



Published in final edited form as:

J Appl Stat. 2009 June ; 36(6): 611–619. doi:10.1080/02664760802478208.

Modeling a Cross-Sectional Response Variable with Longitudinal Predictors:

An Example of Pulse Pressure and Pulse Wave Velocity

Veena Shetty¹, Christopher H. Morrell^{2,3}, and Samer S. Najjar²

¹Medstar Research Institute, 6495 New Hampshire Avenue, Hyattsville, MD 20783

²National Institute Aging, 5600 Nathan Shock Drive, Baltimore, MD 21224 USA

³Mathematical Sciences Department, Loyola College in Maryland, 4501 North Charles St., Baltimore, MD 21210 USA

Abstract

We wish to model pulse wave velocity (PWV) as a function of longitudinal measurements of pulse pressure (PP) at the same and prior visits at which the PWV is measured. A number of approaches are compared. First, we use the PP at the same visit as the PWV in a linear regression model. In addition, we also use the average of all available PP's as the explanatory variable in a linear regression model. Next, a two-stage process is applied. The longitudinal PP is modeled using a linear mixed-effects model. This modeled PP is used in the regression model to describe PWV. An approach for using the longitudinal PP data is to obtain a measure of cumulative burden, the area under the PP curve (AUC). This AUC is used as an explanatory variable to model PWV. Finally, a joint Bayesian model is constructed similar to the two-stage model.

Keywords

Mixed effects; Linear regression; Area under the curve

A. Introduction

In this paper we discuss different methodologies to approach the problem of modeling a response variable that is available at a single time point while the explanatory variable is available at multiple time points. In particular, we wish to model pulse wave velocity (PWV) as a function of pulse pressure (PP = systolic blood pressure — diastolic blood pressure). PWV data is available at only the most recent visit while PP values are available at the same and previous visits.

Methods like panel data analysis, econometric modeling, and time series cross-sectional regressions are some approaches that handle similar situations but the data has to be either balanced or equally spaced.

A class of econometric models called dynamic models explores the relationship when the independent variable is a time series and the dependent variable is measured at a certain time t . i.e. when Y is modeled as a function of the current and lagged X 's. These are called dynamic models because the effect of a unit change in the value of the explanatory variable is felt over

a number of time periods. The equation for this kind of model has the following form (Gujarati, 1999):

$$Y_t = A + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + U_t.$$

The time series cross-sectional regression analysis (TSCSREG) procedure analyzes linear econometric models that often arise when cross-sectional and time series data are combined. The TSCSREG procedure analyzes panel data sets that consist of a number of sets of time series data on each of several individuals. The performance of an estimation procedure for the model regression parameters depends on the error components chosen to be in the model. The TSCSREG procedure estimates the regression parameters in the model under several error structures (SAS/ETS, 1999).

Panel data usually consist of two dimensions: the cross-sectional or between subjects dimension and the time-series or longitudinal dimension. Generally the time series component is measured at regular intervals. In our case we have a number of subjects with repeated measurements on the covariates which are measured periodically. Hsiao (2003) presents a number of approaches for analyzing panel data.

Each of the methods described above requires balanced data (the same number of observations at common time points). Given the nature of our study, the data are unequally spaced and subjects have varying numbers of repeated measurements. Consequently these approaches cannot be applied in this case.

In section 2 we describe the data to be used in this study. In section 3 we list the approaches to be used to estimate the effect of PP on PWV. In section 4 we present the results of the study and section 5 presents a simulation study to investigate the relative merits of the approaches used. We present our conclusions in section 6.

B. Data

The data consists of 183 male participants from the Baltimore Longitudinal Study of Aging (BLSA). Participants in the BLSA are volunteers who return every 2 years for approximately two and a half days of clinical examinations. In addition to the other variables measured on these participants, pulse wave velocity (PWV), a measure of arterial stiffness was measured on these subjects at some point in time. Earlier, arterial stiffness was roughly estimated by using pulse pressure (PP). Pulse pressure is the difference of systolic and diastolic blood pressures. We wish to model the cross-sectional PWV based on longitudinal pulse pressure (PP) values.

C. Methods

We use a number of approaches to address this problem. The approaches are based on classical methods like linear regression, linear mixed-effects models, and area under the curve (AUC) to analyze this data. We have also used a Bayesian approach and we compare the results obtained from the Bayesian approach with the linear mixed effects model approach. In addition to PP, all the models include age as a covariate as PWV is related to age. An assumption of linear regression models is that the residuals exhibit normality. In this case, when PWV is used in the analyses, the residuals were quite strongly skewed to the right with some minor heteroscedasticity. To overcome these problems, we perform all analyses using $\ln(\text{PWV})$. This leads to normal residuals and homoscedastic errors.

C.1 Linear Regression

This is the most naive approach where we model $\ln(\text{PWV})$ using multiple regression. Here we initially use PP values at the same visit as PWV and ignore the PP measurements at all previous visits. The linear regression model is:

$$\ln(\text{PWV}_i) = \beta_0 + \beta_1 \times \text{PP}_i + \beta_2 \times \text{LAge}_i + \varepsilon_i$$

where LAge_i is the last age for participant i , the age at which PWV is measured. β_0 is the model intercept, β_1 is the effect of PP on $\ln(\text{PWV})$ and β_2 is the effect of last age on $\ln(\text{PWV})$. That is, for a unit change in PP, $\ln(\text{PWV})$ is expected to change by β_1 and for a unit change in LAge , $\ln(\text{PWV})$ is expected to change by β_2 . In addition, the average of all the repeated PP values for each participant are used as the predictor in the regression model. There is a mean of 4.62 PP measurements per participant (ranging from 2 to 6). About 50% of participants have six PP measurements while the remaining values (2 to 5) are approximately uniformly distributed.

C.2 Linear Mixed Effects/Regression Approach

Stanek et al. (1999) point out that it is often beneficial to use best linear unbiased predictors (BLUPs) in place of the actual value of a variable, especially when the variable is measured with error. In this case, BLUPs are subject-specific predictions from a mixed-effects model. These predictions are computed using both the fixed- and random-effects to obtain the fitted values. Morrell et al. (2003) use this approach to obtain improved parameter estimates in a logistic regression model. They use the mixed-effects model to predict systolic blood pressure. These predicted blood pressure values are used as explanatory variables in the logistic regression model to predict the occurrence of coronary heart disease. The paper shows that this approach yields more unbiased estimates of the parameters than the naïve approach of just using a single blood pressure value and also produces appropriate inferences. Here we apply their two-step approach to this situation. In the first step, a better estimate of PP is obtained by modeling all available PP data using a linear mixed effects (LME) model. Next, using this LME model, PP is predicted at the time point at which PWV is available. In the second step, $\ln(\text{PWV})$ is modeled using the predicted PP from the first step along with last age. The models used in the two-step approach are:

1. $PP_{LME,ij} = \theta_0 + b_{i0} + \theta_1 \times \text{LAge}_i + \theta_2 \times \text{LAge}_i^2 + (\theta_3 + b_{i1}) \times \text{Time}_{ij} + \theta_4 \times \text{Time}_{ij}^2 + \theta_5 \times \text{LAge}_i \times \text{Time}_{ij} + \delta_{ij}$.
2. $\ln(\text{PWV}_i) = \beta_0 + \beta_1 \times PP_{LME,i} + \beta_2 \times \text{LAge}_i + \varepsilon_i$, where in this step PP_{LME} is the value predicted from step 1.

In the first step, the longitudinal PP data are modeled using a mixed-effects model. The PP depends on the age of the participant at the same visit as the PWV measurement as well as the time of the measurement. Note that to allow for curvature, a quadratic in LAge is used. The longitudinal trends are modeled with a quadratic in Time (measured as time before the PWV visit). Finally, there is an interaction between LAge and Time to allow the longitudinal changes to depend on the participant's age. The θ parameters are the fixed-effects regression parameters. The random effects, b_{i0} and b_{i1} , allow the intercept and Time -slope to vary among participants. Consequently, each participant will have their own intercept and Time -slope which will lead to participant-specific predictions of PP. At the second step, the model for $\ln(\text{PWV})$ is the same as in section C.1 but with the observed PP replaced by the predicted value from the mixed-effects model in step 1, $PP_{LME,i}$, predicted at the visit where PWV is measured. The interpretation of the regression parameters is the same as in section C.1.

C.3. Area under the curve or cumulative burden analysis

The area under the curve (AUC) can be used as a way of assessing the cumulative burden of an explanatory variable on a response variable. For example, Li et al. (2003) use the area under the curve “as a measure of cumulative risk burden from childhood to adult.” In this two-step approach, the longitudinal PP is again modeled using the LME model as in section C.2. Then the area under the curve (AUC) is calculated from the fitted model for each subject. The AUC is the integral of the fitted mixed model with respect to time. Since the fitted model is a quadratic function in time, a formula for the AUC is derived and used to compute an AUC value for each participant (see below). These AUCs are used as the explanatory variables in the regression model along with LAge. Since participants have different lengths of followup, the cumulative burden will depend on this length of followup. Consequently, the AUC is divided by the length of the time interval over which the participant has data. In what follows, AUC means AUC/Time. In addition, the AUC’s are calculated in two ways: for the entire time-span (from the participant’s first visit to the PWV visit); and for a fixed time period (the most recent 2 yrs) to eliminate any effect of different lengths of followup on the results. In the second stage, the transformed PWV is modeled using the AUC calculated from stage 1 along with last age. For this approach, the formula for calculating the AUC and the regression model used to model ln(PWV) in this two-step approach are:

1. $PP_{AUC,i} = -\{(\theta_0 + b_{i0} + \theta_1 LAge_i + \theta_2 LAge_i^2) \text{Time} + (\theta_3 + b_{i1} + \theta_5 LAge_i) \text{Time}^2/2 + \theta_4 \text{Time}^3/3\}$, where the parameter estimates are obtained from the LME model in Step-1 of section C.2 and Time is the length of followup where either the entire followup is used or Time = 2 when the followup is restricted to 2 years.
2. $\ln(PWV)_i = \beta_0 + \beta_1 PP_{AUC,i} + \beta_2 LAge_i + \varepsilon_i$ where $PP_{AUC,i}$ is the AUC value computed in step 1 for participant i . The interpretation of the regression parameters in step 2 is similar to the interpretations given in section C.1.

C.4 Bayesian Approach

Guo and Carlin (2004) describe a Bayesian joint model for longitudinal and survival data. Here we follow their approach to jointly model the longitudinal pulse pressure with the regression modeling of ln(PWV) as a function of pulse pressure. The Bayesian joint model first describes the longitudinal pulse pressure data in an identical way to the first part of the two-step LME/Regression approach. The longitudinal model for pulse pressure is:

$$PP_{i,j} = \theta_0 + b_{i,0} + \theta_1 \times LAge_i + \theta_2 \times LAge_i^2 + (\theta_3 + b_{i,1}) \times Time_{i,j} + \theta_4 \times Time_{i,j}^2 + \theta_5 \times LAge_i \times Time_{i,j} + \delta_{i,j}$$

where $\delta_{ij} \sim N(0, \sigma_\delta^2)$ and $b_i \sim N(0, D)$. Note that the model for PP is the same as the longitudinal mixed-effects model in section C.2. In addition, the normal assumptions on the error term and random effects vector, $b_i = (b_{i,0}, b_{i,1})^T$, are the standard assumptions for the mixed effects model. σ_δ^2 is the variance of the random error term and D is a 2×2 positive-definite covariance matrix of the random effects. Next the regression model for the PWV data is specified as:

$$\ln(PWV_i) = \beta_0 + \beta_1 \times \widehat{PP}_{Bayes,i} + \beta_2 \times LAge_i + \varepsilon_i$$

where, since time = 0 at the PWV visit, the predicted pulse pressure from the longitudinal model at the PWV visit is

$$\widehat{PP}_{Bayes,i} = \theta_0 + b_{i,0} + \theta_1 LAge_i + \theta_2 \times LAge_i^2$$

where the parameters and random effects are replaced by their estimates and $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$. Once more, following Guo and Carlin (2004), the vague proper prior distributions are: $\theta \sim$

$N_6(0, \Sigma\theta)$, $1/\sigma^2\delta \sim \text{gamma}(0.2, 0.2)$, $D^{-1} \sim \text{Wishart}\left(\begin{bmatrix} 1000 & 0 \\ 0 & 10 \end{bmatrix}, 23\right)$, $\beta \sim N_3(0, \Sigma\beta)$, and $1/\sigma_\varepsilon^2 \sim \text{gamma}(0.2, 0.2)$ where $\Sigma\theta = \text{diag}(0.0001, 0.01, 0.01, 0.01, 0.01, 0.01)^{-1}$, and $\Sigma\beta = \text{diag}(0.0001, 0.0001, 0.001)^{-1}$. These distributions provide relatively flat prior information about the mixed-effects and regression parameters, θ and β . In addition, the Gamma and Wishart priors chosen for the inverses of the three variance/covariance parameters in the models, $1/\sigma^2\delta$, D^{-1} , and $1/\sigma_\varepsilon^2$, are standard priors for these parameters. The parameters of the prior distributions are again chosen to provide little influence on the posterior estimates of the parameter values.

D. Results

D.1 Results of Approaches using Classical Statistics Models

The parameter estimates of pulse pressure obtained from the regression analyses are compared along with their standard error and their 95% confidence intervals among the various approaches. In Table 1 we note that the estimates range from 0.00315 m/s for the linear regression method to 0.0072 m/s for the mixed model approach. For the AUC approach, when we use the entire follow up period, the estimate is 0.0053 m/s while for a fixed time period of 2 years the estimate is 0.00681 m/s, close to the mixed model approach. It is not clear from these results whether linear regression provides underestimates or mixed models are providing overestimates. The R^2 for all models are very similar. Figure-1 illustrates the pulse pressure estimates obtained from the mixed model approach (Panel-B) as well as the area under the curve approach (Panel-C) in comparison to the observed pulse pressure measures (Panel-A). The slope of pulse pressure in the mixed model approach is much steeper and is in accordance with the result obtained from the model. To enable us to evaluate which approach is best, we conducted a simulation study (see Section E.).

D.2 Results of Bayesian Approach

WinBUGS (Lunn, Thomas, Best, and Spiegelhalter, 2000) was used to fit the Bayesian model to the data. We used 5 chains and thinned the results by 5. After omitting 1000 iterations for burnin, the next 5000 values are used to obtain the posterior distributions and parameter estimates. To check for convergence of the chains to a stable posterior distribution, we examined the history plots, autocorrelations, posterior densities, Brooks-Gelman-Rubin plots, quantile plots, as well as the impact of the priors on the posterior distributions. Based on these checks, the chains of all parameters appear to exhibit a stationary process so that these chains may be used to estimate the posterior distributions of each parameter. Tables 2 and 3 present the means and standard deviations of the MCMC values and compare them to the LME/ regression estimates and standard errors. For the longitudinal pulse pressure model, the estimates/mean of posterior distributions and standard errors/posterior standard deviations are almost identical except for some slight differences with the components of the random-effects covariance matrix (Table 2) for the intercept and covariance. For the regression model, the values are also very similar (Table 3). Thus the LME/Regression and Bayes approaches provide very similar results for this data.

E. Simulation Study

To assess which of the methods provides the best approach to estimating the regression parameters a simulation study was conducted. For the purpose of the simulation, the regression model was simplified to contain only the pulse pressure variable and the longitudinal model contains only a linear time term. This will not affect the results since the LAge and Age terms

in the longitudinal and regression models, respectively, only affect the level for each subject. Excluding them will not affect the relationship between the longitudinal trend in PP and its association with $\ln(\text{PWV})$. The simulation study used 183 subjects with an average of 4 observations per subjects (with 2, 3, 4, 5, or 6 observations). As described in section C.1, in the actual data used in this study there tended to be more participants with 6 observations while the remaining numbers of observations are approximately uniformly distributed. We do not anticipate that using a uniform distribution of the number of observations in the simulation study will invalidate the results. To generate the data we used the following procedure:

1. Pick “true” PP_0 values, values of pulse pressure at the same visit as the PWV value.
2. Compute $\ln(\text{PWV})$ values from the true model: $\ln(\text{PWV}) = \beta_0 + \beta_1 \text{PP}_0 + \varepsilon$ where $\beta_0 = 0$, $\beta_1 = 0.005$, and $\varepsilon \sim N(0, 0.25^2)$.
3. To make sure the longitudinal model has the correct value at the PWV visit, compute random effect: $b_{i0} = \text{PP}_0 + \theta_0$ where $\theta_0 = 46.67$ (the mean of the true PP_0 values).
4. Generate b_{i1} from conditional distribution of $b_{i1}|b_{i0} \sim N\left(d_{12}/d_{11} + \mathcal{Z} \times \sqrt{d_{22} - d_{12}^2/d_{11}}\right)$ where $z \sim N(0, 1)$ and $D = \begin{bmatrix} 54 & 2.718 \\ 2.718 & 0.38 \end{bmatrix}$
5. Compute the “observed” PP values: $\text{PP} = (\theta_0 + b_{i0}) + (\theta_1 + b_{i1}) \times T + \delta_{iT}$ where $\theta_1 = 0.5$ and $\delta_{iT} \sim N(0, 50)$.

This procedure will create data sets that contain longitudinal PP data and the corresponding $\ln(\text{PWV})$ values that will allow us to compare the properties of a number of estimation procedures.

It is assumed that $\ln(\text{PWV})$ follows a regression model with PP as the explanatory variable. In particular,

$$\ln(\text{PWV}_i) = \beta_0 + \beta_1 \times \text{PP}_i + \varepsilon_i$$

where the values PP_i will be chosen in 5 different ways leading to different estimates of the regression parameter, β_1 . Initially, the true PP_0 values and the “observed” pulse pressures (with error) at the PWV visit are used in the regression models. Using PP_0 will provide the standard against which the other procedures may be compared since the true (but in practice unknown) PP values are used. In this case we know we are fitting the correct model to the data. The other approaches use a number of different ways of replacing the unknown true PP value with an estimate. Fitting the models with the observed PP values (with error) mimics what would happen if the researcher just used a single PP value measured with error at the same visit as PWV to fit the model (as was done in section C.1). Next, predicted values of PP from the mixed effects model are used (as in section C.2) and finally the cumulative burden or AUC value from both the subject’s entire follow-up and restricted to a two-year period is used (as in section C.3).

The results of the simulation are presented in Table 4. The bias due to the estimation procedure is measured by how far the mean of the simulated parameter estimates is from the true parameter value. The mean square error (MSE) measures how much the estimates deviate from the true value by accounting for the variability in the estimates. It is not surprising that using the true PP_0 values leads to the smallest bias and mean squared errors (MSE) for both the intercept and slope. As mentioned above this is the standard against which all other procedures can be compared. The LME/Regression approach is the best among the methods where error is added to the PP values in terms of both bias and MSE. The AUC over a fixed two year time frame is only slightly worse than the LME/Regression approach in terms of MSE. It is perhaps not

surprising that the AUC method performs quite well as the AUC values are calculated from the LME model. Interestingly, the mean of the parameter estimates from the AUC method is on the opposite side of the true parameter value compared to the means obtained using the true PP_0 and LME approaches. Using the “observed” PP at the PWV visit leads to the most biased parameter estimates. In particular the slope estimate is severely biased downward towards 0 while the intercept is biased by having the largest mean intercept among the five approaches. Interestingly, the AUC approach over the entire time span (so that different subjects are integrated over different time spans) provides the worst estimates in terms of MSE. However the bias is much better than the “observed” PP for both the intercept and slope. Consequently, when additional data is available one can obtain improved regression parameter estimates using the LME/Regression approach over only using the explanatory variable at the same visit as the response variable as this removes some of the measurement error that may be present in the data.

F. Conclusions

In this paper we have investigated the problem of choosing how to represent an explanatory variable in a regression model with a response variable when repeated measurements on the explanatory variable are available. With the help of a simulation study we are able to draw some conclusions about which approach leads to the best estimates of the parameter in the regression model. The analyses suggest that if multiple values of the explanatory variable are available it is not advisable to use a single time point. Among the different methods that were used to model the pulse pressure data, using a linear mixed-effects model to estimate the value of the explanatory variable leads to the best estimates of the regression parameters. Area under the curve could also be used as a reasonable approach as long as the AUC is computed from a common time frame. A unified longitudinal-regression model was estimated using a Bayesian approach. The Bayesian analysis produced similar results to the LME/regression approach.

Acknowledgments

This research was supported by the Intramural Research Program of the NIH, National Institute on Aging. A portion of that support was through a R&D contract with MedStar Research Institute. We thank an anonymous referee whose comments greatly improved the paper.

References

1. Gujarati, D. Essentials of Econometrics. Vol. 2nd. Irwin McGraw-Hill; 1999.
2. Guo X, Carlin BP. “Separate and Joint Modeling of Longitudinal and EventTime Data Using Standard Computer Packages”. *The American Statistician* 2004;58(1):16–24.
3. Hsiao, C. Analysis of Panel Data. Vol. 2nd. Cambridge University Press; 2003.
4. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 2000;10:325–337.
5. Li S, Chen W, Srinivasan SR, Bond MG, Tang R, Urbina EM, Berenson GS. Childhood Cardiovascular Risk Factors and Carotid Vascular Changes in Adulthood: The Bogalusa Heart Study. *Journal of the American Medical Association* 2003;290(17):2271–2276. [PubMed: 14600185]
6. Morrell CH, Brant LJ, Pearson JD, Verbeke GNM, Fleg JL. Applying Linear Mixed-Effects Models to the Problem of Measurement Error in Epidemiologic Studies. *Communications in Statistics: Simulation and Computation* 2003;32(2):437–459.
7. Vol. Version 8. SAS Institute Inc; Cary, NC: 1999. SAS/ETS User’s Guide.
8. Stanek EJ, Well A, Ockene I. “Why not routinely use best linear unbiased predictors (BLUPs) as estimates of cholesterol, per cent fat from kcal and physical activity?”. *Statistics in Medicine* 1999;18(21):2943–2959. [PubMed: 10523752]

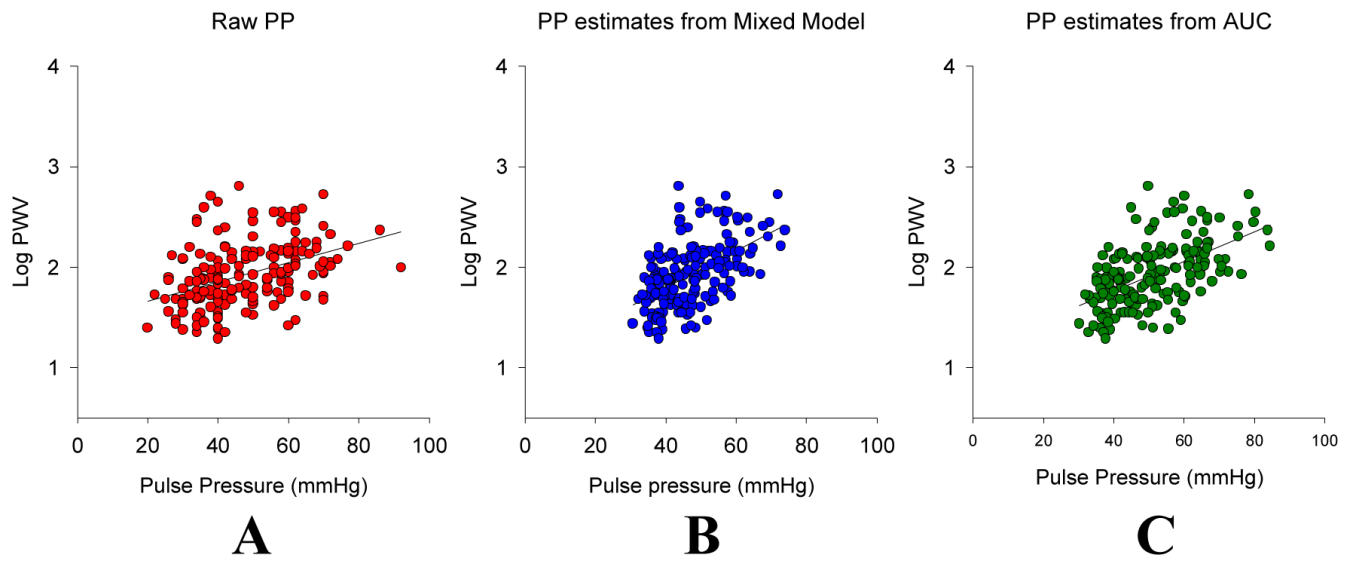


Figure-1.
(Position of the figure should be before part E --the simulation study)

Table I

Comparison of regression estimates obtained from the classical methods.

Method β_1	β_1	$se(\beta_1)$	95% CI	Model R^2
LR (same visit)	0.00315	0.00157	0.000041, 0.00625	0.388
LR (average)	0.00584	0.00232	0.00126, 0.01042	0.388
Mixed Model	0.00723	0.00288	0.00154, 0.01291	0.395
AUC(entire follow-up)	0.00539	0.00225	0.00094, 0.00983	0.387
AUC (2 yrs follow-up)	0.00681	0.00277	0.00135, 0.01227	0.388

(Position of Table-I: Should be in the results section at the end of section D.1 and before the beginning of section D.2)

Table II

Comparison of the LME estimates with the posterior mean and standard deviation of the parameters in the longitudinal pulse pressure model.

	LME/Regression Approach			Bayesian Approach		
	Estimate	se	95% CI	Estimate	se	95% CI
Fixed Effects						
Intercept	52.1	5.9	40.5, 63.8	52.0	6.0	40.3, 63.7
Lage	-0.70	0.22	-1.13, -0.28	-0.69	0.22	-1.13, -0.26
Lage2	0.100	0.019	0.062, .0137	0.099	0.020	0.061, .0138
Time	-1.01	0.40	-1.79, -0.22	-0.98	0.43	-1.81, -0.15
Time2	0.036	0.010	0.016, 0.057	0.040	0.012	0.018, 0.063
Lage×Time	0.033	0.006	0.020, 0.046	0.033	0.007	0.020, 0.046
Random Components						
σ^2_ϵ	75.34			75.3	4.50	66.93, 84.56
Intercept	53.89			44.92	7.79	31.51, 61.61
Time	0.38			0.40	0.10	0.25, 0.62
Covariance	3.28			1.75	0.77	0.43, 3.43

(Position of Table-II: Should be in the results section at the end of section D.2)

Table III

Comparison of the regression estimates (from LME/Regression approach) with the posterior mean and standard deviation of the parameters in the regression model of the predicted pulse pressure on the PWV.

	LME			Bayes		
	Estimate	se	95% CI	Estimate	se	95% CI
Fixed Effects						
Intercept	5.700	0.095	5.513, 5.887	5.683	0.102	5.476, 5.879
LAge	0.00875	0.00162	0.00555, 0.01195	0.00849	0.00177	0.00497, 0.0119
PP	0.00723	0.00288	0.00154, 0.01291	0.00789	0.00327	0.00153, 0.0144
Random Component						
σ^2_e	0.06239			0.06398	0.00699	0.05154, 0.07891

(Position of Table-III: Should be in the results section after table-II)

Table IV

Results of Simulation Study (number of replications = 1000)

Method	β_0 (True value = 0)			β_1 (True value = 0.005)		
	Mean	Std. Dev.	MSE	Mean	Std. Dev.	MSE $\times 10^6$
True PP ₀	0.0019	0.1180	0.01394	0.00496	0.00252	6.34
PP at Time = 0	0.1100	0.0856	0.01943	0.00260	0.00181	9.03
LME	0.0029	0.1343	0.01804	0.00494	0.00289	8.36
AUC over entire time	-0.0150	0.1440	0.02096	0.00544	0.00317	10.02
AUC over 2 years	-0.0059	0.1387	0.01926	0.00518	0.00302	9.13

(Position of Table-IV: Should be after the conclusions)