A RAND NOTE

Merit Rating for Physicians' Malpractice Premiums: Only a Modest Deterrent

John E. Rolph

RAND

The research described in this report was supported by grants from The Medical Trust, Robert Wood Johnson Foundation and by RAND using its own research funds.

This Note contains an offprint of RAND research originally published in a journal or book. The text is reproduced here, with permission of the original publisher.

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of RAND do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published 1991 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

A RAND NOTE

N-3426-MT/RWJ/RC

Merit Rating for Physicians' Malpractice Premiums: Only a Modest Deterrent

John E. Rolph

Supported by
The Medical Trust
Robert Wood Johnson Foundation

RAND

MERIT RATING FOR PHYSICIANS' MALPRACTICE PREMIUMS: ONLY A MODEST DETERRENT

JOHN E. ROLPH

I

INTRODUCTION

A. Background

Medical malpractice has again attracted significant public concern in the 1990s. Three factors seem most responsible: the expense of malpractice coverage for physicians, its possible effects on patients' access to care, and, finally, the quality of medical care.[1] Since the crisis-spawned creation of physician-owned insurers and state joint underwriting associations in the mid-1970s, availability of coverage has not been a pressing issue. Instead, rising premiums[2] have focused concern on affordability, because higher malpractice premiums cause higher physician fees,[3] thus directly affecting patients' access to care. In addition, there is evidence that some physicians in a few specialties, notably obstetrics, have limited their scope of practice and otherwise withdrawn certain medical services in response to malpractice fears.[4] Finally, malpractice rules potentially affect the quality of medical care

---

Reprinted by permission from *Law and Contemporary Problems*, Vol. 54, No. 2, Spring 1991. Copyright © Law and Contemporary Problems 1991.

This investigation was supported by Tufts University under a subcontract funded by The Medical Trust. Additional support was provided by a grant from The Robert Wood Johnson Foundation and by RAND. Kimberly McGuigan and Elizabeth Lewis assisted with computer programming, data management, and statistical analysis. The Medical Inter-Insurance Exchange of New Jersey generously shared its claims data base and expertise in interpreting the data. Comments by Randall Bovbjerg, Robert Brook, Deborah Hensler, Peter Jacobson, Emmett Keeler, Robert Klitgaard, Richard Kravitz, Kimberly McGuigan, William Reilly, Elizabeth Rolph, and Adam Wilczek on earlier drafts were most helpful.

1. See Frank A. Sloan & Randall R. Bovbjerg, *Medical Malpractice: Crises, Response and Effects*, Health Ins Ass'n Am Res Bull 4, 9-10 (May 1989), for a more complete discussion.

2. See US Gen Acct'g Office, *Medical Malpractice: Six State Case Studies Show Claims and Insurance Costs Still Rise Despite Reforms* 4 (1986) ("*Six Case Studies*"), for documentation of the rise in premiums, and Sloan & Bovbjerg, Health Ins Ass'n Am Res Bull at 5, 6 (cited in note 1), for data on the rise in the frequency and severity of claims between 1980 and 1986. Since 1986 the frequency seems to have fallen off and severity has apparently stabilized. See Harry T. Paxton, *Malpractice Premiums Have Leveled Off—For Now*, Med Econ 52, 56 (May 1, 1989); Janet Ocha Weiner, *Malpractice: Calm Before the Storm?*, Med & Health Perspectives (September 18, 1989).

3. For research on how physicians' fees are affected by the malpractice system, see Frank A. Sloan, *Effects of Health Insurance on Physicians' Fees*, 17 J Hum Resources 533 (Fall 1982); Patricia M. Danzon, Mark V. Pauly & Raynard S. Kington, *The Effects of Malpractice Litigation on Physicians' Fees and Incomes*, 80 AEA Papers & Proceedings 122 (May 1990).

4. For discussions of the effects of medical professional liability on the practice of obstetrics, see *Medical Professional Liability and the Delivery of Obstetrical Care* (Nat'l Acad Press, 1989) ("*Medical*

not only by creating incentives for the oft-cited practice of "defensive medicine,"[5] but also through their incentives for avoiding negligence. This article focuses on the latter issue: how the malpractice system is meant to deter poor quality care. Specifically, we consider whether basing liability insurance premiums partly on a physician's malpractice experience is a feasible improvement.

The three objectives of the tort system are (1) to compensate injured patients,[6] (2) to shift losses from the innocent patient victim to the negligent physician,[7] and (3) to motivate doctors to provide more careful treatment to their patients.[8] Legislative tort reforms in the 1970s largely effected the first objective, compensation.[9] No-fault approaches with compensation goals have also been proposed to speed compensation, cut the cost of handling claims, and reduce the uncertainty about compensation of injured patients.[10] In 1986, the federal Health Care Quality Improvement Act mandated the reporting of malpractice and physician discipline to further deter poor care.[11] However, the existence of malpractice insurance to pay tort damages blunts the economic deterrent effect of the tort system.[12]

Professional Liability"). See specifically chapter 4 of Institute of Medicine, 1 *Medical Professional Liability* (addressing the effects of professional liability on the provision of obstetrical care to poor women); Dana Hughes, et al, *Obstetrical Care for Low-Income Women: The Effects of Medical Malpractice on Community Health Centers*, in Victoria P. Rostow & Roger P. Bulger, eds, 2 *Medical Professional Liability* 59 (same).

5. Tests and procedures that are performed to avoid tort liability rather than to improve medical care are believed to occur on a large scale, but the extent of this phenomenon is difficult to measure. See Roger A. Reynolds, John A. Rizzo & Martin L. Gonzalez, *The Cost of Medical Professional Liability*, 257 J Am Med Ass'n 2776 (May 22/29, 1987), for estimates of the cost of defensive medicine, and Stephen Zuckerman, Christopher F. Koller & Randall R. Bovbjerg, *Information on Malpractice: A Review of Empirical Research on Major Policy Issues*, 49 L & Contemp Probs 85 (Spring 1986), for an analysis of earlier similar estimates.

6. See Randall R. Bovbjerg, *Medical Malpractice on Trial: Quality of Care is the Most Important Standard*, 49 L & Contemp Probs 321 (Spring 1986), for the argument that if deterrence cannot be made to work in the tort system, that system should be scrapped because it is also an inefficient compensation system.

7. For an analysis of tort law as corrective justice, see Ernest J. Weinrib, *Understanding Tort Law*, 23 Valp U L Rev 485 (1989).

8. For a thorough treatment of the deterrence objective in medical practice, see Peter A. Bell, *Legislative Intrusions into the Common Law of Medical Malpractice: Thoughts About the Deterrent Effect of Tort Liability*, 35 Syracuse L Rev 939 (1984).

9. Significant tort reforms include caps on awards for noneconomic damages, changes to the collateral source rule, and tightening of the statute of limitations. Sloan & Bovbjerg, Health Ins Ass'n Am Res Bull at 11-24 (cited in note 1); US Gen Acct'g Office, *Six Case Studies* at 3 (cited in note 2); and Hamilton, Rabinowitz & Alschuler, Inc, *Claim Evaluation Project* (Commissioned by Insurance Services Office, Inc, 1987), all briefly describe various tort reforms' possible effects. Randall R. Bovbjerg, *Legislation on Medical Malpractice: Further Developments and a Preliminary Report Card*, 22 UC Davis L Rev 499 (1989) comprehensively reviews medical malpractice legislation.

10. For discussions of these proposals, see Harvard Medical Practice Study, *Patients, Doctors, and Lawyers: Medical Injury, Malpractice Litigation, and Patient Compensation in New York* 2(11)-2(12) (Harvard U Press, 1990) ("*Patients, Doctors, and Lawyers*"); Physician Ins Ass'n Am, *A Comprehensive Review of Alternatives to the Present System of Resolving Medical Liability Claims* 72-77, 113-124 (Physician Ins Ass'n Am, 1989); Ronald S. Latz, *No-Fault Liability and Medical Malpractice: A Viability Analysis*, 10 J Legal Med 479 (1989).

11. Health Care Quality Improvement Act, 42 USC §§ 11101-11151 (1986).

12. See Harvard Medical Practice Study, *Patients, Doctors, and Lawyers* at 2-3, 2-4 (cited in note 10); Bell, 35 Syracuse L Rev at 954-57 (cited in note 8).

One deterrent tool other types of insurance have long used is "experience" or "merit" rating in setting premiums. "Merit rating," as used here, means the use of the individual policyholder's past claims (or events believed to be correlated with his or her propensity to generate claims, such as traffic tickets in automobile insurance) to adjust premium levels. We prefer the term "merit" rating because it suggests using each individual physician's claims history to determine future premiums. In contrast, "experience" rating can and does also refer to the experience of groups of physicians as a whole. Under merit rating, premium differences provide an incentive to reduce claims frequency and injury rates.[13] However, merit rating is unevenly used in medical malpractice insurance. In Part VI, we develop a merit rating scheme and analyze its deterrent potential.

C. Insurers and Tort Law's Deterrence Goal

Medical malpractice insurers also have several other tools to alter physicians' behavior so as to cut malpractice claims. Insurers can restrict physicians' practices, require further training, supervision, or cost-sharing through deductibles or coinsurance, or, ultimately, cancel the insurance policy. However, such measures are used largely by doctor-owned insurers, and then only infrequently. We conducted a survey of physician-owned companies in 1986 through the Physician Insurance Association of American ("PIAA") and found that 87 percent of responding insurers employed restrictions on practice, 42 percent used further training, and 36 percent used supervision.[14] Fewer than 1 percent of policyholders were affected by these measures. On the economic side, only one of thirty-one respondents used coinsurance and four insurers employed deductibles. Industry sources state that malpractice insurers are reluctant to use deductibles and coinsurance because of possible conflicts of interest between the policyholders and the insurer when defending claims.[15] This reluctance is also confirmed by other research.[16]

Notwithstanding these other measures, premium surcharges or merit rating remain the most commonly used ways of deterring individual

---

13. Michael J. Moore & W. Kip Viscusi, *Compensation Mechanisms for Job Risks: Wages, Workers Compensation, and Product Liability* 151-78 (Princeton U Press, 1989), and Elisabeth Landes, *Insurance, Liability, and Accidents: A Theoretical and Empirical Investigation of the Effect of No-Fault Accidents*, 25 J L Econ 49, 49-65 (1982), give limited support to the effect of merit rating on deterring negligent behavior in work-related accidents and automobile accidents, respectively.

14. Additional results from this survey are reported in William B. Schwartz & Daniel N. Mendelson, *The Role of Physician-Owned Insurance Companies in the Detection and Deterrence of Negligence*, 262 J Am Med Ass'n 1342 (September 8, 1989).

15. Interview with William A. Reilly, vice president, Medical Inter-Insurance Exchange of New Jersey (1989).

16. Schwartz & Mendelson, 262 J Am Med Ass'n at 1344-45 (cited in note 14); and Frank A. Sloan, Randall R. Bovbjerg & Penny B. Githens, *Insuring Medical Malpractice* (Oxford U Press, 1991).

negligence. Such methods are used mostly by physician-owned companies.[17] However, few of the merit rating programs are actuarially based; that is, premium surcharges are, for the most part, set subjectively rather than based on statistical estimates of future claims rates for physicians from past claims.[18]

D. The Feasibility of Actuarially Based Merit Rating

Several recent studies suggest that actuarially based merit rating programs can be effectively implemented for medical malpractice insurance.[19] However, actuarially based premium surcharges are rare. Sloan and his colleagues report that most merit rating programs impose surcharges on fewer than 1 percent of the physicians covered and mandate no more than a doubling of premiums for those surcharged physicians.[20] However, the Massachusetts Division of Insurance has proposed statistically based merit rating,[21] and New York State requires merit rating.[22]

The peculiarities of medical malpractice claims distributions raise some difficult design choices for merit rating plans. Most notably one must decide whether to base surcharges on claims filed or to wait many years until claims have closed and then surcharge only if indemnity was paid. In addition, dollar losses are dominated by the "long tail" (seldom occurring, very large claims) of the distribution of indemnity paid which make it problematic to include claims severity in any feasible merit rating plan.

Part II of this article describes the large, comprehensive claims data base upon which we rely to investigate such design issues and to comment more broadly on the feasibility of using malpractice deterrent policies that target individual physicians. To accomplish this task, we carry out statistical analysis of our database to examine four basic issues, each of which is discussed in a later section of the article. Part III addresses the preliminary question of what aspects of the claims history should be used to measure negligence. Part IV

---

17. Sloan, Bovbjerg & Githens, *Insuring Medical Malpractice* at 315 (cited in note 16), report that thirteen of fourteen insurers they surveyed had used merit rating while in our PIAA survey, twenty-eight of thirty-one (90%) of physician insurers reported using premium surcharges. See Schwartz & Mendelson, 262 J Am Med Ass'n at 1345 (cited in note 14).

18. Interviews with industry underwriters (for example, the interview with William A. Reilly (cited in note 15)), suggest that Pennsylvania Medical Society Liability Insurance Company is one of the few PIAA companies that uses an actuarially-based surcharge program (Penn Rate Schedule, 1990). By contrast, most malpractice insurers assess premium surcharges based on what they judge to be an appropriate "slap on the wrist."

19. For statistical models of merit rating in medical malpractice, see John E. Rolph, *Some Statistical Evidence on Merit Rating in Medical Malpractice Insurance*, 48 J Risk & Ins 247 (June 1981); Frank A. Sloan, et al, *Medical Malpractice Experience of Physicians: Predictable or Haphazard?*, 262 J Am Med Ass'n 3291 (December 15, 1989); Blaine F. Nye & Alfred E. Hofflander, *Experience Rating in Medical Professional Liability Insurance*, 55 J Risk & Ins 150 (March 1988); Randall P. Ellis, Cynthia L. Gallup & Thomas G. McGuire, *Should Medical Professional Liability Insurance be Experienced Rated?*, 57 J Risk & Ins 66 (March 1990); Bruce K. Cooil, *Using Medical Malpractice Data to Predict the Frequency of Claims: A Study of Poisson Process Models*, 86 J Am Statistical Ass'n 285 (June 1991).

20. Frank A. Sloan, *Experience Rating: Does it Make Sense for Medical Malpractice Insurance?*, 80 AEA Papers & Proceedings 128, 128 (May 1990); Sloan, Bovbjerg & Githens, *Insuring Medical Malpractice* at 315 (cited in note 16).

21. State Rating Bureau, Commonwealth of Massachusetts Division of Insurance (1989).

22. NY Insurance Law § 2343 (McKinney, 1990).

assesses whether malpractice-prone physicians do in fact exist, or, conversely, whether claims occur at random. Part V, in contrast to Part III, asks a more substantive question concerning claims history: if certain physicians are malpractice-prone, what elements of their claims histories should be used to identify them prospectively. Part VI evaluates the effectiveness of merit rating, which uses claims history to identify and target malpractice-prone physicians. Finally, the article addresses the potential implications of our findings on the deterrent effect of merit rating.

II

DATA: MANY YEARS' EXPERIENCE FROM A POPULOUS STATE

Data came from the Medical Inter-Insurance Exchange of New Jersey ("The Exchange"), which has grown since early 1977 to insure approximately 8,000 physicians, comprising about 70 percent of those practicing in the state. We were fortunate to have data from more than ten years of their claims experience (from February 1, 1977, through June 30, 1987). We analyzed closed claims arising from incidents that occurred on or before June 30, 1984, thereby allowing a minimum of three years for claims to close.

The Exchange groups physicians into premium classes according to their specialties and surgical activities. Table 1 gives the specialties and surgical risks defining each of the eleven premium classes[23] analyzed here with the corresponding 1988 annual premium. Note that premiums vary by a factor of ten from the maximum to the minimum.

Table 2 shows the average frequency of paid claims and the average indemnity per paid claim for the eleven premium classes. Comparing the data from Table 1 and 2 shows that, with a few exceptions, the frequency of paid claims increases monotonically as premiums rise. However, the average indemnity per paid claim is not a strictly increasing function of the premium, presumably, in part, because of the sensitivity of averages to the occasional very large claim.[24] More generally, the higher the premium class, the larger the expected loss per physician (the frequency times the severity of claims). However, the fact that the expected loss increases with rating class does not imply that both frequency (which measures liability) and severity (which measures damages) must increase. Indeed, we show in Appendix I that they are independent.

---

23. We elected to analyze claims histories for physicians who were primarily clinicians and therefore excluded the two lowest risk premium classes. These categories include administrative medicine, insurance company employees, pharmaceutical company employees, physical medicine/rehabilitation, preventive occupational medicine, psychiatry/neuropsychiatry, public health, and retired.

24. For example, Premium Class I (consisting of anesthesiologists and urologists) is not well described by the average indemnity per claim. The class has the highest proportion of all filed claims resulting in indemnity being paid (37%). This higher payout rate is caused by the large number of very small claims for minor injuries such as claims against anesthesiologists for chipped teeth. These, combined with a few major anesthesia mishaps, produce a middling average indemnity per paid claim ($68,000).

TABLE 1

PREMIUM CLASSES AND CONSTITUENT MEDICAL SPECIALTIES

Premium* Class	Annual Premium** (in $000s)	Specialties Within Class***
C	4	Allergy, Forensic, Hematology, Manipulation, Oncology, Pathology
D	6	Dermatology (min), EENT (no) FP (no) GP (no), Geriatrics, Industrial (no), Neurology (min), Nuclear, Opthamology (no), Pediatrics (no), School Physician
F	7	Acupuncture, Cardiology, EENT (min, maj), ER (off), FP (min), GP (min) Gynecology (off, min), Industrial (min), Internal Medicine (General, Cardiology, Endocrinology, Hematology, Nephrology, Oncology, Pulmonary Disease, Rheumatology), Industrial (min), Ophthamology (min, maj), Orthopedics (off), Otolaryngology (off, min), Pediatrics (min), Pediatric Cardiology, Pulmonary Therapy, Scierotherapy, Urology (off)
G	10	ER (H), Gastroenterology, Internal Medicine (Gastroenterology), Radiology, Roentgenology
H	13	Dermatology (maj, asst), ER (asst), FP (maj, asst), GP (asst, maj), Gynecology (min, asst), Internal Medicine (asst), Neurology (min, asst), Orthopedics (asst), Otolaryngology (off, asst), Pediatrics (maj), Proctology, Surgical Assistant, Urology (off, asst)
I	19	Anesthesiology, Urology (maj)
J	25	EENT (maj), General Surgery, Gynecology (maj), Industrial (maj), Otolaryngology (maj), Pediatric Surgery
K	28	Cardio-Thoracic Surgery, Cardio-Vascular Surgery, Hand Surgery, Plastic Surgery, Thoracic Surgery, Vascular Surgery
L	31	Obstetrics and Gynecology
N	35	Orthopedics (maj)
P	42	Neurosurgery

* No significance should be attached to gaps in the lettering of classes. They occur for historical reasons.

** SOURCE: 1988 Premium Rate Schedule, The Medical Inter-Insurance Exchange of New Jersey. Claims Policy, Not Discounted, $1/1/3 million coverage (medical incident/aggregate policy period/aggregate-extended policy period).

*** Abbreviations: maj- major surgery, min - minor surgery, no -no surgery, asst- assisting surgery practice, off - non-hospital or office practice, EENT -ear, eye, nose, throat, FP - family practice, GP - general practitioner, ER - emergency room.

TABLE 2

PAID CLAIMS FREQUENCIES AND AVERAGE PAYOUT BY PREMIUM CLASS*

Premium Class	Annual Number of Paid Claims Per 100 Insured MDs	Average Indemnity per Paid Claim (in $000's)**	Variance to Mean Ratio (p-value)***	Number of Doctors
C	1	84	1.1 (.30)	133
D	2	67	1.2 (.01)	557
F	4	79	1.1 (.00)	1198
G	4	67	1.1 (.14)	318
H	8	86	2.0 (.00)	53
I	7	68	1.3 (.00)	242
J	7	91	1.2 (.03)	248
K	9	66	1.7 (.00)	67
L	11	96	1.1 (.28)	129
N	14	88	1.3 (.03)	135
P	9	164	1.7 (.00)	19
Total	5	85		3099

* Data in this table based on data for physicians insured by the Exchange during the entire 1977-1984 peroid.

** Indemnity in 1987 dollars.

*** p-values are based on a likelihood ratio test using the deviances. See Jerald F. Lawless, *Negative Binomial and Mixed Poisson Regression*, 15 Can J Stat 209 (September 1987).

III

PAID CLAIMS FREQUENCY AS A MEASURE OF NEGLIGENCE

The primary public policy rationale for using merit rating to set malpractice premiums is to deter negligent physicians. This goal raises three distinct questions. First, how should negligence be measured retrospectively? Second, do malpractice-prone physicians exist? And third, if they exist, how should they be identified prospectively? The first question is addressed in this section, the second in Part IV, and the last in Part V.

The only readily available measure of medical negligence is the malpractice claim. However, it should be noted that the connection between negligent care and malpractice claims has been long debated.[25] For the purposes of this article, we assume that deterring malpractice claims also deters negligence.

---

25. In recent works, Patricia M. Danzon, *Medical Malpractice: Theory, Evidence and Public Policy*, (Harvard U Press, 1985), and Harvard Medical Practice Study, *Patients, Doctors, and Lawyers* (cited in note 10), discuss data from the 1970s and 1980s, respectively, that bear on this connection. Harvard Medical Practice Study, *Patients, Doctors, and Lawyers* (cited in note 10), attempts to measure deterrence. While the study admittedly was not designed to assess deterrence, the data showed negligible or at most small deterrent effects for the malpractice claims system. Patricia M. Danzon, *Liability for Medical Malpractice: Incidence and Incentive Effects*, presented at the RAND Health Economics Conference, March 1990, also attempts to measure deterrence with a similar lack of definitive conclusion.

Since we have chosen to treat malpractice claims as our window on physician negligence, we must decide what aspects of those claims best measure medical malpractice. Three aspects of the physician's claims history might be used either alone or in combination: (1) the total number of claims filed against a physician including both paid and unpaid claims (where "unpaid" claims are those for which no indemnity payment is made in the settlement or verdict); (2) the number of paid claims; and (3) the average amount of the indemnity paid on each claim (claims severity). Because the tort system only recognizes negligence in the case where indemnity is actually paid, we chose not to use claims filed as a measure of negligence. Similarly, because the tort system awards damages based upon the degree of injury sustained by the patient, rather than the extent of the physician's negligence, greater claims severity is not necessarily an indicator of greater negligence. Moreover, we show in Appendix I that individual physicians do not have systematically different levels of indemnity payments, so that claims severity is not even a helpful indicator of the damage a physician does when a claim is successful. Thus, since neither claims filed or claims severity is a useful criterion, we measure an individual's negligence solely by the number of paid claims.[26]

IV

OBSERVED VARIATION IN CLAIMS PROPENSITIES: RANDOM OR PREDICTABLE?

Since the point of focusing malpractice reduction policies on individual physicians is to deter future malpractice claims, the next question to ask is: Do malpractice claims-prone individual physicians exist? That is, within a premium class,[27] is there evidence that individual physicians systematically differ in their propensities to generate paid claims? If not, targeting one physician rather than another would serve no purpose. Thus, individual physicians' propensities to generate paid claims must vary in order for there to be a statistical basis for merit rating.

A. Testing Randomness of Paid Claims

Accordingly, the next question we address is whether the statistical evidence shows that paid claims were nonrandom within each premium class. Technically, we investigate whether the number of paid claims per physician follows a Poisson distribution.[28] Applied to the set of physicians in all eleven

26. Paid claims are, of course, an imperfect measure of negligence and are affected by the social and legal factors that affect the tort system.

27. Premium classes of physicians are defined by specialty and surgical activity. These are described in Part II. Premiums, in principle, reflect the expected payout per physician per year, that is, claims frequency times claims severity. Strictly speaking, premiums are based on the discounted present value of expected future claims costs (indemnity, adjustment costs, and overhead).

28. The Poisson distribution is a probability distribution that is often used as a mathematical model of the number of outcomes obtained in a suitable interval of time when the outcomes occur independently. It has the mean equal to the variance. For further details, see William H. Kruskal &

premium classes as a group, the test of "Poissonness" has a significance probability of less than 1 percent. That is, if the number of paid claims per physician were truly random, the probability that the observed distribution of paid claims per physician would be as broad as in our data is less than 1 percent.

However, since we expect variation between premium classes in the claims frequencies, the relevant test is whether the distribution of the number of paid claims per physician is random (Poisson distributed) *within* each premium class. Table 2, above, also presented the "variance-to-mean ratio." This ratio measures the dispersion of the distribution of the number of paid claims per physician for each premium class.[29] Values above 1.0 indicate nonrandomness, and the significance probabilities (p-values) in Table 2 measure Poissonness within each class. The distribution of paid claims per physician differs from a Poisson distribution at the 5 percent significance level for all but three premium classes.[30] Thus, for roughly the 80 percent of all physicians in the remaining eight premium classes, there is good statistical evidence of differences in propensities to generate paid claims.[31]

B. Characterizing the Nonrandomness

Our next step was to determine the magnitude of this difference in claims propensity. For the nonrandom premium classes, we estimated the distribution of the underlying paid-claims propensities. These estimated distributions are characterized in Table 3 (excluding premium classes C, G, and L).[32] The expected number of paid claims per doctor per five years for each premium class is given in the column labeled "average" and ranges from .15 to 1.07 paid claims per five year period. The "Bottom and Top Quartile" values represent the 25th and 75th percentile of this distribution.

It is remarkable how much overlap exists in the middle half of the distribution of paid-claims propensities across all premium classes. The paid-claims propensities for low-risk physicians are even more striking; six of the

---

Judith M. Tanur, eds, 1 *International Encyclopedia of Statistics* 156-57 (The Free Press, 1978). Appendix II describes in more detail the statistical rationale behind this approach and its interpretation.

29. As explained in Appendix II, variance-to-mean ratios for Poisson distributions should be near 1.0. Values greater than 1.0 indicate variability across physicians in their propensity to generate paid claims.

30. These classes are: C (allergy, forensic medicine, hematology, manipulation, oncology, and pathology); G (gastroenterology, radiology, roentgenology, and emergency room medicine); and L (obstetrics and gynecology).

31. If the paid-claims data were consistent with all physicians in a given premium class having the same underlying propensity to generate paid claims, then (in theory) no amount of claims history could differentiate individuals in a statistically valid way within the class. In statistical parlance, "consistent" means that the number of paid claims of physicians in the class follows a Poisson distribution with the same mean.

32. The choice of five years is somewhat arbitrary, but is consistent with that typically used in the industry. See Medical Inter-Insurance Exchange of New Jersey, High Risk Evaluation Program (August 26, 1983). The underlying gamma mixture of Poisson distribution to get the fitted Negative Binomial distribution of claims is based on all data available here—7.4 years, allowing for the 3 years of "runoff" (see Appendix II). For a description of how the Negative Binomial distribution arises in this way, see Kruskal & Tanur, eds, 1 *International Encyclopedia of Statistics* at 157 (cited in note 28).

eight classes show at least a quarter of their physicians estimated with fewer than .04 claims per five years. Moreover, the classes with lower premiums generally have smaller "interquartile ranges" (differences between top quartile and bottom quartile) than the higher premium classes. The higher classes have both higher average frequencies and larger interquartile ranges. Physicians in Premium Class N (orthopedics) have the largest expected paid claims rate (1.07). Among the low premium classes, Class H is most diverse in paid claims propensities; with an interquartile range of .53, reflecting the differences in paid-claims propensities across the many specialties in this class. (See Table 1). The "mid-ranges" (bottom quartile to top quartile) of the propensity distributions overlap across adjacent premium classes. Thus, while the classes differ in their average underlying paid-claims propensities, there are many doctors in higher premium classes who are less prone to generating paid claims than colleagues in lower rate classes.

TABLE 3

VARIABILITY IN PAID-CLAIM PROPENSITIES BY PREMIUM CLASS

	Frequency of Expected Number of Paid Claims for Five Years				
Premium Class*	Average	Bottom Quartile	Top Quartile	Interquartile Range	Ratio of Top to Bottom Quartile (rounded)
D	.15	.03	.20	.17	7
F	.21	.02	.39	.37	20
H	.65	.19	.72	.53	4
I	.51	.02	.62	.60	31
J	.57	.04	.73	.69	18
K	.78	.04	.98	.94	24
N	1.07	.04	1.30	1.26	32
P	.75	.27	1.04	.77	4

* Only premium classes with statistical evidence of heterogeneous paid claims propensities used. See Table 2.

The spread of the quartiles in most premium classes is evidence of the extreme differences in individual physicians' paid-claims propensities. The rightmost column in Table 3 gives the ratio of the top quartile to the bottom quartile and ranges from a minimum of four to a maximum of thirty. Interpreting the top quartile as the paid-claims propensity of the typical high-risk physician and the bottom quartile analogously for low-risk physicians, these typical doctors differ by an order of magnitude in their proneness to generate paid claims.

Thus, the data show that there are systematic differences between physicians in their proneness to generate paid claims. These differences provide a basis for merit rating. Our analysis is consistent with other research[33] suggesting that claims frequencies for physicians in other claims data bases vary more than the randomness assumption would imply.

33. See note 19.

V

FILED CLAIMS, PAID CLAIMS, OR CLAIMS SEVERITY AS A PREDICTOR?

Establishing that physicians differ systematically in their propensities to generate paid-claims raises the question of how best to identify malpractice-prone doctors prospectively for merit rating or other targeted deterrent policies. Claims histories offer useful data for this predictive task. However, the question in this prospective context is which aspects of claims histories—total claims filed, paid-claims or claims severity—has the greatest predictive power?

If the percentage of claims resulting in payment were exactly the same for every physician, a physician's propensity to generate unpaid claims would be a perfect predictor of his or her paid claims. If this were even approximately true, there would be an actuarial basis for insurers to use a physician's unpaid claims rate in addition to his or her paid claims frequency as a predictor of future paid claims. That is, an insurer could use claims filed rather than waiting for claims to be closed and their disposition known. Alternatively, if individual physicians' rates of paid claims and unpaid claims were statistically independent, unpaid claims would have no predictive value for paid claims and should not be used.[34]

The frequency of unpaid malpractice claims (71 percent of the closed claims in our data) is sometimes cited as evidence of the malpractice claims system's capriciousness in its attempt to identify medical negligence.[35] Our analysis in Appendix III shows that the correlation between the number of an individual's paid claims and the number not paid is only modest. The lack of a clear systematic positive relationship between individual physicians' paid and unpaid claims rates suggests that unpaid claims do not predict a physician's tendency toward medical negligence.

In addition, using unpaid claims to help set premiums may lead to other practical problems for the insurer. For example, one insurer encountered substantial resistance from policyholders when it based surcharges exclusively on the number of lawsuits filed against a policyholder irrespective of whether indemnity was ultimately paid. The policyholders complained that this system was unfair because many surcharges stemmed from lawsuits where the insurer paid no indemnity.[36] These types of practical problems, together with the unpaid claim's poor properties as a predictor of paid claims, led us to exclude no-pay claims from our merit-rating calculations.

The third aspect to consider in targeting malpractice-prone doctors is claims severity. Appendix I demonstrates that claims severity and paid-claims frequency are not related, at least in a simple way. Appendix I also shows that

---

34. Strictly speaking, the expense of investigating and litigating unpaid claims is included in developing the premium, but we ignore the small contribution that this consideration would make.

35. Sloan & Bovbjerg, Health Ins Ass'n Am Res Bull (cited in note 1).

36. Closed claims studies have shown that even among malpractice claims in which a lawsuit is filed, almost half result in no payment to the claimant. See US Gen Acct'g Office, *Medical Malpractice: Characteristics of Claims Closed in 1984*, 82 Table V.14 (April 1987).

individual physicians do not systematically differ within a rating class in the size of their claims payments. Therefore, claims severity should not be used as a discriminator between physicians.[37] Accordingly, the lack of relation between claims severity and claims frequency, and our exclusion of unpaid claims, led us to measure an individual's malpractice by paid claims frequency only.

VI

TARGETING PHYSICIANS WITH A MERIT RATING PLAN

Evidence that physicians vary widely in their proneness to generate paid claims suggests that insurers could target physicians with deterrent policies. Realizing this potential depends on accurately differentiating between more and less claims-prone physicians in order to effectively deter negligence. Two empirical questions arise: (1) How much of the differential in physicians' underlying paid-claims propensities is predictable from past paid-claims histories; and (2) how much does a merit rating plan based on a physician's past paid claims affect premiums, and for how many physicians does it do so?

How well one can identify physicians with different propensities to generate paid claims is critical to a merit rating plan, or for that matter, to any policy targeted to reduce malpractice claims. Part IV established that doctors vary in their propensities to generate paid-claims. We now turn to how well past paid claims histories can identify individual physicians with differing propensities.

A. A Statistical Model for Merit Rating

To answer this question, we develop a statistical model predicting future paid claims for each individual physician based on past paid claims. A merit rating plan constructed from these predictions characterizes, in practical terms, how well we can differentiate the high- and low-risk doctors. Physicians' rating classes are based on specialty and performance of surgery, as indicated in Table 1. Both factors measure some aspects of exposure to malpractice claims.[38] As described in Appendix II, for each premium class we

37. The logic here lies in the malpractice insurer's wish to avoid future indemnity costs (we ignore the relatively small contribution of adjustment costs). Since there is no relationship between payment size for the same individual after controlling for premium class and specialty, the number of paid claims is the relevant criterion for predicting future paid claims. Note that this lack of a relationship between severity of claims paid on the one hand and premium class and specialty on the other is contrary to that found by Sloan et al, 262 J Am Med Ass'n 3291 (cited in note 19).

38. However, they do not directly capture the amount of exposure in the same way that "miles driven" measures exposure for automobile insurance risks. Malpractice insurers do not currently collect and use other factors that could be used to measure exposure and help define premium classes. These factors include the number of patient encounters, the amount of surgery, other measures of practice volume, and patient or procedure mix. Therefore, such factors are not currently used as part of merit rating plans. It is worth noting that the disparities in individual physician claims rates reported here and elsewhere are due, in part, to these individual differences. In an ideal world, we would measure and adjust for these factors in order to identify physicians who have higher paid claims propensities than their colleagues with comparable exposure.

fit a statistical model to the frequency of paid-claims. The model captures the diversity in individual paid-claims propensities. Standard actuarial techniques allow us to use a physician's past paid-claims history to predict future claims.[39]

B. A Merit Rating Plan with Modest Accuracy

Table 4 presents the merit rating scheme we developed. It illustrates the scheme's effect on physicians' premiums.

TABLE 4

PERCENTAGE PREMIUM SURCHARGES BASED ON FIVE YEARS OF PAID CLAIMS

Premium Class*	Expected Number of Paid Claims in next 5 Years	Percentage Premium Change, by Number of Paid Claims in Past 5 Years			
		0	1	2	3 or more***
			(% of MDs in each subclass)		
D	.15	−10%	55%	119%	184%
		(.87)**	(.11)	(.02)	(.00)
F	.21	−10%	38%	85%	133%
		(.84)	(.13)	(.03)	(.01)
H	.65	−39%	21%	82%	142%
		(.65)	(.23)	(.08)	(.04)
I	.51	−15%	14%	44%	74%
		(.73)	(.15)	(.06)	(.06)
J	.57	−19%	14%	48%	82%
		(.70)	(.17)	(.07)	(.07)
K	.78	−23%	7%	37%	68%
		(.65)	(.17)	(.08)	(.10)
N	1.07	−27%	−2%	23%	49%
		(.61)	(.16)	(.08)	(.14)
P	.75	−49%	16%	81%	146%
		(.56)	(.26)	(.11)	(.07)

* Only premium classes with statistical evidence of heterogeneous paid claims propensities used. See Table 2.

** Proportion of doctors estimated to be in each category is given in parentheses. Numbers may not add due to rounding.

*** Values based on three claims. Surcharges based on 4 or more claims would be correspondingly higher if such small numbers of physicians were put in a separate class.

The percentage entries give doctors' hypothetical premium changes (as determined from a base premium of a doctor with no experience and thus no claims history) as a function of past paid-claims frequencies over the past five years. The entries in parentheses give the proportion of doctors in the class with the given number of paid claims. For example, for Premium Class D, most doctors (87 percent) have no paid claims. They would receive a ten percent discount compared to the "no experience" premiums, while the

39. Appendix II describes the distributional assumptions and the statistical fitting methods used.

physicians with two, three, or more paid claims (only two percent of the class) would be surcharged over 100 percent (119 percent for two claims and 184 percent for three or more claims). On the other hand, Class N has the highest average paid-claims rate, with 22 percent of the doctors having two or more claims. This results in a smaller surcharge (23 percent for two claims; 49 percent for three or more claims) than those charged in other classes.

Table 4 allows comparisons of doctors with different histories of paid claims in different premium classes. Consider two hypothetical physicians: Doctor Smith is an anesthesiologist (Class I) with no paid claims in the past five years, a trait shared with 73 percent of his premium class, while Doctor Jones is a dermatologist (Class H) who performs major surgery, with two paid claims (along with 8 percent of his premium class) during the same period. Using the Table 1 premium and the Table 4 surcharge, their annual premiums (rounded to thousands of dollars) are:

	Premium without Merit Rating	Premium with Merit Rating
Dr. Smith	$19,000	$16,000
Dr. Jones	$13,000	$24,000

Without merit rating, Dr. Smith pays about 50 percent more than Dr. Jones. With premium surcharges, the reverse is true. Striking reversals are possible under an actuarially based merit rating plan.

However, there are limits to the magnitude of premium surcharges. First, under the merit rating program, doctors in the lowest premium classes (D, F) can never pay more than those in the major surgical classes (J, K, N, P) no matter what their respective paid claims experience. Second, since paid claims rates over five years in the low-risk classes are zero for most physicians (87 percent in D and 84 percent in F), only a small proportion of physicians in these classes pay surcharges; consequently, the discount for no paid claims is only ten percent. This calculation illustrates the principle that a merit rating plan results in premium surcharges and discounts that must, when averaged over the entire class, produce the same average premium collected.

In examining the effect of our simple merit rating scheme, we found that some premium surcharges could exceed 100 percent. Table 4 implies that only one-half of 1 percent of the physicians covered by our insurer during the study period could have such large surcharges. Further, this group accounts for less than 2 percent of the paid claims (calculated from the values given in Table 4). If we define this group as "identifiably claims-prone," its contribution to the incidence of paid claims is statistically significant, but relatively small as measured by the number of physicians affected in a substantial way.

Finally, in order to assess how well the differential in physicians' underlying claims propensities are predictable from past-paid claims histories, we compared the variability of observed claims rates among high- and low-risk physicians identified by the merit rating plan with the variability of claims propensities. In very rough terms, the ratio of the percentage premium

surcharge for physicians with no claims in five years to those with one claim (Table 4) can be thought of as being analogous to the ratio of the lower and upper quartiles of the distribution of propensities for most classes (Table 3). The ratios of surcharges have a maximum of 2.3 (116/51) for Class P and are much smaller for most of the other classes. This is an order of magnitude difference from the maximum quartile ratio in the distribution of paid claims propensity of 32. This rough comparison illustrates the relatively low proportion of variability that can be explained using past paid claims as a predictor of true claims-generating propensity. Therefore, physicians' past paid-claims histories have only modest discriminatory power between high propensity and low propensity physicians with respect to generating paid claims.

VII

CONCLUSION

Our study has several implications. First, while there are malpractice-prone physicians, the potential effect of deterrent policies that target individual physicians using readily available information (paid claims) is modest. That is, our results show that the "targeting" achievable from past paid-claims histories is only moderately accurate.

Second, however, it is possible that gathering more detailed information about physicians in addition to claims history and premium class might lead to more accurate prospective identification of those who will incur future paid claims. Recent work suggests that adding information on an individual physician's demographic characteristics and practice patterns increases predictive accuracy.[40]

Finally, even though we have shown that the contribution of identifiable claims-prone physicians to malpractice is small, this finding does not address the effect, if any, of targeting policies in deterring all physicians from engaging in negligent behavior. Policies in this category (other than merit rating) include targeted education, practice monitoring, and other individual specific interventions. Even if the targets of a merit rating program (that is, those having excessive paid claims) account for only a small share of claims, other physicians may find the potential stigma and economic consequences of premium surcharges substantial enough to cause them to improve their practice habits or suspend their medical practices. This possible deterrent effect is a rationale of many public policies that directly target potential "offenders," such as giving speeding tickets. Unfortunately, measuring the deterrent effect of such policies is usually a daunting task; recent research has once again made this point.[41]

---

40. See Sloan, et al, 262 J Am Med Ass'n at 3295 (cited in note 19). However, definitional differences do not allow direct numerical comparisons to our results. We have research underway that will allow such comparisons.

41. See Harvard Medical Practice Study, *Patients, Doctors, and Lawyers* (cited in note 10). The results are not definitive.

APPENDIX I

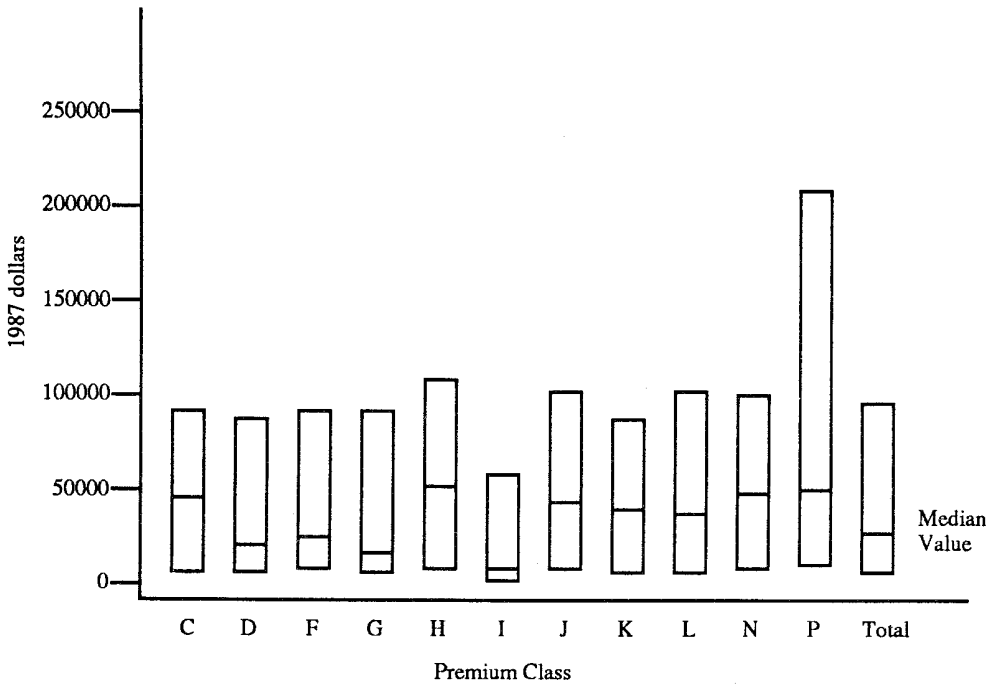
RELATING CLAIMS SEVERITY TO PAID CLAIMS RATES

This Appendix describes the distribution of indemnity payouts by premium class and shows the lack of relation between payout size and the frequency of paid claims. This lack of relationship is why claims severity was not used in the merit rating plan.

Do indemnity payments vary by premium class? Table 2, above, shows that on average they do vary. Figure 1 presents the three quartiles of the distribution of indemnity paid per paid claim for the period 1977-84 using "box plots."[1]

FIGURE 1

DISTRIBUTION OF INDEMNITY PAYMENTS: MIDDLE 50 PERCENT OF DISTRIBUTION AND MEDIAN VALUES BY PREMIUM CLASS



We use quartiles rather than means and standard deviations to describe the distributions in order to eliminate the large influence of a single extremely high indemnity payment. Looking at total paid claims (far right), indemnity payments have skewed distribution, with the median claim being $30,000 in 1987 dollars (the line across the interior of the box), a quarter of the claims being less than $8,830 (the bottom of the box), and a quarter of the claims being more than $90,000 (the top of the box). Because of the skewed

1. For more details on box plots and their uses, see Paul F. Velleman & David C. Hoaglin, *Applications, Basics, and Computing of Exploratory Data Analysis* 65-92 (Duxbury Press, 1981).

distribution, the mean claim is about $85,000 (Table 2)—much closer to the top quartile than to the median of the distribution. The box plots of payouts in Figure 1 depict the considerable variation in the distribution of indemnity paid per claim in the different premium classes. Like the mean, the median indemnity per claim varies by premium class. The distribution of indemnity per claim is skewed for every premium class, and the quartiles fluctuate across classes. The interquartile range tends to increase for the higher-risk physicians (moving from left to right). The difference between classes in the distribution of indemnity per claim is statistically significant.[2]

Some studies suggest that physicians who experience high frequency of claims have higher payouts per claim, but the evidence is not compelling.[3] Figure 2 shows no systematic relationship between physicians' indemnity payouts and paid-claims frequencies in our data base. The distributions of indemnity payments exhibit no particular pattern in their medians or their top and bottom quartiles. The differences in distributions are not statistically significant even at the 10 percent level.[4] Within each premium class, this same lack of relationship holds. A comparable analysis using claims filed rather than claims paid yielded similar results. In sum, physicians with high frequencies of paid-claims do not tend to have systematically higher indemnity payments than their colleagues.

---

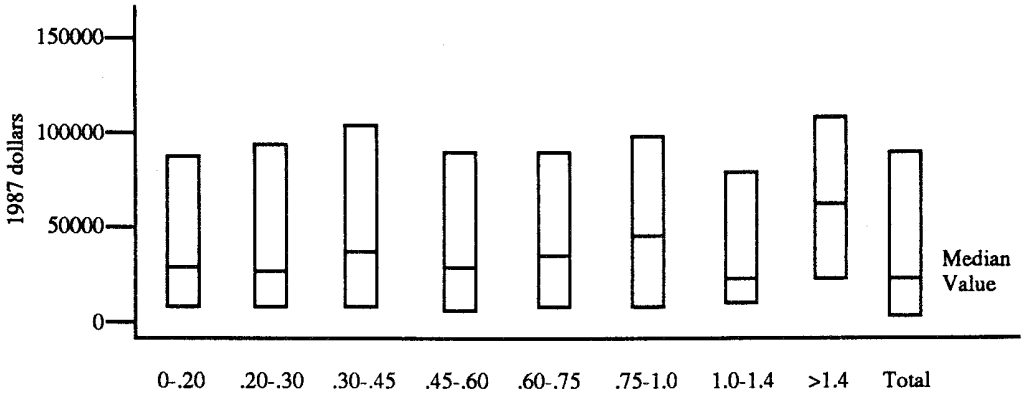
2. The non-parametric Kruskal-Wallis H test was used to test whether there was any statistical difference in the distributions of indemnity payouts by premium class. It was significant at the .01 level. For a description, see Erich L. Lehmann, *Nonparametrics: Statistical Methods Based on Ranks* (Holden-Day, 1975).

3. See John E. Rolph, *Some Statistical Evidence on Merit Rating in Medical Malpractice Insurance*, 48 J Risk & Ins 247 (June 1981); Blaine F. Nye & Alfred E. Hofflander, *Experience Rating in Medical Professional Liability Insurance*, 55 J Risk & Ins 150 (March 1988); Frank A. Sloan, et al, *Medical Malpractice Experience of Physicians: Predictable or Haphazard?*, 262 J Am Med Ass'n 3291 (December 15, 1989).

4. Using the Kruskal-Wallis H test, described in App note 2.

FIGURE 2

DISTRIBUTION OF INDEMNITY PAYMENTS: MIDDLE 50 PERCENT OF DISTRIBUTION AND MEDIAN VALUES BY PAID CLAIMS CATEGORY



APPENDIX II

A STATISTICAL MODEL OF THE CLAIMS GENERATING PROCESS

Our statistical model of paid claims per doctor postulates that each physician in a given premium class has an underlying propensity for generating paid claims. Given this propensity, paid claims occur independently of one another over time, that is, more or less randomly. Hence, the number of paid claims against each physician has a Poisson distribution with mean LT where L is the physician's propensity for generating paid claims each year and T is the number of years of exposure.[1]

For merit rating purposes, the question is whether L varies across individuals or is constant across all physicians in a rate class. So we test whether L is constant across doctors within a given premium class by computing a statistical test of whether the number of physicians with 0, 1, 2, etc., paid claims is consistent with the number sampled from a Poisson distribution.[2] As a summary measure of the spread of the distribution of physicians' propensities to generate paid claims, we give the ratio of the variance to the mean of the number of paid claims per doctor in each premium class.[3] If all doctors in each premium class had the same propensity, the number of paid-claims per physician would be Poisson distributed, and hence the variance-to-mean ratio would be near 1. If not, variation in individual propensities would produce a variance-to-mean ratio greater than 1.

For those premium classes where statistical tests reject the hypothesis of L being constant across physicians, we use a Negative Binomial distribution to model the number of paid claims per physician.[4] This approach incorporates variability in individual propensities by postulating that physician i has underlying paid-claims generation rate $L(i)$ drawn from a Gamma probability distribution. The family of Gamma distributions includes a wide range of shapes and scales of the density and is widely used in actuarial work.[5] The probability of having future paid claims is related to a physician's paid claims history through Bayes' Theorem.[6]

The analysis presented here uses data from the 3,099 physicians who were insured by the Exchange for the entire 7.4-year period (1977-84). Based on these data, our merit rating plan uses predictions of future paid-claims

---

1. See John E. Rolph, *Some Statistical Evidence on Merit Rating in Medical Malpractice Insurance*, 48 J Risk & Ins 247, 249 (June 1981).

2. For the test, which is based on deviance, see Jerald F. Lawless, *Negative Binomial and Mixed Poisson Regression*, 15 Can J Stat 209, 218-19 (September 1987). We test the entire distribution for departures from Poissonness, not just the variance-to-mean ratio referred to below.

3. For the properties of this measure, see James S. Hodges, *Modeling the Demand for Spare Parts: Estimating the Variance-to-Mean Ratio and Other Issues* 8-29 (RAND, 1985) (No N-2086-AF).

4. For detailed derivations regarding automobile insurance, see Joseph Ferreira, Jr., *The Long Term Effects of Merit-Rating Plans on Individual Motorists*, 22 Op Res 954 (1974); Jean Lemaire, *A Comparative Analysis of Most European and Japanese Bonus-Malus Systems*, 55 J Risk & Ins 660 (1988). Others use this approach for medical malpractice claims. See, for example, Rolph, 48 J Risk & Ins 247 (cited in note 1); Blaine F. Nye & Alfred E. Hofflander, *Experience Rating in Medical Professional Liability Insurance*, 55 J Risk & Ins 150 (March 1990).

5. See Kimiko O. Bowman & L. R. Shenton, *Properties of Estimators for the Gamma Distribution* (Marcel Dekker, 1988).

6. Dennis V. Lindley, *Making Decisions* 97 (John Wiley & Sons, 2d ed 1985).

frequencies generated from the above statistical model to set premium surcharges.

APPENDIX III

RELATING PAID CLAIMS TO UNPAID CLAIMS

Across all physicians, there is a positive association between a physician's paid claims and unpaid-claims frequencies as shown in Table A.[1]

TABLE A

AVERAGE PAID CLAIMS PER 100 MD-YEARS BY NUMBER OF UNPAID CLAIMS AND PREMIUM CLASS

Premium Class	All	Number of Unpaid Claims in 7.4 Years					Rank Correlation	Significance Probability	Number of Policyholders
		0	1	2	3	4 or more			
C	2	2	2	0	0	-	.01	.67	(133)
D	3	2	3	3	9	6	.09	.00	(557)
F	4	4	4	5	6	7	.08	.00	(1198)
G	6	7	6	5	10	6	.02	.64	(318)
H	13	8	8	20	2	2	.27	.00	(53)
I	10	11	8	11	13	18	.08	.10	(242)
J	11	13	9	9	13	18	.07	.04	(248)
K	16	7	16	20	11	18	.17	.59	(67)
L	18	12	12	17	20	24	.23	.00	(129)
N	22	6	17	25	16	25	.18	.04	(135)
P	15	-	-	7	18	15	.11	.16	(19)
Total	7	4	6	9	11	17	.24	.00	(3099)

Table A also relates the frequencies of unpaid and paid-claims per physician within premium class. The left side of the table shows how the average paid-claims rates (per 100 physician-years) vary with the number of unpaid claims for each premium class. For example, in Premium Class L, obstetrics and gynecology, the paid claims rates are 12, 12, 17, 20, and 24 for the groups of doctors with 0, 1, 2, 3, and 4 or more unpaid claims respectively. The rank correlation column summarizes the association between physicians' paid and unpaid claims, and the significance probability gives the results of testing whether the rank correlation is zero. (Conventional custom deems "statistically significant" a value of .05 or less.)

A positive relationship between paid and unpaid claims rates exists only within a few premium classes—notably H and L where the rank correlations are above .10 (.27 and .23) and are statistically significant at the .01 level. For the premium classes D and F, which have the largest policyholder populations (557 and 1,198), the rank correlations are smaller (.09 and .08), but are statistically significant at the .01 level because of large sample sizes. Controlling for premium class reduces the strength of the relationship, in

1. The nonparametric rank correlation is used rather than the Pearson linear correlation coefficient because of the lumpiness of count data. The overall value of Spearman's rho of 0.24 is statistically significant at the .01 level. The significance probabilities testing whether the rank correlations are zero for each premium class are listed in Table A.

part, because the rates of both paid and unpaid claims are higher in the classes with larger premiums. There is only a weak relationship between paid claims and unpaid claims in the analysis. This finding was the basis for excluding unpaid claims from the merit rating plan.

This weak positive association between paid and unpaid claims for an individual physician could have one of several explanations:

(1) Suppose that unpaid claims are exclusively nuisance claims. Under this assumption, adverse outcomes to patients of non-negligent physicians account for all these claims. Thus, the number of unpaid claims for a physician should be independent of his number of paid claims, which presumably arise largely from negligence. However, even within the same premium class, physicians have patient mixes that differ in patient condition, complexity of treatment, and volume. If volume and intensity of patient exposure tends to generate more nuisance claims as well as paid claims, these "latent variables" would produce a weak relationship between a physician's frequency of unpaid and paid claims even if rates per patient encounter are statistically independent.

(2) Suppose that unpaid claims are a mixture of nuisance claims and valid claims lacking sufficient evidence or legal talent to prove negligence. A weak positive correlation between the frequencies of unpaid and paid claims would result from a strong correlation caused by individuals having different propensities towards negligence degraded by the "noise" of nuisance claims.

(3) In either case, the patients' ease in detecting adverse outcomes varies with specialty. This would induce a positive correlation within the premium classes comprising more than one specialty.

These scenarios may explain, in part, why there is correlation between the frequencies of paid claims and unpaid claims in some premium classes, but not in others. This lack of a clear systematic relationship between unpaid and paid claims is our basis for using paid claims rather than claims filed as the predictor variable in our merit rating plan.