, in this context, refers to reliance on a standard of care based on average skill and competence, as opposed to individual excellence. For tactical reasons, a defendant to a negligence suit should prefer to be judged by the
lowest standard the law will afford him. Whether this effect translates to the clinical setting is an imponderable. One report that the Medical Defence Union (MDU), a counterpart to the CMPA in the United Kingdom, has "through gritted teeth" welcomed a trend towards civil accountability, could be interpreted to mean that the proliferation of medical litigation has been good for its long-term risk management. Although it is not certain whether this means that doctors are more careful (a desired effect of tort law as corrective justice) or have lowered their collective standards (the contrary effect), one cannot imagine the MDU being pleased about the latter. Medicine is a conservative profession, and its history is littered with examples of discouraged innovators. The same history also lionizes them sufficiently to inspire them to persevere.

Defensive medicine, however, is the practice of exceeding the standard of care by ordering more tests or performing more examinations and procedures than are thought to be necessary. It has been observed that if this were a serious problem, one might see regional variations in practice based on rates of litigation. Excessive caution is probably is probably a hidden inefficiency in the diagnostic processes attributable to tests conducted by physicians who should, as they gain experience, require them less. In the individual case, there is no harm caused by the practice. At the macroeconomic level, however, it has often been cited as a major source of inefficiency in the public health system attributable to the deterrent effect of tort law. According to one cited AMA study, this phenomenon added 5% to the entire U.S. health care bill. The capacity of tort law to affect the conduct of doctors may thus extend beyond the moral purpose of corrective justice of making them more careful. Faced with scarcity, it is possible at least to imagine how a 5% waste of resources can lead to deviation from the standard of care by depriving resources from other areas.

Defensive medicine may be, as any problem in this field, impossible to survey with any degree of authority. The obstetrical cases provide proof that a version of defensive medicine certainly does exist in that it has become part of the practice. Testing is done at the behest of the treating obstetrician solely for a medical-legal purpose when the newborn appears to have a neurological deficit. As disclosed at the Second International Symposium on Perinatal Asphyxia, June 8, 1992, obstetricians, as a matter of course in difficult vaginal births, order testing of cord blood gas in order to obtain proof that perinatal asphyxia did not occur. Before ordering the test, however, they satisfy themselves that the gases will likely be in the normal range. The normal blood gas shields the obstetrician from a malpractice action in which it may be alleged that brain damage was caused by perinatal asphyxia due to inadequate response to deceleration or arrest in labour. The peculiar ex post facto nature of this testing conveniently illustrates defensive
medicine as a creature of the legal system, because the procedure is unlikely to add to the information already available for management of the newborn but is thought to be of assistance to the doctor against an anticipated lawsuit resulting from the poor outcome.

This practice can have a double-edged effect, however. In cases of cervical spinal cord injury, proof of normal cord gas so obtained can take away from a doctor’s defence that cord injury was due to an unavoidable hypoxic-ischaemic pathology for the neurological damage and instead cast suspicion on the degree of force used to manipulate forceps. A doctor’s response to the potential litigation consequences of having delivered a neurologically unresponsive newborn thus can lead to a protean outcome in terms of the risk of litigation. Lawyers are familiar with the capacity of parties engaged in legal self-help to raise more suspicion than it quells. In any system in which the burden of proof of causation remains with the plaintiff, however slight, any measure taken to chart data not required by medical considerations is questionable and arguably favours the plaintiff. Furthermore, now that the practice has been established, failure to conduct the test may further cast suspicion on the practitioner and cause the plaintiff’s lawyer to issue a statement of claim.

The problem of the cost of defensive medicine has led to at least one proposal in the United States which should be repugnant to most Canadian doctors: that the standard of medical care be regulated by Health Maintenance Organizations (HMOs) in accordance with public health and efficiency, and that doctors be given immunity against suits for failing to conduct unnecessary tests and procedures. HMOs, according to the argument, are uniquely placed as experts in managing and assessing medical risk. A similarly large number should also argue that the standard of care has already been compromised as a result of cuts by Canadian governments to the state funding of medicine, and that in the United States the same or worse has been effected by the private health insurance system. In theory, this model for medical standard of care approaches the problem as an exercise in econometrics where medical risk is a weighted factor in a cost-benefit analysis. It presupposes that the public will devote a certain figure or percentage of its economy to health care and that the HMO will advise the physician what services are necessary. No doubt the actuaries’ predictions will be uncannily close to the actual rates of sickness, healing and convalescence. To a certain extent, such an exercise is today conducted by hospital administrators, health insurers and government ministries. What is insidious about such a proposal is that, by reducing the profession to a part of the service sector, it denies both doctors and patients the benefit of the practice of medicine as a way and calling. It is bad enough that a scarcity of resources has rendered doctors unable to do everything they can for
patients. To institutionalize a standard of care determined by professional compromise will inevitably demoralize both doctor and patient. Not only would the blow to the confidence held in the doctor reduce the psychological capacity of the sick to get better, we would also see a marked rise in the number of law suits.

The cost of defensive medicine, therefore, cannot be eliminated by granting an ad hoc legislative immunity without imposing the structure of a command economy and contributing to the size of the underlying problem of the occurrence of malpractice and litigation. There are other ways of looking at the cost of defensive medicine. In the insurance model, a small overall increase in the cost of health care, if it provides comfort to doctors, may be considered a justifiable premium against the risk of litigation arising from a bad result within the normal range of outcomes. What may be unjust about it is that the cost of the premium is borne by the consumer, either personally or by insurance. Moreover, from an insurance perspective, the same level of comfort may be had for a smaller premium. The practice is, strictly speaking, surplus to the standard of care. To the extent that practices become more or less universal, the practice may have the unintended effect of raising the level of the standard of care, at least as seen through the eyes of the court. One can see how this effect can unnecessarily fuel litigation, as opposed to retard its occurrence.

Elimination of the cost of defensive medicine in a tort system would appear to be a lost cause, in view of the many ways in which it possesses the logic of a vicious cycle. Apart from efficiencies and advances from better education and communication among practitioners, it is hard to see how this phenomenon can be completely prevented. It must be considered, in economic terms, the measure of the deterrence element of a legal system founded in corrective justice. A deterrent, by its very nature, can lead to a systemic diseconomy in the delivery of health care services to the public.

**No-Fault Insurance and Tort Law Morality**

As a matter of distributive justice, the principled method of providing for victims of medical misadventure is no-fault insurance. It is important to define what is meant by "no-fault." Fault, the attribution of blame for the cause of harm, is a moral notion whereby the act or omission of a perpetrator confers individual responsibility for the harm suffered by another. In the no-fault insurance model, fault is eliminated only by spreading responsibility among all participants in a given activity. If it is acknowledged by all participants (e.g., all drivers of motor vehicles), that damages are caused by error or lack of vigilance, to which all are susceptible, it is easier for victims to accept a prohibition against casting the proverbial first stone. Other
forms of punishment may be reserved for reckless or intentional torts, such as licence suspension and prison for drunk drivers.

If, from the physicians’ perspective, one is more likely to be sued by chance than for being a bad doctor, patients suffering an adverse outcome might come to view this fate as having been determined as much by the disease or injury as the chance of being treated by a bad doctor. The higher the risk and the greater the stakes, the more immediate and inevitable the force of chance is felt. Obstetricians, orthopaedic surgeons, mothers of brain-injured babies and unsatisfied spinal fusion patients may sooner or later feel that the fates have dealt them unfair hands. By having to draw the line somewhere between negligence and observance of the standard of care, the tort system imposes on the plaintiff an arbitrary 51% onus of proof on the balance of probabilities, although in practice the one percent is very subjective. One percent not only separates the "good doctor" from the "bad doctor," but lumps the one-time negligent doctor with the incompetents queuing up to be struck off. The same percentage separates those who may receive millions in an award or settlement from those who will lose and face financial ruin. Yet the cruelty of such justice should not be mistaken for an absence of justice, or unfairness. One recalls the palpable cruelty of the law as portrayed by two American short-story authors. In Shirley Jackson’s The Lottery (1949), once a year the 300 townsfolk congregate, for a reason long forgotten, to draw lots on who is to be stoned to death. In Frank Stockton’s The Lady or the Tiger? (1882), the convicted nobleman must choose between two doors, behind which await a beautiful young maiden to whom he will instantly be wed, or a ravenous tiger. Both laws, however cruel, cannot be attacked on the basis of unfairness. The law applies impartially to all. If the semi-barbarity of the laws offends our sensibilities, it may be in the allocation of the consequences, because the method of decision is not foreign to us.

In the first example, a reversal of fortune based on a lottery is meant to be troubling because the preordained consequence is bad, extreme, and without amoral. In the other, where the crime of which the nobleman is guilty is high treason, it may be as offensive to reward him for his actions as it would be to have him eaten by the tiger. These stories expose the paradoxes in our own legal systems, if only by exaggerating the constituent elements. The margin by which a plaintiff’s case may be dismissed outright is the same one by which a good doctor’s reputation can be ruined. The narrower the margin, the less a "categorical imperative" is distinguishable from chance. The reality of modern medical malpractice law is that, while the finer points of morality underlying tort law may address a need and a desire to correct and compensate for the conduct of bad doctors, its application to the majority of cases is arbitrary. The reasoning is
important because, in global terms, if the same result can be achieved by flipping a coin or drawing lots as by employing lawyers and judges, the economics start to dictate our choice in the way we deal with doctors’ conduct.

By erasing the line between negligence and no-negligence and pushing the policy of deterrence into the realm of professional discipline, "no-fault" is undeniably different from the tort system. However, it is no less arbitrary and, subject to the questions of economic efficiency discussed later, neither "better" nor "worse" than tort in any meaningful moral sense. One almost naturally associates no-fault with insurance, because insurance is capitalism’s method of hedging bets and preventing, at the level of the individual, all from being lost by chance events. Insurance of itself, however, is neutral between tort and no-fault.

Third-party liability insurance has developed in response to the tort system. Its existence is empirical evidence that random chance is a principal element of tort law. By insulating tortfeasors from the consequences of their negligence, this form of insurance has probably prolonged the existence of tort law and has contributed to the occurrence of the torts themselves. Liability insurance does not spread loss but, rather, responsibility for third party loss. The distinction is important because despite some resemblance to distributive principles, it is an adjunct to corrective justice. The purpose of liability insurance is, on the front lines, contrary to distributive justice.

Much in the same way that modern tort could not exist without liability insurance, one cannot imagine no-fault without insurance. No-fault without insurance has existed; it was the prevailing common law before the general duty of care in negligence was recognized by the English House of Lords in 1932. The fact that medical malpractice involves a "special relationship," and was thus covered by a specific duty of care, allows us to sidestep this historical debate. The reason no-fault is perceived as an insurance-driven system of compensation is a combination of politics and myopia. Today it is hard to imagine any form of compensation system, be it personal injury, casualty, maritime or environmental, which does not rely on insurance as the source of funding. Medical defence associations and their hospital counterparts, however, remain among the few pure mutual defence organizations left in contemporary Western society. Depending on the jurisdiction, they may or may not be insurers at law, but in principle they are not, and this distinction may account for some of the perceived historical unwillingness on their part not to behave like commercial insurers. Nevertheless, the choice not to use commercial insurers does not affect the fact that a substantial institution has been established by doctors for the purpose of spreading the risk of malpractice suits among themselves. The real question between fault and no-fault is not whether an
insurance model should be adopted by doctors; they already have one, of sorts, and it is inescapable. Rather, it is: Why should we draw the line at iatrogenic injuries?

The "Bathtub" Argument

The main criticism of proposals for no-fault medical malpractice has been the same for the no-fault scheme for automobile accidents. The so-called "bathtub" argument put forward by opponents of no-fault automobile insurance posed the question: "If we are to compensate auto accident victims, why not someone who slips and falls in the bathtub?" The argument is not only raised by conservatives but also by those who advocate a more comprehensive insurance regime for all accidents. Before we institute no-fault insurance for medical accidents, "we must decide whether this preferential treatment can be justified." The answer to the "bathtub" argument has been that no part of the compensation system could then be reformed unless the changes apply to all parts of it. A further answer, embraced by New Zealand, has been to enact comprehensive accident and rehabilitation insurance legislation, which includes "medical misadventure." Even there, however, the exclusion of known probable outcomes of previously existing conditions involves a preference of "misadventure" over non-accidental injuries or illnesses. In the case of the contracting of viral diseases, for example, the issues of justice become blurred indeed.

The "bathtub" argument employed by defenders of the existing tort regime is philosophically flawed because it entails the use of pure distributive justice to deny altogether any departure from corrective justice. Such an elegant paradox can and must result from a dialogue between two opposed voices of orthodoxy. The same victims who now complain about tort law for raising the bar for recovery in medical cases too high will complain that no-fault benefits are inadequate. This has been borne out in the "threshold injury" litigation after no-fault automobile accident insurance was introduced in Ontario. The doctors who bemoan the mounting cost of litigation will no doubt be the first to accuse benefit claimants of widespread fraud. Fraud, too, has been a source of business for lawyers in the no-fault auto regime. The "bathtub" argument is blind to the historical reason for the need to reform, namely the existence today of a system which inadequately serves the needs of the parties.

The answer to the above question may seem obvious to the tort lawyer, that one must ask in whose bathtub the fall occurred. This is not helpful. The flaw in the "bathtub" argument is that it is directed not at third party liability but first party liability for accidents which involve only one party, who cannot sue himself. One
cannot reform to no-fault a problem in which the fault of another was not originally included in the equation. The logic of reform must operate forward. (One day we may include the victims of bathtub falls.) Intellectually, this response is no more valid than the bathtub argument itself, because it uses the rules of corrective justice to justify a reform of corrective justice to distributive justice. The historical perspective is confusing because change is the only acceptable constant. We cannot, in the final analysis, say no-fault insurance system for medical accidents is preferable to the present tort system, using ethical or absolutist criteria. From a historical perspective, the complaints regarding the present system are social and economic, and we must address the question at that level.

**Total Social Cost**

The total social cost of the litigated iatrogenic injury can be expressed in terms of a formula: \( F = A + B + C + D + E \):

<table>
<thead>
<tr>
<th>A</th>
<th>Value of Dismissed and Unasserted Claims (A1) less non-iatrogenic injuries (A2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Value of Settlements and Judgments</td>
</tr>
<tr>
<td>C</td>
<td>Defence Legal Costs</td>
</tr>
<tr>
<td>D</td>
<td>Plaintiffs’ Legal Costs</td>
</tr>
<tr>
<td>E</td>
<td>Cost of Defensive Medicine</td>
</tr>
<tr>
<td>F</td>
<td>Total ( A + B + C + D )</td>
</tr>
</tbody>
</table>

In the tort system, the cost of iatrogenic injuries, A, and a large portion of plaintiffs’ legal costs, D, is borne entirely by victims. The purpose of legal costs, C, is to increase the ratio of claims dismissed to payouts, or \( A1:B \). There exists a diminishing marginal return to legal costs, both because of extraneous reasons why throwing more lawyers’ time or expenses at a defence cannot increase the likelihood of securing a dismissal, and because in a finite total of \( B + C \) there will necessarily be a point where it will take increasing amounts of \( C \) to obtain decreasing amounts of \( A1 \). It is at the point where the marginal benefit of \( C \) is at an optimum that a medical defence association should set its target budget for legal costs, although in absolute terms the economically justifiable amount of such costs can be staggering, especially since the value of claims dismissed, \( A1 \), exceeds the value of claims dismissed in which there is an iatrogenic cause, A.
In Canada, according to CMPA figures, awards exceeded legal costs for the first time ever in 1996. This fact alone does not mean the legal costs are excessive, and the obvious comparison is unfair. The measure of the marginal value of legal services is to be measured in the value of dismissed claims plus the value of the difference between settlements and judgments, on the one hand, and plaintiffs' reasonable expectations, on the other. This is perhaps an impossible task because the value of such claims and savings is not to be gauged by the figure quoted in the pleadings but in a neutral and fair assessment. However, the legal cost to the defence organization, D, as a portion of its overall absolute cost (B + C + a portion of D), is also restrained by the extraneous factor of the degree to which physicians are prepared to be levied. Thus, to the extent that a marginal increase in C does not produce an equal or better reduction in B and D, physicians will not stand for it. Nor should be forgotten the capacity for the conduct of a defence to affect the cost of asserting the claim. A less than optimal marginal increase in defence costs will likely inflate the total social cost, both by its own increase and by the increase to the plaintiffs' legal bills.

The total social cost in a pure no-fault system should, at least in theory, be lower than the tort system. It is equal to the total social cost of a pure tort system, less the legal cost of fighting cases on the basis of liability (the cost of Bolam and its variations), and less the cost of defensive medicine. For comparison with the previous table, A and B have not been combined in the following table representing the total social cost of iatrogenic injuries in a no-fault system:

<table>
<thead>
<tr>
<th>A+B</th>
<th>Value of Iatrogenic Injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Defence Legal Costs</td>
</tr>
<tr>
<td>D1</td>
<td>Plaintiffs' Legal Costs</td>
</tr>
<tr>
<td>E</td>
<td>Defensive Medicine = nil</td>
</tr>
<tr>
<td>F</td>
<td>Total A + B + C1 + D1</td>
</tr>
</tbody>
</table>

It has been argued that physicians who have called for a no-fault system should temper their enthusiasm because no-fault may compensate more victims than tort and thereby cost more. It is restating the obvious to say that more will be compensated in a no-fault regime than in a tort system. Seen in terms of total social cost, however, the number or value of iatrogenic injuries should not change, and thus the real question is who shall bear the loss.

In market terms, physicians will not approve or readily participate in a no-fault system unless the new regime reduces the cost to them, including the up-front cost of changing the institutions and retraining the lawyers. (They may, with some
persuasive advocacy, accept a system which costs as much or slightly more, on the basis that there can be a value attached to the release from the peril of a judgment in negligence against them.) The cost to doctors, therefore, must be capped at or near the present level of their participation in their defence organizations. If the victims of iatrogenic injuries are to bear the balance, it is arguable that the only way in which they could be any further ahead than in the tort system is in some expected savings in legal costs. Society (or public coffers) as a whole might benefit from the reduction or elimination of the cost of defensive medicine and could reap some collateral dividends, but the effect of a system of loss-spreading must show substantial economies unless one is content with a simple ideological choice between corrective and distributive justice.

The most significant economies to be achieved by a no-fault system must be in the comprehensive organization and implementation of the system to take advantage of the alternative to the adversarial model. The constraint of the common law is that it can only compensate by way of a money judgment for damages. The basic principle of tort law is to restore the plaintiff to the position he or she was in prior to the accident to the extent that it can be done by taking money from the defendant’s pocket and putting it into the plaintiff’s. If we maintain this principle, to some extent our efforts are constrained by a value which is foreign to the distributive justice model. In a pluralist society valuing compromises between the free market and the social safety net, the distributive model cannot ignore the protection of private expectations. Thus, to use extreme ends of the scale, the corrective model would compensate the family and estate of a rich bank executive for the millions he was expected to earn in his or her lifetime, but the lost income of an infant not expected to live past the age of majority would be discounted for duplication with living expenses and assessed at about $150,000 in 1998 dollars. Meanwhile, someone whose pre-injury disabilities excluded gainful employment might be awarded zero or a nominal amount.

It is already the statutory mandate for provincial medicare to provide insured services to the public, no matter what the cause of the injury or illness. In terms of the medical services required to treat the effects of iatrogenic injuries or illnesses, medical, hospital and home care services do not differ among persons of different incomes or other socioeconomic criteria. One might make a strong argument in favour of considering the abolition of subrogation claims on behalf of state health insurers as a contribution or premium on behalf of the public. Moreover, to the extent that victims may choose to use private home care providers at a saving to institutional care, a co-ordinated cross-funding mechanism could strike a compromise between the two insurance systems to take advantage of the fact that the victim contributes the facility of his or her home to relieve the public of the
capital cost of the institutional equivalent. Because the insurance is already in place and does not discriminate beneficiaries by the cause of the occurrence (even self-inflicted injuries are covered), there is no principled reason to single out iatrogenic injuries as entitling the state to claim against another source of insurance. Although removing state medicare from the equation does not, of itself, reduce the value of the total social cost of iatrogenic injury, in a no-fault regime its exclusion is justified whereas the abolition of subrogation cannot be justified in a model of compensation based on fault. Indeed, if the funding of state-insured services were included in a no-fault system for medical mishap, this would involve inappropriate cross-subsidy from the other sources of funding for the no-fault system, and also an inappropriate reverse discrimination according to the "bathtub" argument.

One possible source of premium for insuring economic expectations is a progressive levy on the awards themselves. The larger the income benefit, the larger the deduction for the purpose of funding future like cases. Many no-fault automobile insurance schemes have built in some form of premium in the compensation system, and so require the insured claimant to bear the first week’s loss of income, deductibles and verbal "thresholds" for non-pecuniary losses. Reducing the pay out on a sliding scale, however effective it may be in helping to facilitate the insurance, does not contribute in real dollars to the funding of the system. However, one source of loss allocation does contribute to the fund, by giving up already existing funding. This is the class of victims who, by virtue of negligence, admitted or found, would be recipients of awards or settlements in the tort system.

An alternative to the progressive levy on income benefits is their total abolition beyond a certain level. The executive mentioned in the above example would, as a market choice, be life-insured against the peril of iatrogenic injuries, and the same would apply for his disability insurance. It would be a severe diseconomy for the no-fault regime to insure someone for a specific peril when the "expectation" type of damages are universally insured for all insurable perils under insurance contracts freely entered into for prudent, market reasons.

At present, pharmaceutical companies in theory enjoy a comparative advantage in Canada because, unlike many of the high-population jurisdictions in the United States, the law of product liability does not render them strictly liable for injuries beyond the normal scope of known and published side effects. The same can be said of medical equipment manufacturers. In reality, however, the tort system is stacked sufficiently in favour of plaintiffs that the exposure of such companies to law suits may as well be strict liability. The only significant area of defence
available to them is causation. But, apart from the onus of proof, the substantive question is identical to the causation question which arises in a no-fault insurance scheme. Whether these companies carry liability insurance, or self-insure and carry reserves for the same purpose, it should be logical to enlist this pool of insurance to the no-fault insurance regime in exchange for immunity from civil liability. Insurers presently involved in underwriting the companies need not be excluded, if it is worthwhile for them to participate.

Despite these savings to the total social cost of iatrogenic injuries and illnesses and sources of additional funding, a system of compensation cannot consist entirely of savings and collateral participants; it must be fully funded. According to basic insurance principles, premium participation must be universal among participants (or at least those who contribute to the risk). This raises the question of who pays the premium on behalf of the patient: the individual or the state? On the one hand, for the same principles upon which the state medicare system should be excluded from the benefit of the system, there is a forceful argument that the state should not subsidize the system. On the other hand, the accepted principle of universal access to medical treatment militates against the imposition of mandatory medical mishap insurance among patients. It looks and feels too much like a user fee and it may detract from universality. Provided that measures be taken to remove or contain the patients’ contribution to the occurrence of claims, such as hypochondria and fraud (as opposed to the risk itself), there is an argument that no premium per se need be paid. Rather, it should be the objective of the legislators to devise a no-fault plan which is funded directly by doctors, hospitals and suppliers (such as pharmaceutical companies and manufacturers of medical devices), and indirectly by the abolition of subrogation by medicare plans. Any state funding should, as it is now, be considered a subsidy and not as a premium per se.

Hybrid Tort/No-Fault Systems

In Ontario, an attempt was made to reduce the cost of automobile accidents to the mandatory insurance system by enacting a series of no-fault insurance regimes which diverted most cases into a codified system of statutory accident benefits. Nevertheless, the legislature preserved tort claims for injuries which were serious enough to overcome certain verbal "thresholds." "Threshold" litigation is now a cottage industry unto itself. Despite almost universal frustration expressed by the practicing bar for the ad hoc changes made over the last decade to the statutory rules, it remains to be seen whether no-fault will get over these teething problems and prove an effective and reasonable solution to the crisis in tort law which came to a head in the 1980s.
The principal institutional problems associated with medical malpractice do not exist in the cause and effect of car crashes, largely because of the identity of roles between plaintiffs and defendants. The activity of driving involves no educational or class boundaries. Criticisms for bad driving are easily hurled and otherwise are, apart from drunk or reckless driving, not reflective of good or bad character. The parties can easily accuse each other of negligence because of the anonymity of the parties. (Indeed, it used to be common for both drivers to sue each other and recover from the other’s insurer, and there was no real advantage to being the first to sue.) The law suits were for the most part formulaic and, compared to medical cases, resolved expeditiously in the court system. Claims ranged from small to very large. The main problem was the mounting cost to insurers of awards for loss of earning capacity, and the commensurate rise in insurance premiums. It remains to be seen, from a more distant historical perspective, whether the insurance crisis of the 1980s was not a transitory problem associated with a bubble economy.

The underlying facts of medical malpractice litigation, which drive the parties into undesirably adversarial camps and spread fear and distrust among them, are: (a) that the allegation of a lapse in judgment or lack of competence is not to be made lightly against a doctor, and (b) most cases involve permanent and serious losses. The very existence of a system for compensating victims of iatrogenic injuries that depends on the fault of the doctor perpetuates the basic problems because the tort regime can affect the relationship between him or her and the patient. Small claims in medical malpractice (if they exist in any meaningful number) are not the cause of the crisis in the CMPA. The type of verbal "threshold" for permanent and serious injuries used to preserve the right of action for tort in automobile cases would no doubt be overcome in the majority of cases, and thus it would be illogical to direct the institutions of no-fault insurance to a small proportion of claims which, in any event, are not responsible for the crisis. Furthermore, a system which keeps the quality of care in check not by the ethical impropriety of offences but by the monetary value of the damage caused can be neither just nor effective.

Early Dispute Resolution

The popular perception among doctors that chance has as much to do with the likelihood of being sued for malpractice as being a poor doctor is palpable yet impossible to prove. Nevertheless, the randomness of the decisions of victims of iatrogenic injury to sue is a testament to the cultural influences which bear upon different types of litigation in disparate ways. The absence of a legal ceiling on jury awards for non-pecuniary personal injuries and wrongful deaths and of cost sanctions following the event have led to great incentives for American plaintiffs
and their lawyers to sue. However, in Canada, the judicial limitation of awards to
conventional or arbitrary figures and the cost exposure to plaintiffs have had the
contrary effect. In this country, no injury of a temporary nature is ever worth suing
the doctor unless the conduct of the doctor or his lawyer after the fact has
transformed the issue into one of "principle."

The fact is that medical mishap, whether or not culpability attaches to the
physician, is in many ways no different from events which strain or rupture many
other relationships of confidence among members of our society. Most cases do
not involve any element of breach of trust or active concealment, nor do they reach
the other end of the continuum, the tort committed between strangers, the most
typical being the two-car collision. It is because the professional medical setting
imports not only controllable human factors (the capacity for error) but also
ungovernable biological factors (healing, suffering and death), that the litigation
paranoia which sets in at the first hint of an adverse outcome beyond the usual
scope is both unjustified and unproductive. Where medical resources are limited,
such as in rural communities, the merits of bringing a law suit is to be measured
against the possibility that the plaintiff may deprive himself from access in more
serious situations.

In the United Church of Canada, part of its written constitution consists of four
levels of procedures entitled "Resolution of Conflicts," ranging from early mediation
without an intermediary, informal mediation with an intermediary, a formal hearing
observing the rules of evidence, and an appeal. The governing assumption is that
those first involved in the conflict will resolve the situation between themselves,
without further resort to the Church courts. Based on a perception by the Church
administration that their ministers cannot be trusted, the Church as been known to
abandon their procedures and proceed directly to formal procedures. The
ecclesiastical experience is that litigation is the product of systemic failure of the
governing intermediary, and not of any lack of good faith on the part of the parties.
The source of the breakdown in these cases appears to have been excessive zeal
on the part of the church hierarchy to define parties as "perpetrators" and "victims"
and to extract a confession from the accused as an uncodified condition of fairness
in the process. The injustice could as easily be the opposite: insensitivity or steps
taken to cover up or ignore the complaint. These are perils which face any self-
governing profession employing an inquisitorial model for self-discipline. No doubt
the sources of zeal or insensitivity are political. A conservative institutional culture
will exclude legitimate complainants and drive them to the courts, whereas a self-
flagellating one will attract crackpots.
If the exposure of doctors to tort liability is to be partly or fully relieved in favour of a no-fault compensation system which includes some cases which will be resolved without compensation, a concerted effort as well as some concessions to self-regulation must be made by the profession. Doctors must be prepared to accept the participation of professional mediators, be they lawyers or otherwise. Perhaps the most significant development of the Alternative Dispute Resolution (ADR) model in recent years has been the procedural formalization of participants in a substantively informal process. So long as a compensation model depends upon a tug-of-war between claimants and insurers, it will be impossible to eliminate the adversarial model, even in no-fault. However, by setting the order in which parties address themselves and permitting opportunities to meet with their legal representatives, the institutional rules of engagement pre-empt the necessity to adopt siege tactics. One systemic problem associated with a court-based ADR model is that there is no change to the compensation model. As long as the only remedy is a cash settlement or award, ADR can be a facsimile of the common-law tort system without a binding record.

If, on the other hand, institutionalized ADR between the patient/claimant and the insurer in cases of temporary or non-serious cases were adopted, the doctor could be brought in at this stage (and no further) as a non-party. The purpose of such a process would be to suspend the claim and instead address the immediate concerns of the patient, such as dealing with an employer (and, if called for, compensating the employer so as to preserve the patient’s employment), arranging for an independent second opinion and accessing necessary rehabilitation and nursing services. The patient could be represented by a lawyer or by an independent adjuster. The level of trust in the system may be enhanced even further if incentives were in place to encourage doctors to initiate the process before a claim is made (the reverse of the present outlook) if he or she suspects an iatrogenic source of the patient’s problem.

**Conclusion**

Were fault eliminated as a basis for compensating victims of iatrogenic injuries, there may prove to be a sea change, not only in law, but in the practice of medicine in fields previously fraught with the perils of malpractice litigation. Some of the reforms necessary to implement a no-fault regime require rethinking our ideas of the purpose of compensation and of the definition of fault itself. There must also be greater openness in the disciplinary process to relieve and fairly keep participants in the health care system accountable. One might also foresee a more widely accessible, albeit less dramatic, role for personal injury lawyers on both sides of the divide between plaintiffs and defendants. It is also easy to see how some
litigation will necessarily survive, not as between doctors and patients but between patients and the no-fault insurer(s). The outcry of doctors participating in high-risk, high-premium fields of practice stems from the very high-stakes litigation that has been encouraged by the law of malpractice itself. Some on the other side might say that those doctors themselves are partly to blame for this predicament. However, closer examination of the tort model shows that much of the energy and resources are devoted to the drawing of an imaginary line between negligence and good care. It is by that thin line that tort law suspends the sword of Damocles over our physicians.