



Published in final edited form as:

Blood Press Monit. 2014 June ; 19(3): 153–163. doi:10.1097/MBP.0000000000000039.

An Analysis of 24-Hour Ambulatory Blood Pressure Monitoring Data using Orthonormal Polynomials in the Linear Mixed Model

Lloyd J. Edwards^{a,*} and Sean L. Simpson^{a,b}

^aDepartment of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7420

^bDepartment of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157-1063

Abstract

Background—The use of 24-hour ambulatory blood pressure monitoring (ABPM) in clinical practice and observational epidemiological studies has grown considerably in the past 25 years. ABPM is a very effective technique for assessing biological, environmental, and drug effects on blood pressure.

Objectives—In order to enhance the effectiveness of ABPM for clinical and observational research studies via analytical and graphical results, developing alternative data analysis approaches using modern statistical techniques are important.

Methods—The linear mixed model for the analysis of longitudinal data is particularly well-suited for the estimation of, inference about, and interpretation of both population (mean) and subject-specific trajectories for ABPM data. We propose using a linear mixed model with orthonormal polynomials across time in both the fixed and random effects to analyze ABPM data.

Results—We demonstrate the proposed analysis technique using data from the Dietary Approaches to Stop Hypertension (DASH) study, a multicenter, randomized, parallel arm feeding study that tested the effects of dietary patterns on blood pressure.

Conclusions—The linear mixed model is relatively easy to implement (given the complexity of the technique) using available software, allows for straight-forward testing of multiple hypotheses, and the results can be presented to research clinicians using both graphical and tabular displays. Using orthonormal polynomials provides the ability to model the nonlinear trajectories of each subject with the same complexity as the mean model (fixed effects).

Keywords

AUC; DASH Study; Graphical Display; Hypertension; Longitudinal analysis; Model Selection; Orthogonal Polynomials

*Corresponding Author: Lloyd J. Edwards, Ph.D., Department of Biostatistics, The University of North Carolina at Chapel Hill, 3105H McGavran-Greenberg, CB# 7420, Chapel Hill, NC 27599-7420, phone: 919-843-6603, fax: 919-966-3804, Lloyd_Edwards@unc.edu.

Conflict of interest: There are no conflicts of interest to declare.

Introduction

Ambulatory blood pressure monitoring (ABPM) is a powerful research tool for examining blood pressure (BP) variability and the physiologic and environmental factors that affect BP [1]. The evidence that ABPM gives information over and above conventional blood pressure measurement (CBPM) has been growing steadily over the past 25 years and the technique is now accepted as being indispensable to good clinical practice [2,3]. ABPM is a non-invasive technique in which a standard cuff, attached to a lightweight, portable data recording unit, is placed around the upper arm and inflated at regular preset intervals during a 24 hour time period.

Traditionally, clinicians have used office blood pressure measurements as the preferred method of monitoring blood pressure and consequently the method of diagnosing hypertension. Unfortunately, physician's office blood pressure measurements often can give rise to what is termed "white-coat hypertension", that is, an artificially high blood pressure reading [4]. ABPM provides a profile of blood pressure away from the medical environment, thereby allowing identification of individuals with a white coat response [5]. It also provides several other important advantages discussed by O'Brien [5]: ABPM can demonstrate the efficacy of antihypertensive medication over a 24 hour period rather than making a decision based on one or a few CBPMs confined to a short period of the diurnal cycle; ABPM can identify patients whose blood pressure does not reduce at night-time (the non-dippers) who are probably at high risk for a number of conditions; and the technique can demonstrate a number of patterns of blood pressure behavior that may be relevant to clinical management (isolated systolic hypertension, hypotension, dipping and non-dipping, etc.).

As noted by O'Brien [5], recent longitudinal studies have shown that ABPM is a much stronger predictor of cardiovascular morbidity and mortality than CBPM. The importance of the technique is further evidenced by the fact that the Centers for Medicare and Medicaid Services (CMS) in the US have approved ABPM for reimbursement. O'Brien [2,5] concluded that ABPM should be available to all hypertensive patients given its international acceptance as an indispensable tool for patients with established and suspected hypertension. This recommendation has important implications for clinical practice. Among the questions that O'Brien [5] posed were: 1. how should the data be presented?; 2. how best can doctors and nurses unfamiliar with the technique be educated in its use and interpretation of the data?

Some analyses of ABPM data have used means and/or medians, either obtaining means over the entire 24-hour period or obtaining means over defined intervals of the 24-hour period, for example daytime and nighttime intervals [6-8]. Other analyses have stressed the use of traditional simple or extended cosinor models [9-10]. Some authors have proposed smoothing splines or other smoothing techniques for the analysis of 24-hour ABPM data [11-13]. Jaccard and Wan [14] proposed using cross-sectional pooled time series designs for the analysis of 24-hour ABPM data. Schwartz et al. [15] proposed a very limited variance-components model which is equivalent to a mixed model using very simple explanatory factors (time is not explicitly used as a factor). Selwyn and DiFranco [16] demonstrated the

use of the general linear mixed model as an approach to analyzing 24-hour ABPM data. Selwyn and DiFranco [16] considered an extended cosinor model and also a cubic spline model with 8 knots.

Lambert et al. [17] noted that there has been relatively little work done on the longitudinal analysis of 24-hour ABPM data. Though some authors have attempted to address longitudinal analysis of 24-hour ABPM data, there is not a generally accepted 'standard' method of analyzing 24-hour ABPM. Lambert et al. [17] states that the most common method of analysis using trigonometric functions has been cosinor analysis [16], with Fourier analysis being an alternative. The authors cite that these techniques have been criticized as they impose too many restrictions on the shape of the profile and have been shown to fit real profiles poorly. Indeed, Edwards et al. [18] demonstrates how severe restrictions can be placed on the 24-hour trajectory when a linear mixed model with regression-splines are used and it is assumed that the trajectory is such that the 0-hour value is forced to equal the 24-hour value as is the case for trigonometric functions and periodicity.

Typically the emphasis of 24-hour ABPM data analyses are on between-subject effects (fixed effects) and less emphasis is placed on within-subject effects (random effects) [16,17]. We propose using a linear mixed model with orthonormal polynomials in both the fixed and random effects as an effective approach to analyzing 24-hour ABPM data, particularly when subject-specific trajectories are of primary emphasis. The general linear mixed model [19,20] has become a very powerful statistical tool in the analysis of longitudinal data with continuous outcomes in both clinical and non-clinical studies [21-23]. The linear mixed model for the analysis of longitudinal data is particularly well-suited for the estimation of, inference about, and interpretation of both population and subject-specific trajectories for ABPM data. Typical mean comparison approaches to the analysis of 24-hour ABPM assume that the difference in BP between groups remains constant over the 24-hour time duration, i.e., that the effect is a main effect. However, there are no guarantees that the difference in BP between groups across time should be a main effect only. If there are interaction effects across the 24-hour period between groups, the effects can be estimated and tested in the fixed effects component of the linear mixed model. Subject-specific trajectories are of great importance in ABPM studies but little emphasis has been placed on this dimension of the problem in data analyses. Using orthonormal polynomials provides the ability to address the nonlinear trajectories represented by individual subjects' trajectories with the same complexity in which the mean model (fixed effects) are addressed. Graphical presentations are proposed that help to address questions posed by O'Brien [5] regarding how the data should be presented. We demonstrate the proposed technique using data from the Dietary Approaches to Stop Hypertension (DASH) study, a multicenter, randomized, parallel arm feeding study that tested the effects of dietary patterns on blood pressure.

In the Methods section, the basic linear mixed model formulation and the details for the orthonormal polynomial statistical methods are presented. We discuss the results of applying the models to the DASH data in Results section and we also discuss the new graphical displays created. A summary and conclusions are given in Discussion.

Methods

As discussed in the Introduction, several statistical methods have been proposed for the analysis of 24-hour ABPM data. In this paper, we consider the linear mixed model with orthonormal polynomials with degree of polynomial ranging from 4 to 9.

With N independent sampling units (often *persons* in practice), the linear mixed model for person i may be written

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i. \quad (1)$$

Here, \mathbf{y}_i is a $p_i \times 1$ vector of observations on person i ; \mathbf{X}_i is a $p_i \times q$ known, constant design matrix for person i , with full column rank q while $\boldsymbol{\beta}$ is a $q \times 1$ vector of unknown, constant, population parameters. Also \mathbf{Z}_i is a $p_i \times m$ known, constant design matrix with rank m for person i corresponding to the $m \times 1$ vector of unknown random effects \mathbf{b}_i , while \mathbf{e}_i is a $p_i \times 1$ vector of unknown random errors. Gaussian \mathbf{b}_i and \mathbf{e}_i are independent with mean $\mathbf{0}$ and

$$V \left(\begin{bmatrix} \mathbf{b}_i \\ \mathbf{e}_i \end{bmatrix} \right) = \begin{bmatrix} \Sigma_{b_i}(\boldsymbol{\tau}_b) & \mathbf{0} \\ \mathbf{0} & \Sigma_{e_i}(\boldsymbol{\tau}_e) \end{bmatrix}. \quad (2)$$

Here $V(\cdot)$ is the covariance operator, while both $\Sigma_{b_i}(\boldsymbol{\tau}_b)$ and $\Sigma_{e_i}(\boldsymbol{\tau}_e)$ are positive-definite, symmetric covariance matrices. Therefore $V(\mathbf{y}_i)$ may be written $\Sigma_i = \mathbf{Z}_i \Sigma_{b_i}(\boldsymbol{\tau}_b) \mathbf{Z}_i' + \Sigma_{e_i}(\boldsymbol{\tau}_e)$. We assume that Σ_i can be characterized by a finite set of parameters represented by an $r \times 1$ vector $\boldsymbol{\tau}$ which consists of the unique parameters in $\boldsymbol{\tau}_b$ and $\boldsymbol{\tau}_e$. Throughout $n = \sum_{i=1}^N p_i$.

We may also need to refer to a stacked data formulation of model (1) given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \mathbf{e}, \quad (3)$$

with $\mathbf{y} = [\mathbf{y}'_1 \cdots \mathbf{y}'_N]'$, $\mathbf{X} = [\mathbf{X}'_1 \cdots \mathbf{X}'_N]'$, $\mathbf{Z} = \text{diag}(\mathbf{Z}_1, \cdots, \mathbf{Z}_N)$, $\mathbf{b} = [\mathbf{b}'_1 \cdots \mathbf{b}'_N]'$, and $\mathbf{e} = [\mathbf{e}'_1 \cdots \mathbf{e}'_N]'$. Here $\mathbf{b} \sim N_{Nm}[\mathbf{0}, \Sigma_{b_i}(\boldsymbol{\tau}_b) \otimes \mathbf{I}_N]$ and $\mathbf{e} \sim N_n(\mathbf{0}, \Sigma_e)$ for $\Sigma_e = \text{diag}[\Sigma_{e_1}(\boldsymbol{\tau}_e), \cdots, \Sigma_{e_N}(\boldsymbol{\tau}_e)]$. In turn $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \Sigma)$ with $\Sigma = V(\mathbf{y}) = \text{diag}(\Sigma_1, \cdots, \Sigma_N)$.

The advantage of reducing bias in covariance estimation has made restricted maximum likelihood (REML) estimation very popular for the linear mixed model. Given our focus on variance estimates, all parameter estimates in this paper are done using REML. However, the formulations also apply to computations based on maximum likelihood estimates.

For readers who may not be familiar with the linear mixed model and the equations used, consider a setup for a simple intercept and slope model for both the fixed and random effects (1) for a longitudinal outcome with 3 observations and no missing values:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{pmatrix} = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \end{pmatrix} \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} + \begin{pmatrix} e_{i1} \\ e_{i2} \\ e_{i3} \end{pmatrix}.$$

Here, \mathbf{y}_i is a 3×1 ; \mathbf{X}_i is a 3×2 with full column rank 2, $\boldsymbol{\beta}$ is 2×1 , \mathbf{Z}_i is 3×2 (for this example, $\mathbf{Z}_i = \mathbf{X}_i$), \mathbf{b}_i is 2×1 and \mathbf{e}_i is a 3×1 . Additional fixed-effect covariates can be added such as race and gender and we would have

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{pmatrix} = \begin{pmatrix} 1 & \delta_i & \gamma_i & t_1 \\ 1 & \delta_i & \gamma_i & t_2 \\ 1 & \delta_i & \gamma_i & t_3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \end{pmatrix} \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} + \begin{pmatrix} e_{i1} \\ e_{i2} \\ e_{i3} \end{pmatrix},$$

where $\delta_i = 1$ if i th subject is female, 0 if male, and $\gamma_i = 1$ if black and 0 if white. Here, \mathbf{y}_i is 3×1 , \mathbf{X}_i is a 3×4 with full column rank 4, $\boldsymbol{\beta}$ is 4×1 , \mathbf{Z}_i is 3×2 (same random effects as before but now $\mathbf{Z}_i \subset \mathbf{X}_i$), \mathbf{b}_i is 2×1 and \mathbf{e}_i is a 3×1 . From (2) we have $\boldsymbol{\Sigma}_i = \mathbf{Z}_i \boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b) \mathbf{Z}_i' + \boldsymbol{\Sigma}_{e_i}(\boldsymbol{\tau}_e)$, where

$$\boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b) = \begin{pmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{01}^2 & \sigma_1^2 \end{pmatrix} \text{ and } \boldsymbol{\Sigma}_{e_i}(\boldsymbol{\tau}_e) = \sigma_e^2 \mathbf{I}_3,$$

and σ_0^2 denotes the variance of the subject-specific intercept, σ_1^2 denotes the variance of the subject-specific slope, σ_{01}^2 denotes the covariance between the random intercept and slope, \mathbf{I}_3 is a 3×3 identity matrix. Here $\boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b)$, 2×2 and unstructured, and $\boldsymbol{\Sigma}_{e_i}(\boldsymbol{\tau}_e)$ (3×3) are the same for all subjects with $\boldsymbol{\tau}_b = (\sigma_0^2, \sigma_1^2, \sigma_{01}^2)'$ (3×1) and $\boldsymbol{\tau}_e = \sigma_e^2$. The model has 2 fixed effect parameters to estimate (β_0 and β_1) and 4 variance parameters to estimate ($\sigma_0^2, \sigma_1^2, \sigma_{01}^2$, and σ_e^2).

Why orthonormal polynomials?

Polynomial models play a prominent role in mathematics. In mathematical analysis, the Weierstrass approximation theorem [24] states that every continuous function defined on a closed interval $[a, b]$ can be uniformly approximated as closely as desired by a polynomial function. Because polynomials are among the simplest functions, and because computers can directly evaluate polynomials, this theorem has both practical and theoretical relevance, especially in polynomial interpolation. The Extended Mean Value Theorem [24] is a remarkable theorem usually associated with Brook Taylor (1685-1731) that enables us to approximate various complicated functions by much simpler natural polynomial functions. Although a power series expansion cannot always be obtained, most of the familiar functions in calculus can be represented as a sum of a convergent power series. For any individual, blood pressure as a function of time on the interval $[1, 24]$ can be assumed to be a continuous function that meets the regularity conditions posed by the Weierstrass approximation theorem and the Extended Mean Value Theorem.

Polynomials are commonly used to describe curved relationships in statistical models where a model must be developed empirically. Due to the heterogeneity of 24-hour ABPM trajectories among subjects, polynomials can be very effective statistical tools for empirically developing models for subject-specific trajectories. The goal of polynomial

regression is to model a non-linear relationship between the independent and dependent variables (technically, between the independent variable and the conditional mean of the dependent variable). An advantage of traditional polynomial regression is that the inferential framework of the linear mixed model can be used.

The use of natural polynomials for the analysis of 24-hour ABPM recordings poses problems, particularly when modeling subject-specific trajectories. The problems include the existence of very large values in the fixed and random effects design matrices, for example $24^3 = 13,824$ and $24^9 = 2.64 \times 10^{12}$; problems of multicollinearity since x, x^2, \dots, x^q are correlated; and mixed model software convergence problems when used in random effects. The lack of convergence in mixed model software can be especially acute when employing natural polynomials. In most cases it is caused by the multicollinearity and large values present in the random effects and their resulting effect on the estimation of the random effects covariance $\Sigma_{b_i}(\tau_b)$. The typical solution is to reduce the number of random effects, but this results in an increasing lack of fit for the subject-specific trajectories. The use of orthonormal polynomials helps alleviate this problem and hence allows for more accurate modeling of the heterogeneity in subject-specific 24-hour ABPM trajectories than natural polynomials.

Orthonormal polynomial predictors in linear regression are transformations of the natural polynomials that provide a new set of predictors that meet the following criteria: 1. The new predictors contain the same information as the original set; 2. The new predictors are linear combinations of the original natural polynomials; 3. The new columns of predictors all have mean zero (except for the constant term); 4. The new columns of predictors are all mutually orthogonal. Combining (3) and (4) implies that the new predictors (except the constant) are mutually uncorrelated. The first new column captures the information in the intercept, the second captures all of the linear term information adjusted for the intercept, the third captures quadratic information above and beyond the linear and intercept information and so forth.

We recognize that polynomial terms of order greater than cubic are not very intuitive from a biological perspective when considered individually. However, in the case of 24-hour ABPM trajectories, the main interest is in the shape of the trajectories that the orthonormal polynomials produce collectively and not in any one single polynomial component. As such, higher order polynomials are to be interpreted collectively as an approximation to the population trajectory and in particular the trajectories of individual subjects. Two big advantages of using orthonormal polynomials are that a) the problem of multicollinearity is greatly reduced; and b) the orthonormal polynomials are “normalized” and bounded in the interval $(-1, 1)$. Thus, orthonormal polynomials have few problems with rounding error or very large (small) regression coefficients.

Using the linear mixed model with orthonormal polynomials, subject-specific 24-hour ABPM trajectories can be represented and parameterized, within- and between- treatment comparisons based on individual trajectories can be made, and discrimination between different trajectories can be obtained. The linear mixed model can account for circadian rhythms, subject effects, and treatment effects.

The Case of Complete, Balanced Data with Equally Spaced Observations

For the analysis of 24-hour ABPM data, we advocate the use of a linear mixed model with orthonormal polynomials used in both the fixed and random effects, particularly when an emphasis is placed on accurate modeling of subject-specific parameters. Here, we assume that the degree of the polynomials used in the fixed effects is equal to that used in the random effects. For demonstration purposes, we first start with a linear mixed model with only time effects and intercept, but no additional covariates. We assume that time is equally spaced 1-hour intervals ranging from 1 to 24 and that each subject has all 24 observations (complete and balanced). We have

$$\mathbf{y}_i = \mathbf{X}_i^* \boldsymbol{\beta} + \mathbf{Z}_i^* \mathbf{b}_i + \mathbf{e}_i. \quad (4)$$

Here, \mathbf{y}_i is as before, a $p_i \times 1$ vector of observations on person i ; \mathbf{X}_i^* is a $p_i \times q$ fixed effects design matrix of orthonormal polynomials for person i ; \mathbf{Z}_i^* is a $p_i \times q$ random design matrix of orthonormal polynomials with rank q for person i . Gaussian \mathbf{b}_i and \mathbf{e}_i are independent with mean $\mathbf{0}$ and variance given in equation 2.

Because the polynomials are orthogonal, mathematically we can model $\boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b)$ as diagonal with heterogeneous diagonal elements, i.e., $\boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b) = \text{diag}(\boldsymbol{\tau}_b)$, where

$\boldsymbol{\tau}_b = (\sigma_0^2, \sigma_1^2, \dots, \sigma_{q-1}^2)'$ is the $q \times 1$ vector of variances for each element of the random

effects vector \mathbf{b}_i . Under this assumption, $\sum_i \mathbf{Z}_i^* \text{diag}(\boldsymbol{\tau}_b) \mathbf{Z}_i^{*'} + \sigma_e^2 \mathbf{I}_{p_i}$. This greatly simplifies the complexity of the random effects covariance and thus the modeling of $V(\mathbf{y}_i)$.

The number of parameters in $\boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b)$ is q , the number of diagonal elements. However, in actual data analysis, the best fitting $\hat{\boldsymbol{\Sigma}}_{b_i}(\hat{\boldsymbol{\tau}}_b)$ may be achieved by using an unstructured $\boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b)$, where the collinearity of the elements of $\boldsymbol{\Sigma}_{b_i}(\boldsymbol{\tau}_b)$ is greatly reduced in comparison to a natural polynomials fit. These empirical covariances are present in the orthonormal fits due to the lack of fit of the fixed effects, $\mathbf{X}_i^* \hat{\boldsymbol{\beta}}$, in equation 2. We address this issue in the example data.

The maximum degree orthonormal polynomial that we consider for the DASH data is a 9th degree. Deciding on using a 9th degree orthonormal polynomial was a process that considered several factors. Our general goal was to model the nonlinear ABPM trajectories without overfitting the data. Also, our prior experience analyzing DASH data guided us to begin with a maximum model with a polynomial degree between 6 and 9 [18,25]. This also included consideration of the number of parameters, 9, used by both Selwyn and Difranco [16] and Lambert et al. [17]. We opted for the maximum degree orthonormal polynomial to be the 9th degree since it was significant in the SBP model (though not in the DBP model). As a check, we also fitted a 10th degree polynomial model to both SBP and DBP and found that it was not significant in either model. Simpson and Edwards [25] used a fixed-effects only model for the DASH data and the order of the orthonormal polynomial was chosen to be six when considering additional covariates and aiming for parsimony. However, for the example analysis in this paper, we also consider a range of polynomials between 4 and 9 with a focus on the 9th degree model since it is the largest polynomial considered. We do

not consider additional covariates but we discuss how the covariates can be included in the model.

In the 9th degree model, we estimate $q = 10$ fixed effects (intercept and 9 orthonormal polynomials), 10 random effects covariance parameters and σ_e^2 (assuming $\Sigma_{bi}(\tau_b) = \text{diag}(\tau_b)$ and $\sum_{e_i}(\tau_e) = \sigma_e^2 \mathbf{I}_{p_i}$) for a total of 11 covariance parameters. We also compare the fit using $\Sigma_{bi}(\tau_b) = \text{diag}(\tau_b)$ to that using an unstructured covariance for $\Sigma_{bi}(\tau_b)$. However, with conventional mixed effects software like SAS Proc Mixed, it would be impossible to use an unstructured $\Sigma_{bi}(\tau_b)$ with natural polynomials for the subject-specific models for the reasons discussed in previously.

The Case of Incomplete, Unbalanced Data

In the DASH Study, about 81% (288/357) of the subjects had complete data at baseline. The additional 19% of subjects had very little missing data. We assumed that the data were missing completely at random for these analyses, i.e., an “intent-to-treat” treatment of missing data by including every observation of the dependent measure (no observations are discarded and no data are imputed).

One of the main advantages of the linear mixed model for longitudinal data is that it uses all available data, rather than excluding cases with missing observations, and accommodates unbalanced designs. This feature of the linear mixed model is of particular relevance to ABPM data since it is common for a significant proportion of participants to be missing at least one observation [26].

In 24-hour ABPM studies, typically there are criteria for controlling the quality of the monitoring data. For example, in the DASH study, if fewer than 14 acceptable readings were obtained, the subject was asked to repeat the monitoring. Among participants with acceptable ABPM measurements, more than 90% of the possible waking and sleeping readings were obtained [27]. Sometimes more rigid criteria are used, e.g., for Hermida et al. [28], BP series were not considered valid for analysis if 30% of the measurements were lacking, if data were missing for > 2-hour spans, if data were collected from subjects while experiencing an irregular rest–activity schedule, or if the nighttime sleep span was < 6 hours or > 12 hours during monitoring.

Missing data was minimal in the hourly DASH data. In this DASH analysis, 69 subjects had fewer than 3 missing values each. Because of the sample size, number of data points for each trajectory, and minimal missing data, we believe that assuming the data was missing completely at random (MCAR) is a very reasonable assumption for the DASH data. However, further research will be required to fully assess the effect of missing observations in 24-hour ABPM studies and whether model fitting results are different if imputation is used.

Results

DASH Study

The Dietary Approaches to Stop Hypertension (DASH) trial was a multicenter, randomized, parallel arm feeding study that tested the effects of dietary patterns on blood pressure [8]. The three diets were a control diet (low in fruits, vegetables, and dairy products, with a fat content typical of the average diet in the United States), a diet rich in fruits and vegetables (a diet similar to the control except it provided more fruits and vegetables and fewer snacks and sweets), and a combination diet rich in fruits, vegetables, and low-fat dairy foods and reduced in saturated fat, total fat, and cholesterol. The combination diet will be subsequently referred to as the DASH diet.

Participants were healthy, community-dwelling adults 22 years of age or older who were not taking antihypertensive medication. Each subject had an average systolic blood pressure of less than 160 mm Hg and a diastolic blood pressure of 80 to 95 mm Hg (mean of six measurements across three screening visits). Study subjects were enrolled sequentially in groups; the first group began the run-in phase of the trial in September 1994, and the fifth and last group started in January 1996.

For each group, data were collected during three phases (screening, run-in, and intervention). Run-in was a three-week period in which all participants were fed the control diet. Toward the end of run-in, 24-hour ambulatory blood pressure monitoring was obtained once. This constituted the “baseline” ABPM reading. During the third week, participants were randomized to one of three diets. Intervention was an eight-week period in which participants were fed their assigned diets. During the last two weeks, one 24-hour ambulatory blood pressure monitoring was obtained (end-of-intervention ABPM).

ABPM was attempted on the 362 participants enrolled in groups 2-5. We use 357 subjects for baseline 24-hour ABPM data analysis, including incomplete data. Five (5) subjects were deleted because they failed to produce a satisfactory baseline 24-hour ABPM recording. We also use the 288 subjects who had complete data at baseline and the 287 subjects who had complete data during the intervention for our AUC example.

DASH Study Data Analysis Results - Estimation, Inference and Goodness-of-fit

Table 1 provides estimates, standard errors (SE), and p-values for the 9th degree orthonormal polynomial fit to the DASH baseline 24-hour ABPM data. Table 2 provides model comparisons using orthonormal polynomial models of degree ranging 4-9. All computations were done using SAS v9.2 and the ORPOL function in the SAS IML procedure, which handles equal and unequal time spacings when generating orthonormal polynomials. Restricted maximum likelihood estimation (REML) was used for estimation and the Kenward-Roger F and adjusted denominator degrees of freedom were used for all fixed effect inference. However, maximum likelihood estimation results were almost identical (results not shown). From Table 1, we can see that the sizes of the absolute values of the orthonormal polynomial regression coefficients indicate that orthonormal polynomials 1-6 have the largest effect on BP, changing from being greater than 4 to less than 2.

Table 2 provides model selection criteria results for the Akaike Information Criterion (AIC) [29] and Bayesian Information Criterion (BIC) [30] for orthonormal polynomial models using 4-9th degree polynomials in both the fixed and random effects. AIC and BIC were computed for both $\Sigma_{b_i}(\tau_b) = \text{diag}(\tau_b)$ and $\Sigma_{b_i}(\tau_b) = \text{Unstructured (UN)}$. Morrell et al. [31] commented that the best way to select among linear mixed-effects models based on various information criteria is still not clearly determined. The conclusion from both the AIC and BIC is that the 9th degree orthonormal polynomial model using $\Sigma_{b_i}(\tau_b) = \text{UN}$ is the best model for DBP and SBP since it yields the smallest AIC and BIC values when compared to all other orthonormal polynomial models. Mathematically, the assumption of $\Sigma_{b_i}(\tau_b) = \text{diag}(\tau_b)$ should have yielded the best covariance fit, but the lack of fit in the fixed effects is carried over into the variance (the estimated $\Sigma_{b_i}(\tau_b)$ was not diagonal). There were no model convergence problems using $\Sigma_{b_i}(\tau_b) = \text{UN}$.

To check the model distribution assumptions for the random effects and residual error, we used an important result from Gurka et al. [32] which states that if $\varepsilon_i = \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i$ is multivariate normal, then both \mathbf{b}_i and \mathbf{e}_i are multivariate normal given the model assumption that \mathbf{b}_i and \mathbf{e}_i are independent. Using this result, we used standard residual analysis for the estimated “stacked” ε_i and concluded that the errors were approximately normally distributed for each of the orthonormal polynomial models considered.

The linear mixed model easily accommodates additional explanatory variables. Typical mean comparison approaches to the analysis of 24-hour ABPM assume that the difference in BP between groups remains constant over the 24-hour time duration, i.e, that the effect is a main effect. However, there are no guarantees that the difference in BP between groups across time should be a main effect only. If there are interaction effects across the 24-hour period between groups, the effects can be estimated and tested in the fixed effects component of the linear mixed model. For example, we can do a simple check for interaction between the 3 diet groups and time by including intercept effects for diet (two indicator variables with control diet as the reference group) and interaction effects in time for diet. Let $\delta_i = 1$ if subject i is assigned the diet rich in fruits and vegetables and $\delta_i = 0$ otherwise. Similarly, let $\lambda_i = 1$ if subject i is assigned the DASH diet and $\lambda_i = 0$ otherwise. We can then add diet fixed effects (both main effects and diet \times time interaction effects) in the linear mixed model as follows

$$\mathbf{y}_i = (\mathbf{X}_i^* \quad \delta_i \mathbf{X}_i^* \quad \lambda_i \mathbf{X}_i^*) \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\alpha} \\ \boldsymbol{\gamma} \end{pmatrix} + \mathbf{Z}_i^* \mathbf{b}_i + \mathbf{e}_i. \quad (5)$$

Here $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, and $\boldsymbol{\gamma}$ are $q \times 1$ vectors of fixed effect population parameters where $\boldsymbol{\beta}$ represents the effect of the control diet (intercept and time) on blood pressure, $\boldsymbol{\alpha}$ represents the differential effect of the diet rich in fruits and vegetables relative to the control diet, and $\boldsymbol{\gamma}$ represents the differential effect of the DASH diet relative to the control diet. Note that the structure of random effects and within-subject errors remain the same as in equation (4) as does the corresponding covariance structure. We can rewrite the model as a standard linear mixed model

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i. \quad (6)$$

Here \mathbf{X}_i is a $p_i \times 3q$ fixed effects design matrix for person i and $\boldsymbol{\beta}$ is the corresponding $3q \times 1$ vector of fixed effect parameters. Tests of interaction and intercept effects (not shown here) were not significant and hence we concluded that there was no evidence of a difference in 24-hour BP profiles at baseline. Using the same model for the intervention period, there was no evidence of interaction effects but there was evidence of diet main effects (again, results not shown here). The result is very important to know since it helps supports the conclusions of analyses using mean 24-hour ABPM for the DASH study data because the effect of diet appears to be constant across the 24-hour duration. An additional example of including several explanatory variables in the analysis of the DASH data can be found in Simpson and Edwards [26], which features a general linear model for longitudinal data with the circular LEAR correlation structure and uses diet groups, race, age, and orthonormal polynomials as fixed effects.

In order to incorporate both baseline and intervention into the model, we use a shared parameter, multivariate linear mixed model [33,34] given by

$$\begin{pmatrix} \mathbf{y}_{0i} \\ \mathbf{y}_{1i} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_{0i} \\ \mathbf{X}_{1i} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \mathbf{Z}_{0i} \\ \mathbf{Z}_{1i} \end{pmatrix} \mathbf{b}_i + \begin{pmatrix} \mathbf{e}_{0i} \\ \mathbf{e}_{1i} \end{pmatrix}, \quad (7)$$

where \mathbf{y}_{0i} is the 24×1 vector of observations for baseline BP, \mathbf{y}_{1i} is the 24×1 vector of observations for intervention BP, \mathbf{X}_{0i} and \mathbf{X}_{1i} the $24 \times q$ fixed effects design matrices for baseline and intervention periods, and \mathbf{Z}_{0i} and \mathbf{Z}_{1i} the $24 \times m$ random effects design matrices for baseline and intervention periods. The shared parameter model is appropriate since the response is the same for baseline and intervention. The scalar representation of the model is given by

$$y_{hij} = \beta_0 + \beta_1 \pi_{hi} + \beta_2 \delta_i + \beta_3 \lambda_i + \beta_4 \pi_{hi} \delta_i + \beta_5 \pi_{hi} \lambda_i + \sum_{k=6}^{14} \beta_k x_{kij}^* + \sum_{l=1}^{10} b_{il} z_{lij}^* + e_{hij}$$

where $h = 0, 1$ represents baseline and intervention BP for subject i , $\pi_{0i} = 0$ if baseline period and $\pi_{1i} = 0$ if intervention period for subject i , and x_{kij}^* and z_{lij}^* are the fixed and random orthonormal time effects. The model assumes that within-subject variation is the same at baseline and intervention and the between-subject effect of intervention is modeled using the period effect, π_{hi} , for the intercept and diet main effects ($\beta_1 \pi_{hi}$, $\beta_4 \pi_{hi} \delta_i$, $\beta_5 \pi_{hi} \lambda_i$).

Table 3 provides the results for the change in SBP and DBP for each diet main effect. Both the DASH and fruit/vegetable diets lowered SBP and DBP significantly compared with the control diet ($p < 0.0001$ for SBP and DBP on both diets: control diet, $-0.3/-0.02$ mm Hg; fruit/vegetable diet, $-3.2/-2.1$ mm Hg; DASH diet, $-4.6/-2.8$ mm Hg). The model results are very similar to the results Moore et al. 28 found using mean ABP where both the fruit/vegetable and DASH diets lowered 24-hour ABP significantly compared with the control diet ($p < 0.0001$ for SBP and DBP on both diets: control diet, $-0.2/+0.1$ mm Hg; fruit/vegetable diet, $-3.2/-1.9$ mm Hg; DASH diet, $-4.6/-2.6$ mm Hg).

Graphical Displays

As discussed in section 1, O'Brien [5] posed the important question regarding how the data should be presented. O'Brien [5] provided depictions of 24-hour ABPM data based on observed values that contained a plot of normal bands across the 24 hours for which he overlaid selected individual subject trajectories to demonstrate examples of normal blood pressure and various instances of deviations from normal blood pressure (above or below normal range). O'Brien [5] noted that as with conventional measurement, normal ranges for ABPM have been the subject of much debate over the years. The focus of the outcomes-based figures is to provide practicing physicians with a graphical approach to evaluating an individual's 24-hour BP. However, for clinical and observational studies that require group comparisons and exploring additional explanatory factors that may impact BP, the outcomes-based graphical display is not as useful. A model-based approach helps overcome some of the limitations of the outcomes-based approach.

The linear mixed model with orthonormal polynomials provides the statistical framework for producing a model-based graphical display which will be extremely useful for clinical and observational studies of 24-hour ABPM. Subject-specific trajectories can be of great importance in ABPM studies, especially in clinical practice and research. In practice, it is the subject-specific values that are used to provide diagnoses and to design treatment regimens. However, little emphasis has been placed on this dimension of the problem in modeling 24-hour ABPM data. The smoothed predicted trajectories will provide researchers with the ability to construct additional measures of 24-hour ABPM and correlate the measures with important cardiovascular disease (CVD) outcomes such as hypertension and electrocardiogram measures such as left ventricle mass index (LVMI).

Figures 1 and 2 for SBP and DBP provide model-based graphical displays of 24-hour ABPM based on O'Brien [2] outcomes-based concepts. Figures 1 and 2 show one subject's predicted curve computed from the linear mixed model for SBP and DBP (solid blue lines), the subject's observed data (connected by dashed blue lines), the 90% prediction interval for DASH study normals at baseline (shaded gray region), and the commonly used static measures of upper limit of normal for daytime and nighttime (red horizontal lines) SBP (135/120) and DBP (85/75). The 90% prediction interval for the 9th degree orthonormal polynomial model provides a model-based definition of normal range. The 90% prediction interval closely matches the outcomes-based ± 2 standard deviations about the mean for each time point. We constructed the 90% prediction interval by selecting the subset of all subjects that had normal SBP (DBP) at baseline, fitting a linear mixed model with orthonormal polynomials, and then computing a 90% prediction interval using a SAS macro developed by By [35]. The prediction intervals are used to provide an interval estimate for a single observation as opposed to confidence intervals which provide an (narrower) interval estimate for an average of observations.

For SBP, the smoothed subject-specific trajectory suggests that the subject is above the upper limit of normal for the majority of the day and night. However, there are observed values for which the smoothed subject-specific trajectory suggests there may be masked hypertension - 10:00-11:00 a.m. and 1:00-4:00 p.m., hours 10-11 and 13-16 in Figure 1, respectively. For DBP, both the smoothed subject-specific trajectory and the observed data

indicate that this person's DBP is above the upper limit of normal for both daytime and nighttime. However, in practice, actual observed BP values will necessarily take precedence over predicted values.

The graphical displays provided in Figures 1 and 2 also visually reveal that the discontinuity at the change point from daytime to nighttime can present a problem for assessing observed values and for computing AUC values for the smoothed subject-specific trajectories using the static cutoffs. It is clear that the static cutoffs do not take into account the natural variability and downward trend of the 24-hour ABPM data near the discontinuity.

Area Under the Curve (AUC)

The DASH diet is now well known to reduce blood pressure. Thus, model-based subject-specific AUC should correlate with the performance of the DASH diet. Table 3 provides a descriptive comparison of the average model-based subject-specific AUC for the control, DASH, and fruits/vegetable diets. Results are for the 288 subjects who had complete 24-ABPM data at baseline and the 287 subjects who had complete data during the intervention. The DASH diet group clearly has the largest drop from baseline of average AUC for both SBP and DBP, followed by the fruits/vegetables group and then the control group having the poorest performance. For the DASH diet, the mean change in the model based AUC divided by 24 hours for SBP is 5.2 mmHg and DBP is 3.0 mmHg. The model-based AUC result is consistent with the DASH study finding that the DASH diet lowered blood pressures by an average of 5.5 and 3.0 mmHg for systolic and diastolic, compared with the control diet [36]. Sacks et al. [36] DASH analyses defined baseline blood pressure as the average of 7 days of measurements, 3 during screening and 4 during run-in. The last five sets of measurements during the intervention were used to compute end of study blood pressure, and the effect of the diets. For mean change in 24-hour ABPM measurements, Sacks et al. [36] reported 4.5 mmHg for SBP and 2.7 mmHg for DBP.

The model-based subject-specific scaled AUC can be shown to be a linear combination of the fixed and random effect regression coefficients. In order to demonstrate, suppose that we were using natural polynomials, then equation 4 assumes $\mathbf{X}_i^* = \mathbf{Z}^*$ and that BP is a polynomial function in time, x , and hence the subject-specific AUC is given by

$$AUC_i = \int_0^{24} (\hat{\beta}_{i,0} + \sum_{j=1}^9 \hat{\beta}_{i,j} x^j) dx = 24\hat{\beta}_{i,0} + \sum_{j=1}^9 \hat{\beta}_{i,j} \frac{24^{j+1}}{j+1}$$

and the corresponding AUC scaled by 24 hours is

$$AUC_i/24 = \hat{\beta}_{i,0} + \sum_{j=1}^9 \frac{24^j}{j+1} \hat{\beta}_{i,j},$$

where $\hat{\beta}_{i,0} = \hat{\beta}_0 + \hat{b}_{i,0}$ is the estimated BP intercept for the i -th subject and $\hat{\beta}_{i,j} = \hat{\beta}_j + \hat{b}_{i,j}$, $j > 0$, is the estimated j -th regression coefficient for the i -th subject. So the model-based subject-specific scaled AUC is a linear combination of the adjusted mean BP (subject-specific

intercept) and a weighted linear combination of subject-specific regression coefficients. It is clear that the scaled AUC is different than a simple mean of 24-hour ABPM data. However, further research is required to assess whether the model-based subject-specific AUC is a better summary measure of 24-hour ABPM than the simple arithmetic mean that is typically used.

These results have very important implications for the use of model-based subject-specific AUC computed from 24-ABPM data in CVD studies. They provide some evidence that model-based AUC computed from 24-ABPM data may potentially be used as a correlate of CVD outcomes. For example, Nobre and Mion [37] calculated areas under systolic and diastolic blood pressure trajectories using the observed data (outcomes-based) and compared with systolic and diastolic blood pressure load and 24-hour systolic and diastolic blood pressure in order to determine which provided the best correlation with left ventricular mass index (LVMI). Nobre and Mion [37] found that the correlations of the parameters obtained by ABPM with LVMI were high and statistically significant except for blood pressure load between 90 and 100% and for 24-hour SBP of 135 mmHg or less, and SBP load higher than 50%.

Discussion

The development of alternative, robust and flexible statistical methods for analyzing 24-hour ABPM data will be important in refining its role in clinical research, clinical practice, and observational studies. Therefore, we were motivated to develop an alternative statistical model for analyzing 24-hour ABPM data that will overcome some of the limitations of conventional analyses and potentially lead to improved 24-hour ABPM correlates of CVD outcomes and electrocardiogram measures such as LVMI. Though the DASH study did not measure electrocardiogram parameters, the Jackson Heart Study [38] is an example of a study that has both baseline 24-hour ABPM and LVMI measurements.

We have shown how the linear mixed model with orthonormal polynomials across time in both the fixed and random effects provides a powerful approach to the analysis of 24-hour ABPM data, particularly when focus is on modeling the subject-specific trajectories. In addition to replicating DASH Study results, the linear mixed model with orthonormal polynomials provides researchers with the ability to explore the simultaneous effects of multiple predictors and their interactions. Typical mean comparison approaches to the analysis of 24-hour ABPM assume that the difference in BP between groups remains constant over the 24-hour time duration, i.e., that the effect is a main effect. For the DASH study, it stands to reason that that the differences in diets remain the same across the 24-hour duration and therefore simple mean differences suffice for testing significance. However, for drug interventions there are no guarantees that the difference in BP between groups across time should be a main effect only. If there are interaction effects across the 24-hour period between groups, the effects can be estimated and tested in the fixed effects component of the linear mixed model. In addition, subject-specific trajectories, which are of great importance in ABPM studies, can also be estimated. Accurately modeling subject-specific trajectories provides a great opportunity to address tailored or individualized treatment regimens.

Since data visualization can be a very useful tool for assessing 24-hour ABPM, we presented model-based graphical presentations of 24-hour ABPM based on the linear mixed model results that can aid in the visual representation and interpretation of 24-hour ABPM data. However, further research will be required to fully understand the robustness of the orthonormal polynomial model to outliers, departures from distributional assumptions, and the effect of more severe missingness, irregularly timed data, and truncated data. In addition, different assumptions and parameterizations for the multivariate linear mixed model should be further explored in order to gain further insight into the statistical properties of the differences between DASH Study 24-hour ABP at baseline and intervention.

The linear mixed model is relatively easy to implement (given the complexity of the technique) using available software, allows for straight-forward testing of multiple hypotheses, and the results can be presented to research clinicians using both graphical and tabular displays. Using orthonormal polynomials provides the ability to model the nonlinear trajectories of each subject with the same complexity as the mean model (fixed effects).

Acknowledgments

Source of Funding: Lloyd Edwards was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Award Number UL1TR000083. Sean Simpson was supported by NIBIB K25 EB012236-01A1.

References

1. Appel LJ, Stason WB. Ambulatory blood pressure monitoring and blood pressure self-measurement in the diagnosis and management of hypertension. *Ann Intern Med.* 1993; 118:889–892. [PubMed: 8480963]
2. O'Brien E. Ambulatory blood pressure measurement: the case for implementation in primary care. *Hypertension.* 2008; 51:1435–1441. [PubMed: 18362225]
3. O'Brien E. Ambulatory blood pressure monitoring: 24 hour blood pressure control as a therapeutic goal for improving cardiovascular prognosis. *Medicographia.* 2010; 32:241–249.
4. Owens P, Atkins N, O'Brien E. Diagnosis of white coat hypertension by ambulatory blood pressure monitoring. *Hypertension.* 1999; 34:267–72. [PubMed: 10454452]
5. O'Brien E. Ambulatory blood pressure monitoring in the management of hypertension. *Heart.* 2003; 89:571–576. [PubMed: 12695477]
6. Weber MA, Cheung DG, Grattinger WF, et al. Characterization of antihypertensive therapy by whole-day blood pressure monitoring. *JAMA.* 1988; 259:3281–3285. [PubMed: 3373659]
7. Ferguson JH, Shaar CJ. The effective diagnosis and treatment of hypertension by the primary care physician: impact of ambulatory blood pressure monitoring. *J Am Board Fam Pract.* 1992; 5:457–65. [PubMed: 1414446]
8. Appel LJ, Moore TJ, Obarzanek E, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH collaborative research group. *N Engl J Med.* 1997; 336:1117–1124. [PubMed: 9099655]
9. Halberg J, Halberg F, Leach CN. Variability of human blood pressure with reference mostly to the non-chronologic literature. *Chronobiologia.* 1984; 11:205–16. [PubMed: 6510118]
10. Gaffney M, Taylor C, Cusenza E. Harmonic regression analysis of the effect of drug treatment on the diurnal rhythm of blood pressure and angina. *Stat Med.* 1993; 12:129–142. [PubMed: 8446808]
11. Streitberg B, Mayer-Sabellek W, Baumgart P. Statistical analysis of circadian blood pressure recordings in controlled clinical trials. *J Hypertens Suppl.* 1989; 7:S11–S17. [PubMed: 2668461]

12. Streitberg B, Mayer-Sabellek W. Smoothing twenty-four-hour ambulatory blood pressure profiles: a comparison of alternative methods. *J Hypertens Suppl.* 1990; 8:S21–S37. [PubMed: 2081996]
13. Dickson D, Hasford J. 24-hour blood pressure measurement in antihypertensive drug trials: Data requirements and methods of Analysis. *Stat Med.* 1992; 11:2147–2158. [PubMed: 1293675]
14. Jaccard J, Wan CK. Statistical analysis of temporal data with many observations: issues for behavioral medicine data. *Ann Behav Med.* 1993; 15:41–50.
15. Schwartz JE, Warren K, Pickering TG. Mood, location and physical position as predictors of ambulatory blood pressure and heart rate: Application of a multi-level random effects model. *Ann Behav Med.* 1994; 16:210–220.
16. Selwyn MR, Difranco DM. The application of large Gaussian mixed models to the analysis of 24 hour ambulatory blood pressure monitoring data in clinical trials. *Stat Med.* 1993; 12:1665–1682. [PubMed: 8248660]
17. Lambert PC, Abrams KR, Jones DR, et al. Analysis of ambulatory blood pressure monitor data using a hierarchical model incorporating restricted cubic splines and heterogeneous within-subject variances. *Stat Med.* 2001; 20:3789–3805. [PubMed: 11782034]
18. Edwards LJ, Stewart PJ, MacDougall JE, et al. A method for fitting regression splines with varying polynomial order in the linear mixed model. *Stat Med.* 2006; 25:513–527. [PubMed: 16158407]
19. Harville DA. Extension of the Gauss-Markov theorem to include the estimation of random effects. *Ann Stat.* 1976; 4:384–395.
20. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics.* 1982; 38:963–974. [PubMed: 7168798]
21. Edward LJ. Modern statistical techniques for the analysis of longitudinal data in biomedical research. *Pediatr Pulmonol.* 2000; 30:330–334. [PubMed: 11015135]
22. Gurka, MJ.; Edwards, LJ. *Handbook of Statistics, Epidemiology and Medical Statistics.* Vol. 27. Elsevier; North-Holland: Amsterdam: 2007. Mixed Models; p. 253-280.
23. Cheng J, Edwards LJ, Maldonado-Molina MM, et al. Real longitudinal data analysis for real people: Building a good enough mixed model. *Stat Med.* 2010; 29:504–520. [PubMed: 20013937]
24. Jeffreys, H.; Jeffreys, BS. *Methods of Mathematical Physics.* 3. Cambridge, England: Cambridge University Press; 1988. Weierstrass's theorem on approximation by polynomials and extension of Weierstrass's approximation theory; p. 446-448.
25. Simpson SL, Edwards LJ. A circular LEAR correlation structure for cyclical longitudinal data. *Stat Methods Med Res.* 2013; 22:296–306. [PubMed: 21216801]
26. Ibrahim JG, Molenberghs G. Missing data methods in longitudinal studies: a review. *Test (Madr).* 2009; 8:1–43. [PubMed: 21218187]
27. Moore TJ, Vollmer WM, Appel LJ, et al. Effect of dietary patterns on ambulatory blood pressure: results from the Dietary Approaches to Stop Hypertension (DASH) trial. DASH Collaborative Research Group. *Hypertension.* 1999; 34:472–477. [PubMed: 10489396]
28. Hermida RC, Ayala DE, Calvo C, et al. Effects of time of day of treatment on ambulatory blood pressure pattern of patients with resistant hypertension. *Hypertension.* 2005; 46:1053–1059. [PubMed: 16087787]
29. Akaike H. A new look at the statistical model identification. *IEEE Trans Automatic Control.* 1974; AC-19:716–723.
30. Schwarz SR. Estimating the dimension of a model. *Ann Stat.* 1978; 6:461–464.
31. Morrell CH, Brant LJ, Ferrucci L. Model choice can obscure results in longitudinal studies. *J Gerontol A Biol Sci Med Sci.* 2009; 64:215–222. [PubMed: 19196902]
32. Gurka MJ, Edwards LJ, Muller KE, et al. Extending the Box–Cox transformation to the linear mixed model. *J R Stat Soc Ser A Stat Soc.* 2006; 169:273–288.
33. Verbeke G, Fieuws S, Molenberghs G, et al. The analysis of multivariate longitudinal data: A review. *Stat Methods Med Res.* 2012.10.1177/0962280212445834
34. Bandyopadhyay S, Ganguli B, Chatterjee A. A review of multivariate longitudinal data analysis. *Stat Methods Med Res.* 2011; 20:299–330. [PubMed: 20212072]
35. By, K. Macro for computing prediction interval. 2005. <http://www.unc.edu/~kby/sas/predVar.sas>

36. Sacks FM, Appel LJ, Moore TJ, et al. The Dietary Approach to Prevent Hypertension: A review of the Dietary Approaches to Stop Hypertension (DASH) Study. *Clin Cardiol.* 1999; 22(Suppl. III):III-6–III-10. [PubMed: 10410299]
37. Nobre F, Mion D. Is the area under blood pressure curve the best parameter to evaluate 24-h ambulatory blood pressure monitoring data? *Blood Press Monit.* 2005; 10:263–270. [PubMed: 16205445]
38. Carpenter MA, Crow R, Steffes M, et al. Laboratory, Reading Center, and Coordinating Center Data management Methods in the Jackson Heart Study. *Am J Med Sci.* 2004; 328:131–144. [PubMed: 15367870]

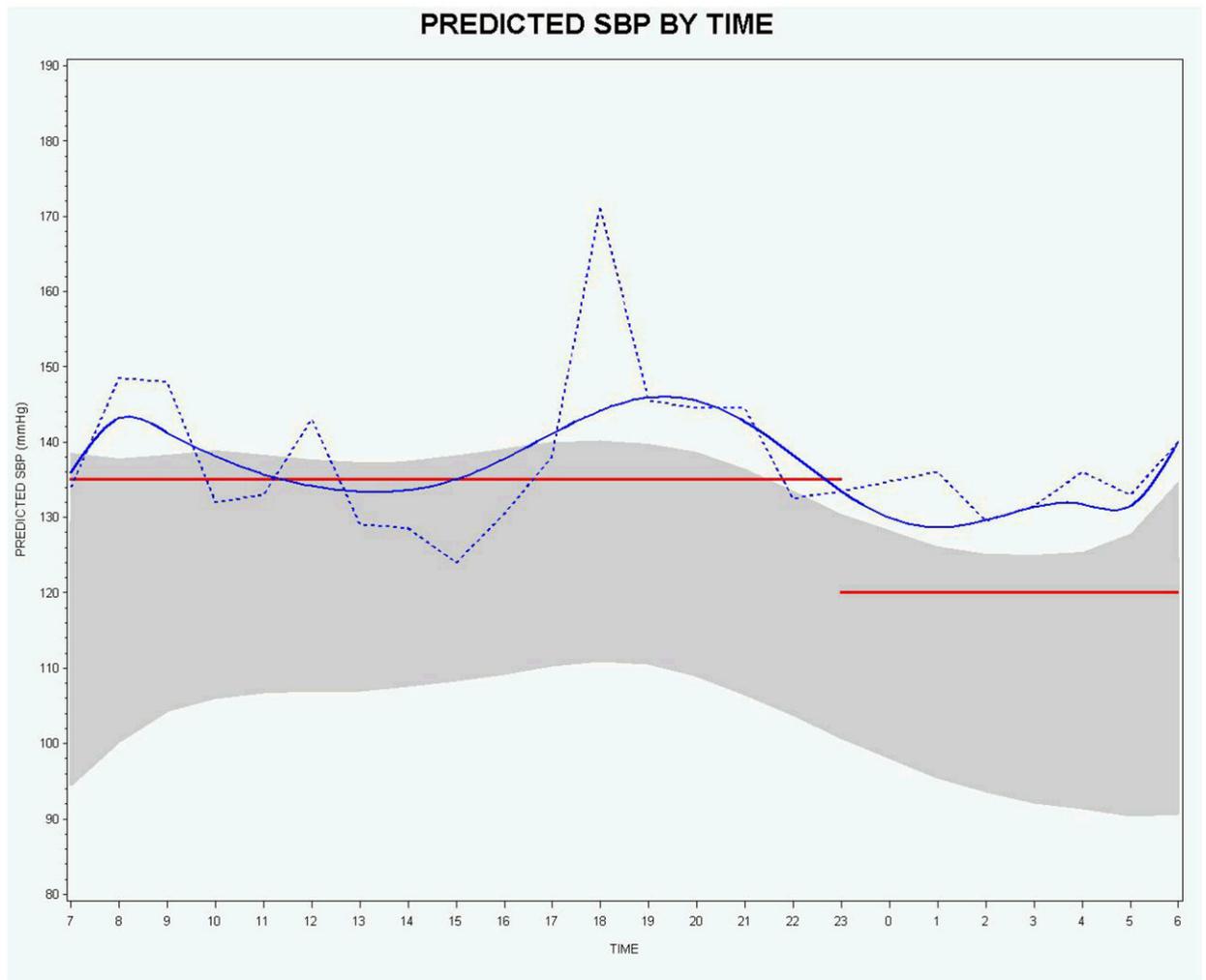


Figure 1. Predicted (blue solid curve) and observed (blue dotted curve) SBP by time for 1 subject with the 90% prediction interval (shaded region) based on normal subjects and commonly used upper normal limit static measures for daytime and nighttime (red horizontal lines).

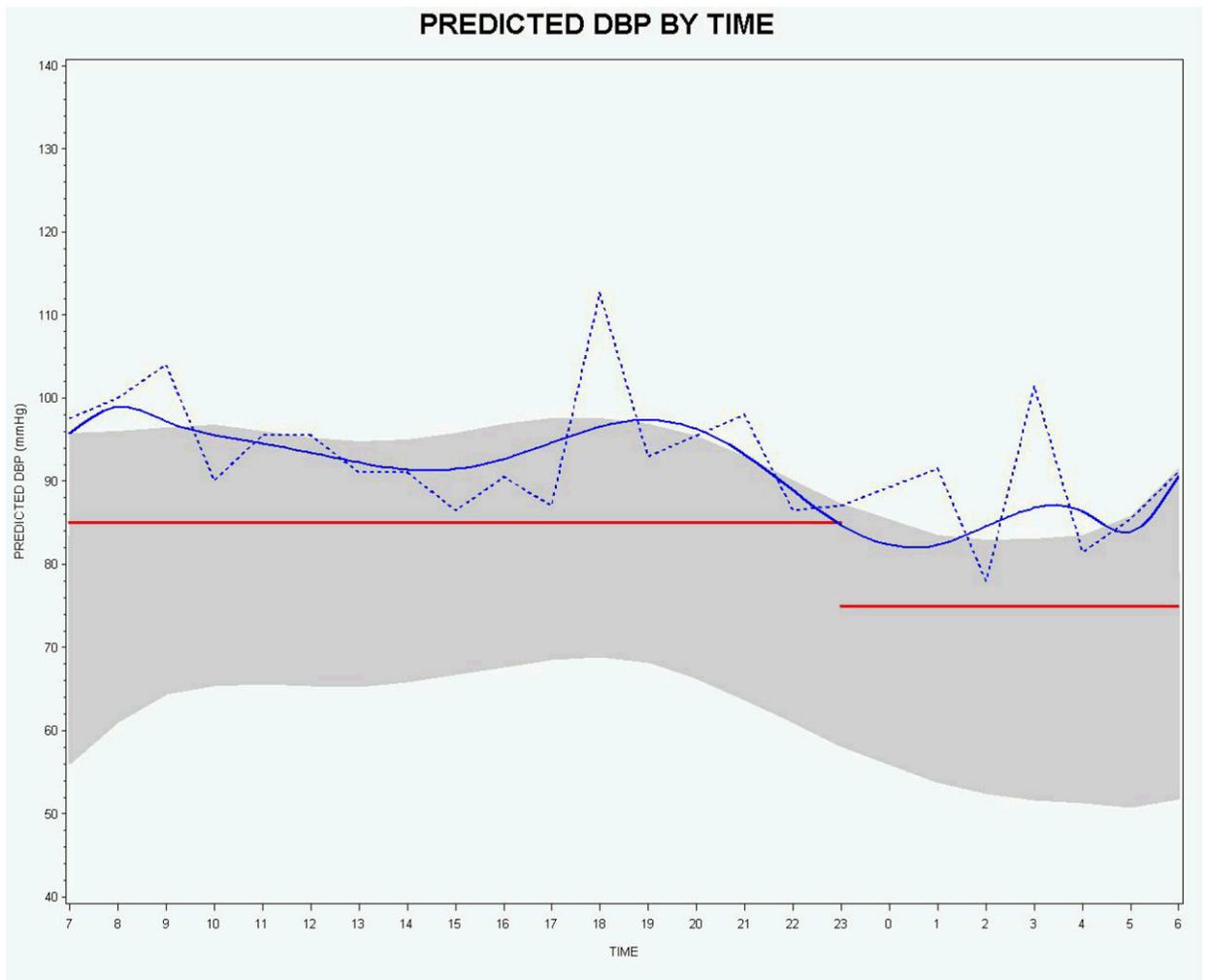


Figure 2. Predicted (blue solid curve) and observed (blue dotted curve) DBP by time for 1 subject with the 90% prediction interval (shaded region) based on normal subjects and commonly used upper normal limit static measures for daytime and nighttime (red horizontal lines).

Table 1

Fixed Effect Estimates, Standard Errors (SE), and P-values for 9th Degree Orthornormal Polynomial Fit to Baseline DASH Study Data, N=357

Outcome	Parameter	Ortho Poly Degree	Estimate	SE	P-value
DBP	β_0	Intercept	83.71	0.394	<0.0001
	β_1	1	-22.98	0.795	<0.0001
	β_2	2	-13.41	0.748	<0.0001
	β_3	3	9.91	0.605	<0.0001
	β_4	4	9.61	0.584	<0.0001
	β_5	5	4.35	0.518	<0.0001
	β_6	6	-4.62	0.436	<0.0001
	β_7	7	-0.51	0.471	0.2768
	β_8	8	1.96	0.420	<0.0001
SBP	β_9	9	0.58	0.428	0.1782
	β_0	Intercept	131.64	0.575	<0.0001
	β_1	1	-25.03	0.972	<0.0001
	β_2	2	-19.71	0.909	<0.0001
	β_3	3	10.18	0.732	<0.0001
	β_4	4	10.74	0.708	<0.0001
	β_5	5	5.19	0.614	<0.0001
	β_6	6	-4.38	0.528	<0.0001
	β_7	7	-0.74	0.542	0.1722
β_8	8	1.13	0.466	0.0153	
	β_9	9	1.43	0.470	0.0025

Table 2

Model Comparisons: AIC and BIC, FE=Fixed Effects, RE=Random Effects, Baseline DASH Study Data, N=357

Outcome	Model	$\Sigma_{diff}(\eta_i) = \text{diag}(\eta_i)$		$\Sigma_{diff}(\eta_i) = UN$	
		AIC	BIC	AIC	BIC
DBP	Ortho Poly (FE_degree = 9, RE_degree = 9)	58,989	59,032	58,710	58,927
	Ortho Poly (FE_degree = 8, RE_degree = 8)	59,031	59,070	58,784	58,963
	Ortho Poly (FE_degree = 7, RE_degree = 7)	59,089	59,124	58,873	59,016
	Ortho Poly (FE_degree = 6, RE_degree = 6)	59,158	59,189	58,962	59,074
	Ortho Poly (FE_degree = 5, RE_degree = 5)	59,344	59,371	59,180	59,266
	Ortho Poly (FE_degree = 4, RE_degree = 4)	59,570	59,593	59,502	59,564
	Ortho Poly (FE_degree = 9, RE_degree = 9)	61,598	61,640	61,336	61,553
	Ortho Poly (FE_degree = 8, RE_degree = 8)	61,637	61,676	61,410	61,588
	Ortho Poly (FE_degree = 7, RE_degree = 7)	61,665	61,699	61,452	61,596
	Ortho Poly (FE_degree = 6, RE_degree = 6)	61,741	61,772	61,543	61,655
SBP	Ortho Poly (FE_degree = 5, RE_degree = 5)	61,901	61,928	61,748	61,833
	Ortho Poly (FE_degree = 4, RE_degree = 4)	62,164	62,188	62,091	62,153

Table 3

Baseline and Intervention Diet Effects (p-values) for 9th Degree Orthornormal Polynomial Model Fit to DASH Data, N=357

Outcome	Change from Baseline to Intervention		
	Control	Fruit/Veg	DASH
SBP	-0.29 (0.1427)	-3.23 (< 0.0001)	-4.55 (< 0.0001)
DBP	-0.02 (0.8906)	-2.12 (< 0.0001)	-2.84 (< 0.0001)

AUC Descriptive Statistics for SBP and DBP: Control, DASH, and Fruit/Veg. Baseline (N=288) and Intervention (N=287) DASH Study, Complete Data

Table 4

Diet	Outcome	Baseline		Intervention		Mean Change	
		N	Mean* (SE)	N	Mean* (SE)	AUC	AUC/24
Control	SBP	101	3026 (23.7)	97	3013 (25.9)	13	0.5
	DBP		1923 (16.2)		1928 (19.3)	-5	0.2
DASH	SBP	89	3045 (25.4)	97	2920 (20.1)	125	5.2
	DBP		1923 (15.8)		1852 (12.7)	71	3.0
Fruit/Veg	SBP	98	3028 (21.4)	93	2978 (24.6)	50	2.1
	DBP		1927 (17.1)		1888 (17.8)	39	1.6
Total	SBP	288	3032 (13.5)	287	2970 (13.8)	62	2.6
	DBP		1924 (9.4)		1889 (9.8)	35	1.5

* Mean of subject-specific AUC computed from LMM