STATISTICS IN MEDICINE, VOL. 15, 1663–1685 (1996)

SIMULTANEOUSLY MODELLING CENSORED SURVIVAL DATA AND REPEATEDLY MEASURED COVARIATES: A GIBBS SAMPLING APPROACH

CHERYL L. FAUCETT

Department of Biostatistics, University of California, Los Angeles, 10833 Le Conte Avenue, Los Angeles, CA 90024, U.S.A.

AND

DUNCAN C. THOMAS

Department of Preventive Medicine, University of Southern California, 1540 Alcazar Street Suite 220, Los Angeles, CA 90033, U.S.A.

SUMMARY

Recent methodologic developments in the analysis of longitudinal data have typically addressed one of two aspects: (i) the modelling of repeated measurements of a covariate as a function of time or other covariates, or (ii) the modelling of the effect of a covariate on disease risk. In this paper, we address both of these issues in a single analysis by modelling a continuous covariate over time and simultaneously relating the covariate to disease risk. We use the Markov chain Monte Carlo technique of Gibbs sampling to estimate the joint posterior distribution of the unknown parameters of the model. Simulation studies showed that jointly modelling survival and covariate data reduced bias in parameter estimates due to covariate measurement error and informative censoring. We illustrate the methodology by application to a data set that consists of repeated measurements of the immunologic marker CD4 and times of diagnosis of AIDS for a cohort of anti-HIV-1 positive recipients of anti-HIV-1 positive blood transfusions. We assume a linear random effects model with subject-specific intercepts and slopes and normal errors for the true log and square root CD4 counts, and a proportional hazards model for AIDS-free survival time expressed as a function of current true CD4 value. On the square root scale, the joint approach yielded a mean slope for CD4 that was 7 per cent steeper and a log relative risk of AIDS that was 35 per cent larger than those obtained by analysis of the component sub-models separately.

1. INTRODUCTION

Many epidemiologic studies consist of follow-up of a cohort of subjects over time to examine the relationship between one or more explanatory variables and the risk of developing a disease. Measurements of the explanatory variables often occur at periodic intervals and result in a data set that consists of repeated measurements of time-dependent covariates, single measurements of fixed covariates, and times of disease development or censoring for each subject. Frequently, analysis of two aspects of these data has interest: (i) the relationship between a repeatedly measured covariate and time or other covariates (*covariate tracking model*), and (ii) the relationship between the time-dependent covariate and the probability of disease development (*disease risk model*).

CCC 0277–6715/96/151663–23

Received June 1994
Revised August 1995

Approaches to each type of analysis have had extensive description in the literature. For the first analysis, ultimate interest is in the description of the relationship between the repeated measurements and time or possibly other covariates. One commonly uses a random effects model (for example, Laird and Ware,[1] Zeger and Karim[2]) to model the level of a covariate. Wu and Carroll[3] and Wu and Bailey[4] have described methods to estimate rates of change of a continuous variable in the presence of informative right censoring. In these models, one does not consider the association between the repeated measurements and time of development of disease. In the second analysis, one models disease risk either parametrically or non-parametrically as a function of a time-dependent covariate, the most familiar and widely used technique being Cox regression.[5] Prentice[6] has also described an approach to survival analysis that accounts for covariate measurement error.

We address both of these components in a single analysis by modelling a continuous covariate over time and simultaneously relating the covariate to disease risk. We use the Markov chain Monte Carlo method of Gibbs sampling[7,8] to generate the joint posterior distribution of all unknown parameters of the comprehensive model given only the observed data. The model allows for unequally spaced or missing covariate data, different numbers of observations per subject, and censoring of the survival times.

There are several advantages to modelling jointly the covariate and disease processes. In a survival analysis setting, where the covariate of interest is time-dependent, either the entire history of the covariate for every subject, or, minimally, measurements of the covariate at each time of disease occurrence for all subjects in the corresponding risk set, are necessary. This extensive measurement of covariates is rarely, if ever, executed and the values obtained are typically subject to measurement error. By modelling the covariates over time, we can enhance the survival analysis since we can interpolate covariate values between the observed measurements to the specific times of disease occurrence, with use of the entire covariate history of the subjects. Modelling the covariate also allows adjustment for covariate measurement error, which is known to result in biased estimates of relative risk parameters.[6] By accounting for measurement error, the standard error of the relative risk estimate will reflect correctly the uncertainty in the measurements of the covariate. Conversely, utilizing the survival data in the covariate tracking model will yield improved covariate tracking parameter estimates by allowing adjustment for informative right censoring of the repeated measurements by the disease process.

Clayton[9] proposed a comprehensive model that combined the covariate tracking and disease risk models, and several other investigators have described methods for estimating parameters of similar models. De Gruttola and Tu,[10,11] and Tsiatis *et al.*[12,13] consider the progression of CD4 lymphocyte counts and survival time in patients with AIDS. De Gruttola and Tu assume that the joint distribution of time-dependent log CD4 counts and some transformation of survival times are multivariate normally distributed. This formulation allows them to use a modified EM algorithm (Laird and Ware[1]) to fit the model. Using Cox's proportional hazards model, Tsiatis *et al.* model the hazard of death as a function of the conditional expectation of 'true' log CD4 counts given the history of observed counts, thus relaxing the normality assumption for the survival time. They propose a two-step procedure for fitting their model. First, they assume a growth curve random components model with normal errors for true CD4 counts and they use the modified EM algorithm for estimation. One then substitutes these estimates into the proportional hazards model and uses Cox regression to obtain estimates of the survival parameters. Self and Pawitan[14] model the relationship between CD4/CD8 ratios and time to AIDS diagnosis. They also propose a two-step method for parameter estimation that differs from Tsiatis *et al.* in that they condition on the survival information when computing expected values of the

covariates. They also use partial likelihood to obtain estimates of the disease risk parameters, but they derive corresponding variances to account for the uncertainty in the expected covariate values. To obtain these variances, they make the simplifying assumption that the variance of the covariate random effects is fixed and known. Finally, Pawitan and Self[15] use maximum likelihood methods to model jointly immunologic markers, time of infection, and time to AIDS. Instead of modelling time of disease as a function of the marker, they model the marker as a function of disease time. They consider fully parametric Weibull regression models for the times of disease and infection.

We assume a random components model with normal errors for the 'true' log and square root CD4 counts and a proportional hazards model for AIDS-free survival time. We utilize Gibbs sampling to fit simultaneously the covariate tracking model and the disease risk model, an approach that has several advantages. First, we adjust estimates of the true covariates for the informative censoring caused by the survival data. Second, we obtain estimates of variability, and in fact the entire marginal or joint posterior distributions, of all model parameters in the Gibbs sampling framework without complex derivations or simplifying assumptions. Thus, the variance estimates for the disease risk parameters correctly reflect the uncertainty inherent in the covariate tracking model parameters, and, conversely, the variance estimates of the covariate tracking model parameters reflect the uncertainty in the parameters of the disease risk model. Also, we can incorporate informative priors if desired for fully Bayesian analyses. Finally, since the Gibbs sampling approach does not require that the component submodels are conjugate, we can modify it easily to accommodate a variety of situations and assumptions.

In this paper, we use simulation studies to compare the analysis of the joint covariate tracking and disease risk model (using Gibbs sampling) to the separate analysis of each component (using standard methods), and then apply our method to a data set from the Transfusion Safety Study. The Transfusion Safety Study is an observational, epidemiologic study that follows a cohort of blood transfusion recipients and controls to determine risk factors for the development of transfusion-related diseases, particularly AIDS. Risk factors examined in this study include immunologic markers measured at periodic intervals. Dates of transfusion, entry into the study, and diagnosis of AIDS have been recorded. The data used for illustration of the methodology consist of absolute CD4 measurements made at each visit for anti-HIV-1 positive recipients of anti-HIV-1 positive blood donations.[16]

2. METHODS

2.1. Combined model and assumptions

The methods described apply to a data set that consists of the following for each of $i = 1, \ldots, I$ subjects: (i) $j = 1, \ldots, J_i$ measurements z_{ij} of a continuous time-dependent covariate, possibly measured with error; (ii) observation times of these measurements t_{ij} relative to some baseline event (for example, transfusion); (iii) time of entry into the study e_i ; (iv) a disease status indicator d_i; and (v) time (since baseline event) of development of disease or total length of follow-up if a subject is censored s_i. The number of observations and the intervals between observations need not be uniform across subjects.

We can specify the combined model in terms of two submodels: (i) the *covariate tracking model* that describes the relationship of the observed covariate measurements (measured with error) as a function of the true, unobserved covariate values (containing no error) and their underlying dependence on time; and (ii) the *disease risk model* that describes the relationship between the risk

of disease and the true, unobserved time-dependent covariate. Although the correct forms of the covariate tracking and disease risk models would depend on the specific variables and application under examination, we assumed the following forms for each model for the purposes of demonstrating the methodology. For simplicity, the following models and derivations assume that the covariate tracking model includes only time as a predictor, and that the disease risk model includes only the time-dependent covariate of interest. One could easily generalize both the covariate tracking model and the disease risk model to accommodate a vector of additional explanatory covariates, both fixed and time-dependent. We do assume, however, that we will model the progression of a single time-dependent covariate, that is, there is only one covariate tracking model.

2.1.1. Covariate Tracking Model

We may subdivide further the covariate tracking model into the measurement error model and the true covariate tracking model. We assumed a classical measurement error model for the observed covariate

$$z_{ij} = x_i(t_{ij}) + \varepsilon_{ij}$$

where $x_i(t_{ij})$ is the value of the true, unobserved covariate at time t_{ij}, and ε_{ij} are error terms assumed independent and normally distributed with mean 0 and variance σ_ε^2.

For illustrative purposes, we assumed that the true covariate tracks linearly as a function of time with subject-specific intercept α_i and slope β_i:

$$x_i(t) = \alpha_i + \beta_i t.$$

We assumed further that the random effects α_i and β_i have a bivariate normal distribution with means μ_α and μ_β, and covariance matrix

$$\mathbf{\Sigma} \equiv \begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{bmatrix}$$

where $\sigma_{\alpha\beta} = \rho\sigma_\alpha\sigma_\beta$ and ρ is the correlation between the random effects α_i and β_i.

2.1.2. Disease Risk Model

We assumed a proportional hazards model with disease risk that depends log-linearly on the true, unobserved covariate. Let $\lambda_i(t) = \lambda_0(t)\exp[\gamma x_i(t)]$ denote the disease hazard for subject i at time t, where $\lambda_0(t)$ represents the baseline hazard of disease evaluated at time t and γ is the regression coefficient for estimation. We assumed further that the baseline hazard was a step function, $\lambda_0(t) = \lambda_k$, defined over some arbitrary partitioning of the time scale into $k = 1, \ldots, K$ intervals, $t_{k-1} < t \leqslant t_k$, not necessarily related to the times of covariate measurement. The step widths should be small enough such that the assumption of a constant baseline hazard within an interval is reasonable. In theory, one could increase the number of steps and make the step widths arbitrarily small to approximate semi-parametric methods.

2.2. Gibbs sampling

The standard likelihood approach to this problem involves integration of the two component models over the distribution of random effects, which requires numerical integration since the two

models are not conjugate. As an alternative, we focus on the posterior distribution of model parameters, which approximates the likelihood function if we use flat priors. We estimate the joint posterior distribution of all unknown model parameters

$$[\{\alpha_i, \beta_i\}, \mu_\alpha, \mu_\beta, \boldsymbol{\Sigma}, \sigma_\varepsilon^2, \lambda_0(t), \gamma | \{z_{ij}\}, \{t_{ij}\}, \{s_i\}, \{d_i\}, \{e_i\}]$$

using Gibbs sampling. This is a Monte Carlo method for generating samples from the joint posterior distribution of unknown parameters in a model, conditional only on the observed data. The method involves iteratively sampling from the full conditional distributions of each parameter given the current assignment of all other parameters and data. Thus, the method is most useful when the joint distribution of parameters is intractable, but the generation of samples from each full conditional distribution is feasible.

Given a set of initial estimates for each of the unknown parameters, we generated in turn samples from the full conditional distributions of each unknown parameter over a large number of cycles. After discarding the early samples to allow the process to converge, we used subsequent realizations of each parameter for summarizing the posterior distributions. We can use means, medians and variances of the Gibbs samples, and graphs of the empirical distributions to describe the results. We can base approximate confidence limits on the means and variances of the generated samples, although we can also use percentiles of the empirical distributions to estimate confidence intervals. We can obtain superior estimates of posterior distributions by averaging the full conditional distributions across Gibbs samples.[8] One can easily incorporate informative priors if a more Bayesian perspective is of interest. We conservatively based the numbers of iterations required for convergence and alternative methods of sampling to reduce autocorrelation on preliminary simulation studies, with use of the method of Gelman and Rubin[17] to assess convergence. Gelman and Rubin compare the within- and between-series variability from multiple sequences of Gibbs samples to evaluate convergence. Alternatively, Geyer[18] advocates the use of a single series of samples and argues that it also allows for adequate assessment of convergence while making the most efficient use of the Gibbs samples for estimation.

2.3. Procedures for sampling from each conditional distribution

We describe below the procedures used to generate samples from each conditional distribution. We can use approximately uninformative priors for parameters to approximate likelihood methods.[19] Specifically, we used flat priors for μ_α, μ_β and γ, $|\boldsymbol{\Sigma}|^{-3/2}$ for $\boldsymbol{\Sigma}$, $1/\sigma_\varepsilon^2$ for σ_ε^2, and $1/\lambda_k$ for baseline hazards λ_k.

2.3.1. Random Effects: α_i and β_i, $i = 1, \ldots, I$

Using Bayes' rule, the posterior distribution of each random effect (for example, α_i) at the $(m+1)$th Gibbs iteration is proportional to the product of three simple conditional distributions. That is,

$$[\alpha_i | \beta_i^{(m)}, \mu_\alpha^{(m)}, \mu_\beta^{(m)}, \boldsymbol{\Sigma}^{(m)}, \sigma_\varepsilon^{2(m)}, \lambda_0(t)^{(m)}, \gamma^{(m)}, \{z_{ij}\}, s_i, d_i]$$

$$\propto [\{z_{ij}\} | \alpha_i, \beta_i^{(m)}, \sigma_\varepsilon^{2(m)}] \times [\alpha_i | \beta_i^{(m)}, \mu_\alpha^{(m)}, \mu_\beta^{(m)}, \boldsymbol{\Sigma}^{(m)}] \times [s_i, d_i | \alpha_i, \beta_i^{(m)}, \lambda_0(t)^{(m)}, \gamma^{(m)}]$$

for $i = 1, \ldots, I$. The superscript denotes the mth sampled value of each parameter. For convenience, we omit the superscripts in the following descriptions. The first two terms are simply the

normal distributions proportional to:

$$\exp\left\{-\sum_{j=1}^{J_i}[z_{ij}-(\alpha_i+\beta_i t_{ij})]^2/(2\sigma_\varepsilon^2)\right\}$$

and

$$\exp\left\{-\left[\alpha_i-\left(\mu_\alpha+\frac{\sigma_{\alpha\beta}}{\sigma_\beta^2}(\beta_i-\mu_\beta)\right)\right]^2\Big/[2\sigma_\alpha^2(1-\rho^2)]\right\}$$

respectively, and the last term is the full likelihood of the survival parameters based on the unobserved, true covariate values

$$\{\lambda_0(s_i)\exp[\gamma(\alpha_i+\beta_i s_i)]\}^{d_i}\exp\left\{-\int_{e_i}^{s_i}\lambda_0(t)\exp[\gamma(\alpha_i+\beta_i t)]\mathrm{d}t\right\}. \tag{1}$$

Since this function of α_i has no simple distributional form, we used a rejection sampling procedure to generate samples from the exact posterior distribution. Owing to the difficulty of consistently finding a guess density that completely covered the true density, we applied derivative-free adaptive rejection sampling applicable to log-concave functions.[20] We generated samples from the posterior distribution of each β_i similarly.

2.3.2. Means of Random Effects: μ_α and μ_β

Again using Bayes' rule, the posterior distribution for μ_α (and similarly for μ_β) is normal with mean $\bar{\alpha}-(\bar{\beta}-\mu_\beta)\sigma_{\alpha\beta}/\sigma_\beta^2$ and variance $\sigma_\alpha^2(1-\rho^2)/I$, where $\bar{\alpha}$ and $\bar{\beta}$ are the means across subjects of the currently sampled values of the random effects.

2.3.3. Random Effects Covariance Matrix: $\boldsymbol{\Sigma}$

We used the procedure of Odell and Feiveson,[21] summarized by Zeger and Karim,[2] to generate the covariance matrix $\boldsymbol{\Sigma}$. We first generated a standardized Wishart variate with $I-1$ degrees of freedom, $\mathbf{W}$, and then computed $\boldsymbol{\Sigma}=(\mathbf{H}'\mathbf{W}\mathbf{H})^{-1}$ where $\mathbf{H}$ is

$$\left(-\sum_{i=1}^{I}\mathbf{B}_i\mathbf{B}_i'\right)^{-1}=\mathbf{H}'\mathbf{H}$$

and $\mathbf{B}_i$ is the vector $[(\alpha_i-\mu_\alpha),(\beta_i-\mu_\beta)]'$.

2.3.4. Error Variance: σ_ε^2

With a prior proportional to $1/\sigma_\varepsilon^2$, the posterior distribution for the error variance is inverse gamma-distributed with parameters

$$N/2=\sum_{i=1}^{I}J_i/2$$

and

$$SS/2=\sum_{i=1}^{I}\sum_{j=1}^{J_i}[z_{ij}-(\alpha_i+\beta_i t_{ij})]^2/2.$$

Thus, we can generate a sample from the posterior distribution of σ_ε^2 by simply computing $SS/\chi_{(N)}^2$ where $\chi_{(N)}^2$ is a random chi-square deviate with N degrees of freedom.

2.3.5. Baseline Hazard: λ_k, $k = 1, \ldots, K$

The posterior distribution for a single step of the baseline hazard λ_k also has a gamma distribution with parameters

$$D_k = \sum_{i=1}^{I} d_i I_{\{t_{k-1} < s_i \leqslant t_k\}}$$

and

$$Y_k = \sum_{i=1}^{I} I_{\{s_i > t_{k-1}\}} \int_{t_{k-1}}^{\min(s_i, t_k)} \exp[\gamma(\alpha_i + \beta_i t)]\,dt$$

where $I_{\{t_{k-1} < s_i \leqslant t_k\}}$ and $I_{\{s_i > t_{k-1}\}}$ are indicators that the ith subject developed disease or was followed during the kth time period, respectively. We generated a random sample from this distribution by dividing a standard gamma deviate with shape parameter D_k by Y_k.

2.3.6. Disease Risk Parameter: γ

The posterior distribution for γ is proportional to the likelihood function given in equation (1). We used traditional rejection sampling[22] with a normal guess density with mean equal to the iteration-specific maximum likelihood estimate of γ and variance equal to twice the inverse of the observed information evaluated at that maximum likelihood estimate to generate a sample from this distribution. We doubled the variance to ensure that the guess density always covered the true density.[2]

3. SIMULATIONS

All simulations consisted of 100 data replications, each composed of $J_i = 1$ to 15 unequally spaced observations for $I = 110$ subjects. The total number of observations per subject was determined by the disease and censoring processes (see Faucett[23] for details). Table I summarizes the true parameter values used in generating the data ($\rho = 0$ and $K = 1$ for this simulation). Other simulations used the same set-up as the base simulation but we varied the disease risk, mean slope, and measurement error variance parameters (see Table II for true parameter values used in these simulations). An additional simulation with $\rho = -0.55$ and $K = 10$ (data not shown) produced results similar to those shown below.

For comparison with the full model, we also fitted each submodel separately using standard methods. For the covariate tracking model, we fitted a random effects model ignoring the survival data with use of the EM algorithm described by Laird and Ware[1] and Laird *et al.*[24] For the disease risk model, we maximized the corresponding full likelihood using the observed covariates (treated as time-dependent step-function covariates).

3.1. Gibbs sampling details

Preliminary simulation studies[23] indicated that elimination of 60 samples for convergence and retaining every third sample, for a total of 100 samples per data set, resulted in reasonable convergence diagnostics and a feasible computation time of approximately 24 hours to simulate and analyse 100 data sets on a 486/25MHz personal computer. We updated the highly correlated parameters α_i and β_i, and γ and λ_k twice before retaining a value to speed convergence,

Table I. Parameter estimates and performance statistics for base simulation

Parameter	True value	Parameter estimates		Performance statistics		Ratio of MSEs
		Comparison	Combined	Comparison	Combined	
Covariate tracking model						
μ_α	1·314	1·279* (0·104)	1·315 (0·105)	− 2·69%† (0·101) 95%	0·05% (0·099) 97%	1·16‡
μ_β	− 0·005	− 0·00475 (0·000167)	− 0·00501 (0·000176)	− 5·08% (0·000185) 66%	0·20% (0·000194) 91%	2·63
σ_α	1·033	1·030 nc	1·037 (0·077)	− 0·27% (0·080) nc	0·39% (0·078) 92%	1·03
σ_β	0·00143	0·00139 nc	0·00146 (0·000143)	− 3·01% (0·000138) nc	2·17% (0·000146) 94%	0·94
σ_ε	0·548	0·550 nc	0·550 (0·016)	0·39% (0·0155) nc	0·39% (0·0151) 96%	1·04
Disease risk model						
γ	− 2·000	− 1·266 (0·100)	− 2·116 (0·210)	− 36·7% (0·091) 0%	5·82% (0·304) 84%	5·22
$\ln \lambda_1$	− 9·835	− 8·159 (0·223)	− 10·101 (0·461)	− 17·0% (0·020) 0%	2·70% (0·659) 85%	5·70

* Parameter estimate = average of the posterior means over the 100 data sets; (standard error = average of the posterior standard deviations over the 100 data sets)
† Bias as per cent of true parameter; (standard deviation of parameter estimates across data replications); coverage of approximate 95 per cent confidence limits
‡ Mean squared error from comparison analysis/mean squared error from combined analysis
nc Not computed

a procedure suggested by Zeger and Karim.[2] Gelman and Rubin scale reduction factors[17] were all less than 1·07 (based on five restarts) suggesting that we would gain little additional precision by allowing the process to run longer (scale reduction factor = 1 indicates complete convergence). We used means and standard deviations of a single series of 100 Gibbs samples as point estimates of the parameters and their standard errors to facilitate simple comparisons with the naive approaches. We calculated approximate 95 per cent confidence limits by taking the mean $\pm 1{\cdot}96$ times the standard deviation.

3.2. Results

Table I gives the results of the first simulation. The combined model resulted in improved estimates of the means of the random effects compared to the model that ignored survival information. The comparison model resulted in an underestimated mean intercept (μ_α) and slope (μ_β) due to informative right censoring. This bias was virtually eliminated with estimates based on the combined model. The covariate tracking model variance parameters from both analyses, however, contained similar amounts of bias.

Table II. Performance statistics for μ_β and γ from multiple simulations

True parameter values			Average number diseased	μ_β			γ		
μ_β	σ_ε^2	γ		Comparison	Combined	Ratio of MSEs	Comparison	Combined	Ratio of MSEs
−0·001	0·30	−2·00	21	−7·1%* (0·000138) 95%	−1·9% (0·000145) 96%	1·1†	−22·9% (0·206) 41%	17·6% (0·650) 91%	0·46
−0·001	0·30	−5·00	37	−17·6% (0·000150) 81%	1·5% (0·000151) 93%	2·3	−61·9% (0·211) 0%	30·6% (2·098) 75%	1·44
−0·001	0·50	−2·00	21	−6·6% (0·000153) 89%	0·9% (0·000155) 93%	1·2	−33·4% (0·175) 9%	18·2% (0·640) 94%	0·88
−0·001	0·50	−5·00	36	−20·2% (0·000150) 72%	2·6% (0·000159) 94%	2·5	−70·0% (0·175) 0%	36·5% (2·540) 76%	1·26
−0·005	0·30	−2·00	97	−5·1% (0·000185) 66%	0·2% (0·000194) 91%	2·6	−36·7% (0·091) 0%	5·8% (0·304) 84%	5·22
−0·005	0·30	−5·00	106	−14·0% (0·000255) 10%	0·2% (0·000234) 91%	10·3	−70·0% (0·118) 0%	15·4% (1·329) 70%	5·24
−0·005	0·50	−2·00	97	−6·7% (0·000186) 48%	0·6% (0·000188) 96%	4·0	−46·5% (0·088) 0%	5·5% (0·326) 86%	7·40
−0·005	0·50	−5·00	106	−18·4% (0·000295) 6%	−0·1% (0·000289) 88%	11·3	−76·4% (0·102) 0%	15·0% (1·418) 70%	5·72

* Bias as per cent of true parameter; (standard deviation of estimates across data replications); coverage of approximate 95 per cent confidence limits

† Mean squared error from comparison analysis/mean squared error from combined analysis

The differences between the models were more striking for the disease risk parameters. The model that ignored the covariate tracking information underestimated the true γ and $\ln \lambda_1$ by 37 and 17 per cent, respectively, due to measurement error; the combined analysis produced less biased estimates (6 per cent and 3 per cent, respectively). This remaining mean bias may be partly due to a skewed posterior distribution for γ: the median bias in γ was only 3 per cent. Other simulation studies,[2,3] using a simpler model, have shown that the remaining bias may be due to sample size, since, for a given sample size, the bias is the same as that found in maximum likelihood estimates, and increasing sample size reduced that bias.

The estimates of standard error were larger for the combined model than for the separate models for every parameter. We expected this since the larger standard error correctly reflects the uncertainty in both the covariate parameters and the survival parameters, as reflected in the superior coverage of approximate 95 per cent confidence limits. Both analyses performed well for all of the covariate tracking parameters except for the mean slope estimate, where the combined analysis (91 per cent coverage) outperformed the comparison analysis (66 per cent coverage). For the survival parameters the differences were more pronounced. For the comparison analysis, the coverage of the confidence limits for both γ and $\ln \lambda_1$ was 0 per cent, whereas the combined analysis improved the coverage to 84 per cent and 85 per cent, respectively.

For the covariate tracking parameters, the combined approach generally yielded smaller MSE than the comparison analysis particularly for the mean slope (μ_β) and intercept (μ_α) parameters. For the survival parameters, the combined analysis yielded much smaller MSE for both the log baseline hazard (ratio = 5·7) and the disease risk parameter (ratio = 5·2). Most of this, however, was due to the reduction in bias since the variance of the estimates was considerably larger for the combined model than for the comparison analysis (0·304 versus 0·091 for γ and 0·659 versus 0·020 for $\ln \lambda_1$). Thus, we gained reduction in bias at the cost of an increase in variability. The increase in variability was not surprising since, for the comparison analysis, we treat the covariates as known, whereas for the combined analysis, we estimate two random effects for all 110 subjects in addition to the survival parameters.

Table II summarizes the results of simulations that investigated the effect of varying the true values of selected parameters (true values for the remaining parameters are those shown in Table I). For the mean slope (μ_β) parameter, the combined analysis yielded estimates with less bias, comparable amounts of variability, smaller MSEs, and better confidence interval coverage than the comparison analysis for all parameter combinations. The reduction in bias and improvement in confidence interval coverage in the disease risk estimates (γ) based on the combined model compared to the separate model was also consistent across all parameter values considered. The advantages of using the combined approach for estimating both parameters were greater for larger true values of μ_β and γ.

Although the coverage of the confidence limits was superior for the combined model analyses, all were lower than 95 per cent. This may be due to the effects of the high autocorrelation in successive samples of the disease risk parameters (more samples may be necessary to obtain less biased variance estimates) or due to skewed posterior distributions of the disease risk parameter (violation of the normality assumption when computing confidence intervals). Use of posterior percentiles might produce confidence limits with better coverage.

Both methods performed well for estimating most of the other covariate tracking parameters (data not shown). For all parameters, however, the combined analysis produced estimates that were less biased but more variable. In general, combined model estimates had lower mean squared error and had better coverage of the 95 per cent confidence limits. Bias tended to increase with increasing mean slopes, increasing disease risk parameter, and decreasing proportion of diseased cases. Changes in measurement error did not affect bias considerably. The variability of the estimates also decreased and the ratio of MSE increased with increasing proportions of diseased cases.

3.3. Model misspecification

We ran another set of simulations to determine whether residual plots are useful in detecting a misspecified covariate tracking model. We generated four sets of simulated data using different underlying covariate tracking processes. For the first simulation, we used a linear model with true mean intercept 1·314, mean slope $-0{\cdot}0015$, and with other true parameter values shown in Table I. The second simulation added a quadratic term with coefficient $-0{\cdot}00000076$ to the above linear model. This corresponds to an average maximum covariate value at about 657 days. For the third simulation, we used a logistic feedback model, with differential equation given by $dy/dt = -\beta_i(\alpha_i - y(t))(y(t) + 8)$. This corresponds to an asymptote of α_i to the left declining to an asymptote of -8 to the right (or close to 0 for a log-transformed covariate). Finally, the fourth simulation used a transition point model with a zero slope until time 550 days, and linear decline thereafter with true parameters used in the first simulation.

For each simulation, we analysed the data assuming a linear decline model. Hence, the model for the first simulation is correctly specified while the model for the remaining three simulations is misspecified. We obtained the predicted covariate value for subject i at each of their observation times from the linear prediction equation with intercept and slope computed as the mean of the Gibbs samples of α_i and β_i, respectively.

Figure 1 shows plots of residuals (observed–predicted) versus predicted covariate values from the four simulated situations. The residuals for the correctly specified model appear as an even band around zero, whereas definite patterns emerged with the model misspecified. We also plotted the residuals versus time of observation (data not shown). These plots were similarly useful in detecting these misspecified models, particularly for the transition point model where an arc was clearly evident, indicating a poor fit.

4. APPLICATION

We illustrate the methodology using data from the NHLBI funded Transfusion Safety Study (TSS), a multi-centre study that examined risk factors for the transmission of disease, particularly AIDS, through blood transfusions. From September 1984 to February 1985, prior to development of the test for the HIV-1 antibody, the TSS collected approximately 200,000 blood samples from donors in high AIDS prevalence areas (Los Angeles, San Francisco, Miami and New York). Once the HIV-1 antibody test became available, the blood samples were tested and anti-HIV-1 positive donors and their recipients were enrolled in the study.[25] Additionally, negative control donors and recipients, congenital haematologic disorder subjects, and household contacts were also enrolled. All subjects were followed longitudinally and their haematologic, immunologic and clinical status was measured at either three or six month intervals depending on the type of subject and level of disease.

For this application, we analysed the subset of data for the anti-HIV-1 positive recipients of positive blood donations. The disease outcome of interest was diagnosis with AIDS and the time-dependent covariate was the immunologic marker absolute CD4. The flow cytometric methods for measuring the CD4 markers have been published previously,[26] as well as a detailed description of these recipients.[16]

The data set analysed consisted of 109 HIV positive recipients known not to have any other risk factors for AIDS (such as intravenous drug use or male homosexuality). The exact date of transfusion with the positive blood product was known for all members of this cohort. Enrollment into the study occurred between 6 and 36 months post-transfusion (median, 20 months). After enrollment, there were repeated measurements obtained of the immunologic marker absolute CD4. Only observations taken before 1 October 1989 (the time of initiation of a different method of measuring absolute CD4) and before the development of AIDS were utilized. There was a total of 642 observations of CD4; each subject had anywhere from $J_i = 1$ to 13 observations (median = 6), and these observations corresponded to a follow-up time of from just less than 1 year to 5·6 years (median = 4·1 years).

All AIDS cases diagnosed within six months past the last allowed observation of CD4 (1 October 1989) were included in the analysis. Of the 109 subjects analysed, there was a total of 24 AIDS cases diagnosed during the period of interest. After this point, there were an additional 5 cases of AIDS diagnosed corresponding to a total of 29 AIDS cases in the entire data set. The Kaplan–Meier estimate of the cumulative probability of developing AIDS at five years post-transfusion was 32 per cent (95 per cent confidence limits: 22 per cent to 44 per cent).

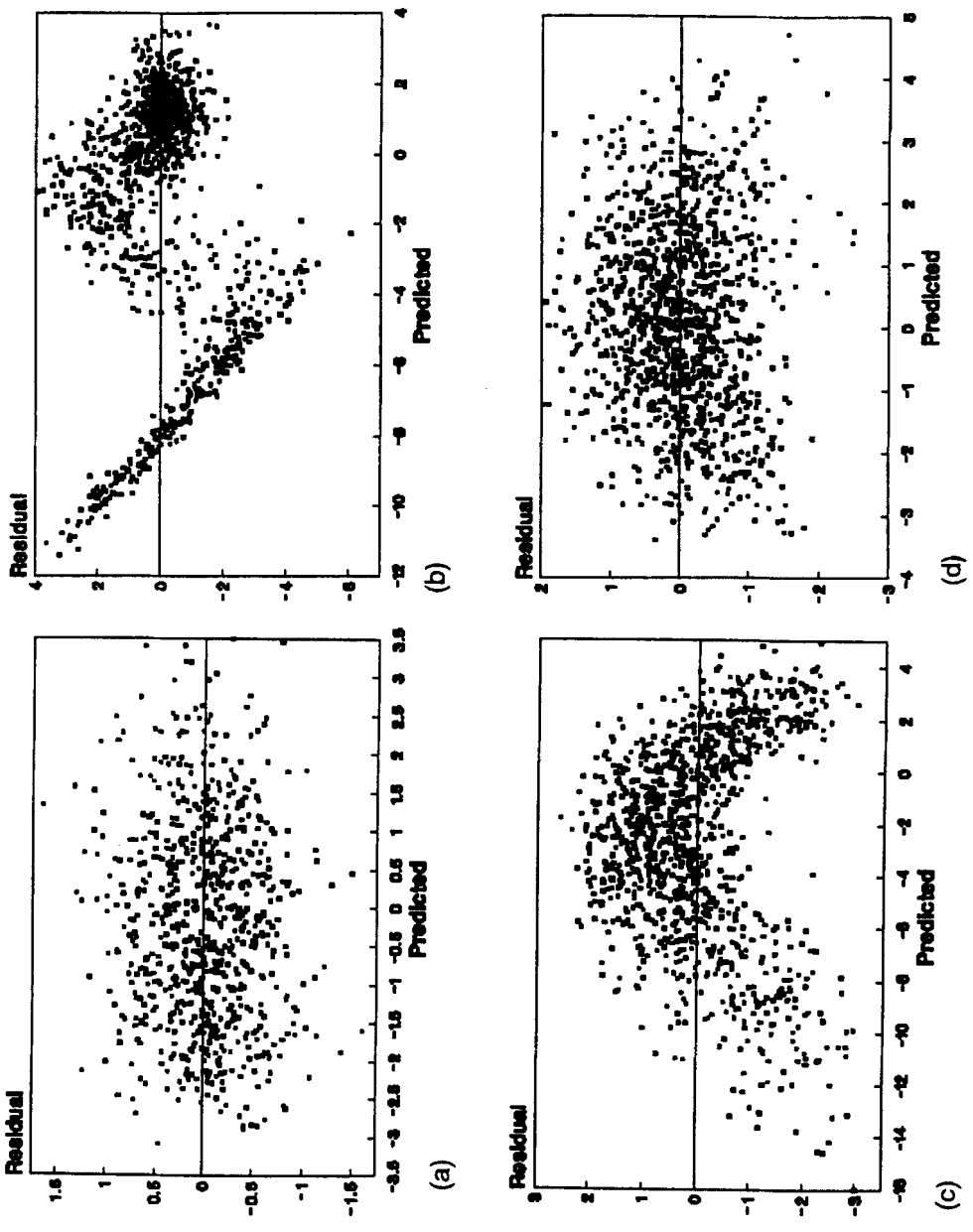

Figure 1. Scatter plots of residuals (observed–predicted covariate measurements) versus predicted covariate values from four simulated data sets analysed assuming a linear decline for the covariate model. The data sets were generated using a: (*a*) linear model; (*b*) logistic feedback model; (*c*) quadratic model, and (*d*) transition point model

Table III. TSS application using combined and comparison analyses

Parameter	Parameter estimates			
	Log CD4 − 5.7		$\sqrt{\text{CD4}} - 18.8$	
	Comparison	Combined	Comparison	Combined
Covariate tracking model				
μ_α	0·80* (0·11) (0·58, 1·02)	0·83 (0·12) (0·60, 1·06)	6·33 (1·07) (4·24, 8·43)	6·64 (1·11) (4·47, 8·82)
μ_β	− 0·00085 (0·000117) (−0·00108, −0·00062)	− 0·00090 (0·000129) (−0·00115, −0·00064)	− 0·0062 (0·00086) (−0·0079, −0·0045)	− 0·0067 (0·00093) (− 0·0085, − 0·0049)
σ_α	0·92 nc	0·95 (0·13) (0·70, 1·21)	9·66 nc	10·05 (1·06) (7·96, 12·14)
σ_β	0·00093 nc	0·00100 (0·00012) (0·00077, 0·00124)	0·0067 nc	0·0074 (0·00086) (0·0057, 0·0091)
$\sigma_{\alpha\beta}$	− 0·00047 nc	− 0·00054 (0·00022) (−0·00097, −0·00011)	− 0·042 nc	− 0·048 (0·017) (−0·81, −0·016)
σ_ε	0·356 nc	0·354 (0·012) (0·331, 0·378)	2·72 nc	2·72 (0·09) (2·53, 2·90)
Disease risk model				
γ	− 1·53 (0·21) (− 1·94, − 1·13)	− 1·81 (0·29) (− 2·39, − 1·24)	− 0·31 (0·047) (− 0·40, − 0·21)	− 0·42 (0·096) (− 0·61, − 0·24)
Average $\ln \lambda_k$	nc	− 8·88	nc	− 10·27
Rescaled CD4: $RR_{(200 vs. 400)}$	2·9	3·5	6·0	12·0

*Parameter estimate; (standard error); approximate 95 per cent confidence limits
nc Not computed

We considered linear models over time for both the natural log and square root transformations of absolute CD4. The log-linear model is based on the premise that the change in CD4 at any time is proportional to the current value of CD4 at that time. The square root transformation is the variance stabilizing transformation for count data. Both transformations force predicted CD4 values to be non-negative. We subtracted the mean of the transformed values (5·7 for log CD4 and 18·8 for square root CD4) from each subject's covariates for analysis, and we measured time in days since transfusion. Lange *et al.*[27] and Pawitan and Self[15] also used linear models for the square root and log CD4, respectively, and found that the more complicated models considered, such as inclusion of a random change-point, did not substantially improve model fit. We assumed a proportional hazards model with log-linear relative risk that depends on the 'true' CD4 for the hazard of developing AIDS as described in Section 2.1. We used steps of width one-month for the baseline hazard, based on the assumption that AIDS risk was constant for a given CD4 value within a month.

For each Gibbs sampling analysis, we eliminated a total of 100 iterations for convergence and we generated an additional 10,000 samples for summarization. Table III shows the results of the analyses of the data (based on both log and square root transformations) using the combined model, together with the parameter estimates obtained from the comparison analyses (the random effects model that ignores survival, and a Cox regression based on the observed covariates). The Gibbs sampling analyses took approximately 17 hours to complete on a 486/66MHz personal computer.

As expected from the stimulation studies, the estimates of mean slope from the joint analyses (log CD4: $-0{\cdot}00090$; $\sqrt{\text{CD4}}$: $-0{\cdot}0067$) were somewhat steeper than those from the corresponding comparison analyses (log CD4: $-0{\cdot}00085$; $\sqrt{\text{CD4}}$: $-0{\cdot}0062$), but these differences were clinically unimportant. The comparison estimate corresponded to a mean decrease in CD4 of 27 per cent per year compared with a decrease of 28 per cent per year for the combined analysis (based on the log transformation). These differences were sometimes larger on an individual basis. The most extreme example was for a subject whose CD4 was estimated to decrease 34 per cent per year based on the model that ignored survival time, but was estimated to decrease by 43 per cent per year when based on the combined model. This subject developed AIDS and had only one CD4 measurement.

Also noteworthy is the difference in the disease risk parameter estimates from the two methods. The comparison analysis that used the log-transformed observed covariates yielded an estimate for γ of $-1{\cdot}5$ (approximate 95 per cent confidence limit: $-1{\cdot}9$, $-1{\cdot}1$) compared to $-1{\cdot}8$ ($-2{\cdot}4$, $-1{\cdot}2$) from the combined model analysis, an 18 per cent increase in magnitude. On an absolute scale, the combined model estimate corresponds to a relative risk of AIDS development of 3·5 for a subject with a CD4 of 200 compared to a subject with a CD4 of 400. The corresponding relative risk based on the comparison model is 2·9. Using the square root transformed data, the difference was more pronounced. The comparison analysis yielded an estimated γ of $-0{\cdot}31$ compared to $-0{\cdot}42$ for the joint analysis, an increase of 35 per cent. The 95 per cent confidence limit from the comparison analysis ($-0{\cdot}40$, $-0{\cdot}21$) did not include the combined model estimate. On an absolute scale, the combined and comparison model estimates corresponded to relative risks of AIDS development of 12·0 and 6·0, respectively, for a subject with a CD4 of 200 compared to a subject with a CD4 of 400.

The difference in relative risk estimates obtained using the log and square root transformations is due to a difference in model assumptions. For example, with log transformed data we assume that the increased AIDS risk for a CD4 of 200 compared to 400 is the same as the risk for a CD4 of 100 compared to 200. For the square root transformed data, we assume that the increased risk for a CD4 of 200 compared to 400 is the same as the risk for a CD4 of 68 compared to 200. For smaller values of CD4, the relative risks for log and square root transformations are more comparable, for example, the relative risks for a CD4 of 100 compared to 150 are 2·1 and 2·3 for the log and square root transformations, respectively.

We performed several analyses to examine model adequacy. Preliminary Cox regression analyses, including an interaction term between last observed CD4 and log of time since transfusion, did not suggest significant deviation from proportionality (p-values for both transformations $> 0{\cdot}45$). In another analysis, we compared the relative risks estimated from a Cox regression on CD4 categorized using quintiles with those obtained from Cox regressions based on the log and square root transformations of CD4 treated continuously. The risk estimates based on the square root transformation more closely matched the estimates based on the categorized CD4, suggesting that the square root transformation yielded a superior fit for the disease risk model.

Preliminary analyses that fit separate regression lines of CD4 versus time showed that residuals for the square root transformed CD4 were more nearly normal than those from the log CD4

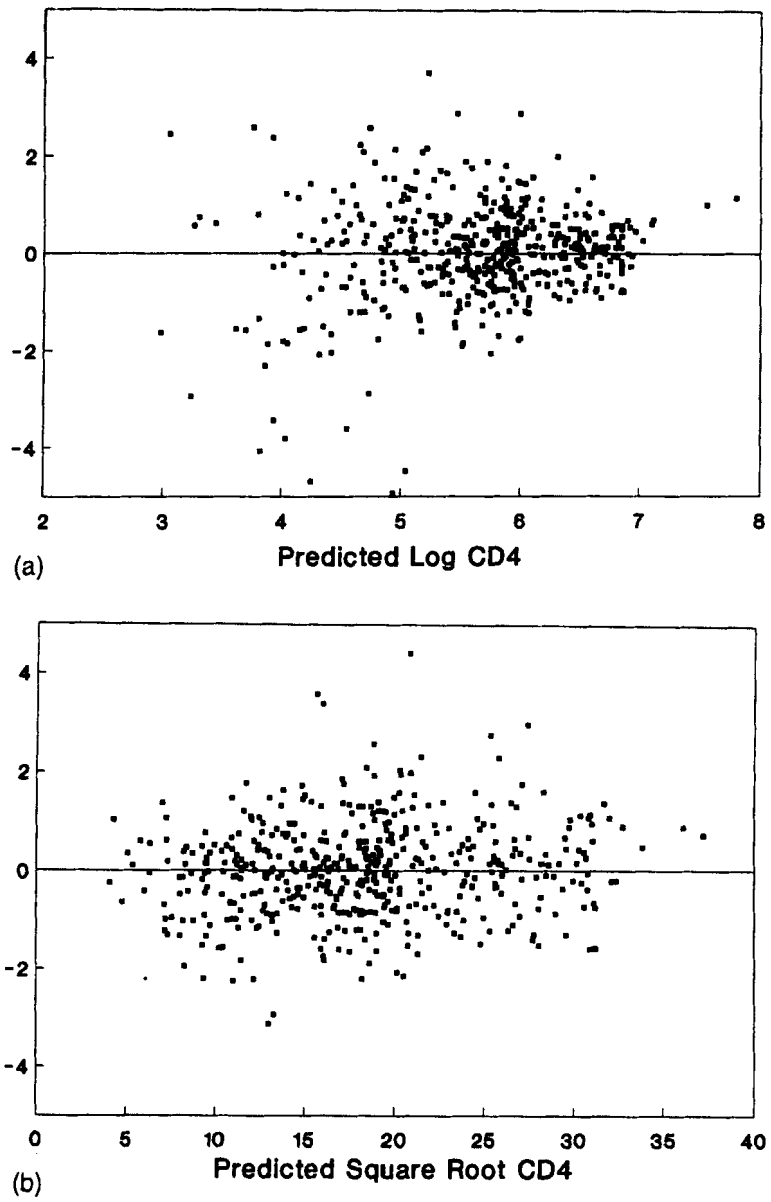

Figure 2. Standardized residuals versus (*a*) predicted log CD4, and (*b*) predicted $\sqrt{\text{CD4}}$ from the TSS joint analysis

analyses. Figure 2 shows plots of standardized residuals from the joint analyses (computed as described in Section 3.3) versus the predicted log and square root CD4. The residuals based on the square root transformation are more evenly distributed around zero and furthermore, no clear patterns emerge to suggest significant deviation from the linear model. We obtained similar findings based on the plots of residuals versus time since transfusion (data not shown).

Figure 3 shows the histogram of standardized residuals for the tracking model fitted using the combined analysis of the square root transformed data. Only six of the 642 CD4 observations produced a residual larger than three in absolute value; these were from six different subjects.

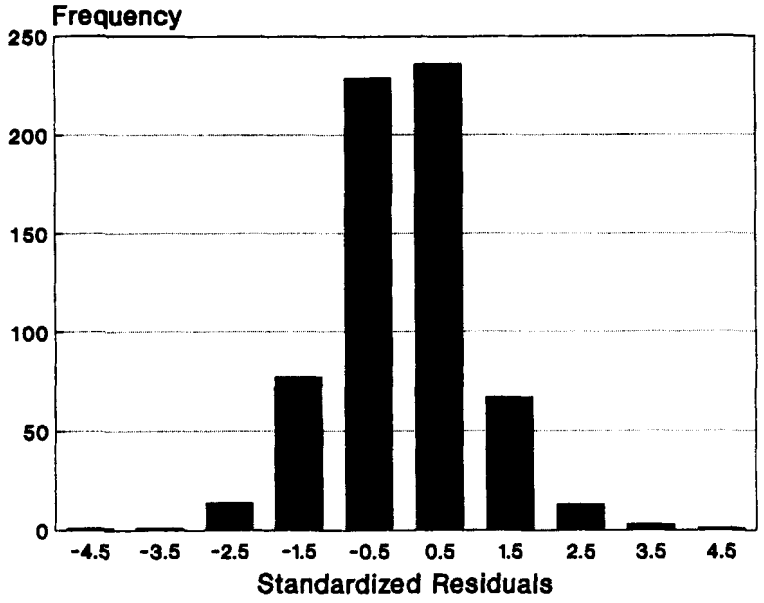

Figure 3. Distribution of standardized residuals for $\sqrt{\text{CD4}}$ from the TSS joint analysis

Examination of these six individuals' CD4 histories showed that the outlying CD4 measurement fell substantially below or above the expected linear decline. The distribution of these residuals did not differ significantly from a normal distribution ($p = 0{\cdot}74$). The distribution of residuals from the log transformed data, however, was significantly non-normal ($p < 0{\cdot}001$) and included 10 of 642 observations with residuals larger than 3. In all but one of these observations, the predicted CD4 was larger than the observed value. All of the above analyses suggested that the square root transformation yielded a superior fit and so we present the remaining results using the analysis based on this transformation.

Figure 4 shows the estimated bivariate distribution of person-specific slopes and intercepts from the joint and separate analyses based on the square root transformation, for those who developed AIDS and for those who did not. For subjects who had not yet developed AIDS, both analyses resulted in similar slope and intercept estimates (Figure 4(a)). For those who had developed AIDS, the joint analysis produced estimates that were steeper than the separate analysis for all but one case (Figure 4(b)). Additionally, the estimated intercepts were generally larger based on the joint analysis.

Figure 5 shows the empirical posterior distributions for μ_β and γ from the joint analysis of the square root transformed data, along with the normal distributions with mean and standard deviation from the estimates obtained from the comparison analyses. This graph shows that the posterior distribution of μ_β from the joint analysis shifts towards steeper values compared to that found with survival information ignored. Similarly, the posterior distribution for γ from the joint analysis shifts towards larger (absolute) values compared to the estimate obtained from the separate analysis. Also evident is the increased variability in the posterior distribution for γ from the joint analysis.

To investigate the sensitivity of the model to the number and width of the baseline hazard steps, we conducted three additional analyses with increases in the baseline hazard step widths to

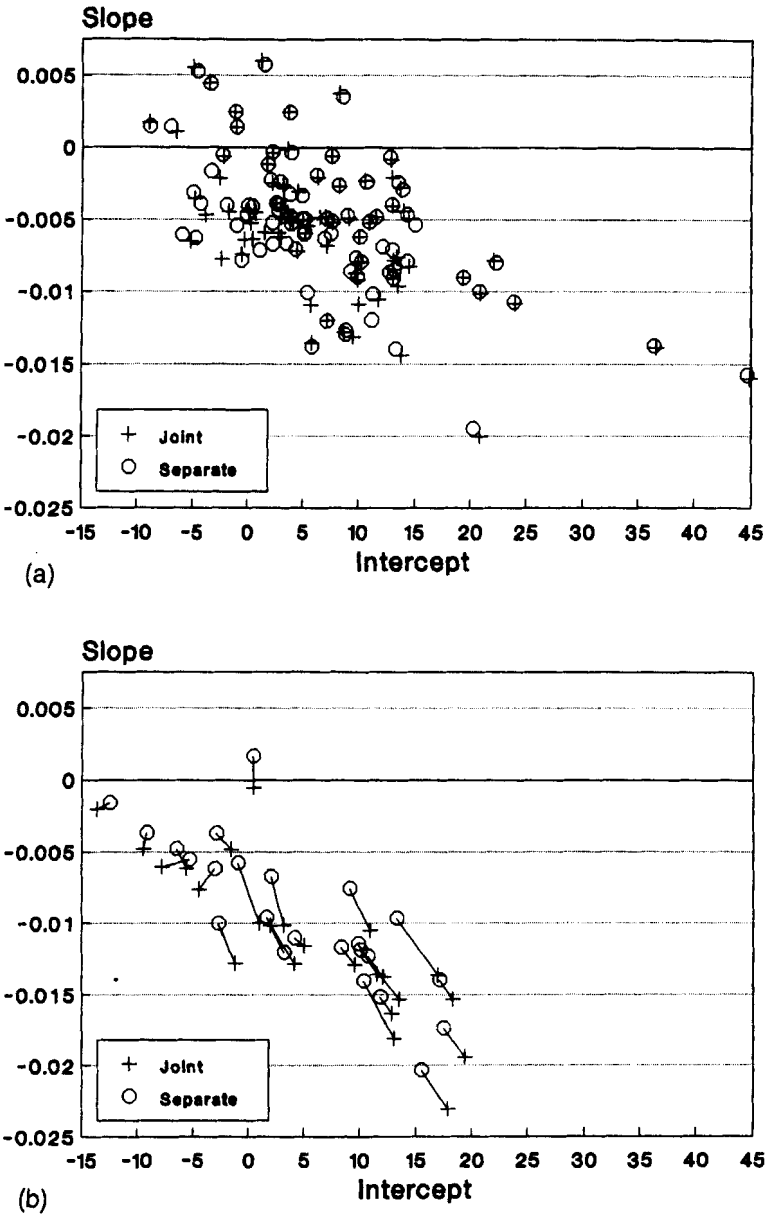

Figure 4. Estimated bivariate distribution of person-specific slopes and intercepts from the joint and separate analyses (based on $\sqrt{\text{CD4}}$) for (*a*) those who had not developed AIDS, and (*b*) for those who had developed AIDS

three-month, six-month and one-year intervals, respectively. Estimates of the disease risk parameter varied from $-0{\cdot}37$ (one-year intervals) to $-0{\cdot}42$ (one-month intervals), and the mean slope parameter estimates varied from $-0{\cdot}0066$ (one-year intervals) to $-0{\cdot}0067$ (one-month intervals). Parameter estimates from the one-month and three-month interval analyses were similar, suggesting that even three-month intervals are adequate, and even the one-year interval analysis did not produce substantially different parameter estimates.

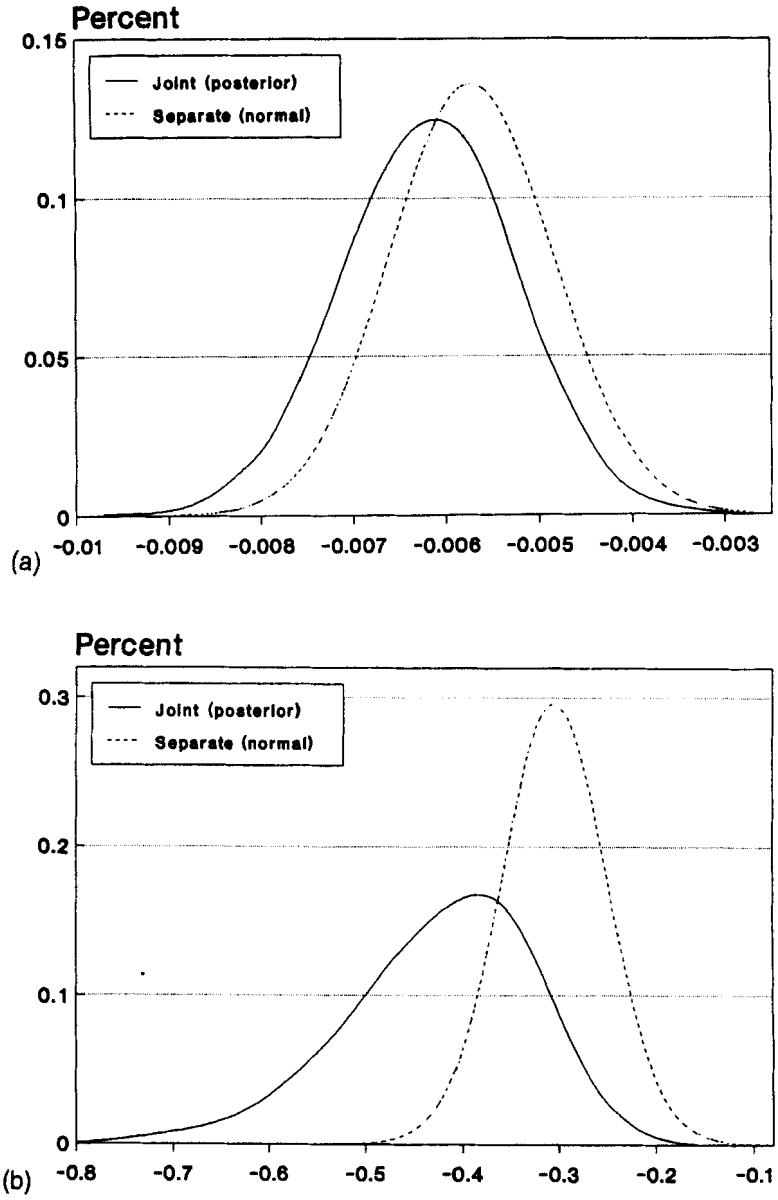

Figure 5. Solid lines are empirical posterior distribution from joint analysis of $\sqrt{\text{CD4}}$ and dotted lines are normal distribution using mean and variance from comparison analyses for (*a*) μ_β and (*b*) γ

Figure 6 shows a plot of the predicted hazard of AIDS per month as a function of CD4 count for each year post-transfusion (based on the joint analysis with one-year baseline hazard intervals). We give the year-specific confidence limits on the hazard for a CD4 of 200 in the figure legend. For a given CD4 value, the risk of AIDS is highest during the third year post-infection and decreases thereafter, although the hazard estimates in the second and sixth years are unstable as indicated by the wide confidence limits.

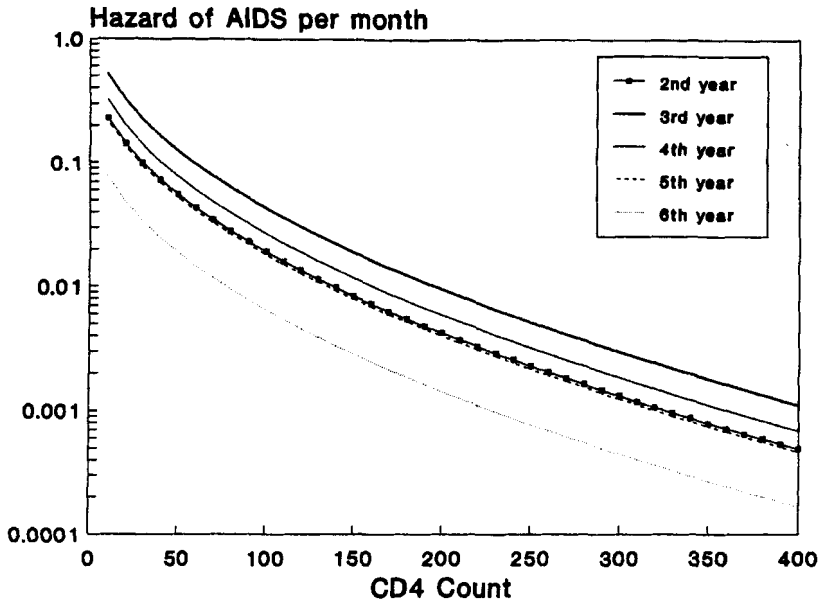

Figure 6. Hazard of AIDS per month as a function of CD4 count for each year post-transfusion. Year-specific 95 per cent confidence limits for a CD4 of 200 are: year 2 (0·0007, 0·025); year 3 (0·0038, 0·023); year 4 (0·0023, 0·015); year 5 (0·0012, 0·013); year 6 (0·0001, 0·024). These are based on the following estimates:

Parameter	Estimate	Standard error	Covariance of γ with $\ln \lambda_k$
γ	− 0·365	0·072	
$\ln \lambda_2$	− 10·59	1·09	0·053
$\ln \lambda_3$	− 9·78	0·67	0·037
$\ln \lambda_4$	− 10·25	0·73	0·044
$\ln \lambda_5$	− 10·65	0·86	0·053
$\ln \lambda_6$	− 11·67	1·58	0·054

Figure 7 shows the cumulative probability of developing AIDS (based on the joint analysis with one-month baseline hazard intervals) for three CD4 paths. The solid line is based on the estimated mean intercept and mean slope ($6{\cdot}64 - 0{\cdot}0067t$), the dashed line is based on the mean intercept and the 25th percentile of the estimated slopes ($6{\cdot}64 - 0{\cdot}0102t$), and the dotted line is based on the mean intercept and the 75th percentile of the estimated slopes ($6{\cdot}64 - 0{\cdot}0036t$). A patient with the average covariate history has approximately a 6 per cent chance of developing AIDS within 5·2 years of infection, whereas a patient with mean intercept and a slope of − 0·0102 has a 44 per cent chance of developing AIDS within 5·2 years. One could generate plots of this type for any CD4 path for use in prediction of time of AIDS development.

In general, the results for our CD4 tracking model are comparable with those of De Gruttola *et al.*[28] who modelled the square root of CD4 using a random effects model with measurement errors on data from a cohort of HIV infected and uninfected men. The data used in our analyses

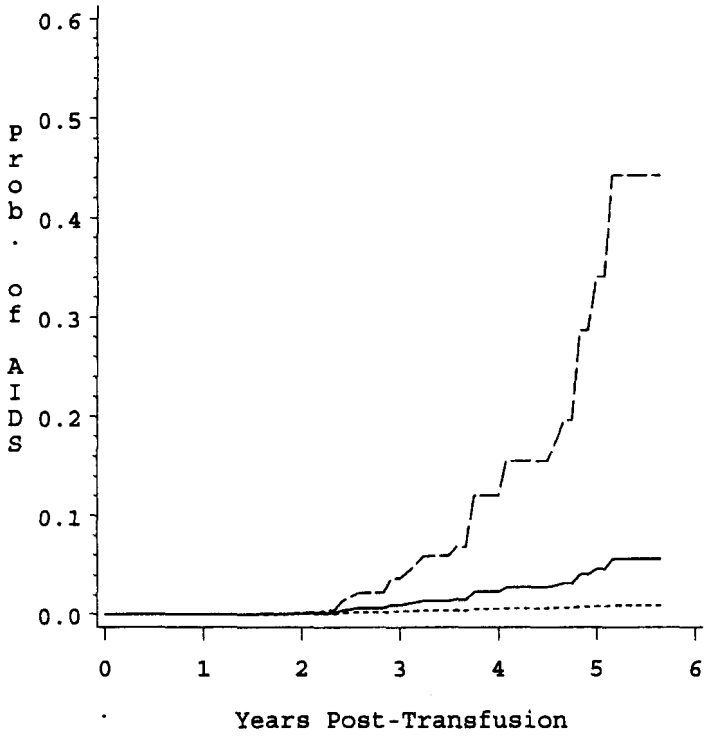

Figure 7. Cumulative probability of developing AIDS for three CD4 paths. The solid line is based on the estimated mean intercept and slope ($6{\cdot}64 - 0{\cdot}0067t$), the dashed line is based on the mean intercept and 25th percentile of estimated slopes ($6{\cdot}64 - 0{\cdot}0102t$), and the dotted line is based on the mean intercept and 75th percentile of estimated slopes ($6{\cdot}64 - 0{\cdot}0036t$)

differ in that all subjects are HIV positive transfusion recipients and we know the date of infection. We estimated the mean intercept from our joint analysis as $6{\cdot}64 + 18{\cdot}8 = 25{\cdot}44$ which is similar to their mean intercept estimate of 25·13 (based on infected men only). We estimated the square root of CD4 to decline by $0{\cdot}0067 \times 365{\cdot}25 = 2{\cdot}45$ per year which is steeper than the steepest estimated decline (2·17 from an analysis of both HIV infected and uninfected men) obtained from the four models they considered.

5. DISCUSSION

We have described an approach for modelling simultaneously repeated measurements of a time-dependent covariate and relating the covariate to disease risk. This approach is ideal when both covariate tracking and disease risk models are of interest as in the analysis of trends of immunologic markers over time and their effect on AIDS risk. Extensive simulation studies and a data application demonstrated that the combined analysis is a feasible approach. Assuming that we specified the underlying model correctly, the parameter estimates for the disease risk model improved considerably with incorporation of the covariate tracking model. This combined approach improved variance estimation and effectively reduced the attenuation of disease risk estimates obtained when one ignores covariate measurement error. Conversely, covariate tracking model parameter estimates improved with incorporation of the survival data. Use of the

survival information had the effect of minimizing the attentuation of the slope estimates caused by the effect of informative censoring of the repeated covariate measurements by the disease process.

Gibbs sampling was an effective approach for fitting the joint model. Other numerical methods would have been computationally intensive, perhaps impossible, and would have required complex numerical integrations and maximization of the likelihood over a large number of parameters. These methods are also unfeasible as one expands the submodels to incorporate additional covariates or makes changes to non-linear or non-normal forms. Gibbs sampling, although also computationally intensive, provides a feasible approach to the fit of this large model without making simplifying assumptions, and it is flexible to accommodate a variety of other expanded models. Even if we doubled the number of parameters (for example, adding a second covariate tracking model), the analysis of the data set would take approximately 34 hours on a 486/66MHz computer. We would shorten run times considerably if we ran the analysis on a faster computer such as a Pentium or SUN workstation. Gibbs sampling also allows incorporation of informative priors, if desired, for a Bayesian approach. We did not use informative priors in this analysis so that we could compare the results from the separate and joint analyses.

The methods of De Gruttola and Tu[10] and Tsiatis *et al.*[13] have the advantage that one can use existing software to fit the models, and, with Tsiatis' method, there is no need to model the baseline hazard since one uses Cox regression. Our method, however, provides estimates of variability that correctly reflect the uncertainty of all model parameters while still relaxing the restrictive assumption made by De Gruttola and Tu that the survival times are jointly normal, an assumption that may not be met in certain circumstances. We assumed a proportional hazards model for the hazard of developing disease with the parametric assumption that the baseline hazard is piecewise constant. This framework, however, does provide a general formulation of the baseline hazard in that we can model a large number of steps, and, in theory, we can make the step widths arbitrarily small to achieve a semi-parametric model.

One must interpret the relative risk e^{γ} as a comparison of AIDS risk based on subjects' 'true' unobserved CD4 values; one cannot interpret it as the risk of AIDS given, say, an observed CD4 of 200 versus an observed CD4 of 400. There are several instances, however, where estimation of this parameter is of ultimate interest. First, with underlying biological mechanisms examined and modelled, the relationship between disease risk and the 'true' value of the covariate is of interest. Second, many covariates, including CD4 are inherently highly variable and it may be more informative to estimate associations between AIDS risk and a subject's long-term average trend of CD4 rather than day-to-day specific values. Finally, even with observed CD4 measurements, we can apply this estimated relative risk in certain circumstances. For example, rather than applying the relative risk estimate to a single CD4 measurement, we could average several measurements at a single time, or we could fit a line to the subject's entire history of CD4 and obtain an estimated 'true' CD4.

The advantage of this combined approach (as in any statistical analysis) hinges on the assumption that we have specified correctly the underlying models and relationships. We have assumed a simple linear model for the log and square root of CD4 decline over time. Other investigators have found that more complicated models, including quadratic terms or random transition-points, did not substantially improve the fit of their datasets.[15,27] However, Taylor *et al.*[29] found a better fit with use of a random effects model for CD4 that incorporates random intercepts, a linear population decline, Brownian motion, and measurement error. In our model the person-specific slopes accounted for the variation in CD4 patterns over time, whereas they found that Brownian motion best accounted for variation over time. Munoz *et al.*[30] considered

a damped exponential correlation structure for repeated measurements of CD4 counts in HIV positive homosexual men. They found that this correlation structure differed significantly from the compound symmetry and AR(1) models. Clearly, more research is necessary to investigate these and other models for patterns of CD4 over time with use of the joint approach.

ACKNOWLEDGEMENTS

This work was supported by NIH grants CA 42949, N01-HB-4-7002, 4-7003, and 9-7074. The authors thank James Mosley, M.D., and the Transfusion Safety Study for providing the data set that motivated this research, and the reviewers for their many helpful comments.

REFERENCES

1. Laird, N. M. and Ware, J. H. 'Random-effects models for longitudinal data', *Biometrics*, **38**, 963–974 (1982).
2. Zeger, S. L. and Karim, M. R. 'Generalized linear models with random effects: a Gibbs sampling approach', *Journal of the American Statistical Association*, **86**, 79–86 (1991).
3. Wu, M. C. and Carroll, R. J. 'Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process', *Biometrics*, **44**, 175–188 (1988).
4. Wu, M. C. and Bailey, K. 'Analysing changes in the presence of informative right censoring caused by death and withdrawal', *Statistics in Medicine*, **7**, 337–346 (1988).
5. Cox, D. R. 'Regression models and life tables' (with discussion), *Journal of the Royal Statistical Society, Series B*, **34**, 187–220 (1972).
6. Prentice, R. L. 'Covariate measurement errors and parameter estimation in a failure time regression model', *Biometrika*, **69**, 331–342 (1982).
7. Geman, S. and Geman, D. 'Stochastic relaxation, gibbs distributions and the bayesian restoration of images', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741 (1984).
8. Gelfand, A. E. and Smith, A. F. M. 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association*, **85**, 398–409 (1990).
9. Clayton, D. G. 'Models for the analysis of cohort and case-control studies with inaccurately measured exposures', *in* Dwyer, J. H., Lippert, P., Feinleib, M. and Hoffmeister, H. (eds), *Statistical Models for Longitudinal Studies of Health*, Oxford University Press, New York, 1991.
10. De Gruttola, V. and Tu, X. M. 'Modelling progression of CD4-lymphocyte count and its relationship to survival time', *Biometrics*, **50**, 1003–1014 (1994).
11. De Gruttola, V. and Tu, X. M. 'Modeling the relationship between progression of CD4-lymphocyte count and survival time', *in* Jewell, N. P., Dietz, K. and Farewell, V. T. (eds), *AIDS Epidemiology: Methodologic Issues*, Birkhauser, Boston, 1992.
12. Tsiatis, A. A., De Gruttola, V. and Wulfsohn, M. S. 'Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS', *Journal of the American Statistical Association*, **90**, 27–37 (1995).
13. Tsiatis, A. A., De Gruttola, V., Strawderman, R. L. Dafni, U., Propert, K. J. and Wulfsohn, M. S. 'The relationship of CD4 counts over time to survival in patients with AIDS: Is CD4 a good surrogate marker?', *in* Jewell, N. P., Dietz, K. and Farewell, V. T. (eds), *AIDS Epidemiology: Methodologic Issues*, Birkhauser, Boston, 1992.
14. Self, S. and Pawitan, Y. 'Modeling a marker of disease progression and onset of disease', *in* Jewell, N. P., Dietz, K. and Farewell, V. T. (eds), *AIDS Epidemiology: Methodologic Issues*, Birkhauser, Boston, 1992.
15. Pawitan, Y. and Self, S. 'Modeling disease marker processes in AIDS', *Journal of the American Statistical Association*, **88**, 719–726 (1993).
16. Donegan, E., Stuart, M., Niland, J. C., Sacks, H. S., Azen, S. P., Dietrich, S. L., Faucett, C., Fletcher, M. A., Kleinman, S. H., Opersklski, E. A., Perkins, H. A., Pindyck, J., Schiff, E. R., Stites, D. P., Tomasulo, P. A., Mosley, J. W. and the Transfusion Safety Group. 'Infection with human immunodeficiency virus type 1 (hiv) among recipients of antibody-positive blood donations', *Annals of Internal Medicine*, **113**, 733–739 (1990).
17. Gelman, A. and Rubin, D. B. 'Inference from iterative simulation using multiple sequences', *Statistical Science*, **7**, 457–472 (1992).
18. Geyer, C. J. 'Practical Markov chain Monte Carlo', *Statistical Science*, **7**, 473–511 (1992).

19. Box, G. E. P. and Tiao, G. C. *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, MA, 1973.
20. Gilks, W. R. 'Derivative-free adaptive rejection sampling for Gibbs sampling', *Bayesian Statistics*, **4**, 1–9 (1992).
21. Odell, P. L. and Feiveson, A. H. 'A numerical procedure to generate a sample covariance matrix', *Journal of the American Statistical Association*, **61**, 198–203 (1966).
22. Ripley, B. *Stochastic Simulation*, Wiley, New York, 1987.
23. Faucett, C. 'Simultaneously modeling censored survival data and repeatedly measured covariates using gibbs sampling', Dissertation, University of Southern California, Department of Preventive Medicine, Division of Biometry, 1992.
24. Laird, N., Lange, N. and Stram, D. 'Maximum likelihood computations with repeated measures: application of the EM algorithm', *Journal of the American Statistical Association*, **82**, 97–105 (1987).
25. Kleinman, S. H., Niland, J. C., Azen, S. P., Operskalski, E. A., Barbosa, L. H., Chernoff, A. I., Edwards, V. M., Lenes, B. A., Marshall, G. J., Nemo, G. J., Norman, G. L., Perkins, H., Pindyck, J., Pitlick, F., Rasheed, S., Shriver, K., Toy, P., Tomasulo, P. A., Waldman, A., Mosley, J. W., and the Transfusion Safety Study Group. 'Prevalence of antibodies to human immunodeficiency virus type 1 among blood donors prior to screening: the transfusion safety study/NHLBI donor repository', *Transfusion*, **29**, 572–580 (1989).
26. Parker, J. W., Adelsberg, B., Azen, S. P., Boone, D., Fletcher, M. A., Gjerset, G. F., Hassett, J., Kaplan, J., Niland, J. C., Odom-Maryon, T., Operskalski, E. A., Prince, H., Scott, D., Stites, D. P., Mosley, J. W. and the Transfusion Safety Study Group. 'Leukocyte immunophenotyping by flow cytometry in a multisite study: standardization, quality control, and normal values in the transfusion safety study', *Clinical Immunology and Immunopathology*, **55**, 187–220 (1990).
27. Lange, N., Carlin, B. P. and Gelfand, A. E. 'Hierarchical Bayes models for the progression of HIV infection using longitudinal CD4 T-cell numbers', *Journal of the American Statistical Association*, **87**, 615–626 (1992).
28. De Gruttola, V., Lange, N. and Dafni, U. 'Modeling the progression of HIV infection', *Journal of the American Statistical Association*, **86**, 569–577 (1991).
29. Taylor, J. M. G., Cumberland, W. G. and Sy, J. P. 'A stochastic model for analysis of longitudinal AIDS data', *Journal of the American Statistical Association*, **89**, 727–736 (1994).
30. Munoz, A., Carey, V., Schouten, J. P., Segal, M. and Rosner, B. 'A parametric family of correlation structures for the analysis of longitudinal data', *Biometrics*, **48**, 733–742 (1992).