

Liability for Medical Malpractice

Patricia M. Danzon

Physicians and other medical providers are subject to a negligence rule of liability. To prevail, a plaintiff must show that he or she sustained damages that were caused by the failure of the physician to take due care, defined as customary practice of physicians in good standing with the profession, or a significant minority of such physicians. In a simple model, with perfect information and homogeneous physicians, a negligence rule of liability with an appropriately defined due care standard should induce complete compliance: there should be no malpractice, no malpractice claims and no demand for malpractice insurance.

The malpractice experience is seriously at odds with this prediction. The incidence of negligent injury is not trivial—roughly one per hundred hospital admissions (Harvard, 1990). From 1975–1985 the frequency of malpractice claims per hundred physicians increased at roughly 10 percent a year, claim severity (average amount per paid claim, including jury verdicts and out-of-court settlements) rose twice as fast as the consumer price index, and malpractice premiums increased sharply. From 1985 to 1989 claim costs and premiums stabilized, but are beginning to increase again. There have been no major changes in the basic common law rules defining medical liability that could explain this increase in claims.

This discrepancy between the simple theory and actual experience raises two related issues. First, what goes wrong? Second, if the system does indeed operate imperfectly, does it yield benefits in terms of injuries deterred that outweigh the high overhead costs of operating a liability system?

■ *Patricia M. Danzon is Professor of Health Care Systems and Insurance, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania.*

If the sole function of liability is to provide compensation, it is extremely inefficient. Tort benefits are unpredictable and therefore provide poor insurance to the victim. Liability also imposes risk on defendants. Moreover, the overhead rate of the tort system per \$1 of compensation is roughly 120 percent (Danzon, 1985a; Kakalik and Pace, 1986), compared to 20 percent for large group private health and disability insurance programs. The primary difference is the litigation expense incurred to determine cause, fault and liability. This cost of assigning liability is worth incurring only if there is a deterrence pay off, in terms of injuries prevented because the threat of liability makes providers more careful.

It is widely believed that the deterrence benefits of the malpractice system are minimal and that the costs of “defensive medicine”—defined as liability-induced changes in practice that would not be desired by an informed patient—are high. Many states have enacted tort “reforms,” including caps on awards for pain and suffering, offset of benefits from other insurance (collateral source offset), limits on plaintiff attorneys’ contingent fees, and so on. But there is growing interest in more radical alternatives. The Commissioner of Health in New York recently proposed a no-fault system of compensation for all injuries arising from medical care, modeled after the workers’ compensation system for work-related injuries (Sack, 1990). Virginia and Florida have enacted no-fault plans for birth-related neurological injuries. The American Medical Association (AMA, 1988) has proposed establishing special administrative agencies to adjudicate medical malpractice claims, replacing the courts, while retaining a (modified) fault-based rule of liability.

This paper begins by reviewing the theoretical issues raised by the application of provider liability in the context of medical care. Next, it reviews the evidence on the incidence of injury due to medical negligence, trends in claims, and insurance costs, as well as the evidence on effects of liability on the practice of medicine. It concludes by briefly discussing policy implications and alternatives to the status quo.

Peculiarities of Medical Markets

Medical care is an infrequently purchased service that is often produced under conditions of severely asymmetric information between producer and consumer. As Spence (1977) has shown, if customers misperceive risks, a rule of caveat emptor leads to nonoptimal levels of risky activities and nonoptimal care per unit of activity. For example, if patients underestimate the risks of cosmetic surgery and underestimate the risk-reduction benefits of the surgeon spending additional time implementing the surgery, there will be too many nose jobs, too little care per nose job and consequently too many botched nose jobs. Since many medical services are infrequently purchased, it is difficult to become an informed shopper until it is too late. Because the product is patient-specific

service, rather than a mass-produced, homogeneous good, the existence of a core of informed shoppers will not necessarily suffice to police quality for everyone, especially if physicians can tell who is and who is not informed. There are coarse screens to protect against incompetence or systematic misconduct: licensure as a condition of practice, peer review, hospital credentialing committees, reputation and referral networks and, more recently, quality monitoring by third party payers. But in theory, provider liability could enhance incentives for care on a case-by-case basis.

A negligence rule of liability for failure to take due care (defined as efficient care) can in theory create incentives for optimal care per unit of activity. However, the level of risky activities may still be nonoptimal if patients misperceive average risk (Shavell, 1980), unless the definition of negligence also extends to performing additional “unnecessary” procedures. But in medical malpractice and other professional liability, due care is defined in terms of professional custom, possibly because the courts cannot at reasonable cost acquire the information necessary to define due care according to a cost-benefit standard. By definition, a custom standard of liability cannot correct any systematic nonoptimality in customary care that may be induced by consumer misperceptions. A custom-based standard in imperfectly informed markets could be either too low or too high. Liability could prevent significant deviations from this standard, which may be a proxy for consumer expectations, but this would not necessarily result in optimal care.

After asymmetric information, a second peculiarity of medical markets—and one that raises further questions about the optimality of the custom standard—is the prevalence of private first party and social insurance that pays for the bulk of medical care. Over 80 percent of the population has some form of health insurance, typically with modest co-payment and premiums unrelated to past use (experience rating). Consequently, only roughly 10 percent of expenditure on hospital care and 26 percent of physicians’ services are paid directly out-of-pocket (Levit and Freeland, 1988). Because insurance drastically reduces the point-of-purchase price of medical care to patients, this distorts customary levels of use of services,¹ relative to a first best optimum, and may also distort customary “quality,” defined as technologies typically used to treat a particular condition.

But the full-information, first-best optimum may not be a relevant standard given the costs to health insurers of controlling moral hazard (Zeckhauser, 1970). In addition, the outcome is further distorted by the fact that employer contributions to health insurance are not taxable income to employees, thus providing the employed population with an incentive to demand excessive insurance coverage. Given the prevalence of health insurance that is either fully subsidized through general tax revenues (Medicare and Medicaid) or heavily subsidized through the tax system (employment-based private health insur-

¹For evidence on the effects of co-payment on use of medical care, see Manning et al. (1987).

ance), physicians acting as good agents would provide medical services as long as the marginal benefit to the patient exceeds the private marginal cost facing that patient, which is typically a small fraction of the social marginal cost.

The prevalence of heavily subsidized health insurance has several implications for the application of liability rules to medical markets. First, the use of a cost-benefit calculus for determining due care would be problematic even if courts had good information on the benefits of alternative medical treatments. If liability is to promote efficient resource allocation, the calculus should use social costs and benefits. But this may differ from private marginal costs and benefits to the patient, which the physician acting as a good agent should use and is under some pressure to use, assuming some degree of competition and information in medical markets. Use of a social cost-benefit calculus could create unbearable tension between legal standards and medical care markets.

But if courts adhere to traditional customary practice, this could undermine recent attempts by private health insurers and social insurance programs to substitute provider incentives for holding down costs for patient incentives. This has been done through use of health maintenance organizations (HMOs), preferred provider organizations (PPOs), utilization review, and so on. In the long run, these innovative contractual forms offer potentially efficient means of reducing the deadweight loss generated by patient moral hazard. But they create a conflict between what the patient wants when sick and what the physician or insurer is willing to provide under the terms of the contract. Although suits for withholding services, alleging that the patient was not fully informed about the service restrictions under the insurance contract, have so far been rare, they could become increasingly important in the future. How the courts resolve such suits may be critical to the feasibility of controlling the growth of medical costs within the context of a private markets for medical care and medical insurance.

Second, the prevalence of health insurance makes it very difficult to distinguish cost-justified precautions from “defensive” responses to liability. Defensive medicine should be defined as liability-induced changes in medical practice that entail costs in excess of benefits and that would not have occurred in the absence of liability. Health insurance leads to the use of many medical services that yield benefits less than full social cost because of moral hazard. This is insurance-induced, not liability-induced defensive medicine, but distinguishing the two empirically is difficult. There is concern that heavily insured patients have little incentive to refuse any treatment with non-negative benefit, and that physicians therefore engage extensively in such defensive practices to reduce their own risk of suit. One widely cited estimate put the cost of defensive medicine related to ambulatory care alone at \$10.7 billion in 1984, or 16 percent of total expenditures on physicians’ services (Reynolds et al., 1987). Both the theoretical argument and the empirical evidence on defensive medicine are discussed further below.

Trends in Malpractice Injuries, Claims, and Insurance Rates

The Frequency of Negligent Injury

The best data on the incidence of injuries due to medical malpractice come from two broad-based surveys of medical records of hospitalized patients, the first in California in 1974 (CMA, 1977) and the second in New York in 1984 (Harvard Medical Practice Study, 1990). The CMA study reviewed records of 20,864 patients in 23 hospitals, while the Harvard study reviewed records of 31,429 patients in 51 hospitals. In both studies, a stratified random sample of hospital records were reviewed by experts in legal medicine, to determine the incidence of any injury due to medical care and the incidence of injury due to negligence. The findings were quite similar. The California study reported that 4.65 percent of hospitalized patients suffered an injury caused by health care management; of these, 17 percent (or one in 126 patients) involved negligent injury.² In the New York study, 3.7 percent of patients sustained an injury that resulted in measurable disability. Of these injuries, 28 percent (or 1 percent of all hospital discharges) involved negligence.

In both studies, most injuries were relatively minor. In the New York study, 57 percent of patients recovered fully within a month and 70 percent within six months. But 14 percent of medical injuries were fatal, and over half of these were attributed to negligence. This is almost certainly a downward-biased estimate of the total number of negligent injuries, because hospital records may be inadequate to detect such injuries and because injuries occurring in ambulatory settings are excluded, unless they resulted in hospitalization.

The California study did not track the injured patients to determine the number who subsequently filed suit. However, a comparison of the total number of malpractice claims filed in California in 1975–1978 to the number of negligent injuries implies that at most one in ten of the victims filed a claim, and only 40 percent of these claimants received compensation through the tort system (Danzon, 1985a). The New York study actually tracked the injured patients. Although the total number of malpractice claims against doctors and hospitals was about 12 percent of the number of negligent injuries, only 6 percent of patients identified by the study as having sustained an injury due to negligence filed a claim. These data suggest that many who sustain an injury because of negligence do not file a claim, and that many claims are filed by people who have not sustained an injury caused by negligence, at least relative to the standards of negligence used in this study. Because many of these cases were still unresolved at the time of the study, it could not assess the ability of the litigation system to screen out claims where the study had not found negligence.

²The data are described in Danzon (1985a). An injury was defined as negligent based on an evaluation by the reviewers of the likelihood of a jury finding of liability under prevailing negligence law.

The data available are not ideal for understanding the causes of negligent medical injuries and the ability of the malpractice system to detect different types of error. If certain mistakes are easier than others to discover and prove to a jury, or if small stakes cases are not worth filing because of the costs of litigation, then the malpractice system will impose uneven penalties on different types of error and may distort the delivery of care. In the CMA (1977) study, 82 percent of all injuries were adverse effects of treatment; only 15 percent were effects of incomplete diagnosis or treatment. The study concludes (p. 62), "Problems of performance, rather than purely judgmental issues were the overwhelmingly responsible mechanisms." Similarly, errors in performance were the most common allegation in claims filed in 1975–1978. Whether or not this reflects bias in the ability to detect different types of error, it does suggest that preventing injuries and claims requires being more careful, not simply performing more tests or X-rays. The ratio of claims to injuries tends to be lower for minor relative to major injuries, and lower for persons over 65. This is consistent with the hypothesis that the expected award is an important factor in determining whether a claim is filed.³

After the problem of ascertaining the causes of negligent medical injuries, a second question is whether malpractice claims are largely a problem of a minority of incompetent physicians and low quality hospitals or of occasional inadvertent lapses by normally competent providers. Studies consistently show that both factors play a role. Adverse event rates varied 10-fold between individual hospitals, when standardized for patient age and diagnosis (Harvard, 1990, p. 4). Analysis of claims experience of groups of physicians indicate that, after controlling for medical specialty, the distribution of claim frequency is more concentrated than would be expected based on chance alone (Rolph, 1981; Nye and Hofflander, 1988; Ellis, Gallup and McGuire, 1989; Sloan et al., 1989).

However, while some physicians who have a disproportionate number of claims may be of below average competence, other physicians may be sued relatively frequently because they treat more difficult cases or a larger volume of patients, in which case they may be of above average competence in some regards. Consistent with this hypothesis, Sloan et al. (1989) report that board-certified physicians and physicians working longer hours tended to face more malpractice claims than physicians who are not board-certified.

³The partial data on causes of injuries and claims reported so far from the more recent New York study are less conclusive. Surgical complications accounted for 47 percent of all adverse events, but the percent of these attributed to negligence was lower than for non-surgical adverse events (17 percent vs. 37 percent). Over 75 percent of adverse events resulting from errors in diagnosis and in non-invasive treatment were judged to be due to negligence; for falls, 45 percent were attributed to negligence (Harvard, 1990, p. 5). Without knowing the population at risk in each category, these percentages cannot be transformed into rates of negligent injury. Ratios of claims to injuries, by type of injury, are not yet available.

But even if it could be shown that the great majority of physicians take appropriate care most of the time and that most negligent injuries are the result of occasional mistakes by otherwise competent physicians or the incompetence of a small atypical minority, it certainly would not follow that liability is superfluous for purposes of deterrence. It is possible that in the absence of liability, most physicians would be less careful and the occasional mistakes would be more frequent. Although it is often argued that liability is unnecessary because of state professional review mechanisms, Sloan et al. (1989) report that fewer than 10 percent of physicians with adverse claims experience were disciplined in any manner. If one assumes that most adverse claims experience does reflect suboptimal care, that seems a low figure. Moreover, the fact that hospital and other peer review procedures have been strengthened in response to liability suggests that, although in theory these mechanisms could be substitutes, in practice they may be complements.

Trends in Malpractice Claims

The frequency of malpractice claims (number of claims per 100 physicians) has increased at roughly 10 percent a year for the last two decades, with sharp increases in the early 1970s and 1980s and slower growth in the second half of both decades. Claim severity has risen at roughly twice the rate of increase of the consumer price index, with some evidence of disproportionate growth for the highest stakes cases (Shanley and Peterson, 1983).

Some increase in claim frequency can be attributed to changes in the practice of medical care, in particular, to the increased frequency of surgical procedures. Explicit changes in legal doctrine can only account for a small fraction of this growth in claims. Some of the surge in claims in the early 1970s was associated with pro-plaintiff shifts in law. For example, the abolition of the locality rule substituted a statewide or national standard for a local standard of due care; the abolition of charitable and government immunity exposed voluntary and government hospitals to suit; the doctrine of *respondeat superior* extended the liability of hospitals for actions of their employees; and informed consent was defined as requiring information that a reasonable patient would want, rather than what was customary for physicians to provide. But these factors had run their course by the mid-1970s and cannot explain claim growth in the 1980s (Danzon, 1984b, 1986).

In response to the malpractice insurance crisis of the mid-1970s, most states enacted one or more tort reforms. Some of these changes slowed the growth of claim costs. In particular, as of 1985 caps on awards and collateral source offset had reduced claim severity (by 23 percent and 18 percent, respectively), relative to what it would have been in the absence of the reforms. Shorter statutes of limitations reduced claim frequency—one year off the statute of limitations for adults is estimated to reduce claim frequency by 8 percent. Collateral source offset has also reduced claim frequency (by 14

percent) presumably because of the feedback effect from lower expected award to reduced incentive to file claims. Other reforms do not seem to have had significant effects (Danzon, 1986).⁴ The sharp increase in claim costs and insurance rates in the early 1980s, after a lull in the late 1970s, prompted a new round of tort reforms in the mid-1980s. Following several years of no or slow growth, claim frequency started to rise again in 1989.

Although medical and legal factors account for some of the trends in claim costs, much of the growth remains unexplained. The New York evidence of a very low rate of bringing suit based on valid claims indicates that claim growth over the last 15 years cannot simply be attributed to the “catch-up” filing of claims for a larger percentage of negligent injuries. Growth in the number of lawyers per capita is not a statistically significant explanatory factor, after controlling for other attributes of lawyer-dense areas. Danzon (1984b) found that urban areas tend to have much higher claim frequency and severity, but that this urban phenomenon could not be explained by specific observable characteristics of urban areas such as income, unemployment, welfare reciprocity, or population turnover rates (intended as a proxy for the “physician-patient relationship”). The growth in frequency and severity of malpractice claims parallels to some extent the growth in other types of tort litigation, notably product liability, but pro-plaintiff changes in legal doctrine (see, for example, Henderson and Eisenberg, 1988) provide a ready explanation of product liability claim growth.

Another unanswered question is whether the increase in claim frequency is largely a response to higher expected awards, leaving the increase in awards as the main factor to be explained. Evidence from workers’ compensation (for example, Butler and Worrall, 1983) shows that claim rates do respond to benefit levels. A similar supply response is plausible for medical malpractice, although much harder to measure because award levels are not statutorily determined. The finding that collateral source offset reduces claim frequency as well as claim severity is also consistent with the supply response hypothesis. However, caps on awards affect only the few very high stakes cases, so should not be expected to reduce claim frequency, despite a significant effect on claim severity.

The rate of growth of malpractice claim frequency and severity has been as high in Canada and the United Kingdom over the last two decades as in the United States, although levels remain higher in the United States (Dewees et al., 1989; Danzon, 1990). This growth has occurred in Canada and the United Kingdom despite the fact that these countries do not permit contingent fees for plaintiff attorneys, have limits on awards for pain and suffering, have

⁴These estimates are the average effect over the period 1976–1985. Since some reforms were under challenge during part of this period, these estimates may understate full long-run effects of the reforms; on the other hand, these effects may overestimate the long-run impact if creative lawyers find ways around the constraints.

lower medical costs and lower rates of growth of medical costs—all factors commonly cited to explain the growth of U.S. malpractice claims.

Since measured changes in the medical and legal environment do not fully account for the growth in malpractice claims, it appears that factors that are harder to quantify play an important role. First, as mentioned earlier, higher potential awards tend to increase the expected value from marginal claims. Second, there may have been numerous small and subtle changes in case law and rules of evidence that make it easier for a plaintiff to establish a cause of action and get to a jury, which effectively reduces the expected cost and raises the expected value of bringing marginal claims. Such hypotheses are plausible but are very hard to test.

Malpractice Insurance

The cost of malpractice insurance has risen roughly in step with rising claim costs, but more erratically. In the early 1970s, premium rates initially lagged behind rising claim costs, necessitating sharp premium increases in the mid-1970s of over 300 percent in some states. In the late '70s, claim costs stabilized and insurance rates fell in real terms. A similar pattern occurred in the 1980s except that the rate increases were spread over a longer period. Table 1 shows average (across states) annual rates of increase in the cost of basic limits of coverage (\$100,000 per claim, \$300,000 aggregate for all claims in the policy year). This understates the increase in the cost of a constant level of protection, which requires purchasing higher limits of coverage as awards increase. The majority of physicians now buy coverage of at least \$1 million per claim.

Over half of the total dollar volume of physicians' malpractice insurance is now written by physician-owned mutual companies, and some form of self-insurance is even more common for hospital coverage. The immediate impetus to the formation of mutuals was the withdrawal of commercial carriers in several states in the mid-'70s when regulators disallowed the requested rate increases. However, the survival and growth in market share of these mutual carriers suggests that they have advantages relative to stock companies in this market. Possible areas of advantage are superior ability to distinguish good and bad doctors and valid from invalid claims, which facilitates settlement and control of policy-holder moral hazard. They may also be superior bearers of the common component of risk that derives from uncertainty as to social and legal trends that affect all policyholders in the pool. To the extent mutuals can assess or pay dividends to their members, depending on the realization of this common risk, this form of insurance may be less costly to policy-holders than commercial insurance through stock companies, at least if stock companies cannot costlessly diversify the common insurance risk through equity markets (Danzon, 1984a; Doherty and Dionne, 1988) and must hold capital reserves to protect against such uncertainty. By purchasing reinsurance, mutuals can

Table 1

Average Yearly Increase in Premium Rates for Basic Limits Coverage by Specialty (Percent), 1976–1987

<i>Year</i>	<i>GP</i>	<i>IM</i>	<i>GS</i>	<i>OBG</i>	<i>NS</i>
1976–77	– 1.7	– 3.0	– 6.1	– 2.0	– 2.0
1977–78	– 1.4	– 0.4	1.5	– 1.3	– 2.2
1978–79	4.9	9.4	12.3	6.1	7.7
1979–80	11.8	15.1	12.5	13.6	12.9
1980–81	25.1	14.1	14.8	22.7	33.8
1981–82	8.6	12.1	11.5	13.0	19.7
1982–83	17.4	22.0	11.0	30.0	20.2
1983–84	22.6	19.0	19.3	19.1	21.8
1984–85	26.7	11.5	14.3	23.7	13.4
1985–86	85.9	54.0	72.7	97.4	75.2
1986–87	27.3	26.5	27.2	27.9	27.3

Source: Calculated from HCFA Survey of Malpractice Insurance Premiums. Unweighted average across states.

GP = General Practice
 IM = Internal Medicine
 GS = General Surgery
 OBG = Obstetrics/Gynecology
 NS = Neurosurgery

obtain some of the advantages of diversifying the nonsystematic component of risk through equity markets.

Given the dominant role of mutuals in the medical malpractice insurance market, the allegation that the sharp premium increases reflect collusive behavior by insurers is even less plausible than in the market for product liability insurance.

Effects on Patterns of Medical Practice

Understanding the effect of liability on physicians' practice patterns and ultimately on the rate of negligent injury is critical to measuring the costs and benefits of the malpractice system. But the evidence is inconclusive because the existing data are limited. As already noted, the number of iatrogenic injuries (that is, injuries caused by medical care) is not reported in a sufficiently systematic manner to permit use of either cross-state or time-series variation in litigation rates to estimate deterrence effects.⁵

⁵The New York study attempts to measure deterrence by analyzing the effect on injury rates of claim filing rates in a prior year, exploiting the within-state cross-sectional variation in claim rates as a measure of variation in threat of malpractice suit. With these data it is not possible to reject or confirm the hypothesis that malpractice has a deterrent effect (Harvard, 1990).

Several studies have attempted to measure changes in physician practice patterns in response to liability, but again the evidence remains inconclusive because of data limitations. The most detailed data are from surveys in which physicians were specifically asked about their response to liability, which introduces obvious possibilities for bias (for example, Reynolds, Rizzo and Gonzalez, 1987). Surveys designed for other purposes are less likely to be biased, but typically contain less detailed information on pertinent characteristics of physicians' practices. Further, it is not possible to distinguish cost-justified measures—the "prevention" that the tort system is intended to induce—from any "defensive" measures that are not cost-justified and would not have been chosen by a fully informed patient, given their health insurance coverage. If traditional fee-for-service reimbursement constrains physicians' ability to charge for additional time spent per procedure, in the absence of liability this may create incentives for physicians to spend less than the optimal amount of their own time on the average patient encounter but perform at least the optimal number of tests and procedures. If so, then an increase in physician time in response to liability could be interpreted as positive deterrence, but an increase in tests and X-rays could be interpreted as wasteful defensive medicine.

With these limitations in mind, I will summarize some recent work on these issues. The full analysis, including attempts to distinguish effects of claim frequency and severity, is reported in Danzon (1991), Danzon, Pauly and Kington (1990), and Danzon and Aiuppa (1990). The evidence summarized here comes from surveys of a nationally representative sample of physicians conducted in 1976, 1978 and 1983. Since different physicians were surveyed each year and sample size in each locality is small, these data permit analysis of cross-sectional variation in practice patterns in response to differences in liability across states, but not time-series response to the increase in malpractice costs over time. The liability climate is measured by the rate for basic limits of insurance coverage.

In response to increased liability, physicians have increased their expenditure on insurance but less than in proportion to expected costs. Physicians therefore bear more uninsured risk in states with high liability costs.

Liability cost increases appear to be passed along promptly to fees charged and rates of insurance reimbursement. The elasticities of routine office and hospital visit fees with respect to insurance rates are between 0.1 and 0.2. This is more than sufficient to pass on the cost of increased expenditure on insurance, assuming no change in volume, since on average physicians spend roughly 4 percent of gross revenues on insurance. Elasticities of reimbursement rates paid by third party health insurers are similar to fee elasticities in the 1970s, but somewhat lower in the 1980s, probably reflecting the increasingly aggressive attempts by third party payers to control medical care costs. Indeed by 1983, the ratio of Medicaid reimbursement to usual fees is negatively related to liability costs, possibly because Medicaid reimbursement has lagged most in urban areas which also tend to have high malpractice costs. This suggests that

if, as is often alleged, liability has made physicians less willing to treat Medicaid patients, the relatively tighter constraints on cost pass-through to Medicaid is an important contributing factor. But in general, this pass-through of malpractice costs into higher fees and reimbursement levels appears to have been quite rapid and direct, without requiring an adjustment in physician stocks. The number of physicians per capita, by county, and the rate of change between these years is unrelated to either levels or rates of growth of liability costs.

The excess of the fee elasticities over the level required to fully pass-through the costs of malpractice insurance may reflect several factors: increased physician time per patient encounter; a compensating differential for exposure to uninsured claim costs; uninsurable time and non-monetary costs associated with the risk of suit; and a reduction in volume in response to higher fees. There is weak evidence that liability induces physicians to spend more time per patient visit—a possible indicator of more prevention effort.⁶ By contrast, the frequency of lab tests or procedures is significantly negatively related to liability costs; the frequency of X-rays or fluoroscopies is positively related to malpractice costs in the 1970s but the relationship is negative in 1983. Total number of office visits is negatively related to liability costs, which is consistent with standard constraints on demand and not consistent with unlimited ability or willingness of physicians to shift demand for defensive purposes, or with demand shifting outward in response to perceived improvement in quality of care or higher expected compensation in the event of injury. This evidence is thus not consistent with defensive ordering of a lot of extra tests and visits, at least in connection with ambulatory care.

In a survey conducted immediately after the 50–100 percent increases in malpractice rates that occurred in 1985, roughly 20 percent of physicians reported that they had stopped performing high risk procedures in response to liability. This response is significantly related to the cost of liability insurance, which gives it added credibility. Older physicians and generalists are more likely to cut back on high risk cases than younger physicians and board-certified physicians (Danzon and Aiuppa, 1990). Since recent training and board certification may both be indicators of technical competence, these patterns are potentially consistent with an efficient increase in specialization, although some patients may be inconvenienced, particularly in rural areas which have fewer specialists per capita than urban areas (Newhouse et al., 1982).

On average, physicians' net money incomes were not adversely affected by liability costs through 1983, which is consistent with a rapid pass-through of cost increases to fees and no effects on the geographic distribution of physicians on average. However, several caveats are in order. First, even if net money

⁶This is probably a downward-biased estimate of the effect of liability on total time per patient episode. The data refer to time for an "intermediate" office visit, but this type of visit might be reclassified and billed as a "comprehensive" visit if significantly more time were spent. Further, estimates of effects on mean length of visit cannot reflect reallocation of time from high risk to lower risk patients or procedures.

incomes have been maintained, physicians' real utility may be lower, both because of slightly longer hours of work and because of increased exposure to uninsured risk in high liability areas. Second, as medical care markets have become more competitive in the 1980s, the ability to pass through the sharp premium increases of the mid-1980s may have become much more limited, resulting in negative effects on net income and a fortiori on utility. (Unfortunately, the 1986 survey did not report the physician's net income.) Third, these estimates of mean effects may obscure significant distributional effects, with losses to some physicians offset by gains to others. A combination of these factors may explain why physicians press for changes in the malpractice system, even if their fees and net money incomes have, on average, risen to keep pace with malpractice insurance costs.

Evaluations of Alternatives

The evidence reported here indicates that previous estimates of defensive medicine cost were probably upward-biased and some potentially efficient changes in practice have occurred in response to liability, at least in the ambulatory care settings for which data are available. Casual evidence indicates that hospitals have also strengthened their quality assurance and risk management programs in response to liability, although such changes have not been subject to systematic analysis. But the incidence of negligent injury shows that liability is not a perfect deterrent. The fundamental policy questions remain unanswered—first, whether the malpractice system yields benefits that outweigh its costs, and second, whether alternatives would yield higher net benefits. Two alternatives will be briefly discussed: mandatory experience rating of liability insurance premiums and no-fault compensation programs.

Experience Rating

A puzzle of medical malpractice insurance is the relatively small role of co-payment, explicit experience rating (that is, premiums based on prior claims), or other features to control moral hazard. Medical malpractice policies typically do not have deductibles or co-insurance. Rates are based on medical specialty; geographic location; whether the physician performs certain high-risk procedures, such as surgery by general practitioners (but not the volume of such procedures); and part-time practice. Rating related automatically to prior experience of malpractice claims is relatively uncommon. However, bad claims experience may lead to restrictions on coverage or to nonrenewal by companies that are more selective and offer lower rates.⁷

⁷Experience rating by market segmentation, with companies specializing in writing good or bad risks, is common in other lines of insurance, like automobile liability.

Companies that do base rates on prior experience usually impose a temporary surcharge related to paid claims or “valid” paid claims, as determined by review of the claims by medical experts, rather than an automatic adjustment for all claims filed. This is consistent with the hypothesis that demand for experience-rated policies is constrained by demand for insurance against the risk of error by claimants, the courts and the settlement process in filings and findings of negligence (Danzon, 1985b). The New York data suggest that there is a significant rate of invalid claims. If this finding carries over to settlements and verdicts, it would create a valid demand for insurance against the risk of judicial error.

Several studies have confirmed that the distribution of claims against physicians within a given specialty is highly skewed, with a small number of physicians accounting for a larger number of claims than would be expected if the probability of a claim were constant and the judicial process entirely random (Rolph, 1981; Nye and Hofflander, 1988; Sloan, 1989). Based on these findings, these authors have argued for more experience rating. But Ellis, Gallup and McGuire (1989) show that rating based on Bayesian conditional means with five years of experience would move premiums only modestly towards actuarially fair rates on average, while introducing inequities between physicians with identical underlying risk and exposing physicians to considerable financial risk. They estimate that under such a rating scheme, a single paid claim results in a four-fold increase in premiums for most medical specialties. Thus, risk aversion in the presence of judicial error may reduce demand for experience-rated policies.

Several factors may limit the demand for policies with formal deductibles and co-payment. First, being sued entails uninsurable costs of time, in addition to anxiety and threat to reputation. The dollar-equivalent measure of such costs is probably several thousand dollars per claim and acts as a per claim deductible. Second, the potential for a claim in excess of the policy limit implies additional uninsured risk. Third, deductibles undermine the malpractice insurer’s incentives to defend claims that could be settled within the limits of the deductible (Danzon, 1985b). Fourth, as noted earlier, the more selective, lower cost insurers do respond to a persistent record of adverse claims by imposing restrictions on the physician’s practice as a condition of coverage, policy exclusions and ultimately nonrenewal. Whether these hidden uninsured costs add up to less or more than the socially optimal degree of co-payment is an important but unanswered question. Imposing more experience-rating or co-payment would entail real social costs in uninsured risk to physicians and possibly defensive responses, like a refusal to take high risk patients.

No-Fault Programs for Iatrogenic Injuries

No-fault programs for iatrogenic injuries would provide compensation for injuries caused by medical care, without regard to the fault or negligence of the medical provider. However, such proposals are typically not simply a shift from

negligence to strict liability on medical providers. In addition to changing the rule of liability, claims would be adjudicated by a special administrative agency rather than the courts, and benefits would be payable according to a schedule, with little or no payment for pain and suffering and full offset of compensation available from collateral sources. The intent is to reduce delay and expense in claims adjudication and permit more even payment to more victims of iatrogenic injury, most of whom are automatically excluded from potential compensation by the fault requirement. The model is the workers' compensation program, which adjudicates claims with lower litigation expense and shorter delays than tort liability, at least for traumatic injury cases. However, whereas workers' compensation is financed by experience-rated premiums paid by employers,⁸ some medical no-fault proposals propose financing through a broad-based tax.

Virginia and Florida have enacted programs to provide compensation for severe birth-related neurological injuries caused by medical care, regardless of negligence or fault of the medical providers. Claims are administered through workers' compensation agencies, with scheduled benefits. Financing is partly through levies on physicians and hospitals, which are not experience rated, and an assessment on insurers writing other lines of insurance in the state. Thus effectively these are social insurance programs for a very restricted set of beneficiaries. North Carolina has considered a similar plan. More comprehensive programs for all iatrogenic injuries have also been proposed, including a recent proposal by the New York State Commissioner of Health (Sack, 1990).

The efficiency effects of a broad-based no-fault scheme for medical injuries must be measured in terms of effects on total social costs of injuries, including the real cost of injuries, prevention and overhead costs. It is also important to identify which dimensions of the proposed program generate savings, specifically, to distinguish the effects of the no-fault rule of liability from other changes. Of course, equity considerations are also relevant.

Proponents claim three sources of savings from no-fault programs. First, benefit payments would be for economic loss (wage loss, medical and other monetary costs) only, with offset of other private and social insurance programs where possible. To the extent compensation for pain and suffering is not optimal compensation (Cook and Graham, 1977, Danzon, 1984a), elimination of such payment is an efficiency gain. But this can be accomplished by scheduled tort awards for pain and suffering, without changing the fault-based rule of liability. Offset of collateral sources is largely a shift in costs to other sources, rather than a real reduction. The net efficiency effect depends on effects of

⁸Under the workers' compensation system, employers are strictly liable for compensation for injuries "arising out of or in the course of employment," without regard to employer or employee negligence. Claims are administered through special agencies; benefits are limited to "economic loss" i.e. medical, rehabilitation and wage loss. Employers are required to carry insurance or self-insure. Premiums are experience-rated to a degree that increases with firm size, reflecting statistical credibility. Large firms are fully self-rated.

collateral source offset on incentives to litigate rather than settle claims and on incentives for injury prevention.⁹

Second, it is argued that eliminating the issue of fault or negligence as a condition for compensation would reduce litigation costs. This question cannot be resolved by analogy to the workers' compensation system. The occurrence of an on-the-job injury is a well-defined event, at least for single-event traumatic injuries. Litigation costs are considerably higher for occupational disease claims, where the role of occupational exposure must be distinguished from life-style, genetic and other contributing factors. Similarly, in the medical context, it would be costly to determine whether the injury was caused by medical care, rather than being the unavoidable outcome of the underlying condition for which treatment was sought. Litigating over cause could be as costly as litigating over fault, and in many cases the issue would probably reduce to showing some "defect" in treatment, which could be operationally similar to proving negligence (Epstein, 1976).

Third, the effect on cost-effective deterrence and defensive responses is highly uncertain and depends on the financing and administration of the program. Imposition of strict liability on physicians with experience-rated premiums would expose them to significant uninsurable costs of spending time defending a much larger number of claims and possibly significant risk of being erroneously rated, in the likely event that claimants and courts often err in distinguishing iatrogenic injuries from bad luck. Because all iatrogenic injuries would be compensable, regardless of fault, between a three- and ten-fold increase in the number of claims might be anticipated, based on the New York and California data, even under the optimistic assumption of no invalid claims. Defensive responses could be considerable.

A second possible financing mechanism is no-fault (strict) enterprise liability on hospitals or large medical groups, rather than individual physicians. Enterprise liability has the advantage of reducing variance of expected loss to defendants, because the hospital effectively pools over a much larger volume of cases than the individual physician. Enterprise liability could also improve deterrence, if hospitals or large groups have better information on true care and hence can more accurately impose sanctions on individual physicians than can the courts or liability insurers under physician-based liability, or if pooling liability increases incentives for efficient risk reduction.¹⁰ But if these arguments for enterprise liability are valid, they apply regardless of whether the liability rule is strict liability or negligence. Such transfers of liability from physicians to hospitals could be adopted by voluntary contract. The fact that this has not occurred widely with the physicians who are independent contractors with admitting privileges at more than one hospital, although it is the

⁹These and other tort reforms are evaluated in Danzon (1984a, 1985a).

¹⁰Stiglitz (1990) shows that peer monitoring in a competitive credit market improves borrowers' welfare.

norm in hospitals where the physicians are salaried employees, suggests that there may be other disadvantages.

A third alternative for financing a no-fault program is a broad-based tax on medical providers, insurance companies, or general revenues, as in Virginia and Florida. Financing by a tax on medical providers retains the internalization of costs to the industry causing the injuries (assuming accurate adjudication of claims), but financing by a broad-based tax eliminates all internalization of costs and deterrence. Such a program is a pure social insurance program. It is arguably neither efficient nor equitable to single out victims of medical injury for special compensation, unless there is a deterrent benefit. Compensation can be provided more cheaply through broad-based private insurance and social insurance programs, like Social Security Disability, Medicare and Medicaid, which provide compensation without regard to cause. As noted at the outset, incurring the cost of determining that a particular condition was caused by medical care, rather than genetic or other factors, is worthwhile only if there is some deterrent benefit.

An element of deterrence could be preserved with broad-based funding if the no-fault program brought suits for negligence against medical providers (subrogation). Decoupling compensation and deterrence could in theory provide prompt and efficient compensation to victims of medical injuries, regardless of fault, while preserving deterrence. For injuries involving nonmonetary loss, the optimal compensatory award to the victim is not necessarily equal to the optimal deterrent penalty on the defendant (Spence, 1977).

However this solution raises several problems. First, if subrogation were only for negligent injuries, significant additional tax financing would be necessary to cover the costs of the majority of nonnegligent injuries. The availability of tax revenues to cover deficits would undermine the incentives of the program to bring claims against providers and hence reduce deterrence. On the other hand, if subrogation claims were filed for all claims that would be brought in a negligence-based tort system, total overhead costs on such claims could increase, since one hearing to determine medical cause and another to determine negligence would be required. Second, as noted earlier, there is no efficiency or equity rationale for that part of the program that would be a pure social insurance program confined to victims of medical injuries.

Concluding Comments

The basic rationale for medical malpractice suits is to improve incentives for safety in the presence of asymmetric information between patients and physicians. The evidence of a significant rate of negligent injury, invalid claims and physicians' preference for insurance policies with minimal explicit co-payment or experience rating indicate that the efficiency of the malpractice system is severely constrained by imperfect information on the part of courts

and insurers. But just as imperfect information undermines the efficient functioning of the market, imperfect information undermines the efficient functioning of the liability system. The fundamental problem is that changing the liability rule does not correct the information asymmetry.

If the tort system's deterrence is considered too uncertain to warrant imposing liability on individual physicians or hospitals—and the evidence on this point is inconclusive—then there is no strong case for singling out victims of iatrogenic injury for a special compensation program. The adequacy of compensation for iatrogenic injury is simply part of the broader question of the efficiency and equity of the existing network of private and social insurance programs.

References

- American Medical Association (AMA)**, "A Proposed Alternative to the Civil Justice System for Resolving Medical Liability Disputes: A Fault-based Administrative System." AMA/Specialty Society Medical Liability Project, Chicago, 1988.
- Butler, Richard J., and John D. Worrall**, "Workers' Compensation: Benefits and Injury Claims Rates in the Seventies," *Review of Economics and Statistics*, November 1983, 65, 580–589.
- California Medical Association (CMA)**, *Medical Insurance Feasibility Study*, San Francisco: Sutter, 1977.
- Cook, Philip J., and Daniel A. Graham**, "The Demand for Insurance and Protection: The Case of Irreplaceable Commodities," *Quarterly Journal of Economics*, February 1977, 91, 143–156.
- Danzon, Patricia**, "Tort Reform and the Role of Government in Private Insurance Markets," *Journal of Legal Studies*, 1984a, 13, 517.
- Danzon, Patricia**, "The Frequency and Severity of Medical Malpractice Claims," *Journal of Law and Economics*, 1984b, 27, 115–148.
- Danzon, Patricia**, *Medical Malpractice: Theory, Evidence and Public Policy*. Cambridge: Harvard University Press, 1985a.
- Danzon, Patricia**, "Liability and Liability Insurance for Medical Malpractice," *Journal of Health Economics*, December 1985b, 4, 309–331.
- Danzon, Patricia**, "New Evidence on the Frequency and Severity of Medical Malpractice Claims," R-33410-ICJ, The RAND Corporation, Santa Monica, 1986.
- Danzon, Patricia**, "The 'Crisis' in Medical Malpractice: A Comparison of Trends in the United States, Canada, the United Kingdom and Australia," *Law, Medicine & Health Care*, May 1990, 18.
- Danzon, Patricia**, "Liability for Medical Malpractice: Incidence and Incentive Effects," unpublished paper, The Wharton School, University of Pennsylvania, Health Care Department, 1991.
- Danzon, Patricia, and Thomas A. Aiuppa**, "Effects of Liability on Physicians' Practice Patterns," unpublished paper, The Wharton School, University of Pennsylvania, Health Care Department, 1990.
- Danzon, Patricia, and Lee Lillard**, "Settlement out of Court: The Disposition of Medical Malpractice Claims," *Journal of Legal Studies*, 1983, 12.
- Danzon, Patricia, Mark V. Pauly, and Raynard S. Kington**, "The Effects of Malpractice Litigation on Physicians' Fees and Incomes," *American Economic Review*, May 1990, 80, 122–127.
- Deweese, Don, Peter Coyte, and Michael Trebilcock**, *Canadian Medical Malpractice Liability: An Empirical Analysis of Recent Trends*, University of Toronto, June 1989.

Doherty, Neil A., and Georges Dionne, "Risk Pooling, Contract Structure and Organizational Form of Insurance Firms," working paper, Department of Insurance, The Wharton School, University of Pennsylvania, August 1988.

Ellis, Randall P., Cynthia L. Gallup, and Thomas McGuire, "Should Medical Professional Liability Insurance Be Experience Rated?" *Journal of Risk and Insurance*, 1989.

Epstein, Richard A., "Medical Malpractice: The Case for Contract," *American Bar Foundation Research Journal*, 1976, 76.

Harvard Medical Practice Study, "Patients, Doctors, Lawyers: Medical Injury, Malpractice Litigation and Patient Compensation in New York," 1990.

Henderson, James A., Jr., and Theodore Eisenberg, "The Quiet Revolution in Products Liability: An Empirical Study of Legal Change," *UCLA Law Review*, 1988, 37, 479-553.

Kakalik, James S., and Nicholas M. Pace, "Costs and Compensation Paid in Tort Litigation," R-3391-ICJ, The RAND Corporation, Santa Monica, 1986.

Levit, Katharine R., and Mark S. Freeland, "National Medical Care Spending," *Health Affairs*, Winter 1988.

Manning, Willard G., et al., "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *American Economic Review*, June 1987, 77, 251-277.

Newhouse, Joseph, et al., "Does the Geographical Distribution of Physicians Reflect Market Failure?" *Bell Journal of Economics*, 1982, 13.

Nye, Blaine F., and Alfred E. Hofflander, "Experience Rating in Medical Malpractice

Insurance," *The Journal of Risk and Insurance*, 1988, 48, 247-260.

Reynolds, Roger, John Rizzo, and Martin Gonzalez, "The Cost of Medical Professional Liability," *Journal of the American Medical Association*, 1987, 257, 2776-2779.

Rolph, John E., "Some Statistical Evidence on Merit Rating in Medical Malpractice Insurance," *Journal of Risk and Insurance*, 1981, 48, 247-260.

Sack, Kevin, "No-Fault for Doctors is Called Feasible," *The New York Times*, March 1, 1990, p. 83.

Shanley, M. G., and M. A. Peterson, "Comparative Justice: Civil Jury Verdicts in San Francisco and Cook Counties, 1959-1980," R-3006-ICJ, The RAND Corporation, Santa Monica, 1983.

Shavell, Steven, "Strict Liability vs. Negligence," *Journal of Legal Studies*, 1980, 9.

Sloan, Frank A., et al., "Medical Malpractice Experience of Physicians: Predictable or Haphazard?" *Journal of the American Medical Association*, 1989, 23.

Sloan, Frank A., Paula M. Mergenhagen, and Randall R. Bovbjerg, "Effects of Tort Reforms on the Value of Closed Medical Malpractice," *Journal of Health Politics, Policy & Law*, 1989, 14.

Spence, A. Michael, "Consumer Misperceptions, Product Failure, and Product Liability," *Review of Economic Studies*, October 1977, 44, 561-572.

Stiglitz, Joseph E., "Peer Monitoring and Credit Markets," *The World Bank Economic Review*, September 1990, 4, 351-366.

Zeckhauser, Richard, "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, 1970, 2, 10-26.