A Random-Effects Probit Model for Predicting Medical Malpractice Claims

Robert D. GIBBONS, Donald HEDEKER, Sara C. CHARLES, and Paul FRISCH*

We use Oregon state data (1981–1990) on medical malpractice claims to develop a random-effects probit model for vulnerability to a medical malpractice claim in practice year k ($k = 1, 2, \ldots, n_i$) for physician i ($i = 1, 2, \ldots, N$ physicians in the sample) conditional on an $n_i \times p$ covariate matrix $\mathbf{W}_i$ that contains a mixture of p time-varying and time-invariant covariates. In this application, time-invariant covariates were physician sex and specialty (surgical versus nonsurgical). Time-varying covariates were age, the cumulative amount of risk management education (i.e., number of courses) taken by physician i to year k, and prior claim history. In addition, the model incorporates a random effect of "claim vulnerability" assumed to be normally distributed in the population of physicians. This random effect represents unobservable and/or unmeasured characteristics that place one physician at greater risk for experiencing a medical malpractice claim than another physician. In addition, we also determine if the effects of risk management training on claim vulnerability differ before and after the physician's first malpractice claim. Results of the analysis reveal that (1) there is a sizable random physician effect; (2) risk increases between age 40 to 60; (3) physicians in a surgical specialty are at increased risk; (4) male physicians are at greater risk than female physicians; (5) risk increases following an initial claim, particularly in the year subsequent to the initial claim, and (6) some beneficial effects of risk management education are observed in physicians with a prior claim history, particularly those in anesthesiology and obstetrics and gynecology.

KEY WORDS: Actuarial statistics; Binary data; Longitudinal data; Marginal maximum likelihood; Risk management education.

1. INTRODUCTION

The continuing effort to create a reasonable and equitable solution to the problem of medical malpractice litigation has remained unsuccessful for almost an entire generation. The relative failure of the tort liability system in individual states to control and regulate medical malpractice law, for example, has led to an increased demand for federal legislation. To assure that such legislation addresses the problem directly, those who formulate health policy need to better appreciate the nature of the individual elements within this complex system (Robert Wood Johnson Foundation, 1991). Central to this endeavor is a better understanding of the role of the individual physician in generating medical malpractice claims.

Are some doctors more likely than others to experience medical malpractice claims? If so, what personal or professional characteristics that render them vulnerable to such claims can be identified? Because a medical malpractice claim is a rare event in a physician's practice, Nye and Hofflander (1988), Cooil (1991), Rolph (1991), and Rolph, Kravitz, and McGuigan (1991) have attempted to refine the predictive capability by proposing the application of a random-effect Poisson process model. These investigators applied a nonhomogeneous intensity model that provides a means of predicting the number of claims within a given time interval dependent on a fixed number of time-invariant physician characteristics and a single random effect that governs the claim rate for a specific individual. Insurers utilizing this model could accurately predict and analyze an individual's claim vulnerability. In this model an individual's claim potential is a random variable in the population of physicians; hence a random rate parameter is included in the parameterization to represent physician characteristics beyond those included as fixed covariates in the model. This research is an important contribution and sets the stage for the application of other models that attempt to include such considerations.

There are important physician characteristics known to affect medical malpractice claims that do not vary over time. For example, Charles et al. (1992) found that in addition to specialty (i.e., surgical versus nonsurgical; also see Sloan 1989 and Brennan 1991), age is a critically important predictor of claims. Prospective longitudinal studies would be helpful in clarifying claim vulnerability, because the putative predictors of medical malpractice claims (physician and practice characteristics) are not static but vary over time.

One intervention designed to influence physician risk is risk management education (RME). Risk pertains to the chance of losing assets by incurring liability for injury. Courses for physicians are designed to minimize liability subsequent to injuries and to reduce the probability of injuries occurring to patients (Ashby 1977). A recent study by Morlock and Malitz (1991) found that only in-hospital RME programs were related to lower numbers of claims, but there were little data to either substantiate or disprove the widespread assumption that RME programs yielded benefits in terms of a reduced number of claims filed (Kapp 1990).

To evaluate the efficacy of risk management courses or to assess the cumulative effect of additional courses, the comparison of claim history before and after such instruction provides a more natural method of testing efficacy than does the alternative cross-sectional approach. This is particularly true in naturalistic settings in which the intervention is not made on the basis of random assignment.

At an even more basic level, longitudinal malpractice data provide a vehicle by which hypotheses regarding the nature

* Robert D. Gibbons is Professor of Biostatistics, Departments of Psychiatry and Biometry, Donald Hedeker is Assistant Professor of Biostatistics, Department of Biometry, and Sara C. Charles is Professor of Psychiatry, Department of Psychiatry, University of Illinois, Chicago, IL 60612. Paul Frisch is Director of Medical Legal Affairs, Oregon Medical Association, Portland, OR 97201. This study was done with support of the Doctors Company, Napa, CA, and the Oregon Medical Association, Portland. The authors are indebted to the applications editor, an associate editor and reviewers for bringing numerous issues to their attention that greatly improved the manuscript. Finally, the authors thank Roberta Baruch for her editorial assistance.

Journal of the American Statistical Association
September 1994, Vol. 89, No. 427, Applications & Case Studies

of claim history may be examined. Repeated malpractice claims may occur because underlying, perhaps unobservable, physician characteristics may increase the risk of a claim. The correlation over time in claim occurrence may be simply due to the fact that this underlying claim "vulnerability" is a random variable in the population of physicians. This type of malpractice "claim proneness" explanation has been termed "heterogeneity" in the econometric literature by Heckman and Borjas (1980). In this context, heterogeneity is modeled as a random effect that allows for variation in each physician's "threshold" or propensity for experiencing a claim over the 10-year period.

The model described here examines the effects of measured and unmeasured physician characteristics on the probability of a claim in practice year t. Claim propensity is allowed to vary randomly over physicians. Because we focus on the probability of a claim in each individual year and not on the rate over T years, we can directly examine the effects of physician characteristics that vary from year to year (e.g., the effect of having a claim in the prior year on the probability of having a claim in the current year). Perhaps of greatest importance in this article is determining the effect of RME courses taken prior to year t on claim propensity in year t. In this way we compare claim-proneness before and after RME in the same physician. Using previously cited Poisson process models, we could have examined only the relationship between total amount of RME and total number of claims for the entire study period, regardless of whether or not RME preceded the claim.

This article is organized as follows. It begins by presenting the random-effects probit regression model and some general details of parameter estimation. Description of the data set and presentation of the principal analysis follow. Alternate model specifications in terms of prior distribution (i.e., normal versus rectangular) and response function (i.e., normal versus logistic) are then described and illustrated. Finally, the effects of subspecialties are examined.

2. RANDOM-EFFECTS PROBIT REGRESSION MODEL

There has been considerable interest in random-effects models for longitudinal and hierarchical, clustered, or multilevel data in the biological (Laird and Ware 1982; Ware 1985; Jennrich and Schluchter 1986; Waternaux et al. 1989), educational (Goldstein 1987; Bock 1989), psychological (Bryk and Raudenbush 1987; Willett 1991) and biomedical (Gibbons et al. 1988; Hedeker et al. 1989; Gibbons et al. 1993) statistical literatures. This work cited above has focused on continuous and normally distributed response measures. In contrast, there has been less focus on random-effects models for discrete data. Gibbons and Bock (1987) have presented a random-effects probit regression model to estimate trend in a binary variable measured repeatedly in the same subjects. Similar models based on the assumption of a logistic response function were developed by Stiratelli, Laird, and Ware (1984), Wong and Mason (1985), and Conaway (1989). Using quasi-likelihood methods in which no distributional form is assumed for the outcome measure, Liang and Zeger (1986) and Zeger and Liang (1986) have shown that consistent estimates of regression parameters and their variance estimates can be obtained regardless of time dependence. Koch, Landis, Freeman, Freeman, and Lehnen (1977) and Goldstein (1991) have illustrated how random effects can be incorporated into log-linear models. Finally, generalizations of the logistic regression model in which values of all regression coefficients vary randomly over individuals have also been proposed by Wong and Mason (1985).

The work of Gibbons and Bock can be modified to provide a random-effects probit regression model applicable to the problem at hand. For this, consider the following model for practice year k (where $k = 1, 2, \ldots, n_i$) for physician i ($i = 1, \ldots, N$ physicians in the sample):

$$x_{ik} = \beta_i + \alpha_1 w_{1_i} + \alpha_2 w_{2_{ik}} + \varepsilon_{ik}, \qquad (1)$$

where x_{ik} is the unobservable "claim vulnerability" on practice year k for physician i, β_i is the random effect due to physician i, α_1 is the fixed effect of the physician level covariate w_{1_i}, α_2 is the fixed effect of the practice year level covariate $w_{2_{ik}}$, and ε_{ik} is an independent residual distributed $\mathcal{N}(0, \sigma^2)$. As such, a case in this regression model is a physician practice year; physician i contributes n_i cases. Again, β_i is considered to be a coefficient for the random physician effect, and the assumed distribution for β is $\mathcal{N}(\mu, \sigma_\beta^2)$. This random effect represents the level of heterogeneity for physician i. With the foregoing assumptions, the $n_i \times 1$ vector of practice year claim vulnerabilities for physician i, $\mathbf{x}_i$, are multivariate normal with mean $\mathcal{E}(\mathbf{x}_i) = \mathbf{1}_i\mu + \mathbf{W}_i\alpha$ and covariance matrix $\mathcal{V}(\mathbf{x}_i) = \sigma_\beta^2 \mathbf{1}_i \mathbf{1}_i' + \sigma^2 \mathbf{I}_i$, where $\mathbf{1}_i$ is an $n_i \times 1$ unity vector, $\mathbf{I}_i$ is an $n_i \times n_i$ identity matrix, α is the $p \times 1$ vector of covariate coefficients, and $\mathbf{W}_i$ is the $n_i \times p$ covariate matrix.

To relate the manifest dichotomous response with the continuous latent claim vulnerability x_{ik}, Gibbons and Bock (1987) utilized the "threshold concept" (Bock 1975, p. 513). This concept assumes that the underlying variable is continuous and that in the binary response setting, one threshold value (γ) exists on the continuum of this variable. The presence or absence of at least one claim in practice year k for physician i is then determined by whether or not claim vulnerability exceeds the threshold value. When the claim vulnerability exceeds the threshold, a positive response is given (coded $y_{ik} = 1$); otherwise, a negative response is given ($y_{ik} = 0$). In the following analysis, $y_{ik} = 1$ if one or more claims were made in practice year k for physician i.

The use of the threshold model allows the expression of probability of a claim in terms of the value $1 - \Phi(z_k)$; that is, the area under the standard normal distribution function above the point z_k, where z_k is the normal deviate given by $(\beta + \alpha_1 w_1 + \alpha_2 w_{2_k} - \gamma)/\sigma$. Additionally, the origin and unit of z may be chosen arbitrarily; so for convenience, let $\sigma = 1$ and $\gamma = 0$. Then the probability of a particular pattern of claims (over n_i practice years) for a given physician i, denoted $\mathbf{y}_i$, is the product of the probability of each of the n_i binary responses—namely,

$$l(\mathbf{y}_i \mid \beta, \alpha) = \prod_{k=1}^{n_i} [\Phi(z_{ik})]^{1-y_{ik}} [1 - \Phi(z_{ik})]^{y_{ik}}. \qquad (2)$$

Thus the marginal probability of this pattern is given by

$$h(\mathbf{y}_i) = \int_\beta l(\mathbf{y}_i \mid \beta, \alpha) g(\beta)\, d\beta,$$

where $l(\mathbf{y}_i \mid \beta, \alpha)$ is given in (2) and $g(\beta)$ represents the distribution of β in the population—namely, a normal distribution with mean μ and variance σ_β^2. The parameters μ and σ_β^2 are estimated from the data.

2.1 Orthogonalization of the Model Parameters

In the estimation of parameters for the random-effects probit model, Gibbons and Bock (1987) orthogonally transformed the response model to use the marginal maximum likelihood estimation procedure for the dichotomous factor analysis model discussed by Bock and Aitkin (1981). Orthogonalization can be achieved by letting $\beta = \sigma_\beta \theta + \mu$, where σ_β is the standard deviation of β in the population. Then $\theta = (\beta - \mu)/\sigma_\beta$, and so $\mathcal{E}(\theta) = 0$ and $\mathcal{V}(\theta) = 1$. The reparameterized model is then written as

$$z_k = \mu + \sigma_\beta \theta + \mathbf{w}_k' \alpha, \tag{3}$$

and the marginal density becomes

$$h(\mathbf{y}_i) = \int_\theta l(\mathbf{y}_i \mid \theta, \alpha, \mu, \sigma_\beta^2) g(\theta)\, d\theta, \tag{4}$$

where $g(\theta)$ represents the distribution of the θ vector in the population; that is, the standard normal density. The log-likelihood for the patterns from the N physicians can be written as

$$\log L = \sum_i^N \log h(\mathbf{y}_i), \tag{5}$$

and the likelihood equations for an arbitrary parameter vector η are

$$\frac{\partial \log L}{\partial \eta} = \sum_{i=1}^N \frac{1}{h(\mathbf{y}_i)} \frac{\partial h(\mathbf{y}_i)}{\partial \eta}. \tag{6}$$

2.2 Solution to the Likelihood Equations

One approach to solving the aforementioned likelihood equations involves numerical integration. Because the assumed distribution of β is normal and the distribution of θ is standard normal, Gauss–Hermite quadrature can be used to approximate the integrals to any practical degree of accuracy. In Gaussian quadrature, integration is approximated by a summation on a specified number of quadrature nodes Q for each dimension of the integration. Integration terms are then evaluated at these quadrature points and multiplied by corresponding quadrature weights. For the standard normal univariate density, optimal points and weights, denoted B_q and $A(B_q)$, were given by Stroud and Sechrest (1966).

The derivatives are now approximated as

$$\frac{\partial \log L}{\partial \eta} \approx \sum_{i=1}^N \frac{1}{h(\mathbf{y}_i)} \sum_{q=1}^Q \sum_{k=1}^{n_i} \frac{y_{ik} - \Phi(z_{iqk})}{\Phi(z_{iqk})(1 - \Phi(z_{iqk}))} \times \phi(z_{iqk}) \frac{\partial z_{iqk}}{\partial \eta'} l(\mathbf{y}_i \mid B_q, \alpha, \mu, \sigma_\beta^2) A(B_q), \tag{7}$$

where

$$\frac{\partial z_{iqk}}{\partial \alpha} = \mathbf{w}_{ik}', \qquad \frac{\partial z_{iqk}}{\partial \mu} = 1, \qquad \frac{\partial z_{iqk}}{\partial \sigma_\beta} = B_q,$$

with the marginal likelihood of the response vector $\mathbf{y}_i$ approximated as

$$\hat{h}(\mathbf{y}_i) \approx \sum_q^Q l(\mathbf{y}_i \mid B_q, \alpha, \mu, \sigma_\beta^2) A(B_q), \tag{8}$$

where

$$z_{iqk} \approx \sigma_\beta B_q + \mu + \mathbf{w}_{ik}' \alpha.$$

Fisher's method of scoring can be used to provide estimates using these approximated derivatives. In practice, a vector of 1s is appended to α to estimate μ.

This approach contrasts with previous work (see, for example, Stiratelli et al. 1984; Wong and Mason 1985) in which the integrals in (6) were approximated by a multivariate normal distribution with the same mode and curvature at the mode as the true posterior.

An advantage of numerical integration is that a normal prior distribution for the random effect(s) is no longer required, and the assumption of normality and its influence on the structural effects of interest can be empirically tested (e.g., comparison of a normal versus a rectangular prior). The major disadvantage is that as the number of random effects increases, so does the dimensionality of the integral in (6); thus this approach is computationally restricted to problems with small numbers of random effects (i.e., 5 or less). This is not a problem in the present context, where the physician is the only random effect in the model (i.e., the effects of sex, age, specialty, and risk management training are considered fixed and do not vary systematically over individual physicians). Anderson and Aitkin (1985) have developed a similar model for examining interview variability that utilizes numerical integration to obtain maximum likelihood parameter estimates.

2.3 Estimating Claim Vulnerability

In some cases it may be desirable to estimate the level of the random effect, β_i, for a particular physician. A good choice for this purpose (Bock and Aitkin 1981; Gibbons and Bock 1987) is the expected a posteriori (EAP) value (Bayes estimate) of θ_i, given the claim history vector $\mathbf{y}_i$ and covariate matrix $\mathbf{W}_i$ of physician i:

$$\hat{\theta}_i = \mathcal{E}(\theta_i \mid \mathbf{y}_i, \mathbf{W}_i) = \frac{1}{h(\mathbf{y}_i)} \int_{-\infty}^{\infty} \theta_i l(\mathbf{y}_i \mid \theta, \alpha) g(\theta)\, d\theta. \tag{9}$$

Similarly, the posterior standard deviation of $\hat{\theta}_i$, which may be used to express the precision of the EAP estimator, is given by

$$\mathcal{V}(\hat{\theta}_i \mid \mathbf{y}_i, \mathbf{W}_i) = \frac{1}{h(\mathbf{y}_i)} \int_{-\infty}^{\infty} (\theta_i - \hat{\theta}_i)^2 l(\mathbf{y}_i \mid \theta, \alpha) g(\theta)\, d\theta. \tag{10}$$

These quantities can be evaluated using Gauss–Hermite quadrature as previously described. Estimates of β_i can be recovered by $\hat{\beta}_i = \hat{\sigma}_\beta \hat{\theta}_i + \hat{\mu}$, using the marginal maximum likelihood estimates of the parameters. Because the prior distribution $g(\theta)$ is normal, these linear transformations of EAP estimates are also EAP estimates. The probability of a claim for a physician with claim history vector $\mathbf{y}_i$ and covariate vector $\mathbf{w}_{ik}$ is, therefore, $1 - \Phi(\hat{\beta}_i + \mathbf{w}'_{ik}\hat{\alpha})$.

All computations described here were performed using the computer program MIXOR (Hedeker 1993), which is available from the National Institutes of Mental Health, Division of Services Research. Convergence was generally obtained in 10 iterations, which required approximately 10 minutes on a 486, 50 MHz DX computer running under MS-DOS 5.0. No computational problems were encountered.

3. DESCRIPTION OF THE DATA SET

The study sample consisted of 1,903 physicians representing 16,083 practice years. The data set consists of all Oregon physicians who were insured between 1981 and 1990 inclusively by the CNA Insurance Company under the sponsorship of the Oregon Medical Association's (OMA) Physician Protection Program. Because the OMA's malpractice claims database could be tied to its membership database, accurate data on each physician's age, gender, specialty, and dates of attendance at four different and sequential OMA-sponsored risk management courses during that decade were readily available.

The population of OMA-sponsored insured physicians remained relatively constant during this time period. Missing data in this study are due to physicians changing insurers, moving out of state, dying, or retiring, or to new physicians coming into the program. Although approximately eight to ten physicians in the state lost their medical licenses during this period, we have no indication that our missing data include members of this group.

Since 1979, physicians have been granted a 3-year premium discount of 7.5% per year for voluntary participation in RME. Approximately 98% availed themselves of the opportunity. Since 1987, participation has been mandatory (on a 3-year cycle), and the discount remains. The fact that some physicians in the study have had no RME exposure is a function of them having entered and left the program during the 3-year cycle, thereby missing the opportunity to take the basic RME course.

For the purpose of this study, the data on which a claim arose is the date of the alleged injury, as distinguished from the date when the physician reported it to the liability insurance carrier. Relating the date of injury to the date when the RME course was taken more accurately reflects the relationship of the claim to the educational intervention.

4. PRINCIPAL ANALYSIS

In the present context of predicting medical malpractice claims, we considered a model with a single random effect that governs the underlying claim vulnerability of physician i. This model is referred to as the *random intercept model.* In terms of overall trends over time, it is possible that an overall shift in claim incidence has occurred over the 10-year period; therefore, we included a linear time trend as a fixed covariate in the model. In addition, based on the previous work of Charles et al. workers (1992), we also included the fixed effects of physician sex, age, and specialty (i.e., surgical = 1 versus nonsurgical = 0). The linear effect of age on claim incidence was examined, and age was treated as a time-varying covariate. In this way, the direct effect of the physician's age on claim risk, not a cohort effect, could be evaluated. Finally, the cumulative linear effect of risk management education (i.e., number of courses taken up to year k), was also included as a time-varying covariate.

Results of applying the random-effects probit model to the 10 years of malpractice data, along with parameter estimates and standard errors for each of the effects in the model, are displayed in Table 1. The ratio of maximum likelihood parameter estimates to their standard error has an asymptotically normal distribution; therefore, the "p value" for the null hypothesis that the coefficient equals 0 can be obtained from the corresponding normal tail probability, which are the values displayed in Table 1.

Inspection of Table 1 reveals the main effects of year (i.e., claim incidence significantly decreased over the period, from .093 in 1980 to .058 in 1989), sex (i.e., women physicians were less likely than men to have a claim, .038 versus .081 per year), age (i.e., risk increases with age), and specialty (i.e., physicians in surgical specialties were more likely to have a claim; .105 versus .056). With respect to RME, a significant age-by-RME interaction was observed, indicating that as physicians age, they benefit from RME courses more than when they were younger; however, this interaction was not significant following the addition of other explanatory variables to the model (see the following sections). Table 1 also reveals that the first RME course is associated with an increase in the likelihood of malpractice claims. This somewhat odd result is further explicated by examining individual specialties in Section 6.

The standard deviation due to the random physician effect (i.e., heterogeneity) was $\hat{\sigma}_\beta = .110$, which corresponds to an intraclass correlation of .099, or approximately 10% of the total variation.

One effect not tested in this model is whether the effects of RME are different based on prior claim history. To examine this question, a time-varying covariate was constructed having value 0 or 1, depending on whether or not a claim had been made prior to that year. Both the main effect of prior claim (MLE = .2597, SE = .0498, $p < .0000$) and the

Table 1. Parameter Estimates, Standard Errors (SE) and Probabilities (P<) for Medical Malpractice Claims

Effect	Estimate	SE	P<
Overall mean (μ)	−2.9824	.6229	.0000
Random physician effect (σ^2_β)	.3321	.0279	.0000
Linear time trend	−0.0585	.0074	.0000
Linear risk management education (RME)	.4887	.1106	.0001
Physician sex (female = 1)	−.2766	.0972	.0044
Linear physician age	.0052	.0025	.0346
Specialty (surgical = 1)	.3040	.0448	.0000
Linear (# RME courses) × sex	.0154	.0855	.8570
Linear (# RME courses) × age	−.0078	.0022	.0004
Linear (# RME courses) × specialty	.0307	.0383	.4223

NOTE: Log likelihood = −4202.525.

Table 2. Relationship of Year, Age, Sex, Specialty, Prior Claims, and RME with Marginal Yearly Claim Propensity Overall and Claims Resulting in Payout (1980–1989)

Predictor	All claims	Payouts	Physician years
# of RME courses			
0	.068	.020	8,286
1	.094	.033	4,403
2	.077	.021	2,872
3	.065	.021	522
Specialty			
Nonsurgical	.056	.017	9,287
Surgical	.105	.033	6,796
Age			
<40	.065	.017	5,577
40–49	.088	.026	6,021
50–59	.081	.031	3,104
60–69	.065	.027	1,180
≥70	.065	.016	189
Sex			
Male	.081	.025	14,683
Female	.038	.012	1,400
Any prior claim			
No	.060	.018	10,230
Yes	.107	.035	5,853
Claim last year			
No	.071	.022	14,861
Yes	.142	.049	1,222
Year			
1980	.093	.036	1,336
1981	.087	.035	1,390
1982	.094	.032	1,451
1983	.107	.034	1,502
1984	.078	.026	1,569
1985	.068	.022	1,625
1986	.067	.017	1,706
1987	.065	.015	1,766
1988	.065	.017	1,835
1989	.058	.014	1,903

linear RME by prior claim interaction (MLE = −.1277, SE = .0406, $p < .0017$) were significant. These results indicate that physicians with prior claims are at increased risk of future claims (a result that is consistent with the nonzero random effect) and that the effect of RME depends on prior suit history. Further analysis by specialty explicates this association (see Sec. 6).

Table 2 indicates that having a claim immediately prior to the current year produced one of the greatest increases in claim incidence in the current year. Adding prior year claim history to the model that included any prior claim did not, however, produce a significant improvement in fit of the model to the observed data. Observed proportions in Table 2 also reveal that the effect of RME is quadratic, with increases in claim risk seen following one or two RME courses but decreasing somewhat after three RME courses. Addition of a quadratic term in the model produced a significant improvement in fit ($\chi_1^2 = 9.4$, $p < .0022$), but the quadratic term was not involved in the interactions with age and prior claim history ($\chi_2^2 = 5.1$, $p < .08$). Similarly, Table 2 also reveals that the effect of age is also quadratic, in that observed claim incidence is highest between age 40 and 60. Addition of a quadratic age effect into the model produced a significant improvement in fit ($\chi_1^2 = 5.0$, $p < .0253$). Interestingly, with the addition of prior claim history and the quadratic effects of age and RME, the RME-by-age interaction was no longer significant. Parameter estimates, standard errors, and associated significance levels for the final model are presented in Table 3.

5. ALTERNATIVE RESPONSE FUNCTIONS AND PRIOR DISTRIBUTIONS

The numerical solution to the likelihood equations provided here makes it possible to substitute any distribution for the previously assumed normal prior distribution for claim vulnerability, requiring only a suitable set of nodes and weights for the quadrature. For example, a uniform or rectangular prior distribution could be used by selecting Q equally spaced points between −4 and 4 and selecting uniform weights $1/Q$. This rectangular prior provides a good check on the dependence of the solution on the assumption that claim vulnerability is normally distributed in the population of physicians. Applied to these data, the log-likelihood decreased slightly from −4,180.26 to −4,181.27. Parameter estimates were virtually unchanged, except for the random-effects variance term, which decreased from .2724 to .0880. These results suggest that the normal prior is a reasonable choice, but that results regarding structural effects show some robustness to specification of the prior distribution.

Selection of a normal response function may also be changed. A common choice is the logistic distribution (Stiratelli et al. 1984; Wong and Mason 1985). In this case, the claim probability is

$$\Psi(z_k) = \frac{1}{1 + \exp[-z_k]}. \tag{11}$$

As in the normal case, we let $\gamma = 0$; however, the residual variance is now $\pi^2/3$. Application of the logistic response function is attractive here, because there is greater tail probability consistent with an overall yearly claim rate of approximately 5%. Results of applying the logistic response function with a normal prior distribution to these data (using the final parameterization given in Table 3) revealed a slight decrease in log-likelihood from −4,180.26 to −4.180.39. Parameter estimates are now in a different metric; however, ratios of estimates to their standard errors, signs, and substantive interpretations were unchanged (Table 4). These results are encouraging, because they demonstrate that the

Table 3. Parameter Estimates, Standard Errors, and Probabilities for Medical Malpractice Claims: Final Model

Effect	Estimate	SE	P<
Overall mean (μ)	−2.1139	.7466	.0046
Random physician effect (σ_β^2)	.2724	.0384	.0000
Linear time trend (1980–1989)	−.0586	.0074	.0000
Physician sex (female = 1)	−.2221	.0734	.0025
Physician age—linear	.0467	.0164	.0043
Physician age—quadratic	−.0005	.0002	.0026
Specialty (surgical = 1)	.2905	.0366	.0000
RME—linear	.3454	.0573	.0000
RME—quadratic	−.0734	.0228	.0013
Prior claims	.2386	.0507	.0000
RME × prior claim	−.1304	.0405	.0013

NOTE: Log likelihood = −4180.26.

Table 4. Parameter Estimates, Standard Errors, and Probabilities for Medical Malpractice Claims: Final Model—Logistic Response Function

Effect	Estimate	SE	P<
Overall mean (μ)	−4.8635	1.5280	.0015
Random physician effect (σ_β^2)	.5222	.0725	.0000
Linear time trend (1980–1989)	−.1202	.0153	.0000
Physician sex (female = 1)	−.4643	.1538	.0027
Physician age—linear	.0915	.0335	.0063
Physician age—quadratic	−.0010	.0003	.0039
Specialty (surgical = 1)	.5721	.0729	.0000
RME—linear	.6959	.1158	.0000
RME—quadratic	−.1435	.0454	.0016
Prior claims	.4909	.0991	.0000
RME × prior claim	−.2623	.0806	.0011

NOTE: Log likelihood = −4180.39.

parameter estimates and fit of the model are reasonably robust to model specification.

6. SPECIALTY EFFECTS

To this point the results indicate that overall, taking one or two RME courses increases the likelihood of a claim. One possible explanation is that this result could have arisen due to the omission of specialties from the analysis. After reviewing previous studies (Brennan et al. 1991; Kravitz et al. 1991; Nye 1988; Rolph 1991; Sloan et al. 1989), we assigned study subjects to one of eight specialty designations: (1) anesthesiology (anesthesiology and obstetrical anesthesiology); (2) radiology (diagnostic radiology, therapeutic radiology, ultrasound); (3) obstetrics (obstetrics and obstetrics and gynecology); (4) primary care (family practice, general practice, internal medicine, pediatrics); (5) medical subspecialties; (6) surgery (general, all subspecialties exclusive of group 7); (7) special surgery (neurosurgery, thoracic, cardiac, orthopedics, vascular, sports medicine); or (8) all other specialties.

The same model was used as in Table 3, except that seven contrasts between the eight specialties replaced the 1 degree of freedom contrast between surgical and nonsurgical specialties. Note that apart from the intercept and specialty terms, the other coefficients displayed in Table 5 changed very little from the analysis with a single dummy-coded specialty (see Table 3).

Inspection of Table 5 reveals a considerable difference among the specialties in claim propensity; however, accounting for these differences does not change the positive association between RME courses and increased risks, at least for one or two courses. The effect of prior claims on the beneficial effects of RME courses is also unchanged with the addition of specialty into the model. To further highlight this association at the level of specialty, Table 6 presents observed claim incidence, overall as well as claims resulting in payout, for each specialty before and after the first claim as a function of number of RME courses. Table 6 reveals that prior to the first claim, RME courses have limited effect, and that for most specialties, the likelihood of a claim increases after one course. Following an initial claim, the overall claim incidence dramatically increases, however, and RME courses now appear to decrease claim incidence for anesthesiology, radiology, obstetrics, and other specialties. The effects for anesthesiology and obstetrics are reflected in a decrease in payouts as well.

7. DISCUSSION

The statistical model developed here provides a new approach to the analysis of longitudinal binary data, of which the medical malpractice claim data described here provide an excellent example. The random-effects probit model provides considerable flexibility in terms of permitting varying numbers of measurements per subject, inclusion of time-varying and time-invariant fixed covariates, and a method by which effects of varying physician vulnerability can be estimated. EAP estimates of person-specific effects can in turn be estimated following estimation of model parameters. These estimates and their corresponding standard errors may be useful in predicting future medical malpractice claim incidence for a given physician.

In terms of medical malpractice claims, the results of these analyses revealed that being male, working in a surgical specialty, and having a prior claim history are associated with increased claim vulnerability. In terms of RME, there appears to be minimal benefit, except perhaps for obstetricians and anesthesiologists who have had a prior claim.

In terms of the random effect, the within-physician correlation was r = .099, indicating that claim incidence is modestly correlated within individual physicians, presumably because of personal characteristics and/or practice behaviors that increase their vulnerability to claims.

A major limitation of the model developed here is that it is capable of modeling only the presence or absence of a claim in each time interval, regardless of the actual number of claims during the interval. In the present application this is not of major consequence, as the incidence of two claims in 1 year was 0.4% and the incidence of three claims in 1 year was only 0.04%. These rates are out of a total of 16,083 physician practice years. Had the incidence of multiple yearly claims been larger, however, a random-effects ordinal probit

Table 5. Parameter Estimates, Standard Errors, and Probabilities for Medical Malpractice Claims, Eight Subspecialties

Effect	Estimate	SE	P<
Overall mean (μ)	−1.4428	.7721	.0617
Random physician effect (σ_β^2)	.2565	.0397	.0000
Linear time trend (1980–1989)	−.0546	.0075	.0000
Physician sex (female = 1)	−.2316	.0765	.0025
Physician age—linear	.0520	.0166	.0018
Physician age—quadratic	−.0006	.0002	.0013
Anesthesiology vs. other	.4767	.1181	.0001
Radiology vs. other	.4074	.0970	.0000
Obstetrics vs. other	.8304	.1057	.0000
Primary care vs. other	.2484	.0796	.0018
Medical subspecialties vs. other	.1044	.0926	.2597
Surgery vs. other	.4839	.0874	.0000
Special surgery vs. other	.6632	.0931	.0000
RME—linear	.3286	.0572	.0000
RME—quadratic	−.0728	.0227	.0015
Prior claims	.1960	.0512	.0001
RME × prior claim	−.1324	.0406	.0011

NOTE: Log likelihood = −4137.23.

Table 6. Marginal Claim Incidence per Specialty, RME Courses, and Prior Claim History

		Prior to first claim			Following first claim		
Specialty	RME courses	All claims	Payouts	Physician years	All claims	Payouts	Physician years
Anesthesiology	0	.092	.029	206	.188	.146	48
Anesthesiology	1	.080	.045	88	.070	.047	86
Anesthesiology	≥2	.123	.015	65	.091	.050	121
Radiology	0	.076	.017	582	.093	.004	226
Radiology	1	.120	.033	150	.071	.027	183
Radiology	≥2	.059	.015	68	.058	.014	139
Obstetrics	0	.115	.035	200	.233	.116	129
Obstetrics	1	.173	.074	81	.198	.070	172
Obstetrics	≥2	.135	.081	37	.152	.042	165
Primary care	0	.051	.015	2728	.089	.031	671
Primary care	1	.054	.022	1176	.092	.036	631
Primary care	≥2	.055	.014	715	.057	.015	649
Medical	0	.033	.008	875	.067	.011	179
Medical	1	.064	.010	405	.074	.016	190
Medical	≥2	.036	.008	248	.045	.017	176
Surgery	0	.064	.011	653	.140	.041	365
Surgery	1	.114	.044	228	.150	.042	354
Surgery	≥2	.078	.023	128	.089	.025	394
Special surgery	0	.095	.031	359	.138	.059	253
Special surgery	1	.109	.048	147	.202	.052	252
Special surgery	≥2	.131	.016	61	.157	.037	268
Other	0	.018	.005	740	.083	.014	72
Other	1	.057	.031	194	.061	.015	66
Other	≥2	.042	.000	96	.031	.016	64

model (Hedeker and Gibbons 1994) could have been used to preserve the ordinal nature of the yearly claim incidence data. In this case, separate thresholds between each yearly claim frequency (i.e., 0–1, 1–2, and 2–3) would be estimated and covariates modeled in terms of their relationship to increasing number of yearly claims instead of simply the presence or absence of a claim.

Some discussion of missing data is appropriate. The model described here, as well as the other full-likelihood approaches (see, for example, Stiratelli et al. 1984 and Wong and Mason 1985), are appropriate under the assumption of "ignorable nonresponse;" that is, missing data are ignorable as long as they are explained by terms in the model and/or the available outcome data for each subject (see Laird 1988). In the present context, missing data arise either from physicians who began practicing after 1981 or retired or moved out of Oregon before 1990. In this case, ignorable nonresponse implies that the observed yearly claim rate would have been observed had the physician begun practicing earlier or postponed retirement. Given that in most cases missing data are restricted to a few years, the assumption seems reasonable (i.e., distribution of unobserved outcomes is known conditional on distribution of available outcomes). It should be noted, however, that some alternative approaches, such as the quasi-likelihood approach of Liang and Zeger (1986) and Zeger and Liang (1986), assume no distributional form for the outcome measures and thus can be applied to a wide variety of data (i.e., binary, ordinal, and continuous). The disadvantage, however, is that missing data are ignorable only if they are completely explained by the covariates in the model. Because no distributional form is assumed for outcomes, distribution of the missing data conditional on the observed outcomes is unknown and thus cannot be used to justify statistical inferences in the presence of missing data. If the missing data are not completely explained by the covariates, then the quasi- or partial-likelihood approaches become even more restrictive than the full-likelihood procedure described here, in that consistency of the quasi-likelihood estimates is now guaranteed only if the true correlation among repeated outcomes is known for each subject. This information, of course, is never available. But because missing data are typically related to the age of the physician and age is an important covariate, these models may have application here as well.

This model offers a new resource in estimating the vulnerability of a given physician to a malpractice claim. In addition to being male, being a surgical specialist, and having a prior claim history, this research suggests that the effect of RME on subsequent claims and payouts has stronger effects for some specialists than for others.

Of all specialists, obstetrician-gynecologists and anesthesiologists who have had a previous claim appear to benefit most, in that an increased number of RME courses for these physicians is associated with decreased claim vulnerability and decreased payout. Risk management efforts for both these groups is often focused on the appropriate use of monitoring and the introduction of checklists to be applied in specific instances. The use of these devices and forms decreases both incentives to pursue claims and the likelihood of payouts, because they fill a previous void in documentation of patient care. This deprives the plaintiff lawyer of the argument that inadequate monitoring during a surgical procedure or throughout a pregnancy led to a bad outcome. Proof to the contrary is in the medical record for the jury to see. Such strategies may be easier to implement and control than other physician behaviors.

Although RME does not reduce the vulnerability to claims of physicians who perform special surgery, it does appear to be associated with slightly fewer payouts. The failure to positively influence claim vulnerability in this group may be a function of the nature of this specialized surgical work rather than of the quality of their work per se. The fact that payouts are decreased, although less so than for anesthesiologists and obstetrician-gynecologists, lends some credence to the contention of many risk managers that even though RME may not prevent claims, it may equip the physician to be a better defendant in the event of a claim.

Among all groups of specialists, prior to their first claim, risk increases after one RME course. This may well be a result of the fact that with increasing years in practice, one's exposure increases with an increased number of patients and the consequent increased chance of a negative outcome. This suggestion is supported by the finding that the peak years of vulnerability to risk are between age 40 and 60. Before age 40, exposure is limited; after age 60, many physicians alter their practices, reducing exposure.

A subsequent result of this study is that a claim in the previous year doubles not only the likelihood of a claim in the following year (from .071 to .142; see Table 2), but also the likelihood that the claim is associated with a payout (from .022 to .049; see Table 2). Strategically focused and specialty specific RME courses taken shortly after a claim may well exert a positive benefit on subsequent claims and payouts.

[Received August 1991. Revised December 1993.]

REFERENCES

Anderson, D., and Aitken, M. (1985), "Variance Component Models With Binary Response: Interviewer Variability," *Journal of the Royal Statistical Society,* Ser. B, 47, 203–210.

Ashby, J. L., Stephens, S. K., and Pearson, S. B. (1977), "Elements in Successful Risk Reduction Programs," *Hospital Progress,* 58, 60–75.

Bock, R. D. (1975), *Multivariate Statistical Methods in Behavioral Research,* New York: McGraw-Hill.

——— (1989), "Measurement of Human Variation: A Two-Stage Model," in *Multilevel Analysis of Educational Data,* ed. R. D. Bock, New York: Academic Press, pp. 319–342.

Bock, R. D., and Aitkin, M. (1981), "Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM Algorithm," *Psychometrika,* 46, 443–459.

Brennan, T. A., Leape, L. L., Laird, N. M., Hebert, L., Localio, A. R., Lawthers, A. G., Newhouse, J. P., Weiler, P. C., and Hiatt, H. H. (1991), "Incidence of Adverse Events and Negligence in Hospitalized Patients," *New England Journal of Medicine,* 324, 379–376.

Bryk, A. S., and Raudenbush, S. W. (1987), "Application of Hierarchical Linear Models to Assessing Change," *Psychological Bulletin,* 101, 147–158.

Charles, S. C., Gibbons, R. D., Frisch, P. R., Pyskoty, C. E., Hedeker, D. R., and Singha, N. K. (1992), "Predicting Risk for Medical Malpractice Claims Utilizing Quality of Care Characteristics," *Western Journal of Medicine,* 157, 433–439.

Conaway, M. R. (1989), "Analysis of Repeated Categorical Measurements With Conditional Likelihood Methods," *Journal of the American Statistical Association,* 84, 53–61.

Cooil, B. (1991), "Using Medical Malpractice Data to Predict the Frequency of Claims: A Study of Poisson Process Models With Random Effects," *Journal of the American Statistical Association,* 86, 285–295.

Gibbons, R. D., and Bock, R. D. (1987), "Trend in Correlated Proportions," *Psychometrika,* 52, 113–124.

Gibbons, R. D., Hedeker, D. R., Elkin, I., Waternaux, C. M., Kraemer, H. C., Greenhouse, J. B., Shea, M. T., Imber, S. D., Sotsky, S. M., and Watkins, J. T. (1993), "Some Conceptual and Statistical Issues in Analysis of Longitudinal Psychiatric Data," *Archives of General Psychiatry,* 50, 739–750.

Gibbons, R. D., Hedeker, D., Waternaux, C. M., and Davis, J. M. (1988), "Random Regression Models: A Comprehensive Approach to the Analysis of Longitudinal Psychiatric Data," *Psychopharmacology Bulletin,* 24, 438–443.

Goldstein, H. (1987), *Multilevel Models in Educational and Social Research,* New York: Oxford University Press.

——— (1991), "Nonlinear Multilevel Models, With an Application to Discrete Response Data," *Biometrika,* 78, 45–51.

Heckman, J., and Borjas, G. (1980), "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers From a Continuous Time Model of Heterogeneity and State Dependence," *Economica,* 47, 247–283.

Hedeker, D., and Gibbons, R. D. (1994), "A Random-Effects Ordinal Regression Model for Multilevel Analysis," submitted to *Biometrics.*

Hedeker, D., Gibbons, R. D., Waternaux, C. M., and Davis, J. M. (1989), "Investigating Drug Plasma Levels and Clinical Response Using Random Regression Models," *Psychopharmacology Bulletin,* 25, 227–231.

Jennrich, R. I., and Schluchter, M. D. (1986), "Unbalanced Repeated-Measures Models With Structured Covariance Matrices," *Biometrics,* 42, 805–820.

Kapp, M. B. (1990), "Health Care Risk Management: The Challenge of Measuring Costs and Benefits," *QRB,* 16, 166–169.

Koch, G., Landis, J., Freeman, J., Freeman, H., and Lehnen, R. (1977), "A General Methodology for the Analysis of Experiments With Repeated Measurements of Categorical Data," *Biometrics,* 33, 133–158.

Kravitz, R. L., Rolph, J. E., and Mcguigan, K. (1991), "Malpractice Claims Data as a Quality Improvement Tool," *Journal of the American Medical Association,* 266, 2087–2092.

Laird, N. M., and Ware, J. H. (1982), "Random Effects Models for Longitudinal Data," *Biometrics,* 38, 963–974.

Laird, N. M. (1988), "Missing Data in Longitudinal Studies," *Statistics in Medicine,* 7, 305–315.

Liang, K. Y., and Zeger, S. L. (1986), "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika,* 73, 13–22.

Morlock, L. L., and Malitz, F. E. (1991), "Do Hospital Risk Management Programs Make a Difference?: Relationships Between Risk Management Program Activities and Hospital Malpractice Claims Experience," *Law and Contemporary Problems,* 43, 1–22.

Nye, B. F., and Hofflander, A. E. (1988), "Experience Rating in Medical Professional Liability Insurance," *The Journal of Risk and Insurance,* 150, 151–157.

Robert Wood Johnson Foundation (1991), *Abridge: A Special Edition on Medical Malpractice,* Princeton, NJ: Robert Wood Johnson Foundation.

Rolph, J. E. (1991), "Merit Rating for Physicians' Malpractice Premiums: Only a Modest Deterrent," *Law and Contemporary Problems,* 54, 65–86.

Rolph, J. E., Kravitz, R. L., and McGuigan, K. (1991), "Malpractice Claims Data as a Quality Improvement Tool. Is Targeting Effective?" *Journal of the American Medical Association,* 566, 2093–2097.

Sloan, F. A., Mergenhagen, P. M., Burfield, R. P., Bovbjerg, R. R., and Hassan, M. (1989), "Medical Malpractice Experience of Physicians: Predictable or Haphazard?" *Journal of the American Medical Association,* 262, 3291–3297.

Stiratelli, R., Laird, N. M., and Ware, J. H. (1984), "Random-Effects Models for Serial Observations With Binary Response," *Biometrics,* 40, 961–971.

Stroud, A. H., and Sechrest, D. (1966), *Gaussian Quadrature Formulas,* Englewood Cliffs, NJ: Prentice-Hall.

Ware, J. (1985), "Linear Models for the Analysis of Longitudinal Studies," *The American Statistician,* 39, 95–101.

Waternaux, C. M., Laird, N. M., and Ware, J. H. (1989), "Methods for Analysis of Longitudinal Data: Blood Lead Concentrations and Cognitive Development," *Journal of the American Statistical Association,* 84, 33–41.

Willett, J. B., Ayoub, C. C., and Robinson, D. (1991), "Using Growth Modeling to Examine Systematic Differences in Growth: An Example of Change in the Functioning of Families at Risk of Maladaptive Parenting, Child Abuse, or Neglect," *Journal of Consulting and Clinical Psychology,* 59, 38–47.

Wong, G. Y., and Mason, W. M. (1985), "The Hierarchical Logistic Regression Model for Multilevel Analysis," *Journal of the American Statistical Association,* 80, 513–524.

Zeger, S. L., and Liang, K. Y. (1986), "Longitudinal Data Analysis for Discrete and Continuous Outcomes," *Biometrics,* 42, 121–130.