RESEARCH ARTICLE

WILEY Statistics in Medicine

Estimating the association between blood pressure variability and cardiovascular disease: An application using the ARIC Study

Jessica K. Barrett[1,2] | Raphael Huille[2,3] | Richard Parker[4] | Yuichiro Yano[5] | Michael Griswold[6]

[1]MRC Biostatistics Unit, University of Cambridge, Cambridge, UK

[2]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

[3]École Nationale de la Statistique et de l'Administration Économique, Malakoff, France

[4]School of Social and Community Medicine, University of Bristol, Bristol, UK

[5]Department of Preventive Medicine, University of Mississippi Medical Center, Jackson, Mississippi

[6]Center of Biostatistics and Bioinformatics, University of Mississippi Medical Center, Jackson, Mississippi

Correspondence
Jessica K. Barrett, MRC Biostatistics Unit, University of Cambridge, Cambridge CB2 0SR, UK.
Email: jessica.barrett@mrc-bsu.cam.ac.uk

The association between visit-to-visit systolic blood pressure variability and cardiovascular events has recently received a lot of attention in the cardiovascular literature. But, blood pressure variability is usually estimated on a person-by-person basis and is therefore subject to considerable measurement error. We demonstrate that hazard ratios estimated using this approach are subject to bias due to regression dilution, and we propose alternative methods to reduce this bias: a two-stage method and a joint model. For the two-stage method, in stage one, repeated measurements are modelled using a mixed effects model with a random component on the residual standard deviation (SD). The mixed effects model is used to estimate the blood pressure SD for each individual, which, in stage two, is used as a covariate in a time-to-event model. For the joint model, the mixed effects submodel and time-to-event submodel are fitted simultaneously using shared random effects. We illustrate the methods using data from the Atherosclerosis Risk in Communities study.

KEYWORDS

cardiovascular disease, heteroscedasticity, joint model, mixed effects model, repeated measurements, survival analysis

1 | INTRODUCTION

Systolic blood pressure (SBP) is universally recognised as an important risk factor for cardiovascular disease (CVD) and is routinely included in risk scores for CVD risk prediction.[1,2] The prognostic value of SBP is primarily based on the mean of measurements over multiple outpatient visits, whereas substantially less attention has been given to the variability of SBP across visits (ie, visit-to-visit SBP variability). However, there is increasing evidence of associations between greater visit-to-visit SBP variability and CVD outcomes in community-based and clinical settings.[3-6] Therefore, in addition to assessing mean clinic SBP levels over time, measurements of visit-to-visit SBP variability may improve the accuracy of CVD risk prediction, which is crucial for the optimisation of patient care.

The standard deviation (SD), coefficient of variation (CV), average real variability (ARV), and SBP variability independent of the mean across multiple visits have been widely used to quantify visit-to-visit SBP variability.[3-7] However, these measures are estimated on a person-by-person basis and are therefore subject to considerable measurement error. This measurement error causes regression dilution bias in the estimated association between visit-to-visit SBP variability and CVD.[8]

Mixed effects models have been proposed, which allow the within-individual variability to differ between individuals. For example, Hedeker et al introduced the mixed location scale model, where the within-individual variances may be assumed to depend on an additional random effect, as well as time-constant and time-varying covariates.[9,10] More recently, Goldstein et al have explored the mixed location scale model using a Bayesian framework in the context of growth curve models.[11]

We consider extensions of the mixed location scale model to estimate the association of SBP variability with CVD. Repeated SBP measurements are modelled using a mixed effects model with a random component on the within-individual SD, thus allowing the borrowing of information across individuals. The model allows for each individual to have different within-individual variation, which describes the variability in their SBP measurements. We propose two methods for estimating the association of this SBP variability with CVD. The first is a two-stage method, where, in stage one, we obtain estimates of the usual level of SBP and the SBP variability for each individual from the mixed effects model, and in stage two, these are included as covariates in a survival model for the time to the first CVD event. The second method is a joint model where the repeated SBP measurements and the time to first CVD event are modelled simultaneously using shared random effects.[12]

There are advantages and disadvantages in the use of two-stage versus joint models. The heavy computational burden associated with simultaneous analysis of repeated measurements and time-to-event data makes the two-stage method a more feasible option for large datasets. The two-stage method could also be readily incorporated into a landmarking approach for dynamic risk prediction,[13] for example, in the dynamic model for CVD risk prediction recently proposed by Paige et al.[14] However, two-stage methods for the analysis of repeated measurements and time-to-event data have been shown to lead to bias in estimated covariate effects compared to joint models.[15] This bias is due to two factors: (1) regression dilution bias caused by residual measurement error in the estimates of random effects from the mixed effects model and (2) bias caused by informative truncation of the repeated measurements by the event of interest. In the two-stage approach, this second source of bias can be removed by separating longitudinal follow-up from survival follow-up if data allows, as we describe in Section 2.

The outline of this paper is as follows. In Section 2, we describe the methods, including the naive method that has predominantly been used in the CVD literature, and our proposed two-stage and joint model methods. In Section 3, we compare the methods using simulation studies and explore the accuracy of all methods in a variety of scenarios. We apply the two-stage and joint models to data from the Atherosclerosis Risk in Communities (ARIC) Study in Section 4, and we present a final discussion in Section 5.

2 | METHODS

The association of SBP with CVD has previously been investigated using mixed effects models to analyse repeated measurements of SBP.[16-18] In this section, we first present the naive methods for estimating SBP variability that have been commonly used in the literature, and then, we explore how the mixed effects model can be extended to allow SBP variability to differ between individuals. We propose a two-stage method and a joint modelling method for estimating the association between SBP variability and CVD events. Although we focus here on the association between SBP variability and CVD, the methods and results presented would be relevant for any application where the association between the

variability in a longitudinal outcome and a time-to-event is of interest. We consider estimating the association of the variability of a longitudinal outcome with a time-to-event, adjusted for the "usual level" of the longitudinal outcome. This "usual level" could be the average value, which may be assumed constant in time, or the current value or baseline value if time-dependence is to be taken into account.

2.1 | Notation

Given a dataset with N individuals and n_i SBP measurements from the ith individual, $i = 1, \dots, N$, let $Y_{ij}, j = 1, \dots, n_i$, be the jth measurement of individual i taken at measurement time t_{ij}. Each individual is followed up from baseline until time $T_i^* = \min(T_i, C_i)$, where T_i is the time of the event for individual i and C_i is the censoring time.

2.2 | Estimating longitudinal variability

2.2.1 | Naive method

In the cardiovascular literature, the usual level of SBP is typically estimated as the mean of an individual's repeated measurements

$$\hat{BP}_i^{naive} = \frac{\sum_j Y_{ij}}{n_i}$$

and the SBP variability as their SD

$$\widehat{BPSD}_i^{naive} = \sqrt{\frac{\sum_j \left(\hat{BP}_i^{naive} - Y_{ij} \right)^2}{n_i}}.$$

The repeated measurements may have been taken prior to follow-up of the time-to-event, or concurrent with the time-to-event follow-up, in which case the repeated measurement process is terminated by an event. Other measures of visit-to-visit SBP variability have been proposed, such as the CV (defined as the SD/mean), the maximum and minimum difference, and the ARV (defined as the average change between successive visits),[7] but all are usually estimated using within-individual information only.

2.2.2 | Mixed effects models

An alternative method to estimate the usual level and variability for each individual is to use a mixed effects model for the repeated measurements from all individuals, allowing information to be borrowed across individuals. The mixed effects model also allows longitudinal trajectories to be modelled by including time-dependent terms in the model.

Consider a standard linear mixed effects model

$$Y_{ij} = \beta^T X_{ij} + b_i^T Z_{ij} + \epsilon_{ij}, \tag{1}$$

where X_{ij} is a covariate vector for the fixed effects β and Z_{ij} is a covariate vector for the random effects b_i, assumed normally distributed $b_i \sim N(0, \Sigma_b)$. The residual errors ϵ_{ij} are assumed independent and normally distributed, $\epsilon_{ij} \sim N(0, \sigma^2)$. We can allow variability in the repeated measurements to differ between individuals by replacing the residual SD σ with an individual-specific residual SD σ_i and assuming that the σ_i are randomly distributed. Note that, in this model, we do not distinguish between true variability in an individual's repeated measurements and measurement error. We assume a log-normal distribution for the residual SD distribution, ensuring positivity of the SDs, $\sigma_i \sim \text{logN}(\mu_\sigma, \tau_\sigma^2)$. The choice of log-normal distribution also allows a natural extension of the model to incorporate correlation between the usual level and the residual SD by assuming a multivariate normal distribution for the random effects and log residual SD

$$\epsilon_{ij} \sim N\left(0, \sigma_i^2\right), \quad \begin{pmatrix} b_i \\ \log \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \Sigma_b & \Sigma_{b\sigma} \\ \Sigma_{b\sigma}^T & \tau_\sigma^2 \end{pmatrix} \right), \tag{2}$$

where $\Sigma_{b\sigma}$ is a vector of covariances between the random effects and the random residual errors. Alternative distributions have been proposed for the residual SD, such as the half-Cauchy distribution.[19] (See also the work of Hedeker et al[9] and recent work of Goldstein et al,[11] where a log-normal distribution was assumed for the residual variances σ_i^2.)

Specific examples of the model specified by Equations (1) and (2) that we will consider are a random-intercept model

$$Y_{ij} = \beta^T X_i + b_{0i} + \epsilon_{ij}, \tag{3}$$

where the random intercept b_{0i} can be interpreted as the usual level of the repeated measurements, and adjustment is made for baseline covariates X_i. We consider a model (LMM1) that does not account for the correlation between $b_0 i$ and σ_i

$$\text{LMM1:} \quad \begin{pmatrix} b_{0i} \\ \log \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \tau_0^2 & 0 \\ 0 & \tau_\sigma^2 \end{pmatrix} \right), \tag{4}$$

and a model (LMM2) that does account for the correlation

$$\text{LMM2:} \quad \begin{pmatrix} b_{0i} \\ \log \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \rho \tau_0 \tau_\sigma \\ \rho \tau_0 \tau_\sigma & \tau_\sigma^2 \end{pmatrix} \right). \tag{5}$$

We also consider a random intercept and slope model

$$Y_{ij} = X_i \beta + \beta_t t_{ij} + b_{0i} + b_{1i} t_{ij} + \epsilon_{ij},$$

where the random intercept b_{0i} is now the baseline value (minus the population average), and σ_i now measures the variability around the individual's average trajectory. Again, we can either ignore or allow for correlations between the random effects and the residual SD's

$$\text{LMM3:} \begin{pmatrix} b_{0i} \\ b_{1i} \\ \log \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \rho_{01} \tau_0 \tau_1 & 0 \\ \rho_{01} \tau_0 \tau_1 & \tau_1^2 & 0 \\ 0 & 0 & \tau_\sigma^2 \end{pmatrix} \right), \quad \text{LMM4:} \begin{pmatrix} b_{0i} \\ b_{1i} \\ \log \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \rho_{01} \tau_0 \tau_1 & \rho_{0\sigma} \tau_0 \tau_\sigma \\ \rho_{01} \tau_0 \tau_1 & \tau_1^2 & \rho_{1\sigma} \tau_1 \tau_\sigma \\ \rho_{0\sigma} \tau_0 \tau_\sigma & \rho_{1\sigma} \tau_1 \tau_\sigma & \tau_\sigma^2 \end{pmatrix} \right).$$

2.3 | Estimating the association between SBP variability and CVD

We consider two approaches for estimating the association between SBP variability and the time to the first CVD event, ie, a two-stage approach and a joint model.

2.3.1 | Two-stage approach

We fit the models in two stages. In stage one, we either (i) estimate the usual level and SD of the repeated measurements for each individual using the naive method or (ii) fit a mixed effects model and estimate the usual level, residual SD, and other random effects for each individual, as described in Section 2.2. In the second stage, we use the estimated usual level and residual SD as covariates in a standard proportional hazards Cox regression for the time to the first event

$$h_i(t) = h_0(t) \exp(\alpha^T U_i + \gamma^T W_i),$$

where U_i is a vector containing the estimated usual level, residual SD and other random effects for each individual, α is a vector of parameters describing the association between the mixed effects model and the time-to-event model and W_i is a vector of baseline covariates.

We estimate mixed effects model parameters using Bayesian Markov chain Monte Carlo (MCMC) implemented using JAGS Version 3.4.0[20] and the R2jags package in R.[21] We assume diffuse priors for all model parameters. The usual level $\widehat{BP}_i^{LMM}$ is estimated by the posterior mean of the random intercept b_{0i} and the variability $\widehat{BPSD}_i^{LMM}$ by the posterior mean of the residual SD σ_i. Additional random effects can also be estimated by posterior means.

The SBP measurement process is truncated informatively by the event of interest, as illustrated in Figure 1A. In order to avoid this informative truncation, we can divide all follow-up into follow-up of the repeated measurements and follow-up of the survival outcome, separated at separation time t_{sep}, as shown in Figure 1B. Only repeated measurements taken before t_{sep} are used in estimation of the mixed effects model, and only events taking place after the separation time are used in estimation of the time-to-event model conditional on surviving until t_{sep}. Separating the longitudinal follow-up from the survival follow-up in this way involves discarding information from repeated measurements following t_{sep} and individuals experiencing events before t_{sep}. In practice, t_{sep} should be chosen to minimise this information loss.

2.3.2 | Joint model

In the joint model, we use a shared parameter approach to link the mixed effects and time-to-event models. The mixed effect submodel is defined by Equations (1) and (2). The time-to-event submodel is defined by

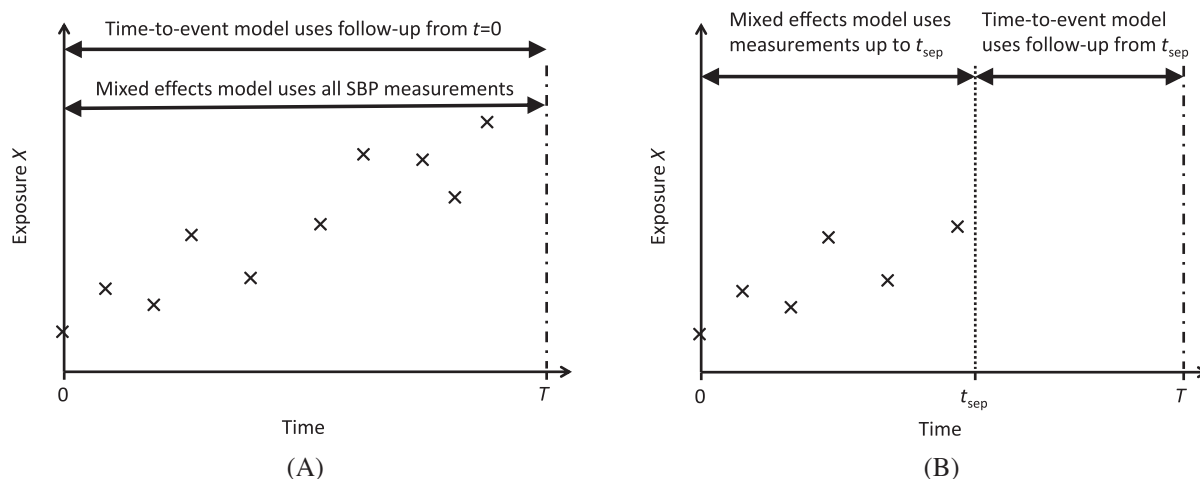$$h_i(t) = h_0(t) \exp(\alpha^T b_i + \gamma^T W_i),$$

FIGURE 1 Diagrams illustrating the use of systolic blood pressure (SBP) measurements and time-to-event data when (A) time-to-event follow-up and SBP measurement follow-up run concurrently from time $t = 0$ and (B) repeated measurement follow-up takes place prior to time-to-event follow-up

where b_i are the random effects shared with the mixed effects submodel. We assume a piecewise constant baseline hazard; we choose cut-points t_k, $0 \leq k \leq K$, at the K-quantiles of the observed event times and assume that the baseline hazard is constant in the time intervals between cut-points, $h_0(t) = h_{0,k}$, $t_{k-1} < t \leq t_k$. The piecewise-constant assumption ensures that the hazard function can be integrated analytically, thus avoiding the need for numerical integration when evaluating the likelihood function. In practice, the number of cut-points can be chosen by assessing model fit, for example, using the deviance information criterion.[22]

We again estimate model parameters using Bayesian MCMC implemented using JAGS Version 3.4.0[20] and the `R2jags` package in R[21] and assuming diffuse priors for all model parameters.

2.4 | Regression dilution bias in association parameters

In the two-stage approach, because SBP variability is estimated with error, we expect bias in the estimated association with CVD due to regression dilution. We aim to reduce regression dilution bias by using a mixed effects model that allows information to be borrowed across individuals. The joint model accounts for this measurement error by including the underlying random effects in the time-to-event model, rather than estimates of those random effects. However, misspecification of the mixed or joint models could introduce other sources of bias. In Section 1 of the Supplementary Materials, we show for the simplified case of linear regression that ignoring the correlation between the random effects and the residual SD could also lead to bias in the estimated effects of both the usual level and the variability on the outcome.

3 | SIMULATION STUDIES

We use simulation studies to explore the bias in parameter estimates due to imprecise estimation of SBP variability and due to ignoring correlation between SBP and SBP variability in the mixed effects model. We consider two scenarios. In Scenario 1, each individual has the same number of repeated measurements, thus bias in the two-stage approach arises solely from measurement error in the estimates of the usual level and variability of SBP (cf Figure 1B). In Scenario 2, the repeated measurement process is truncated by the event times, so that bias in the two-stage approach arises from both measurement error and informative truncation of the repeated measurements (cf Figure 1A). We generated simulated datasets to explore the effect of the number of measurements per individual, the value of the association parameters and the extent of the correlation between the usual level and the variability.

We take a simple model for the repeated measurements with a random intercept only and random residual error SD

$$BP_{ij} = b_{0i} + \epsilon_{ij},$$

where the residual error $\epsilon_{ij} \sim N(0, \sigma_i^2)$, and we take a bivariate normal distribution for b_{0i} and $\log \sigma_i$

$$\begin{pmatrix} b_{0i} \\ \log \sigma_i \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \rho\tau_0\tau_\sigma \\ \rho\tau_0\tau_\sigma & \tau_\sigma^2 \end{pmatrix} \right),$$

Here, b_{0i} is the usual level and σ_i represents the variability in the longitudinal outcome. Based on results from the ARIC study, we take $\mu_0 = 120$, $\tau_0 = 15$, $\mu_\sigma = 2$, and $\tau_\sigma = 0.5$. Event times are drawn from a Weibull distribution with shape parameter k and scale parameter λ. A log-linear effect of the covariates on λ gives a proportional hazards model

$$\lambda = \exp(\gamma_0 + \alpha_0 b_{0i} + \alpha_\sigma \sigma_i).$$

We introduce administrative censoring at 20 years. To maintain an event rate of 20% before censoring, we take $\gamma_0 = -10.26$ and $k = 2$ with default association parameters $\alpha_0 = 0.02$ and $\alpha_\sigma = 0.05$, similar to those found for the ARIC dataset. For Scenario 1, we used all repeated measurements for all individuals. For Scenario 2, we assumed repeated measurements were taken equidistantly between baseline and 18 years follow-up and discarded all repeated measurements following event times. In each simulation setup, we generated 1000 datasets, each consisting of 1500 individuals.

We analysed each generated dataset using the two-stage method from Section 2.3.1, with (i) naive and (ii) mixed model estimates of SBP usual levels and variabilities, and the joint models from Section 2.3.2. For the mixed effects and joint models, we fitted models ignoring the correlation between SBP and SBP variability (mixed effects model LMM1 and joint model JM1 used Equations (3) and (4)), and allowing for the correlation (mixed effects model LMM2 and joint model JM2 used Equations (3) and (5)). The joint models were fitted using 15 time intervals for the baseline hazard. Model convergence was checked using the Gelman-Rubin statistic as modified by Brooks and Gelman.[23] We also fitted the survival models using the true values of b_{0i} and σ_i as covariates, representing the best achievable results from the time-to-event model when all covariates are known precisely. For all scenarios, we calculated the mean and SD of the estimated log hazard ratios and the root mean squared errors (RMSEs) and coverage probabilities at the 95% level.

Table 1 shows results for $n = 4$, $n = 7$, and $n = 10$ measurements per individual for Scenario 1. With 4 measurements per individual, using the naive method leads to a slight bias in the log hazard ratio (logHR) for the usual level, but substantial negative bias and associated loss of coverage in the logHR for variability, which is most likely caused by

TABLE 1 Scenario 1 results with different numbers of repeated measurements. Presented are the true values, mean (SD) of estimated log hazard ratios, RMSE, and coverage over 1000 simulated datasets. Methods of analysis are (1) True values, where the true usual levels and variabilities are used as covariates, (2) Naive method, (3) LMM1, the two-stage approach with no correlation between the usual level and the variability, (4) LMM2, the two-stage approach with correlation, (5) JM1, the joint model with no correlation, and (6) JM2, the joint model with correlation

| | **Usual Level logHR α_0** | | | | **Variability logHR α_σ** | | | |
	True	**Mean (SD)**	**RMSE**	**Coverage**	**True**	**Mean (SD)**	**RMSE**	**Coverage**
$n = 4$								
True values	0.02	0.02 (0.0043)	0.0043	0.94	0.05	0.0501 (0.0117)	0.0117	0.94
Naive	0.02	0.0208 (0.0037)	0.0038	0.947	0.05	0.0284 (0.0094)	0.0235	0.31
LMM1	0.02	0.0227 (0.0042)	0.005	0.892	0.05	0.0531 (0.0151)	0.0154	0.929
LMM2	0.02	0.0198 (0.0049)	0.0049	0.948	0.05	0.0511 (0.0175)	0.0175	0.931
JM1	0.02	0.023 (0.0042)	0.0052	0.883	0.05	0.0521 (0.0155)	0.0156	0.939
JM2	0.02	0.02 (0.0051)	0.0051	0.949	0.05	0.0502 (0.0183)	0.0183	0.936
$n = 7$								
True values	0.02	0.0202 (0.0042)	0.0042	0.946	0.05	0.0496 (0.0113)	0.0113	0.951
Naive	0.02	0.0209 (0.0039)	0.004	0.943	0.05	0.0354 (0.01)	0.0177	0.684
LMM1	0.02	0.022 (0.0042)	0.0047	0.913	0.05	0.0519 (0.0133)	0.0134	0.943
LMM2	0.02	0.0202 (0.0046)	0.0046	0.95	0.05	0.0502 (0.0146)	0.0146	0.944
JM1	0.02	0.0221 (0.0042)	0.0048	0.91	0.05	0.0509 (0.0135)	0.0136	0.952
JM2	0.02	0.0203 (0.0047)	0.0047	0.945	0.05	0.0493 (0.015)	0.015	0.946
$n = 10$								
True values	0.02	0.0199 (0.0041)	0.0041	0.953	0.05	0.05 (0.0111)	0.0111	0.956
Naive	0.02	0.0206 (0.004)	0.004	0.948	0.05	0.039 (0.0102)	0.015	0.818
LMM1	0.02	0.0212 (0.0042)	0.0044	0.938	0.05	0.0516 (0.0124)	0.0125	0.95
LMM2	0.02	0.0199 (0.0045)	0.0045	0.944	0.05	0.0501 (0.0134)	0.0134	0.95
JM1	0.02	0.0213 (0.0042)	0.0044	0.931	0.05	0.0508 (0.0125)	0.0125	0.956
JM2	0.02	0.02 (0.0045)	0.0045	0.943	0.05	0.0494 (0.0136)	0.0136	0.946

Abbreviations: logHR, log hazard ratio; RMSE, root mean squared error; SD, standard deviation.

regression dilution due to imprecise measurement of longitudinal variability. Using the mixed effects model LMM1 leads to a positive bias in both the usual level logHR and the variability logHR, as we would expect from the arguments given in Section 1 of the Supplementary Materials, because correlation between these quantities was not accounted for in their estimation. Using method LMM2, which accounts for the correlation, results in minimal bias, even with only 4 measurements per individual. For both two-stage models, the coverage is a little below the nominal level, but the RMSE is larger for the variability logHR in LMM2 due to greater variation in effect estimates. Results for the joint models are similar to the two-stage models in terms of the bias, but coverage is closer to the nominal levels. Results for $n = 7$ and $n = 10$ are similar, with the naive model giving reduced but still substantial bias in the variability logHR with a greater number of measurements.

Table 2 shows results for $n = 4$, $n = 7$, and $n = 10$ measurements per individual for Scenario 2. The two-stage model LMM2 now also gives negative bias in estimating the variability logHR due to informative truncation of the repeated measurements by the CVD event. The joint model JM2, which accounts for this informative truncation, gives consistent results. Our results suggest that the bias incurred in the two-stage approach by informative truncation is more considerable here than the bias incurred by error in the estimates of SBP variability from the linear mixed models (LMMs).

In Section 2 of the Supplementary Materials, we present further results investigating performance of the models for different levels of association between the longitudinal trajectories and the CVD event (Supplementary Table 1) and for different levels of correlation between the usual level and the variability (Supplementary Table 2). All results are given for datasets with $n = 4$ measurements. In brief, all methods performed well with no association between times-to-events and longitudinal variability. But, when the associations between the time-to-event and the usual level and variability of the longitudinal trajectories was substantially stronger, all the methods struggled to give consistent results with only 4 repeated measurements. The joint models performed the best in this case with the least bias and highest coverage (Supplementary Table 1). When there is negative correlation between usual levels and variabilities, the direction of bias is reversed. Bias in the usual level logHR using methods LMM1 and JM1 increases with increasing ρ. But, for the variability logHR, the pattern is less clear because of the interplay between regression dilution bias and the bias incurred by ignoring correlation in the mixed effects model (Supplementary Table 2).

TABLE 2 Scenario 2 results with different numbers of repeated measurements. Presented are the true values, mean (SD) of estimated log hazard ratios, RMSE, and coverage over 1000 simulated datasets. Methods of analysis are (1) True values, where the true usual levels and variabilities are used as covariates, (2) Naive method, (3) LMM1, the two-stage approach with no correlation between the usual level and the variability, (4) LMM2, the two-stage approach with correlation, (5) JM1, the joint model with no correlation, and (6) JM2, the joint model with correlation

| | Usual Level logHR α_0 | | | | Variability logHR α_σ | | | |
	True	Mean (SD)	RMSE	Coverage	True	Mean (SD)	RMSE	Coverage
$n = 4$								
True values	0.02	0.02 (0.0042)	0.0042	0.938	0.05	0.0496 (0.0115)	0.0115	0.942
Naive	0.02	0.0229 (0.0041)	0.005	0.86	0.05	0.0118 (0.0122)	0.0401	0.036
LMM1	0.02	0.0229 (0.0042)	0.0051	0.892	0.05	0.0348 (0.0153)	0.0216	0.868
LMM2	0.02	0.0203 (0.0047)	0.0047	0.952	0.05	0.0385 (0.0166)	0.0202	0.934
JM1	0.02	0.0236 (0.0047)	0.0059	0.852	0.05	0.0409 (0.0211)	0.0229	0.908
JM2	0.02	0.0199 (0.0055)	0.0055	0.94	0.05	0.0493 (0.022)	0.022	0.944
$n = 7$								
True values	0.02	0.0203 (0.004)	0.0041	0.96	0.05	0.0498 (0.011)	0.011	0.964
Naive	0.02	0.0225 (0.0039)	0.0047	0.89	0.05	0.0242 (0.011)	0.028	0.28
LMM1	0.02	0.0221 (0.004)	0.0045	0.92	0.05	0.0418 (0.0126)	0.015	0.928
LMM2	0.02	0.0205 (0.0042)	0.0042	0.96	0.05	0.0437 (0.0133)	0.0147	0.96
JM1	0.02	0.0223 (0.0043)	0.0049	0.904	0.05	0.0466 (0.0153)	0.0157	0.952
JM2	0.02	0.0202 (0.0046)	0.0046	0.952	0.05	0.0502 (0.0156)	0.0156	0.968
$n = 10$								
True values	0.02	0.0203 (0.0041)	0.0041	0.954	0.05	0.0506 (0.012)	0.012	0.938
Naive	0.02	0.0222 (0.004)	0.0045	0.902	0.05	0.031 (0.0121)	0.0226	0.544
LMM1	0.02	0.0216 (0.004)	0.0043	0.93	0.05	0.0449 (0.0136)	0.0145	0.918
LMM2	0.02	0.0205 (0.0042)	0.0042	0.958	0.05	0.0458 (0.0139)	0.0145	0.94
JM1	0.02	0.0216 (0.0043)	0.0045	0.926	0.05	0.0487 (0.0155)	0.0155	0.918
JM2	0.02	0.0202 (0.0045)	0.0045	0.956	0.05	0.0506 (0.0159)	0.0159	0.932

Abbreviations: logHR, log hazard ratio; RMSE, root mean squared error; SD, standard deviation.

In summary, the method that overall resulted in the least bias was JM2. For parameter values similar to those observed in the data example, method LMM2 performed well. But, for all methods, biases were observed with strong association parameters with $n = 4$ measurements per individual, so more measurements per individual would be required in this scenario. The interplay between biases due to multiple causes is complex and depends on the correlation parameter ρ. The RMSE, however, is generally slightly higher for LMM2 than for LMM1 and for JM2 than for JM1, suggesting a trade-off between the bias caused by ignoring the correlation for models LMM1 and JM1 and the higher variance induced by the increased complexity of models LMM2 and JM2.

4 | EXAMPLE: ARIC STUDY

We illustrate our methods using data from the ARIC study.[24] Briefly, 15 792 mostly black and white adults aged 45-64 years were enrolled into the ARIC study between 1987 and 1989 via probability sampling from 4 US communities: Washington County, Maryland; Forsyth County, North Carolina; Minneapolis, Minnesota, suburbs; and Jackson, Mississippi. Participants underwent five examinations during 25 years of follow-up (ie, Visit 1, 2, 3, 4, and 5 examinations), with an annual contact by telephone. In the current analysis, we used SBP measurements from Visit 1 (1987-1989) through Visit 4 (1996-1998).

We analysed both the full ARIC dataset and a reduced dataset where longitudinal follow-up was truncated at the last SBP measurement taken at Visit 4 (at $t_{sep} = 11.9$ years from baseline) and the time origin for the survival follow-up was taken at the same time-point (Figure 1B). For the reduced dataset, to have enough measurements to estimate the SD of repeated measurements using the naive method, we restricted to individuals with three or four nonmissing measurements recorded and also to individuals who did not experience an event before t_{sep}. In total, our full dataset consisted of 13 161 individuals and our reduced dataset consisted of 10 019 individuals. Supplementary Table 3 shows the baseline characteristics of the reduced and full datasets.

We analysed both the full and reduced datasets using two-stage and joint models. For the two-stage approaches, we used the repeated measurements of SBP to estimate the usual SBP level and the SBP variability using the naive method and the linear mixed effects models without and with accounting for the correlation between SBP and SBP variability (we again label these models LMM1 and LMM2, respectively). We also fitted joint models both without and with correlation between SBP and SBP variability (again JM1 and JM2, respectively). We applied the mixed effects models/submodels in four ways.

1. With a random intercept only, not adjusting for baseline CVD risk factors.
2. Including a slope term in the model as both a fixed and random effect. The covariance matrix between the random effects and the residual log-standard deviation was taken to be unstructured, allowing for correlations between the random slopes and the random intercept and between the random slopes and the SBP variability.
3. Adjusting the random intercept mixed effects model for baseline CVD risk factors (age, sex, diabetes status, smoking status, baseline total cholesterol, and baseline HDL cholesterol).
4. Adjusting the random intercept and slope mixed effects model for baseline CVD risk factors.

Time-to-event models were adjusted for age, diabetes status, smoking status, baseline total cholesterol, baseline HDL cholesterol, and sex.

For the Bayesian estimation, we used diffuse uniform prior distributions $U[0, 100]$ for SDs, uniform $U[-1, 1]$ prior distributions for correlation parameters, and diffuse normal prior distributions $N(0, 100^2)$ for all other parameters. Priors were specified for the bivariate and trivariate normal distributions by expressing them as two and three conditional univariate normal distributions, respectively. We used a burn-in of 1000 MCMC updates for the mixed effects models and 2000 MCMC updates for the joint models. We calculated results from 1000 sampled updates and checked convergence using the Gelman-Rubin statistic as modified by Brooks and Gelman.[23] For the joint models, there was a high degree of autocorrelation for the usual SBP and SBP variability hazard ratios, so MCMC chains were thinned by 4 updates.

Figure 2 shows histograms of the estimated SBP usual level and SBP variability using the naive method. Also plotted are fitted normal and log-normal distributions, respectively. The plots suggest that the assumption of a normal distribution for the random intercepts and a log-normal distribution for the residual SDs is appropriate. The naive usual level estimates are plotted against the naive variability estimates in Figure 3. There is a positive correlation with those with higher SBP tending to have greater SBP variability. The correlation between the naive estimates is 0.42, but the true correlation without measurement error is likely to be higher.
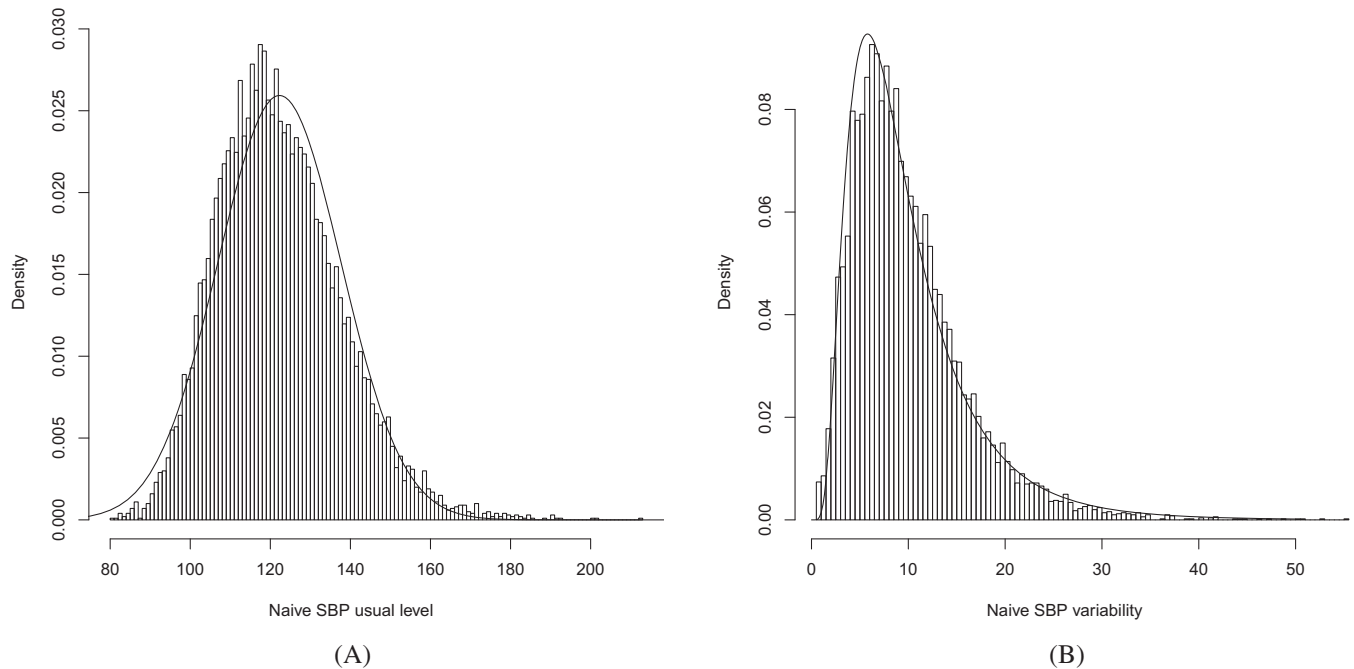
FIGURE 2 Histograms of the distribution of (A) naive usual systolic blood pressure (SBP) estimates and (B) naive SBP variability estimates in data from the ARIC study. Also plotted are probability density functions of (A) a normal distribution fitted to the data and (B) a log-normal distribution fitted to the data
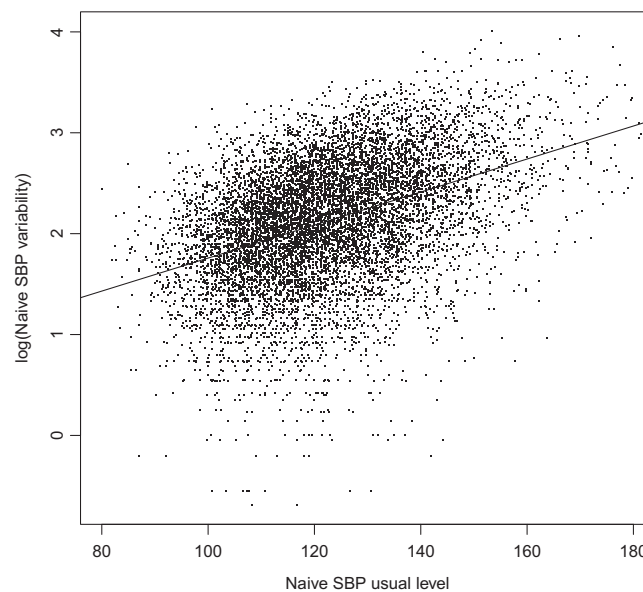

FIGURE 3 Scatter plot of log naive systolic blood pressure (SBP) variability estimates against naive SBP usual level estimates with a fitted regression line

Estimates of usual levels and variabilities from the mixed effects models are compared with the naive estimates in Figure 4. The plotted lines are the lines of agreement between the two methods. Note that there is more difference between the methods for SBP variability estimates than for usual level estimates, with lower SBP variabilities being underestimated by the naive method and higher variabilities being overestimated.

Results from all models applied to the reduced ARIC dataset are shown in Table 3. As expected, the joint models took considerably longer to run than the two-stage models with runtimes of between 17 and 20 hours compared to between 4 and 10 minutes, respectively. All methods find an increased risk of CVD events with higher SBP variability. The methods
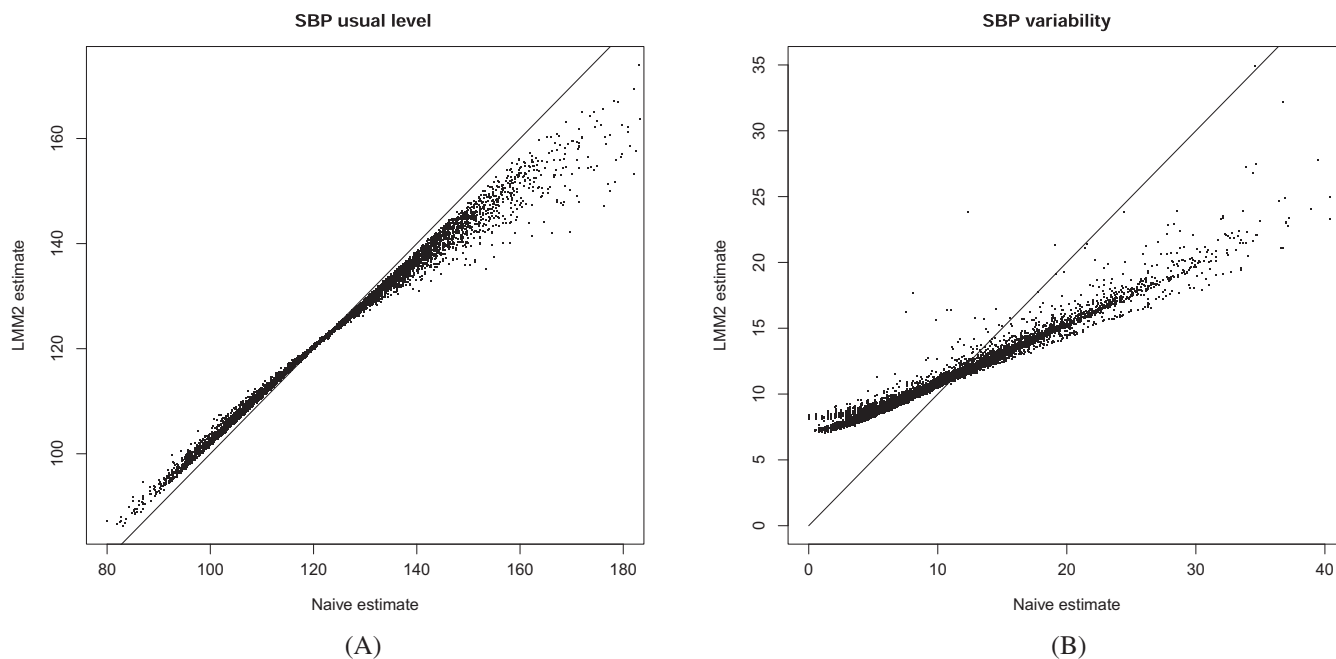
FIGURE 4 Scatter plots of estimates using the LMM2 method against estimates using the naive method of (A) systolic blood pressure (SBP) usual level and (B) SBP variability with lines of agreement between the two estimates

which include random slopes find no evidence for an association between SBP gradient and CVD risk after adjusting for the usual level of SBP and SBP variability.

The estimated association of SBP variability with CVD events is smaller for the naive model, suggesting regression dilution bias in this parameter estimate. Method LMM1, which does not account for the correlation between the SBP usual level and variability, gives a higher estimate of the association with the usual level of SBP than method LMM2, which does account for the correlation, as would be expected from the arguments given in Section 1 of the Supplementary Materials. The estimate of the association with SBP variability is generally lower for LMM2 than LMM1, suggesting a slight bias in the LMM1 estimates. The standard errors are, however, considerably larger for LMM2 than for LMM1. There is little difference between the results from the joint models and the results from the two-stage LMMs. Adjusting for other CVD risk factors in the mixed effects model makes little difference to the estimated association with SBP usual level and variability.

Results from all models applied to the full ARIC dataset are shown in Table 4. For the models without random slopes, the logHRs for the SBP usual level and SBP variability are similar to those from the truncated dataset. Standard errors are slightly smaller reflecting the increased sample size. But, for the models with random slopes, the logHRs for the effect of slope on the hazard of CVD are very different for the full dataset, with more positive slopes now leading to a lower risk of a CVD event when adjusted for the SBP usual level and SBP variability. This is counterintuitive to what we would expect; as higher SBP is associated with greater risk of CVD, we would expect those with increasing SBP over time to also have a higher risk.

5 | DISCUSSION

We have proposed two-stage and joint modelling methods to estimate the association between visit-to-visit SBP variability and CVD, both of which use extensions of standard mixed effects models to allow different individuals to have different SBP variabilities. Both methods reduce regression dilution bias in the estimated hazard ratio compared to naive methods that have previously been used where SBP variability is estimated on an individual-by-individual basis. In addition, the joint modelling methods allow for informative truncation of the repeated measurement process by the event of interest. In practice, the method of choice will depend on the likely extent of the biases incurred by the two-stage approach vs the computational tractability of the joint modelling approach.

TABLE 3 Results (logHRs with standard errors in brackets) from all models applied to reduced data from the ARIC study. Methods LMM1, LMM2, LMM3, and LMM4 are two-stage methods respectively with (1) random intercept only and no correlation between the random intercept and the variability, (2) random intercept only and correlation, (3) random intercepts and slopes and no correlation between the random effects and the variability, and (4) random intercepts and slopes and correlation. JM1, JM2, JM3, and JM4 are the corresponding joint models. Adjusted LMMs adjust the longitudinal models/submodels for baseline CVD risk factors

Model	Usual Level α_1	Variability α_σ	Slope α_3	Total Chol.	HDL Chol.	Age	Smoker	Diabetic	Male Sex
Naive Model									
	0.019 (0.002)	0.017 (0.005)		0.114 (0.029)	−0.369 (0.085)	0.072 (0.006)	−0.503 (0.069)	−0.67 (0.084)	0.273 (0.067)
Unadjusted LMM									
LMM1	0.023 (0.003)	0.045 (0.011)		0.113 (0.029)	−0.37 (0.085)	0.072 (0.006)	−0.505 (0.069)	−0.671 (0.084)	0.273 (0.067)
LMM2	0.015 (0.006)	0.053 (0.021)		0.113 (0.029)	−0.37 (0.085)	0.072 (0.006)	−0.506 (0.069)	−0.668 (0.084)	0.272 (0.067)
JM1	0.023 (0.003)	0.045 (0.011)		0.114 (0.029)	−0.380 (0.087)	0.072 (0.006)	−0.498 (0.068)	−0.670 (0.086)	0.274 (0.065)
JM2	0.014 (0.007)	0.058 (0.023)		0.114 (0.029)	−0.371 (0.085)	0.073 (0.006)	−0.503 (0.070)	−0.657 (0.081)	0.277 (0.068)
Adjusted LMM									
LMM1	0.023 (0.003)	0.047 (0.011)		0.129 (0.029)	−0.405 (0.085)	0.084 (0.006)	−0.447 (0.068)	−0.754 (0.083)	0.297 (0.067)
LMM2	0.012 (0.007)	0.062 (0.022)		0.125 (0.029)	−0.402 (0.085)	0.081 (0.006)	−0.437 (0.068)	−0.749 (0.083)	0.313 (0.067)
JM1	0.023 (0.003)	0.045 (0.012)		0.129 (0.028)	−0.408 (0.085)	0.085 (0.006)	−0.452 (0.066)	−0.766 (0.085)	0.297 (0.069)
JM2	0.010 (0.007)	0.071 (0.023)		0.127 (0.028)	−0.404 (0.084)	0.081 (0.006)	−0.440 (0.070)	−0.754 (0.090)	0.318 (0.069)
Unadjusted LMM with slope									
LMM3	0.023 (0.003)	0.047 (0.009)	0.024 (0.074)	0.112 (0.029)	−0.37 (0.085)	0.072 (0.006)	−0.507 (0.069)	−0.657 (0.084)	0.264 (0.067)
LMM4	0.016 (0.006)	0.053 (0.018)	0.01 (0.089)	0.113 (0.029)	−0.371 (0.085)	0.072 (0.006)	−0.507 (0.069)	−0.657 (0.084)	0.263 (0.067)
JM3	0.023 (0.003)	0.047 (0.011)	0.020 (0.077)	0.114 (0.027)	−0.370 (0.086)	0.073 (0.006)	−0.517 (0.072)	−0.659 0.088	0.265 (0.069)
JM4	0.015 (0.005)	0.056 (0.015)	0.004 (0.089)	0.114 (0.029)	−0.374 (0.087)	0.072 (0.006)	−0.512 (0.069)	−0.662 (0.085)	0.265 (0.068)
Adjusted LMM with slope									
LMM3	0.022 (0.003)	0.049 (0.009)	0.033 (0.071)	0.129 (0.029)	−0.403 (0.085)	0.084 (0.006)	−0.442 (0.068)	−0.747 (0.084)	0.296 (0.067)
LMM4	0.014 (0.006)	0.057 (0.017)	0.006 (0.085)	0.124 (0.029)	−0.406 (0.085)	0.081 (0.006)	−0.437 (0.068)	−0.723 (0.084)	0.303 (0.067)
JM3	0.023 (0.003)	0.049 (0.010)	0.021 (0.071)	0.130 (0.029)	−0.408 (0.085)	0.084 (0.006)	−0.458 (0.067)	−0.766 (0.082)	0.292 (0.067)
JM4	0.014 (0.006)	0.060 (0.017)	−0.003 (0.101)	0.126 (0.029)	−0.406 (0.083)	0.081 (0.006)	−0.441 (0.067)	−0.744 (0.081)	0.304 (0.067)

Abbreviations: ARIC, Atherosclerosis Risk in Communities; CVD, cardiovascular disease; logHRs, log hazard ratios; LMM, linear mixed model.

TABLE 4 Results (logHRs with standard errors in brackets) from all models applied to full data from the ARIC study. Methods LMM1, LMM2, LMM3, and LMM4 are two-stage methods respectively with (1) random intercept only and no correlation between the random intercept and the variability, (2) random intercept only and correlation, (3) random intercepts and slopes and no correlation between the random effects and the variability, and (4) random intercepts and slopes and correlation. JM1, JM2, JM3, and JM4 are the corresponding joint models. Adjusted LMMs adjust the longitudinal models/submodels for baseline CVD risk factors

Model	Usual Level α_1	Variability α_σ	Slope α_3	Total Chol.	HDL Chol.	Age	Smoker	Diabetic	Male Sex
Naive Model									
	0.02 (0.001)	0.016 (0.003)		0.118 (0.02)	−0.461 (0.061)	0.06 (0.004)	−0.62 (0.046)	−0.8 (0.054)	0.395 (0.048)
Unadjusted LMM									
LMM1	0.023 (0.002)	0.042 (0.007)		0.124 (0.019)	−0.436 (0.057)	0.06 (0.004)	−0.634 (0.044)	−0.853 (0.051)	0.382 (0.045)
LMM2	0.015 (0.004)	0.053 (0.014)		0.124 (0.019)	−0.437 (0.057)	0.06 (0.004)	−0.632 (0.044)	−0.849 (0.051)	0.381 (0.045)
JM1	0.024 (0.002)	0.045 (0.007)		0.126 (0.019)	−0.440 (0.055)	0.060 (0.004)	−0.636 (0.044)	−0.855 (0.050)	0.391 (0.045)
JM2	0.012 (0.005)	0.065 (0.016)		0.125 (0.020)	−0.445 (0.058)	0.061 (0.004)	−0.633 (0.045)	−0.862 (0.050)	0.392 (0.047)
Adjusted LMM									
LMM1	0.023 (0.002)	0.043 (0.006)		0.137 (0.019)	−0.453 (0.057)	0.073 (0.004)	−0.59 (0.044)	−0.956 (0.051)	0.409 (0.045)
LMM2	0.013 (0.004)	0.058 (0.013)		0.134 (0.019)	−0.456 (0.057)	0.07 (0.004)	−0.579 (0.044)	−0.949 (0.051)	0.42 (0.045)
JM1	0.024 (0.002)	0.046 (0.007)		0.139 (0.019)	−0.456 (0.061)	0.073 (0.004)	−0.603 (0.045)	−0.988 (0.051)	0.415 (0.045)
JM2	0.009 (0.005)	0.074 (0.016)		0.137 (0.020)	−0.462 (0.058)	0.070 (0.004)	−0.575 (0.043)	−0.943 (0.050)	0.439 (0.049)
Unadjusted LMM with slope									
LMM3	0.025 (0.002)	0.045 (0.006)	−0.129 (0.049)	0.123 (0.019)	−0.441 (0.057)	0.059 (0.004)	−0.635 (0.044)	−0.823 (0.051)	0.363 (0.045)
LMM4	0.02 (0.004)	0.048 (0.011)	−0.212 (0.06)	0.123 (0.019)	−0.442 (0.057)	0.059 (0.004)	−0.632 (0.044)	−0.821 (0.051)	0.359 (0.045)
JM3	0.026 (0.002)	0.049 (0.006)	−0.183 (0.066)	0.125 (0.019)	−0.453 (0.060)	0.060 (0.004)	−0.647 (0.046)	−0.845 (0.048)	0.366 (0.047)
JM4	0.018 (0.004)	0.061 (0.012)	−0.302 (0.085)	0.128 (0.019)	−0.458 (0.058)	0.061 (0.004)	−0.646 (0.047)	−0.847 (0.055)	0.374 (0.046)
Adjusted LMM with slope									
LMM3	0.026 (0.002)	0.047 (0.006)	−0.118 (0.051)	0.138 (0.019)	−0.458 (0.057)	0.072 (0.004)	−0.599 (0.044)	−0.953 (0.051)	0.398 (0.045)
LMM4	0.019 (0.004)	0.051 (0.01)	−0.185 (0.06)	0.132 (0.019)	−0.461 (0.057)	0.069 (0.004)	−0.598 (0.044)	−0.96 (0.051)	0.39 (0.045)
JM3	0.027 (0.002)	0.052 (0.006)	−0.163 (0.061)	0.141 (0.019)	−0.469 (0.057)	0.074 (0.004)	−0.608 (0.046)	−0.968 (0.052)	0.405 (0.045)
JM4	0.017 (0.004)	0.065 (0.011)	−0.253 (0.083)	0.136 (0.019)	−0.480 (0.058)	0.070 (0.004)	−0.587 (0.045)	−0.939 (0.052)	0.416 (0.047)

Abbreviations: ARIC, Atherosclerosis Risk in Communities; CVD, cardiovascular disease; logHRs, log hazard ratios; LMM, linear mixed model.

An important assumption of the two-stage method is that truncation of the observation process is noninformative; otherwise, the number of observations may depend on the underlying risk level, which may lead to bias. Informative truncation can be avoided in the two-stage approach if distinct periods of follow-up for longitudinal and survival data can be defined, but this may only be possible at the cost of discarding information from individuals with events during longitudinal follow-up and repeated measurements taken during survival follow-up. In the clinical literature, analyses have been conducted both using survival follow-up recorded subsequently to the repeated observations of SBP and using survival follow-up recorded concurrently with SBP measurements. As an example of the former, Rothwell et al explored the use of 7 compared with 10 SBP measurements but commented that, when 10 SBP measurements were used, survival follow-up was shorter.[5]

Another important assumption of our method is that the measurement schedule is noninformative, which will not be satisfied in some studies. Some clinical studies investigating the association between SBP variability and CVD have used data, which has not been collected for research purposes, such as data from electronic health records. For example, Gosmanova et al used data from around 3 million US veterans receiving healthcare from the US Veterans Health Administration[25] and Hippisley-Cox et al used SBP variability as a risk predictor in the most recent QRISK tool for predicting CVD based on around 8 million individuals attending general practices in the UK.[2] In such cases, additional bias may be incurred by informative observation of the repeated measurements, where the act of measurement is more (or less) likely in individuals at greater risk of disease. In these cases, it may be necessary to extend the joint model to include an additional submodel for the recurrent event process of an SBP measurement being taken. However, for cohort studies with a pre-specified measurement schedule, such as the ARIC study analysed in this paper, such additional model complexity is unlikely to be warranted. We therefore leave this investigation for future research.

Results from our simulation studies suggest that the best model giving the least bias and coverage closest to nominal levels is the joint model that allows for correlation between the SBP usual level and the SBP variability. The naive method gives considerable bias in the association between SBP variability and CVD, as expected. Results previously reported in the CVD literature using the naive method could be subject to considerable bias in both the estimated association and its standard error. Our simulation results also indicate that there may be a complex interplay between different sources of bias. While the use of mixed effects models and joint models can reduce regression dilution bias, model misspecification, such as not accounting for correlation between these variables, may introduce additional sources of bias.

Our analyses of the ARIC data found evidence of a positive association between higher SBP variability and the risk of CVD events. For this dataset, results between the two-stage LMM approach and the joint modelling approach were similar, suggesting that the biases incurred by use of the two-stage approach were minimal in this example. However, use of the joint model comes at a cost of considerably longer computational time. This might be reduced by running multiple MCMC chains in parallel. If computational time is an issue, then a possible modelling strategy would be to use the two-stage method for model selection and the joint model for inference. If there was also clinical interest in identifying explanatory variables associated with the longitudinal variability, then the mixed effects models could be extended to incorporate a linear predictor in the mean of the $\log \sigma_i$ distribution, $\mu_{\sigma,i} = \zeta^T X_i^\sigma$. Similar models have previously been considered by Hedeker et al[9,10] and Goldstein et al.[11]

*The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

ORCID

Jessica K. Barrett https://orcid.org/0000-0003-1889-9803

REFERENCES

1. D'Agostino RB, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care. *Circulation*. 2008;117(6):743-753.
2. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ*. 2017;357:j2099.
3. Muntner P, Whittle J, Lynch AI, et al. Visit-to-visit variability of blood pressure and coronary heart disease, stroke, heart failure, and mortality a cohort study. *Ann Intern Med*. 2015;163(5):329-338.
4. Parati G, Ochoa JE, Lombardi C, Bilo G. Assessment and management of blood-pressure variability. *Nat Rev Cardiol*. 2013;10(3):143-155.
5. Rothwell PM, Howard SC, Dolan E, et al. Prognostic significance of visit-to-visit variability, maximum systolic blood pressure, and episodic hypertension. *Lancet*. 2010;375(9718):895-905.
6. Shimbo D, Newman JD, Aragaki AK, et al. Association between annual visit-to-visit blood pressure variability and stroke in post-menopausal women: data from the Women's Health Initiative. *Hypertension*. 2012;60(3):625-630.
7. Yano Y. Visit-to-Visit Blood Pressure Variability What is the current challenge. *Am J Hypertens*. 2017;30(2):112-114.
8. Hughes MD. Regression dilution in the proportional hazards model. *Biometrics*. 1993;49(4):1056-1066.
9. Hedeker D, Mermelstein RJ, Demirtas H. An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics*. 2008;64(2):627-634.
10. Hedeker D, Mermelstein RJ, Demirtas H. Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. *Statist Med*. 2012;31(27):3328-3336.
11. Goldstein H, Leckie G, Charlton C, Tilling K, Browne WJ. Multilevel growth curve models that incorporate a random coefficient model for the level 1 variance function. *Stat Methods Med Res*. 2018;27(11):3478-3491. PMID: 28459180.
12. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. Boca Raton, FL: CRC Press; 2012.
13. van Houwelingen H, Putter H. *Dynamic Prediction in Clinical Survival Analysis*. Boca Raton, FL: CRC Press; 2012.
14. Paige E, Barrett J, Stevens D, et al. Landmark models for optimizing the use of repeated measurements of risk factors in electronic health records to predict future disease risk. *Am J Epidemiol*. 2018;187(7):1530-1538.
15. Sweeting MJ, Thompson SG. Joint modelling of longitudinal and time-to-event data with application to predicting abdominal aortic aneurysm growth and rupture. *Biom J*. 2011;53(5):750-763.
16. Sweeting MJ, Barrett JK, Thompson SG, Wood AM. The use of repeated blood pressure measures for cardiovascular risk prediction: a comparison of statistical models in the ARIC study. *Statist Med*. 2017;36(28):4514-4528.
17. Barrett JK, Sweeting MJ, Wood AM. Dynamic risk prediction for cardiovascular disease: an illustration using the ARIC study. *Handb Stat*. 2017;36:47-65.
18. Yang L, Yu M, Gao S. Prediction of coronary artery disease risk based on multiple longitudinal biomarkers. *Statist Med*. 2016;35(8):1299-1314.
19. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter DJ. *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Boca Raton, FL: CRC Press; 2013.
20. Plummer M. JAGS version 4.3.0 user manual. 2017.
21. Su Y-S, Yajima M. R2jags: Using R to run "JAGS". R package version 0.5-7. 2015.
22. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol*. 2002;64(4):583-639.
23. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat*. 1998;7(4):434-455.
24. The ARIC investigators. The atherosclerosis risk in communities (ARIC) study: design and objectives. *Am J Epidemiol*. 1989;129(4):687-702.
25. Gosmanova EO, Mikkelsen MK, Molnar MZ, et al. Association of systolic blood pressure variability with mortality, coronary heart disease, stroke, and renal disease. *J Am Coll Cardiol*. 2016;68(13):1375-1386.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Barrett JK, Huille R, Parker R, Yano Y, Griswold M. Estimating the association between blood pressure variability and cardiovascular disease: An application using the ARIC Study. *Statistics in Medicine*. 2019;38:1855–1868. https://doi.org/10.1002/sim.8074