

**Reconsidering the Harvard Medical Practice Study
Conclusions About
The Validity of Medical Malpractice Claims:
A Literature Review**

Tom Baker
University of Connecticut

Over fifteen years after first reporting to the State of New York, the Harvard Medical Practice Study (HMPS) continues to have a significant impact in medical malpractice policy debates. In those debates the HMPS has come to stand for four main propositions. First, “medical injury ... accounts for more deaths than all other kinds of accidents combined” and “more than a quarter of those were caused by substandard care.” Second, the vast majority of people who are injured as result of substandard care do not file a claim. Third, “a substantial majority of malpractice claims filed are not based on provider carelessness or even iatrogenic injury.” Fourth, “whether negligence or a medical injury had occurred ... bore little relation to the outcome of the claims.” In light of this continuing reliance on the HMPS and the follow up closed claim study, this article reviews the evidence regarding their findings about the validity of medical malpractice claims. The results of this review are as follows: First, the finding that most eligible people do not bring medical malpractice claims is well supported and confirmed by other studies using both similar and very different research methods. Second, the finding that most medical malpractice claims are not based on either iatrogenic injury or provider negligence stands on a small and precarious empirical base. Indeed, the HMPS data are as likely to support a very different finding, namely that most malpractice claims are reasonably related to medical management injuries and provider negligence. Finally, the finding from the follow-up closed claim study rests on an even weaker base and is contradicted by a large body of research on closed medical malpractice claims. In fact, the research reviewed in part IIIB of this article suggests that the legal system filters out most of the weaker claims.

**Reconsidering the Harvard Medical Practice Study Conclusions About
The Validity of Medical Malpractice Claims: A Literature Review**

Tom Baker, University of Connecticut¹

Over fifteen years after first reporting to the State of New York, the Harvard Medical Practice Study (HMPS) continues to have a significant impact in medical malpractice policy debates. In those debates the HMPS has come to stand for four main propositions. First, “medical injury ... accounts for more deaths than all other kinds of accidents combined” and “more than a quarter of those were caused by substandard care.”² Second, the vast majority of people who are injured as result of substandard care do not file a claim.³ Third, “a substantial majority of malpractice claims filed are not based on provider carelessness or even iatrogenic injury.”⁴ Fourth, “whether negligence or a medical injury had occurred ... bore little relation to the outcome of the claims.”⁵

Medical malpractice researchers have long known that the HMPS provides far stronger support for the first two of these propositions than for the last two; the HMPS was not designed or powered to reach strong conclusions about the validity of medical malpractice claims.⁶ Despite the fact that these latter conclusions were the weakest aspects of the HMPS, many policymakers and even some well regarded researchers give them the same “gold standard” deference that appropriately is given to the findings regarding the rate of medical malpractice and the rate of medical malpractice claims.

For example, in a 2005 White Paper on medical liability reform, the Joint Commission on Accreditation of Healthcare Organizations cited the HMPS for the conclusion that “only 17

¹ Thank you to Bernard Black, Michelle Mello, Peter Siegelman, Charles Silver and an anonymous reviewer for helpful comments on an earlier draft and to Yan Hong and Kathleen Sorrentino for excellent research assistance.

² Weiler et al 1993 at 55.

³ Id at 75.

⁴ Id at 140.

⁵ Brennan et al 1996.

⁶ E.g. Saks 1994; Bovbjerg 1993; White 1994. Cf. Mehlman 1991.

percent of claims appeared to involve negligent injury.”⁷ Similarly, in a March 2005 report on medical liability, the Joint Economic Committee of the U.S. Congress cited the same finding and also referred to the well known HMPS analogy: “it is similar to a situation in which a traffic officer is giving tickets to large numbers of motorists who are not speeding.”⁸ In a December 2004 fact sheet, the American Medical Association cited the HMPS follow-up closed claim study for the proposition that “medical liability payments correlate with disability not negligence.”⁹ In addition, in the 2002 volume of the Journal of Health Economics, two respected economists cited the HMPS for the proposition that “five-sixths of the cases that receive compensation have no evidence of negligence.”¹⁰ In these and similar circumstances, neither the acknowledged weaknesses of the HMPS findings nor the contrary evidence are disclosed or discussed.

In light of this continuing reliance on the HMPS and the follow up closed claim, this article reviews the evidence regarding their findings about the validity of medical malpractice claims. Much of the analysis tracks that of earlier researchers. The value-added lies in collecting this analysis in one, easily accessible place, in setting out for the first time a detailed critique of the follow-up study of the 47 closed claims that were matched to the HMPS hospital records, and in comparing the results of the follow up study with the results of other research on medical malpractice claims.

The results of this review are as follows: First, the finding that most eligible people do not bring medical malpractice claims is well supported and confirmed by other studies using both similar and very different research methods. Second, the finding that most medical malpractice claims are not based on either iatrogenic injury or provider negligence stands on a small and precarious empirical base. Indeed, the HMPS data are as likely to support a very different finding, namely that most malpractice claims are reasonably related to medical management

⁷ Joint Committee 2005.

⁸ Joint Economic Committee 2005.

⁹ American Medical Association 2004.

¹⁰ Kessler & McClellan 2002.

injuries and provider negligence. Finally, the finding from the follow-up closed claim study rests on an even weaker base and is contradicted by a large body of research on closed medical malpractice claims. In fact, the research reviewed in part IIIB of this article suggests that the legal system filters out most of the weaker claims.

I. The HMPS Finding Regarding the Low Rate of Claiming

The HMPS included an impressive effort to determine the extent of medical malpractice and compare it with the extent of medical malpractice claiming. The researchers put 31,000 randomly selected hospital records from 51 hospitals in New York through a two-stage review process that identified the rate of negligent medical injuries among that sample. They applied that rate to the total number of hospitalizations in New York to estimate the total number of negligent medical injuries in New York during the research period and then compared that number to the total number of malpractice claims filed in New York during the same period.

The HMPS estimated that there were about 27,000 injuries that resulted from medical negligence in hospitals in New York during 1984 and that only about 3800 claims were filed under malpractice insurance policies covering the year 1984.¹¹ This means that there were more than seven negligent injuries for every medical malpractice claim and, accordingly, that most patients injured as a result of negligent medical management do not make a claim. This research finding is consistent with the comparable findings reached in other large population-based studies using similar record-review research methods,¹² as well as ethnographic research in which researchers closely observed hospital practice over an extended period.¹³

Taken together, these studies demonstrated that medical malpractice is real and that it is widespread. Over 75,000 people die every year from medical malpractice, more than the total

¹¹ Weiler et al 1993 at 70.

¹² Danzon 1985 (using data from Mills et al 1977) (California); Wilson et al 1995 (Australia); Thomas, Studdert, Burstin et al 2000 (Colorado and Utah).

¹³ Andrews et al 1997.

number of deaths from automobile and workplace accidents combined.¹⁴ These results also demonstrated that most medical malpractice victims do not file claims and that the real medical malpractice litigation problem is “the malpractice system is too inaccessible, rather than too accessible, to the victims of negligent medical treatment.”¹⁵

II. The HMPS Finding Regarding the Validity of Filed Claims

The HMPS also attempted to evaluate the validity of medical malpractice claims. The researchers collected all of the medical malpractice claims filed in New York during the study period and subsequent years and then matched those claims against the 31,000 hospitalizations that they studied. This match identified a surprisingly small number of claims – only 47 – that arose out of these hospitalizations.¹⁶ Out of these 47 claims, HMPS reviewers had identified a negligent medical management injury in only 8 of the matched hospital records.¹⁷ This 8 out of 47 number captured the imagination of advocates of tort reform, leading them to assert that most medical malpractice claims have no basis.¹⁸

In fact the HMPS does not support that assertion. The HMPS was well designed to produce a conservative, lower bound estimate of the rate negligent medical management injuries in a large population of hospitalizations, but it was not well designed to make judgments about the validity of a small number medical malpractice claims or to draw significant policy conclusions based on those claims. Understanding this point requires more careful attention to the HMPS research design than even most informed observers have given to this matter.

¹⁴ Weiler et al 1993 at 55.

¹⁵ Weiler et al 1993 at 76.

¹⁶ Id. They later found four additional claims but did not include those claims in the analysis. As explained in note 60, including those four additional claims would not have materially changed the results.

¹⁷ Localio et al 1991 at 248, table 3.

¹⁸ Recent references to this HMPS finding and advocating tort reform include Brody 2004 and Greater New York Hospital Association 2005.

*II.A. The HMPS Review Process*¹⁹

The HMPS review was a two-stage process. In the first stage, medical record analysts reviewed the record of each hospitalization using a detailed screening form that guided them in deciding whether the hospitalization contained one or more clinical indications of a possible iatrogenic injury (i.e. an “injury caused by medical management”). If the record did not contain one of these indications, the hospitalization was classified as not involving a medical management injury, and the record related to the hospitalization was not reviewed further. The first stage review determined that there was an indication of a possible iatrogenic injury in thirty percent of the hospitalizations.²⁰

If the record did contain an indication of a possible iatrogenic injury, the hospitalization advanced to the second stage of the review process. In the second stage, two physicians independently reviewed the record. Each physician first confirmed that the medical record analyst had appropriately interpreted the hospital record. If not, the hospitalization was classified as not involving a medical management injury, and the record was not reviewed further. If the medical record analyst had appropriately interpreted the record, each physician reviewer then assigned the hospitalization an “adverse event” score from 1 to 6 indicating the strength of the evidence of a medical management injury.²¹ If that score was 2 or more, the physician answered a series of questions regarding the nature of the injury and then completed a section of the review form captioned “Is there evidence for negligence?”

The first step in the negligence section directed the reviewer to answer “yes” or “no” to the following question: “Was this AE possibly due to a reasonably avoidable error, or carelessness by either an individual or medical care system, or both?” If the answer was “no,” the

¹⁹ The description of the review process is derived from HMPS 1990 chapter 5.

²⁰ HMPS 1990 at 6-7.

²¹ 1 = “little or no evidence for management causation.” 2 = “slight to modest evidence for management causation.” 3 = “management causation not quite likely; less than 50-50 but close call.” 4 = “management causation more likely than not; more than 50-50 but close call.” 5 = “strong evidence for management causation.” 6 = “virtually certain evidence for management causation.” From form appended to chapter 5 of HMPS 1990.

review was over, and the hospitalization was classified as not involving negligence. Importantly, a “no” answer by either reviewer was sufficient to classify the hospitalization as one that did not involve negligence, even if the other reviewer was very confident that there was negligence.

If the answer to the first negligence question was “yes,” the physician then answered a series of specific questions that ended by asking the reviewer to re-assess the opinion that the injury was possibly due to negligence.²² If the physician did change his or her opinion, the review was over, and the hospitalization was classified as not involving negligence. If the physician did not change his or her opinion, the physician then rated the degree of confidence in the evidence for negligence on a one to six scale that was comparable to the confidence scale for the medical management injury determination.²³

A supervising physician then examined the forms filled out by the two physicians, verified that the forms related to the same hospitalization, and computed the averages of the reviewers’ confidence scores. The hospitalization was classified as involving a medical management injury only if the average of the confidence scores for that determination was greater than 3.5, and it was classified as involving negligence only if the average of the confidence scores for that determination was greater than 3.5. (The practical effect of the 3.5 cutoff is that it a decision by either physician reviewer that there was not an medical management injury is sufficient to rule out a medical management injury, even if the other reviewer was highly confident that a medical management injury had occurred, and likewise with the negligence classification.) The second stage determined that there was strong evidence of a medical

²² “After having considered the factors in 8.1-8.7 you might have reassessed whether negligence occurred. If you feel that there is NO negligence, CHECK THE SPACE ON THE RIGHT AND GO TO Q. 11.” Id.

²³ 1 = “little or no evidence for negligence.” 2 = “slight to modest evidence for negligence.” 3 = “negligence not quite likely; less than 50-50 but close call.” 4 = “negligence more likely than not; more than 50-50 but close call.” 5 = “strong evidence for negligence.” 6 = “virtually certain evidence for negligence.” From form appended to chapter 5 of HMPS 1990.

management injury in just under four percent of hospitalizations and strong evidence that the injury resulted from negligence in one percent of hospitalizations.²⁴

The process just described represents a very conservative approach to identifying negligent medical management injuries. This approach was well suited to the task of providing a solid, lower bound estimate of the rate of those injuries, an estimate that would be trusted by medical providers. Physicians designed and supervised the study. Physicians trained the medical record analysts and conducted the second stage review. The scoring process meant that, in effect, either physician reviewer could veto the medical management injury or negligence finding of the other. And the review process contained a variety of safeguards against a mistaken conclusion that there was a medical management injury or negligence, and essentially no safeguards against a mistaken conclusion that there was not.²⁵ This means that, if anything, the HMPS estimates about the rate of medical management injuries and medical malpractice are low (as the HMPS researchers were aware).²⁶ As a result, the HMPS, together with an earlier California study and a later Colorado/Utah study using similar research methods,²⁷ provides persuasive evidence that the real medical malpractice problem is medical malpractice, not medical malpractice litigation.

II.B. HMPS and the Merits of Individual Medical Malpractice Claims

The HMPS was not well suited to the task of evaluating the merits of a small sample of medical malpractice claims. There are three main reasons for this: the lack of controls to avoid false negatives discussed in the preceding section, the problem of reviewer reliability, and the inherent limits of relying exclusively on hospital records.

²⁴ HMPS 1990 at 6-10.

²⁵ For example, the first stage review was verified only if the reviewer sent the record on to the second stage; the physician reviewers were asked to reconsider their opinion only if they concluded that there was evidence of negligence; and the physician reviewers were asked to provide a confidence rating only for an opinion that there was an iatrogenic injury and that the injury was caused by negligence, but not for an opinion that there was not.

²⁶ HMPS researchers later estimated that making comparatively modest changes to the review process would increase the rate of negligent medical management injuries by 70%. Runciman et al 2000.

²⁷ Danzon 1985 (using data from Mills 1977) (California); Thomas, Studdert, Burstin et al 2000 (Colorado and Utah). See also Andrews et al 1997 (ethnographic approach).

II.B.1. The Reliability Problem

The reviewer reliability problem refers to the fact that repeated reviews of the same records showed that physician reviewers did not always agree about which records contained good evidence of a medical management injury, and they agreed less often about which records contained good evidence of negligence.²⁸ Because the team reviewed so many hospital records, the reliability of judgments in individual cases does not undercut their finding about the rate of errors in hospitalizations, but it does undercut their finding that most medical malpractice claims lack merit.

HMPS researchers were cognizant of the limits of making negligence judgments based on hospital record reviews. The researchers conducted a pilot study that measured the degree of agreement between independent reviewers examining the same records.²⁹ The pilot study found a high level of agreement among reviewers regarding the presence or absence of a medical management injury, but a lower rate of agreement among reviewers regarding the existence of substandard care.³⁰ This lower rate of agreement reflects the difficulty of determining whether substandard care was given, as well as differences of opinion regarding what constitutes substandard care. Like other researchers, the HMPS researchers decided that the lower rate of agreement with regard to substandard care was acceptable for drawing conclusions about the rate of substandard care in a population because later independent reviews consistently found the same rate of substandard care, even though the reviewers did not always agree about which specific cases fell into this category.³¹

²⁸ HMPS 1990 at 5-28, noting that there was a “lack of reliability regarding judgments of negligence.” See also Brennan et al 1989; Brennan et al 1991.

²⁹ Brennan et al 1989.

³⁰ Id at 1156-57

³¹ Brennan et al 1991 at 375; Localio et al 1991 at 249. There is an extensive literature examining the problem of reviewer reliability in studies of medical care. E.g., Goldman 1992; Hayward et al 1998; Hayward & Hofer 2001; Hofer et al 2000; Posner et al 1996; Localio et al 1996; Thomas et al 2002; Wilson et al 1992. Cf. O’Neil et al 1993. In general, this research supports the HMPS researchers’ judgment that reviewing a very large sample of hospital records is an acceptable way to draw conservative conclusions about the rate of injuries caused by substandard care in a population. Nevertheless, the research questions

The HMPS researchers also conducted a similar test of reviewer reliability at the beginning of the New York study.³² The New York test illustrates why the HMPS record review process is not a very reliable way to form strong conclusions about the validity of individual medical malpractice claims.

In the New York test, the HMPS research team took all the records that passed through the medical record analyst screen at two hospitals and put them through the physician review and classification process twice, using two different reviewers each time.³³ (None of the four physicians that reviewed each record knew that the record was receiving any more attention than any other record.) The two teams examined a total of 318 records. Out of that group, the two sets of physician reviewers identified the same medical management injuries 35 times and different medical management injuries 34 times. In percentage terms, this means that the review teams agreed that there was not a medical management injury in 78 percent of the cases, but they disagreed as often as they agreed about whether there was a medical management injury in the remaining 22 percent of the cases. Moreover, the two sets of reviewers disagreed far more often than they agreed about which of the medical management injuries resulted from negligence. They identified the same negligent injuries only four times and different negligent injuries 21 times.³⁴

the reliability of particular judgments made by particular reviewers in particular cases. E.g. Hayward & Hofer 2001.

³² Brennan et al 1991 at 374. See also Localio et al 1991 at 249

³³ The description of the method comes from HMPS 1990 at 5-27 to 5-28.

³⁴ Since both teams concluded that there was evidence of a medical management injury in only 35 cases, and a single team concluded that there was such evidence in only an additional 34 cases, this means that the teams only considered the question whether there was evidence of negligence in 79 cases. With regard to those 79 cases, the teams disagreed 21 times. This calculation is based on the table in Brennan et al 1991 at 375. The table suggests that the two sets of reviewers agreed in 293 out of 318 cases that there was no negligence, and the text of the article reports “the overall agreement on judgments of negligence was excellent (93 percent).” But the reviewers only considered the presence or absence of substandard care if they first identified a medical management injury. Thus the rate of agreement was 73%, not 93%. Because the researchers did not report the number of times that one reviewer from each team or one review from a single team concluded that there was negligence in the seventy nine medical management injury cases, the true rate of disagreement about the negligence determination is likely to be even higher.

In summary, the New York reliability test showed that, although the two teams largely agreed about which records did not contain evidence of a medical management injury, they agreed less often about which cases did have a medical management injury. And when it came to the question of which medical management injuries resulted from negligence, they disagreed five times more often than they agreed. This is an especially high rate of disagreement given that the reviewers only considered whether there was evidence of negligence or not if they first concluded that there was evidence of a medical management injury.

Considering the particular 47 cases the HMPS identified through the medical malpractice match increases concerns about the reliability of the classifications of those cases. As discussed below, many of those cases either were unusually hard for the medical record analysts to evaluate or involved disagreements among the physician reviewers. This means that there is good reason to believe that these particular 47 cases would produce an even higher rate of disagreement among any given set of reviewers than the “average” case put through the duplicate review in the New York test.

The unreliability of individual decisions would not be as troubling if the HMPS had studied a very large number of medical malpractice claims. But when the number is so small, changing the results of only a few cases could have a very substantial change in the rate of cases with findings of negligence, as the HMPS team explained in their report to the State of the New York. Indeed, they identified a few plausible differences in the classifications that might have tripled the percentage of the 47 cases with good evidence of negligence – from 17% to 50%.³⁵ As a result, it is clearly a mistake to draw strong conclusions about the validity of medical malpractice claims generally on the basis of the HMPS.

³⁵ HMPS 1990 at 7-34: “For example, suppose that the following cases not judged to be negligent according to the Study protocols proved in fact to be instances of negligence: the 5 cases with allegations of failure to diagnose and not caught by the MRA screen; the 4 cases in which physicians disagreed on causation; the 1 low-threshold AE for which one reviewer found negligence; the 1 AE with a single, low confidence finding of negligence; and the 6 AEs with a single finding of negligence. The addition of these 17 cases to the 8 with clear evidence of negligence would triple the percentage of claims with findings of negligence from 17% to about 50%.”

II.B.2. The Limits of Hospital Record-Based Research

The third reason that the HPMS was not well suited to the task of evaluating the merits of individual malpractice claims relates to the fact that the HPMS only examined hospital records, not the full range of evidence available in a medical malpractice case. Hospital records provide at best only a partial picture of what occurs in a hospitalization. Sometimes adverse events are deliberately omitted from hospital records; other times they are omitted inadvertently.³⁶

For example, one important ethnographic study concluded that adverse events occur far more often than they are documented in hospital records.³⁷ The ethnographic research team closely observed hospital care in three surgical units in a large tertiary care hospital affiliated to a university medical school and recorded all “adverse events” discussed by health care providers at any regularly scheduled daytime meeting.³⁸ For purposes of this research, “adverse events” were defined as “situations in which an inappropriate decision was made when, at the time, an appropriate alternative could have been chosen.” The researchers found that there was an adverse event in nearly half of the hospitalizations (not all of which resulted in injury) and that the adverse event was “serious” in eighteen percent of the hospitalizations. (A “serious” event produced at least a temporary physical disability). These results are consistent with other observation-based hospital research.³⁹ The researchers noted that this is a much higher rate than that found in the HMPS, but they concluded that even this higher rate is an underestimate because they “only studied discussions of adverse events at regularly scheduled daytime meetings.”⁴⁰

Like the HMPS researchers, the ethnographic research team matched the studied hospitalizations with malpractice claims filed in the jurisdiction. With access to a more complete range of evidence, it is not surprising that this research team found that a much higher percentage of the matched malpractice claims arose out of hospitalizations in which there had been an

³⁶ Saks 1994 at 709 n. 89. See also O’Neil et al 1993.

³⁷ Andrews et al 1997. See also Andrews 1993 and 2005.

³⁸ Andrews et al 1997.

³⁹ Schimmel 1964; Steel et al 1981; Jahnigen et al 1982; Chaudry et al 2003.

⁴⁰ Id. at 312.

“inappropriate decision” noted in their research records (85%). Like the HMPS, their sample of medical malpractice claims is very small (13). Nevertheless the classifications of the ethnographic researchers gave to individual hospitalizations are more reliable than that of the HMPS reviewers for two reasons. First, the ethnographic researchers considered a broader range of evidence. Second the ethnographic researchers did not make independent evaluations of the presence or absence of an “inappropriate decision”; they only noted whether one of the doctors or hospital staff stated that there had been one. By contrast, the HMPS researchers made an independent evaluation of the hospital record, introducing the possibility of bias.

These are not criticism of the HMPS research methods. All research involves tradeoffs. The HMPS researchers chose a large sample and thin data in order to identify a robust, credible lower bound estimate of the rate of preventable errors in hospitals. Given that choice, however, it is a mistake to conclude that the HMPS could determine that the care was appropriate in a particular case. The most that the HMPS researchers could determine was whether there was evidence of a medical management injury and negligence in the hospital record.

In their early articles, the HMPS researchers acknowledged that they could not draw reliable conclusions about the validity of individual claims. Their NEJM article states:

Our study was not designed to evaluate the merits of individual claims. ... [W]ithout access to the full [liability] insurance records, we cannot assess the prospects of individual cases.

... Our reviewers sometimes disagreed about causation and negligence; when only one found negligence, the case did not qualify as an adverse event due to negligence (except in the rare case when there was only a single reviewer). In a lawsuit a single expert opinion might be sufficient to support a finding of negligence; under our protocol it would not. ... Thus our findings are not directly comparable to the results of civil litigation.⁴¹

Unfortunately, this note of caution was less prominent in the books published for general audiences,⁴² and it disappeared almost entirely in the report of the follow up study of the medical malpractice claims.⁴³

⁴¹ Localio et al 1991 at 249.

⁴² Weiler et al 1993; Weiler 1991.

II.B.3. An Australian Comparison

One final piece of evidence regarding the inappropriateness of using the HMPS to draw strong conclusions about the validity of medical malpractice claiming comes from research in Australia that was modeled on the HMPS.⁴⁴ Australians were dismayed when the results of the Quality in Australian Health Care Study indicated that nearly 17 percent of hospital admissions involved an “adverse event” that many researchers initially determined to be equivalent to a finding of a medical management injury in the HMPS.⁴⁵ After closely comparing their research methods with those used in a recent Utah/Colorado study modeled on the HMPS, however, the researchers found several previously unnoticed, small methodological differences. Applying the U.S. approach to the Australian data resulted in a reduction of hospitalizations involving medical management injuries to about 11 percent.⁴⁶ Similarly, the researchers concluded that applying the small differences in the Australian approach to the US data would have resulted in an astounding 70 percent increase in the number of hospitalizations determined to have resulted in medical management injuries.⁴⁷

This does not mean that the Australians were “right” and the HMPS was “wrong.” Instead, it simply indicates the sensitivity of hospital record reviews to small changes in the research method and, therefore, that it is inappropriate to use the HMPS to make strong conclusions about the *absence* of a medical management injury in a particular case. This is not a criticism of the HMPS research methods, simply a limit on what the HMPS results mean.

⁴³ Brennan et al 1996.

⁴⁴ Wilson et al 1995.

⁴⁵ Runciman et al 2000 at 379.

⁴⁶ Thomas, Studdert, Runciman et al 2000.

⁴⁷ Id. at 376. Note that there are some differences in the methods of the HMPS and the later U.S. study that was directly compared to the Australian study, so the 70% increase is not directly applicable to the HMPS.

II.C. Reconsidering the HMPS Finding Regarding the Validity of Filed Claims

As Professor Michael Saks has explained, the HMPS research actually suggests that most of the claims did have *some* basis in fact.⁴⁸ First, there was evidence of a possible medical management injury in all but twelve of the forty-seven hospitalizations. Moreover, the claims relating to five of the twelve hospitalizations that were screened out by the medical record analysts alleged a negligent failure to diagnose a condition in an earlier outpatient visit, making the claims unusually difficult for the analysts to evaluate.⁴⁹ For the thirty-five hospitalizations that advanced to the second stage review, at least one physician concluded that there was a medical management injury in all but ten. For the twenty-one cases for which negligence determinations were made,⁵⁰ the reviewing physicians voted unanimously that there was not negligence in only six, they agreed that there was negligence in eight, and they reached a split decisions in seven.

Putting all this together, the hospital records, alone, provided some factual support for three quarters of the claims (35 out of 47 claims), evidence of substandard care in three quarters of the claims for which that determination was made (at least one reviewer found evidence of negligence in 15 out of 21 claims), and strong evidence of substandard care in over one third of those claims (both reviewers found evidence of negligence in 8 out of 21 claims).⁵¹

⁴⁸ Saks 1994 at 716 et seq. This explanation has not received the attention that it deserves. Indeed it was not acknowledged or addressed in the follow-up study of the HMPS hospitalization malpractice claims. Brennan et al 1996. See also Bovbjerg 1993 at 2166 (noting that the conclusion that malpractice litigation is erratic “arguably stretches the evidence a bit thin”).

⁴⁹ HMPS 1990 at 7-32: “When the management occurred at an outpatient facility, screeners looking at the hospital record might have difficulty in spotting the shortcomings in medical management, because the events during hospitalization might be consistent with optimal treatment of the underlying disease (e.g., cancer).”

⁵⁰ The HMPS reports a finding regarding the appropriateness of care only for cases determined to involve a significant likelihood of medical management injury. Thus, there is no finding regarding the appropriateness of care for the twelve cases screened out in the first stage or for the fourteen cases screened out during the second stage. HMPS 1990 at 7-33 report that seven of the fourteen cases that the physician reviewers determined not to involve a medical management injury were based on an alleged failure to diagnose. If any of these seven cases were among the four in which one reviewer determined that there was a medical management injury they would be likely to have been classified by the review as also involving negligence.

⁵¹ These calculations can all be made using the table in Localio et al 1991 at 248.

Indeed, given that claimants and their lawyers generally must file suit in order to obtain all the information that they need to determine whether the claimants' injuries resulted from negligence,⁵² the percentage of claims with strong evidence of negligence in the hospital record is higher than would be expected in world in which medical providers typically do not disclose that an injury is iatrogenic, let alone the result of negligence.⁵³ For this reason, the analogy to a situation in which "police regularly gave out more tickets to drivers who go through green lights than to those who go through red lights" is fundamentally flawed.⁵⁴ A traffic officer has access to all the information about the color of the traffic light at the time the driver went through. The HMPS reviewers had access only to the hospital record – an important part of the information about patients' treatment, but not the whole. In addition, a traffic officer is for all practical purposes the controlling legal authority able to impose a legally binding judgment regarding a traffic offense. An injured patient is just a complaining witness, and a medical malpractice complaint only begins a legal process; it is not a legally binding judgment.

Another way of evaluating what the HPMS tells us about the validity of medical malpractice claims is to examine how the strength of the evidence of malpractice in the hospital record affected the likelihood that a medical malpractice claim would be filed. Evidence of a possible medical management injury sufficient to satisfy the first screening criteria raised the likelihood of claim eight times. Evidence that was sufficient to support a unanimous finding of a medical management injury by the physician reviewers raised the likelihood of claim an additional six times. Evidence that was sufficient to support a unanimous finding that the medical management injury resulted from negligence raised the likelihood of claim an additional two and a half times. Put another way, a patient whose hospital record provided the strongest evidence of medical malpractice was twenty two times more likely to make a claim than the average patient, eight times more likely to make a claim than all other patients with sufficiently strong evidence of

⁵² Farber & White 1990 and 1994.

⁵³ E.g. Andrews 1993; Farber & White 1994. Cf. Mazor et al 2004; Lamb et al 2003.

⁵⁴ Weiler et al 1993 at 75.

a medical management injury to reach the second stage of the HMPS review, and two and a half times more likely to make a claim than all other patients the HMPS physicians determined to have suffered a medical management injury.⁵⁵ This is hardly a picture of malpractice claiming run amok.⁵⁶

III. The Follow Up Study's Conclusion about the Validity of Paid Claims

In 1996, some HMPS researchers reported the results of a ten-year follow up study of the medical malpractice claims that arose out of the hospitalizations reviewed in the HMPS.⁵⁷ The researchers concluded that “[t]he results of these follow up assessments of whether negligence or a medical injury had occurred 10 years earlier bore little relation to the outcome of the claims, just as in our earlier study they were found to bear little relation to the initial decision to file the claims.”⁵⁸

For the reasons just explored, this characterization of the results of the earlier study is questionable. The earlier study did not find that the assessments bore “little relation to the initial decision to file the claims.” In fact the strength of the evidence of negligent medical injury in the hospital record was related to the likelihood that a malpractice claim would be filed. This characterization of the results of the follow up study is questionable as well.

The methods for conducting the follow up study were as follows.⁵⁹ Starting in 1991 a member of the HMPS team contacted each of the medical malpractice insurers handling one or more of the claims every six months and inquired into the status of the claims. Once a claim was closed, the researchers obtained a claim summary from the insurer and used that summary to identify any “discrepant” cases – i.e. claims that were paid but that the HMPS had classified as

⁵⁵ Calculations from Table 3 in Localio et al 1991 at 248. See also Mello & Hemenway 2004.

⁵⁶ It is worth noting that the HMPS team reported to the State of New York that “the presence of a claim might be an efficient screen for both adverse events and negligence,” because “both negligence and adverse events were dramatically higher among the claimants, as one would expect.” HMPS 1990 at 7-36. This conclusion is inconsistent with later statements in the book published for a popular audience suggesting that the legal system is erratic. Weiler et al 1993 at 139.

⁵⁷ Brennan et al 1996.

⁵⁸ Id at 1966.

⁵⁹ Id. at 1964.

not involving a negligent injury, or claims that were not paid but that the HMPS had classified as involving a negligent injury. At the time the follow up study was concluded, forty-six of the fifty-one claims that had been matched to the HMPS hospitalizations had been closed.⁶⁰ A researcher who was unaware of how the HMPS had classified the hospitalizations in question reviewed the litigation file of twelve of the thirteen discrepant cases to determine whether there was a medical management injury and, if so, whether the injury resulted from negligence.

III.A. Problems with the Follow Up Study Research Methods and Analysis

The follow up study shares the small sample problems of the early HMPS analysis of the 47 claims, plus some new problems. The first significant new problem is the “observer effect” introduced by the fact that the research team repeatedly contacted the medical malpractice insurers about these cases. By 1991 the HMPS study was well known in the medical malpractice insurance world. Executives in medical malpractice insurance companies can be expected to have had intense interest in whatever judgments the HMPS researchers would make about the results of those cases.⁶¹ That interest may well have affected how the insurance companies handled those cases. It is impossible to know exactly what this observer effect was, but it is not difficult to speculate about ways that it could have biased the handling so as to support the insurance industry’s perceived interest in proving that the tort litigation process is erratic.⁶² This is a potentially serious flaw in the research design.

A second significant new problem with the follow up study concerns the analysis used to support the conclusion that the claims were paid without regard to negligence. The researchers used two statistical analyses. Both raise concerns.

⁶⁰ Brennan et al 1991 explained that the number of matched claims increased from the 47 claims discussed earlier in this article to 51 as a result of a second review of hospital records. These additional cases were mentioned in the earlier HMPS reports but not incorporated into the analysis. It does not appear that including those 4 claims in the earlier analysis would have changed the results materially.

⁶¹ Evidence of this intense interest is provided by Anderson 1997, who is identified as an official of a medical malpractice insurance company..

⁶² E.g. Kielhoz & Murray 2004.

In the first analysis, the researchers compared the mean (average) payment amounts in cases according to the presence or absence of a medical management injury and negligence. They reported that the mean payment in non-negligent medical management injury cases was significantly larger than the mean payment in negligent medical management injury cases (\$98,192 as compared to \$66,944), supporting their conclusion that negligence did not matter to the outcome of the claim.⁶³

The problem with this analysis follows from the weakness of using means to analyze a very small number of cases. A single “outlier” can bias the results, which is exactly what happened here. Careful examination of the follow up study data shows that the difference in the means is attributable to one very large payment in one close case that the HMPS classified as a non-negligent medical management injury case.⁶⁴ Reclassifying that case (and two other cases for which the researchers noted that there was reason to believe that the HMPS classification may have been incorrect), results in a dramatic reduction of the mean payment in non-negligent medical management injury cases (from \$98,192 to \$31,375) and a correspondingly dramatic increase in the mean payment in negligent medical management injury cases (from \$66,944 to \$162,750).⁶⁵ With this reclassification, the means for the three categories of cases are as follows:

No medical management injury:	\$ 23,552
Non-negligent management injury:	\$ 31,375
Negligent management injury:	\$162,750

⁶³ Brennan et al 1996 at 1966, Table 2.

⁶⁴ This is one of the cases in which the Harvard researchers read the claims file. They report that this case involved a very serious neurological injury from a vascular procedure and that the insurer thought that “the jury would compensate the patient even though the medical care met the expected standard.” Brennan et al 1996 at 1966. There are reports that neurological injury cases often involve serious differences of opinion regarding the standard of care. See, e.g. Sussman 2003.

⁶⁵ Brennan et al 1996 at 1966.

Significantly, a substantial percentage of the payments in each of the first two categories is attributable to a single large payment within that category, suggesting that even these adjusted means are biased upwards.⁶⁶

More strikingly, even without this reclassification the median payment in the first two categories is \$0, because most of the claims in those categories were closed without any payment.⁶⁷ Among paid claims in the three categories, the median payments are as follows:

No medical management injury:	\$ 25,000
Non-negligent management injury:	\$ 25,000 – 49,999
Negligent management injury:	\$100,000 – 249,000. ⁶⁸

These comparisons suggest that – contrary to the conclusions stated in the published account of the follow up closed claim study – the medical malpractice claiming system is appropriately distinguishing among cases according to their merits.

In the second statistical analysis, the researchers used a logistic regression approach that compared how different factors predicted whether a claim would be paid.⁶⁹ The factors included the severity of the injury, the HMPS classifications regarding medical management injury and negligence, and income, race and age. The researchers reported that the only factor that affected the payment of a claim significantly was the presence or absence of a disabling injury and, thus, concluded that negligence did not matter to the outcome of the claims. This is the finding that is repeated in the medical malpractice debates and the economic literature.⁷⁰

There are two main problems with this logistic regression analysis and, accordingly, with this finding. The first problem follows from the fact that the HMPS team only assigned a

⁶⁶ Based on a comparison of Brennan et al 1996 Table 1 with the reclassification discussion on p.1966.

⁶⁷ The “median” is the midpoint – the point at which half of the settlement amounts are higher. Thus a 0 means that more than half of the claims in the first two categories closed without even a write off of the hospital bills. Medians are not affected by the presence of outlier cases.

⁶⁸ Exact medians cannot be calculated because the settlement amounts are presented in ranges. For the first category the median is given as \$25,000 because the amounts in half of the paid claims were less than \$25,000 and half were greater.

⁶⁹ Brennan et al 1996 at 1966 Table 3.

⁷⁰ See, e.g., American Medical Association 2004 and Kessler & McClellan 2002.

disability score to hospitalizations determined to involve a medical management injury.⁷¹ As a result, their finding about the predictive effect of disability means something other than it first seems. What the follow up study really found was that a plaintiff with a permanently disabling *medical management injury* was more likely to be paid than a plaintiff without a permanently disabling medical management injury. While that still leaves out the question of negligence, it is a far less troubling finding than the HMPS researchers suggested, especially in light of what the pilot study and the New York test tell us about the high level of disagreement between their reviewers about the presence or absence of negligence in medical management injury cases. (Recall that the two teams in the New York test disagreed five times more often than they agreed about which medical management injury cases contained good evidence of negligence.)

The second, more serious problem with the logistic regression analysis relates to the treatment of cases in which the only “payment” was a “write off” of the unpaid hospital bill (meaning that the plaintiff did not actually receive any money). On the one hand, the researchers regarded these cases as so insignificant that they did not need to review the litigation file to determine whether the HMPS classifications assigned to those cases were accurate.⁷² On the other hand, the logistic regression treated these cases the same as those in which the plaintiff received a large payment. This is because a logistic regression analysis evaluates how factors predict the outcome to a “yes/no” question. In the logistic regression here, the “yes/no” question was whether there was a settlement in favor of the plaintiff, and the regression treated a “write-off” as a settlement in favor of the plaintiff.

Write off cases almost certainly were among the least severe or the least negligent,⁷³ so including these cases with the plaintiff wins reduces the estimated effect of the merits in the

⁷¹ Id at 1965, Table 1. For the logistic regression analysis the 0 to 8 disability scale was reduced to a dichotomous variable of permanent disability, yes or no.

⁷² Id. at 1967.

⁷³ Five of the write off cases were classified as not involving a medical management injury, with the result that their disability score was a 0. Two of the remaining three were classified as not involving negligence. Brennan et al 1996 at 1965, Table 1.

logistic regression. In other words, treating the write off cases as plaintiff wins made it much more likely that negligence would not be a significant factor in the prediction of the outcome.

Especially in light of their initial determination that they did not need to review the claim files in the write off cases, the researchers should have treated those cases as settlements in favor of the defendant (which is how a plaintiff and his or her lawyer would have understood a case in which they did receive any money) or at the very least the researchers should have run the logistic regression without those cases. Nearly half of the cases that the researchers classified as payments to undeserving claimants appear to have been “write off” cases (7 out of 16).⁷⁴ In a sample of only 46 cases, the effect of changing the classification of 6 cases, or of eliminating 6 similarly classified cases, could have a very large effect on the logistic regression results. Thus, like the analysis based on the means, the logistic regression analysis falls apart when small, and very reasonable adjustments are made.

A third problem with the follow up study is the fact that the researchers did not address either the criticisms Professor Michael Saks lodged in his review of the earlier HMPS analysis of the validity of medical practice claims or the statements in the HMPS report to the State of New York regarding the unreliability of the assessments among some of the matched cases. As Professor Saks had highlighted, many of the physician reviews of the hospitalizations that lead to the malpractice claims were not unanimous.⁷⁵ In addition, the HMPS report disclosed that 5 of the claims related to alleged negligence in earlier outpatient visits and that those hospitalizations had not been reviewed by physicians (because they were screened out during the first stage of the reviewing process).⁷⁶

As discussed earlier, there is substantial reason to doubt the HMPS classification of a large percentage of the hospitalizations that led to the malpractice claims. Yet the follow up study treats those classifications as if they were certain. Moreover, the article publishing the

⁷⁴ Id.

⁷⁵ Saks 1994.

⁷⁶ HMPS 1990 at 7-32.

results of the follow up study does not disclose which of the cases involved these more doubtful classifications or whether they were among the 12 cases for which the researchers reviewed the claims file. As just noted, changing the classification of only a few cases in such a small sample could have a very substantial effect on any and all of the follow up study results.

III. B. The Contrary Findings of a Large Body of Research

It would not be necessary to be as sharply and publicly critical of the HMPS follow up closed claim study if its results were consistent with the results of other studies that examined larger numbers of medical malpractice claims under conditions in which there were no observer effects. But the results of the follow up study are diametrically opposed to the results of other studies, which find that the medical malpractice claim handling and litigation system is appropriately filtering out most non-meritorious cases.

All of the other academic research on the outcomes of medical malpractice claims involved the analysis of malpractice claims that were already closed; most of that research analyzed a substantially larger number of claims. This is a superior research method for two main reasons. First, because the claims were already closed when the research began there could have been no observer effects. Second, analyzing a larger number of claims provides more confidence in the reliability of the results, especially because the sampled population (closed claims) is the appropriate body of cases for studying the claim closing process.

The paragraphs that follow briefly describe the other studies and their results. The studies are presented in chronological order.

St. Paul Obstetrics Claims Study. Ogburn et al reviewed the claims files of the St. Paul Company and identified all birth injury claims filed from 1980 to 1982 that alleged malpractice on the part of an obstetrician.⁷⁷ They found 600 such claims. Of those 600, 380 were closed with no indemnity payment and a defense payment of less than \$1000. Although the study is not clear on this point, it is likely that these 380 “claims” involved situations in which a physician reported

⁷⁷ Ogburn et al 1988.

a potential claim but no suit was ever filed.⁷⁸ From the 220 cases with either an indemnity payment or a defense payment greater than \$1000, the researchers identified 153 claims that alleged serious permanent injury or death to the baby (the remaining 67 cases apparently involved injury to the mother or less serious injury to the baby). Five obstetricians reviewed those 153 claims in detail and determined that payments to the plaintiff were made in 90% of the cases in which the physician was negligent. They also reported that there was either an indemnity payment or a defense cost greater than \$1000 in 55% of the cases in which the physician was not negligent. Because the researchers did not distinguish between defense and indemnity expenditures in the non-negligent cases, the study does not permit any conclusions regarding indemnity payments in non-negligence cases, other than the general sense that indemnity payments are much more unusual in non-negligence cases.⁷⁹ Nevertheless, the study does strongly support the conclusion that negligence strongly predicts an indemnity payment in medical malpractice cases.

Physicians Mutual Obstetrics Claims Study. Rosenblatt and Hurst reviewed the claims files of an unspecified physician-sponsored medical malpractice insurer and identified all obstetric claims closed during the 1982 to 1988 period.⁸⁰ They found 54 such claims. In contrast to Ogburn et al, they reported defense and indemnity payments separately and, thus, provide a better basis for evaluating indemnity payments in non-negligence cases. They found that 21 of the 54 claims were closed before suit with no payment. Rosenblatt reviewed the claims files for the 33 cases put into suit and determined that no indemnity payments were made in any of the non-meritorious cases and that “for those cases in which payments were made, there was general consensus among insurance company staff, medical experts, defense attorneys and the physician

⁷⁸ Rosenblatt & Hurst 1989 studied obstetric claims closed during 1980-1988 and reported that the defense costs for all claims put into litigation were greater than \$1000. At p. 711, Table 1.

⁷⁹ Intuitions here point in opposite directions. On the one hand, any seriously contested case would be likely to produce defense costs in excess of \$1000. On the other hand there are reasons to expect that there would be payments in a higher percentage of serious birth injury cases than in malpractice cases generally.

⁸⁰ Rosenblatt and Hurst 1989.

defendants that some lapse in standard of care contributed to the observed outcome.”⁸¹ Because the number of claims reviewed was even smaller than in the Harvard study, the Rosenblatt and Hurst study does not provide strong evidence for a general pattern. Nevertheless, the results do support the conclusions reached by other researchers.

Anesthesiologist Closed Claim Review. Cheney et al reported interim results of the American Society of Anesthesiologist’s closed claim review.⁸² The ASA reviewed the claims files of 17 insurers operating in 40 states and identified 1175 claims against anesthesiologists that were closed before January 1988. Anesthesiologists reviewed each claim and evaluated whether there was “inappropriate care” (i.e. below “the standard of care for a prudent anesthesiologist practicing anywhere in the United States at the time of the event”). In addition the reviewers recorded information regarding whether a lawsuit was filed, the amount of any award or settlement, the nature of the injury, and various other factors that were relevant to the risk management goals of the study. The reviewers were able to make a judgment regarding the appropriateness of the care in 869 of the cases. In those 869 cases, payments were made in 82% of the instances in which the reviewer determined that the care was inappropriate and in 42% of the instances in which the reviewed determined that the care was appropriate.⁸³ Among the cases in which payments were made, the amount of the payment was strongly correlated to the presence of negligence and to the type of injury. Among paid cases, the median payment in negligent, disabling injury cases was \$463,000 while the median payment in non-negligent disabling injury cases was \$93,000, and the median payment in negligent death cases was \$200,000 while the

⁸¹ Id at 712.

⁸² Cheney et al 1989.

⁸³ Cheney et al 1989 at 1601, figure 2. Note that two-thirds of the “appropriate care” cases that the insurance companies paid were cases that involved non-disabling injuries and tiny payments. Because the published article uses percentages rather than absolute numbers, this is difficult, but not impossible, to figure out. Using Figure 1 and the percentages given in the text on page 1601, one can compute that payments were made in 34 appropriate care death cases, 25 appropriate care disabling cases, and 104 appropriate care non-disabling cases. The median payment in the anesthesiologists’ non-disabling cases was less than \$15,000. All these cases date back to a time before physicians and insurance companies were required to report all medical malpractice payments on behalf of doctors. Since the U.S. Congress enacted that requirement, nuisance payments have become much less common, so we would not expect to find so many paid appropriate care cases today. Waters et al 2003.

median payment in non-negligent death cases was \$115,000. The median “payment” in non-negligent cases was \$0.⁸⁴ This study supports the conclusion that (a) the presence or absence of negligence most strongly explains both the likelihood of payment and the amount of payment and (b) the severity of the injury affects the size of a payment given the presence or absence of negligence.

Economists’ Single Hospital Study. Farber and White analyzed all malpractice claims relating to care provided in a single hospital initiated in 1977 or later and resolved by 1989 and that involved claims solely against the hospital or a member of its staff. There were a total of 252 such cases.⁸⁵ For each of these cases the hospital obtained expert reports on the quality of care. Quality of care was “good” if the professional standard of care was met, “bad” if it was not, and “ambiguous” if the experts disagreed or could not make an evaluation. Farber and White found that there was a strong relationship between the outcome of the case and the quality of care. Claimants were paid in 89% of the “bad” care cases, 25% of the “good” care cases, and 69% of the “ambiguous” cases. Based on a regression analysis, Farber and White determined that the hospital’s expected liability for damage was 25 times as high on average in a “bad” care case as in a “good” care case.

Florida Closed Claims I. Sloan and Hsieh used the Florida Department of Insurance database of all medical malpractice claims closed in Florida from October 1985 to March 1988.⁸⁶ After eliminating the “claims” that never resulted in an insurer expense, they were left with a database of 6,612 cases. The database contained information about the defense and indemnity payments, the nature of the allegations, and the nature of the injury. They obtained physician liability reviews for a subsample of claims filed against obstetrician-gynecologists, general

⁸⁴ Cheney et al 1989 at 1601, figure 2.

⁸⁵ Farber & White 1990.

⁸⁶ Sloan and Hsieh 1990.

surgeons and orthopedic surgeons and added that information to the database.⁸⁷ Their analysis of that subsample indicated that “cases involving a higher appearance of fault were more likely to be paid” and that the amount of the payment was positively correlated with the severity of the injury.⁸⁸

Florida Closed Claims II. Sloan et al conducted a more intensive analysis of a randomly selected sample of Florida closed claims filed against obstetricians and emergency room physicians.⁸⁹ Four panels of physicians reviewed each of 127 birth injury cases, using the Department of Insurance closed claim forms, information from interviews of the claimants, and an abstract of the hospital record.⁹⁰ A majority of the reviewers determined that negligent care caused an injury in 22 of the cases and that the injury was not caused by negligence in 28 of the cases.⁹¹ Of the remaining 77 cases, the opinions of the reviewers were split in 42 cases and were uncertain in 35 cases due to the availability and quality of the medical records.⁹² Four panels of physicians also reviewed each of 60 emergency room cases using similar information. A majority of the reviewers determined that there was negligence in 32, no negligence in 18, and insufficient information in 10.⁹³ When the researchers compared these liability assessments with the outcomes of the claims, they found a close correlation between the liability assessment and the outcome. “No negligence” claims were most likely to have been voluntarily dropped by the plaintiff with no payment, “negligence” claims were most likely to have been settled with a payment to the plaintiff, and claims with split or uncertain determinations were more mixed.⁹⁴

⁸⁷ They did not report the number of claims in the subsample. Apparently the analysis was done for another project. Id at 1010.

⁸⁸ Id at 1014 and 1019.

⁸⁹ Sloan et al 1993.

⁹⁰ Id. at 98-99.

⁹¹ Id.at 100, 105.

⁹² Id. at 107.

⁹³ Id. at 111-113.

⁹⁴ Id. at 166-67. Because the published study does not provide exact numbers I cannot be more specific.

New Jersey Closed Claim Study. Taragin et al analyzed 8231 claims closed between 1977 and 1992 by the largest medical malpractice insurer in New Jersey.⁹⁵ Based on a review process conducted immediately after a claim was filed, the malpractice insurer had rated each of the claims as “defensible,” “indefensible,” or “unclear,” according to the prevailing standard of care. The review process involved an interview with the defendant physician, an evaluation by a claim representative employed by the insurer, and peer review for any case that the claims representative did not conclude was clearly defensible. The researchers compared those evaluations with the outcomes in the closed claims and found that there was a strong correlation. Claimants received payments in 91% of the indefensible cases, 21% of the defensible cases, and 59% of the unclear cases.⁹⁶ The severity of the injury did not affect the likelihood of payment (though it did affect the size of the payment in cases in which payments were made, as is appropriate). The researchers noted that because the defensible/indefensible determination was made very early in the claim process and because the peer review process was likely to be biased against determining that the physician was liable, the “defensible” cases that resulted in payments may have been incorrectly classified. In sum the researchers concluded, “Further efforts to clarify the frequency of unjustified payments are needed, but our data suggest that such payments are uncommon.”⁹⁷

Michigan Single Hospital Study. Spurr and Howze analyzed 165 closed claims from a single hospital in Michigan constituting all of the claims brought against the hospital or its staff and closed between 1987 and 1995.⁹⁸ They reviewed the entire claim file for each of the claims and assigned a “fault” score on a continuous scale from 0 (beyond reproach) to 1 (conclusive evidence of negligence) based on “the opinion of the hospital’s risk management staff of the quality of the defendant’s medical care, as disclosed by internal memoranda, reports and

⁹⁵ Taragin et al 1992.

⁹⁶ Id at 781.

⁹⁷ Id. at 784.

⁹⁸ Spurr & Howze 2001.

correspondence.”⁹⁹ They used a Probit analysis to estimate the relative effect of fault and a variety of other factors on (a) whether the case was settled or dropped by the plaintiff and (b) size of a payment, if any.¹⁰⁰ The other factors included the severity of the injury, age of the claimant, whether the claimant had private insurance or Medicaid, and the population density of the county of the court or arbitration in which the claim was filed. In their analysis of the settlement vs. drop decision, only the fault factor was significant, meaning “[t]here is no evidence of a tradeoff between potential damages and the plaintiff’s ability to prove negligence, in the determination of the issue between drop and settlement.”¹⁰¹ In their analysis of the settlement payment amount, both fault and severity were significant, but fault was more powerful (consistent with the results in anesthesiologists’ closed claim study).

North Carolina Study. Peeples et al analyzed 87 closed claims files from malpractice lawsuits filed in North Carolina between 1991 and 1995 that were directed to mediation prior to trial and involved either a “major teaching hospital” or a physician insured by a “principal liability insurer of physicians” that had agreed to cooperate in the research.¹⁰² The researchers believe that this sampling method resulted in an unusually high percentage of “durable” cases and, thus, would not be reflective of cases dropped early in the litigation process. The researchers reviewed the claim file and coded the cases using several indications of fault: the initial and last assessment by the insurer of liability (probable, uncertain or likely), the consensus on liability of outside expert reviewers engaged by the insurer (probable, uncertain or likely), whether the insurer believed the standard of care had been breached (yes, no or uncertain), and whether the insurer believed that causation existed between the defendant’s conduct and the alleged injuries (yes, no or uncertain). They also coded the cases for outcome variables: defense made offer (yes

⁹⁹ Id. at 496.

¹⁰⁰ Unlike a logistic regression, a Probit analysis is well-suited to dependent variables like settlement amounts in which there are a large number of \$0 payments and a large spread among the other payments. Using a logistic regression requires collapsing the range of payments into a yes/no variable. It would be interesting to reanalyze the Brennan et al 1996 data using a Probit.

¹⁰¹ Id at 507.

¹⁰² Peeples et al 2002 at 881.

or no), settled (yes or no), trial (yes or no), who won (plaintiff or defendant), money paid to plaintiff (yes or no). This study contains a variety of interesting findings. For example, the insurer evaluated the liability as probable or uncertain in 86% of the cases handled by experienced plaintiffs' lawyers, but only 42% of the cases handled by less experienced lawyers. With regard to the effect of fault on the outcome, the results are consistent with the other research. The insurer made an offer in 96% of the cases in which it concluded that the standard of care was breached, and money was paid to the plaintiff in 93% of those cases.¹⁰³ Money was paid to plaintiffs in only 15% of the cases in which the insurer concluded that the standard of care was not breached and in 37% of the cases in which the insurer was uncertain.¹⁰⁴

Summary. The picture presented by these other closed claim studies is very different than that presented in the report of the HMPS follow up closed claim study. Negligence matters a great deal to the outcome of a medical malpractice claim, and the litigation process weeds out most of the weaker claims.

IV. Conclusion

In taking an epidemiological approach to the incidence of medical malpractice, the Harvard Medical Practice Study performed an enormous service. The HMPS team deserves much of the credit for the fact that medical malpractice has come to be understood as a public health problem.¹⁰⁵ Unfortunately, the HMPS law reform legacy is decidedly more mixed. Policymakers have ignored the HMPS researchers' call for a broad no-fault approach to compensation for medical injuries. Instead, policymakers have seized on the weakest aspect of the HMPS, the analysis of the validity of medical malpractice claims, and used that analysis to justify imposing caps on damages in medical malpractice cases and additional procedural hurdles for medical malpractice claimants. For that reason the practical impact of the HMPS may well have been to expand the gap between the large number of people who are injured by medical

¹⁰³ Id. at 886.

¹⁰⁴ Id. at 888.

¹⁰⁵ E.g. Kohn et al 2000.

malpractice and the few people who are compensated and to increase the likelihood that the compensation that is received will be inadequate.

This does not have to be the legacy of the HMPS. Advocates have pushed the medical malpractice claim data well beyond their limits. A good part of what is reported here can be found in caveats and qualifications contained in the HMPS report to the State of New York and in publications by HMPS researchers, as well as by early critics of the malpractice claiming aspects of the HMPS. But those caveats and qualifications have not received adequate attention in the public debate. What the public heard was that medical malpractice claims are frivolous and that the wrong people get paid. The HMPS and the follow up study prove no such thing and the rest of the literature suggests exactly the opposite conclusion.

References

- American Medical Association, *Medical Liability Reform – NOW! A compendium of facts supporting medical liability reform and debunking arguments against reform* (Chicago: AMA 2004).
- Anderson, Richard, Correspondence, *New England J. Med* 336:1680-1681 (1997).
- Andrews, Lori, *Medical Error and Patient Claiming in a Hospital Setting* (American Bar Foundation 1993).
- Andrews, Lori, *Studying Medical Error In Situ: Implications for Malpractice Law and Policy*, *DePaul Law Review* 54:357-92 (2005).
- Andrews, Lori B., Carol Stocking, Thomas Krizek, Lawrence Gottlieb, Claudette Krizek, Thomas Vargish, Mark Siegler, *An Alternative Strategy for Studying Adverse Events in Medical Care*, *Lancet* 349:309-313 (1997).
- Bovbjerg, Randall R., *Medical Malpractice Research and Reform*, *Virginia Law Review* 79:2155-2208 (1993).
- Brennan, Troyen A., Russell J. Localio, Nan Laird, *Reliability and Validity of Judgments Concerning Adverse Events Suffered by Hospitalized Patients*, *Medical Care* 27:1148-1158 (1989).
- Brennan, Troyen A., Lucian L. Leape, Nan M. Laird, Liesi Hebert, A. Russell Localio, Ann G. Lawthers, Joseph P. Newhouse, Paul C. Weiler, Howard H. Hiatt, *Incidence of Adverse Events and Negligence in Hospitalized Patients*, *New England J. of Med.* 324:370-6 (1991).
- Brennan, Troyen A., Colin M. Sox, Helen R. Burstin, *Relation Between Negligent Adverse Events and the Outcomes of Medical-Malpractice Litigation*, *New England J. Med.* 335:1963-67 (1996).
- Brody, William R., *Dispelling Medical Malpractice Myths*, *Washington Post*, November 14, 2004.
- Chaudry, Sarwat I., Kolawole A. Olofinboba, Harlan M. Krumholz, *Detection of Errors by Attending Physicians on a General Medical Service*, *Journal of General Internal Medicine* 18:595-600 (2003).
- Cheney, Frederick W., Karen Posner, Robert A. Caplan, Richard J. Ward, *Standard of Care and Anesthesia Liability*, *JAMA* 261:1599-1603 (1989).
- Danzon, Patricia, *Medical Malpractice: Theory, Evidence, and Public Policy* (Harvard U. P. 1985).

Farber, Henry S., & Michelle J. White, "Medical Malpractice: An Empirical Examination of the Litigation Process," 22 *RAND Journal of Economics* 199 (1990).

Farber, Henry S., & Michelle J. White, A Comparison of Formal and Informal Dispute Resolution in Medical Malpractice, *Journal of Legal Studies* 23:777-806 (1994).

Goldman, Robert L., The Reliability of Peer Assessments of Quality of Care *JAMA* 267:958-60 (1992).

Greater New York Hospital Association, Medical Malpractice Insurance Costs and Coverage (2005).

Harvard Medical Practice Study. 1990. Patients, Doctors, And Lawyers: Medical Injury, Malpractice Litigation And Patient Compensation In New York: The Report Of The Harvard Medical Practice Study To The State Of New York. Cambridge: Harvard University.

Hayward, Rodney A., Laurence F. McMahon, Annette M. Bernard, Evaluating the Care of General Medicine Inpatients: How Good is Implicit Review? *Annals of Internal Medicine* 118:550-56 (1998).

Hayward, Rodney A., Timothy P. Hofer, Estimating Hospital Deaths Due to Medical Errors: Preventability is in the Eyes of the Beholder, *JAMA* 286:415-20 (2001).

Hofer, Timothy P., Steven J. Bernstein, Sonya DeMonner, Rodney A. Hayward, Discussion Between Reviewers Does Not Improve Reliability of Peer Review of Hospital Quality, *Medical Care* 38:152-61 (2000).

Jahnigen, D., C. Hannon, L. Laxson, F.M. La Force, Iatrogenic Disease in Hospitalized Elderly Veterans, *Journal of American Geriatric Society*, 30:387-90 (1982).

Joint Commission on Accreditation of Healthcare Organizations, Health Care at the Crossroads: Strategies for Improving the Medical Liability System and Preventing Patient Injury (White Paper 2005).

Joint Economic Committee, The Perverse Nature of the Medical Liability System, Research Report #109-2 (March 2005).

Keilholz, Walter B. and Richard H. Murray, Insurance Economics and Legal Systems, *Geneva Papers on Risk and Insurance* 29:1-5 (2004).

Kessler, Daniel P. and Mark B. McClellan, How Liability Law Affects Medical Productivity. *Journal of Health Economics* 21:931-55 (2002).

Kohn, Linda T., Janet M Corrigan, and Molla S. Donaldson, eds. 2000. *To Err Is Human: Building A Safer Health System*. Washington, DC: National Academy Press.

Lamb, Rae M., David M. Studdert, Richard M.J. Bohmer, Donald M. Berwick, Troyen A. Brennan, Hospital Disclosure Practices: Results of a National Survey: Most Hospitals Disclose Harm to Patients at Least Some of the Time, this 2002 Survey Finds, Health Affairs, March –April 2003.

Localio, A. Russell, Ann G. Lawthers, Troyen A. Brennan, Nan M. Laird, Liest E. Hebert, Lynn M. Peterson, Joseph P. Newhouse, Paul C. Weiler, Howard H. Hiatt, Relation Between Malpractice Claims and Adverse Events Due to Negligence, Results of the Harvard Medical Practice Study III, NEJM 325:245-251 (1991).

Localio, A. Russell, Susan L. Weaver, J. Richard Landis, Ann G. Lawthers, Troyen A. Brennan, Liesi Hebert, Tonya Sharp, Identifying Adverse Events Caused by Medical Care: Degree of Physician Agreement in a Retrospective Chart Review, Ann Intern Med 125:457-64 (1996).

Mazor, Kathleen M., Steven R. Simon, Jerry H. Gurwitz, Communicating with Patients About Medical Errors: A Review of the Literature, Arch Intern Med 164:1690-1697 (2004).

Mehlman, Maxwell J., Saying “No” to No-Fault: What the Harvard Malpractice Study Means for Medical Malpractice Reform (New York State Bar Association Special Committee on Medical Malpractice 1991).

Mello, Michelle M., David Hemenway, Medical Malpractice as an Epidemiological Problem, Social Science & Medicine 59:39-46 (2004).

Mills, Don H., John S. Boyden, and David S. Rubjam, California Medical Association & California Hospital Association, Report on the Medical Insurance Feasibility Study (San Francisco: Sutter Publications 1977, sponsored jointly by the California Medical Association and California Hospital Association), summarized in Mills, Don Harper, “Medical Insurance Feasibility Study: A Technical Summary,” West. J. Med. 128:360-365 (1978).

Ogburn, Paul L., Thomas M. Julian, Doris C. Booker, Marilyn S. Joseph, Julius C. Butler, Perston P. Williams, Mark L. Anderson, Ann C. Shepard, Susan L. Ogburn, William C. Preisler, Mark D. Wood, Perinatal Medical Negligence Closed Claims from the St. Paul Company, 1980-1982, Journal of Reproductive Medicine 53:608-11 (1988).

O’Neil, Anne C., Laura A. Petersen, E. Francis Cook, David W. Bates, Thomas H. Lee, Troyen A. Brennan, Physician Reporting Compared with Medical-Record Review to Identify Adverse Medical Events, Ann Intern Med 119:370-76 (1993).

Peebles, Ralph, Catherine T. Harris, Thomas B. Metzloff, The Process Of Managing Medical Malpractice Cases: The Role Of Standard Of Care, 37 Wake Forest L. Rev. 877 (2002).

Posner, Karen L., Robert A. Caplan, Frederick Cheney, Variation in Expert Opinion in Medical Malpractice Review, *Anesthesiology* 85:1049-54 (1996).

Rosenblatt, Roger A., Andy Hurst, An Analysis of Closed Obstetric Malpractice Claims, *Obstetrics & Gynecology* 74:710-14 (1989).

Runciman, William B., Robert K. Webb, Stephen C. Helps, Eric J. Thomas, Elizabeth J. Sexton, David M. Studdert, Troyen A. Brennan, A comparison of iatrogenic injury studies in Australia and the USA II: reviewer behavior and quality of care, *International Journal for Quality in Health Care* 12:379-88 (2000).

Saks, Michael, Medical Malpractice: Facing Real Problems and Finding Real Solutions, *William & Mary L. J.* 35:693-726 (1994) (a review of Weiler et al 1993).

Schimmel, Elihu, The Hazards of Hospitalization, *Annals of Internal Medicine* 60:100-10 (1964).

Sloan, Frank A. & Chee Ruey Hsieh, Variability in Medical Malpractice Payments: Is the Compensation Fair? 24 *Law & Soc'y Rev.* 997 (1990).

Sloan, Frank A., Penny B. Githen, Ellen Wright Clayton, Gerald B. Hickson, Douglas A. Gentile, David F. Partlett, *Suing for Medical Malpractice* (U. Chicago P. 1993).

Spurr, Stephen J., Sandra Howze, The Effect of Care Quality on Medical Malpractice Litigation, *Q. Rev. Econ. & Fin.* 41:491-513 (2001).

Steel, Knight, Paul Gartman, Caroline Crexenzi, Jennifer Anderson, Iatrogenic Illness on a General Medical Service at a University Hospital, *New England Journal of Medicine* 304:638-42 (1981).

Sussman, David, *Pupils: An Eye Opening Account of Medical Practice Without Standards* (Bartleby 2003).

Taragin, Mark I., Laura R. Willett, Adam P. Wilczek, Richard Trout, Jeffrey L. Carson, The Influence of Standard of Care and Severity of Injury on the Resolution of Medical Malpractice Claims, *Annals of Intern Med* 117:780-84 (1992).

Thomas, Eric J., David M. Studdert, Troyen A. Brennan, The Reliability of Medical Record Review for Estimating Adverse Event Rates, *Annals of Intern Med* 136:812-816 (2002).

Thomas, Eric J., David M. Studdert, Helen R. Burstin, E. John Orav, Timothy Seena, Elliott J. Williams, K. Mason Howard, Paul C. Weiler & Troyen A. Brennan, "Incidence and Types of Adverse Events and Negligent Care in Utah and Colorado," *Medical Care* 38: 261-71 (2000).

Thomas, Eric J., David M. Studdert, William B. Runciman, Robert K. Webb, Elizabeth Sexton, Ross McL Wilson, Robert W. Gibberd, Bernadette T. Harrison, Troyen A. Brennan, A Comparison of iatrogenic injury studies in Australia and the USA I: context, methods, casemix, population, patient and hospital characteristics, *International Journal for Quality in Health Care* 12:371-78 (2000).

Waters, Teresa M., David M. Studdert, Troyen A. Brennan, Eric J. Thomas, Orit Almagor, Martha Mancewicz, and Peter P. Budetti, Impact of the National Practitioner Data Bank on the Resolution of Malpractice Claims, *Inquiry* 40:283-94 (2003).

Weiler, Paul C., Howard H. Hiatt, Joseph P. Newhouse, William g. Johnson, Troyen A. Brennan, Lucian L. Leape, *A Measure of Malpractice* (Cambridge, Mass: Harvard U. P. 1993).

White, Michelle J., The Value of Liability in Medical Malpractice, *Health Affairs* 13:75-87 (1994).

Wilson, David S., John McElligott, L. Peter Fielding, Identification of Preventable Trauma Deaths: Confounded Inquiries? *Journal of Trauma* 32:45-51 (1992).

Wilson, Ross McL, William B. Runciman, Robert W. Gibberd et al, The Quality in Australian Health Care Study, *Med. J. Aust.* 1995 163:458-71 (1995).