LA Nurse BP

Case Study 1*

Kunwu Lyu

Katie St. Clair, STAT 330
Feburary 13, 2025

Executive Summary

Here is what I did.

1 Introduction

Certain traits such as family history and mood are expected to increase one's ambulatory blood pressure (BP). Goldstein and Shapiro (2000) studied potential factors that contribute to hypertension. They collected information about the participants' BP, activity levels, work status, and mood ratings throughout the day, as well as relevant family history and information about their menstrual phases, to establish links that lead to elevated BP. Towards the end, they sought to uncover preventative measures for individuals who may be at a higher risks of developing hypertension. The objectives of this project is much simpler. I am interested in exactly what traits are *associated* with elevating one's BP, given the longitudinal structure and various time-dependent metrics in the dataset (given by Goldstein & Shapiro, 2000; cited by Roback & Legler, 2021).

1.1 Methods

The dataset includes repeated measures over the course of two work and off-work days on 203 registered nurses between the ages of 24 and 50 years working in Los Angeles, in the year 2000. Of those 203 nurses, 172 has complete data on all of the variables recorded.[1] BP of the participants were measured 30 minutes before their normal start of work, and measured measured repeatedly every 20 minutes for the rest of the day. This led to around 40-60 observations per nurse (9573 total observations). Each time the BP was taken, participants were asked to give several mood ratings including happiness, stress, and tiredness. In addition, participants wore an actigraph on their waist to record frequency of movements in one-minute intervals; the researchers obtained an activity measure for the ten-minute periods before each BP reading.

*Code used for this project is available at this GitHub Repository.

[1] In the original paper by Goldstein and Shapiro (pp. 228–29), they reported 171 nurses who completed all sessions. It is likely that this is due to Roback and Legler (2021) or Weiss (2005) excluding other personality variables used in the original study. Because I am not investigating personality traits in this project, 172 suffices being my total subject counts. Additionally, Goldstein and Shapiro mention that "[s]imilar patterns of findings were obtained in the sample of 171 as in the total sample" (p. 229).

Variable Name	Variable Descrption
ID (Cluster)	Unique identification number for each participant
BP (Response)	Systolic blood pressure (in mmHg)
Act	Activity level (frequency of movements in 1-minute intervals, over a 10-minute period)
Phase	Menstrual phase (follicular or luteal)
Day	Workday or non-workday
Posture	Position during BP measurement (sitting, standing, or reclining)
HAP	Self-ratings of happiness by each nurse at the time of each BP measurement on a 5-point scale (5 strongest and 1 weakest)
STR	Self-ratings of stress by each nurse at the time of each BP measurement on a 5-point scale (5 strongest and 1 weakest)
TIR	Self-ratings of tiredness by each nurse at the time of each BP measurement on a 5-point scale (5 strongest and 1 weakest)
Age	Age (in years)
Full FH	Family history, coded as either NO (no family history of hypertension), YES (1 hypertensive parent), or YESYES (both parents hypertensive)
Timepass	Number of minutes since the first measurement
Stand	Indicator variable for standing, where it equals 0 if Posture is either sitting or reclining and 1 when Posture is standing
Mood	Combined mood ratings: HAP - (STR + TIR)/2
FH	Indicator variable for having family history of hypertension (1 for YES and YESYES, 0 for NO)
Age24	Recentered Age: Age - 24

Note: *emphasis* added to re-parameterized variables; colors represent levels.

Table 1: Variable Descrptions

2 Exploratory Data Analysis

The variables (original and re-parameterized) used in this project are given in Table 1. Variable missingness was explored in Figure 1. We see that the data were missing primarily mood ratings and some activity levels from several participants. Because I am interested in how both of these variables relate to BP, I will proceed the analysis with the missing rows removed, leaving us with a total of 172 participants and 7877 observations. The visuals and statistics below will be given in terms of that subset of the data, unless otherwise specified.

Because of the longitudinal nature of our data, it is worth noting our levels of analysis and to which our variables belong. I will call time-dependent measurements *Level 1* data (variables) and the rest of the subject-level measurements *Level 2* data (variables). Each participant has a unique identification number; this will distinguish between different clusters. Our primary response is systolic blood pressure (BP), and I will use our exploratory data analysis below to guide our model selection process from the bottom-up.

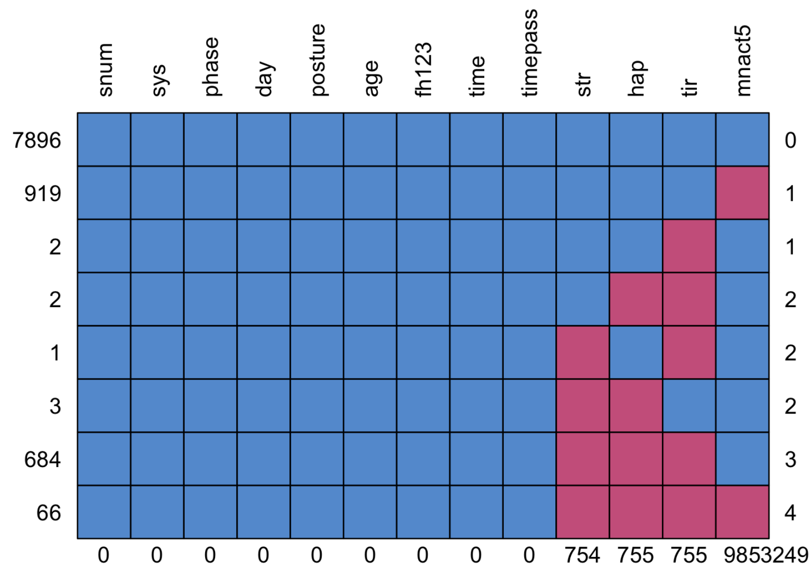


Figure 1: Data Missingness

Note that I created three additional variables from the original dataset to help simplify the modeling process. In the original dataset, `Posture` is a factor variable with three levels: Recline ($n = 530$), Sit ($n = 3644$), and Stand ($n = 3703$). Because the relatively small number of observations where participants were reclining (presumably sleeping at night or resting), I collapsed the three levels into either Standing or non-Standing. Similarly, because of the relatively small number of participants have both parents with hypertension ($n = 13$), I collapsed it together with having one parent with hypertension ($n = 66$) to compare against those without any family history ($n = 103$). Lastly, a general `Mood` measurement was created by subtracting the average of tiredness and stress from happiness.

2.1 Level 1 by Clusters

I will first explore how BP vary among participants and how it relates to the level 1 predictors to examine if there is any sign of clustering between different participants (which is the premise of a longitudinal study). From Figure 2, we see that BP readings do tend to vary among different participants, though variations within each participant do not seem to be pronounced. This suggests that a Linear Mixed Model (LMM) might be more appropriate than a regular Multiple Linear Regression (MLR).

Since our data is ordered in time, one might be interested in answering how the participants' BP evolve over time. For sake of brevity, I show a random subsets of subjects and their BP readings over time in Figure 3a. From this plot alone, not much could be deciphered; some trends (e.g., participant 1116) seem to decrease over time while others either slightly increase (e.g., participant 1267) or do not see much fluctuation overall. Indeed, when I fit separate linear smoothers for each participant, I see various different intercepts and slopes for each

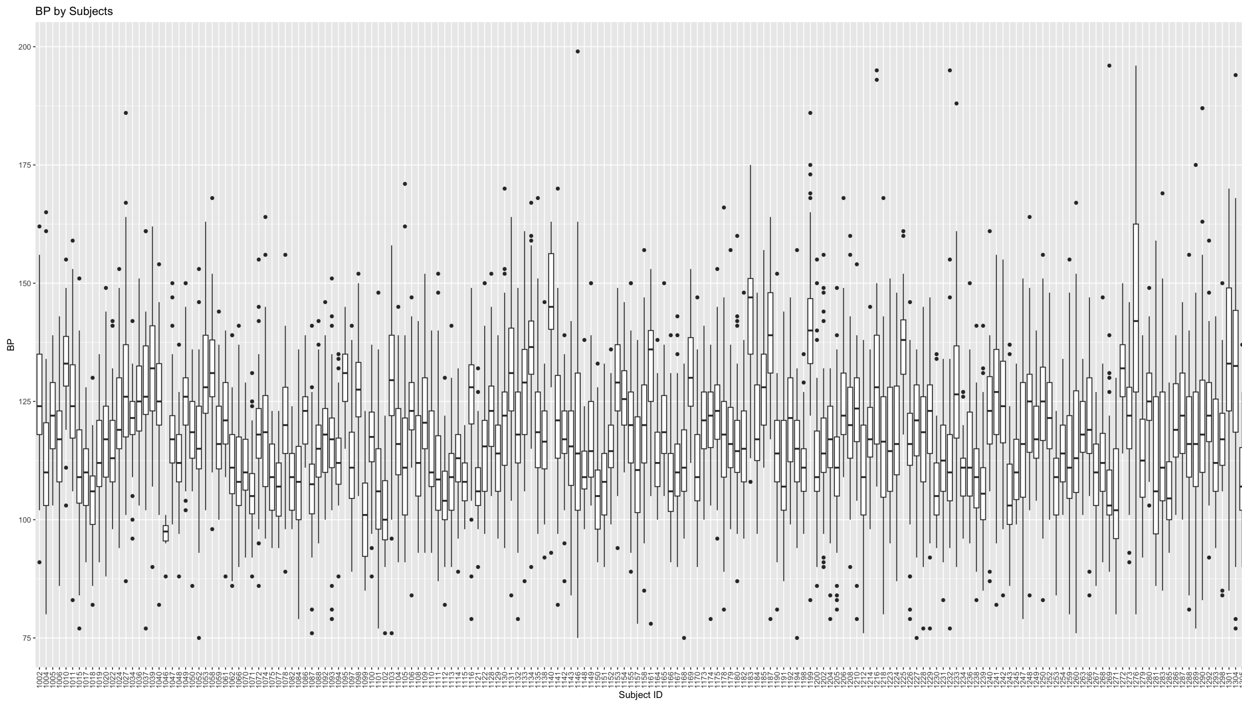


Figure 2: BP by Subjects

participant (see Figure 3b). This suggests that a random effect for both the intercept and slope of time.

We perform similar analysis for other level 1 covariates (see Figure 4). We see the most prominent effect in `Mood`, where the slopes and intercepts for each participant differs the most. They all show, to a certain extend, signs of varying slopes and intercepts, but I cannot definitely conclude anything at this point.

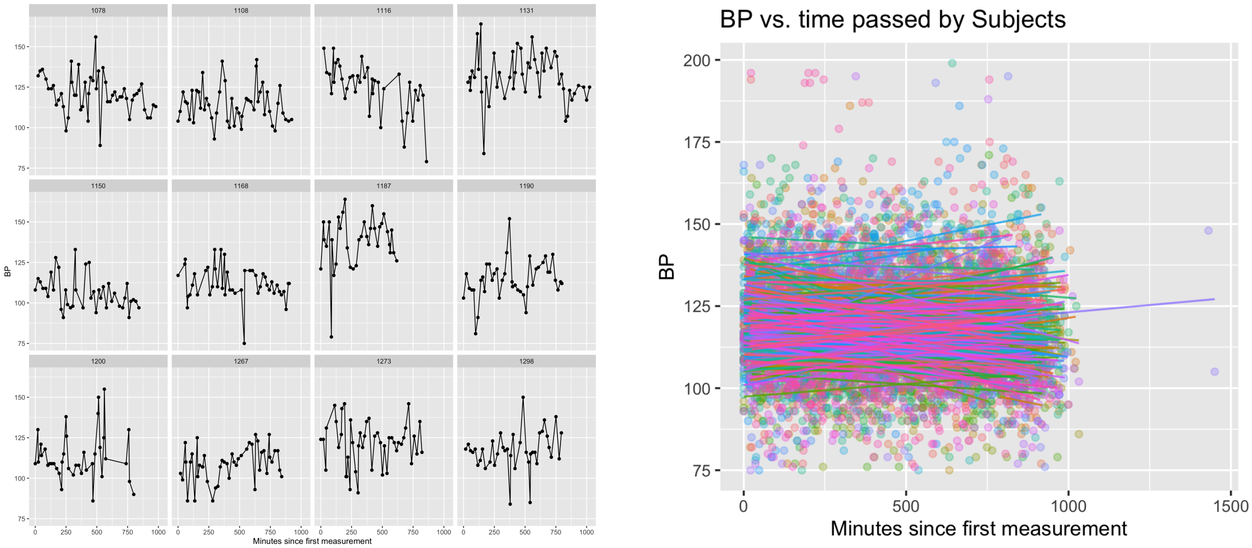
2.2 Level 2 Covariates

We then want to understand the distributions of our level two covairates before we look at any interactions between the two levels. Because our level 2 covariates contains the same measurement over time (whereas our primary response `BP` is level 1 and hence time-dependent), we will compare the level 2 covariates with the *average* `BP` of a given participant.

From Figure 5a, we do not see that being on a different menstural phase impacts the participants' average `BP` much at all, suggesting that we likely do not need to control for `Phase` as a level 2 predictor. We see that in a workday, as opposed to a non-workday, the average `BP` of the participants appears to have a higher mean (see Figure 5b).

Similarly for family history; having at least one parent with hypertension bumps the participants' average `BP` by a marginal amount (Figure 5d). This suggest that we might want to look into having `Day` and `FH` in our model as control or predictors; we will return to their interaction with level 1 predictors in the next section.

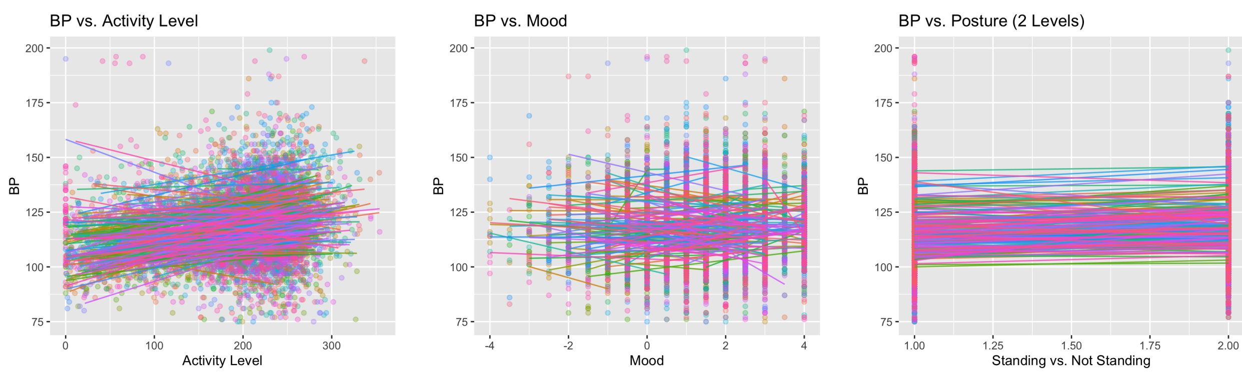
Lastly, it is unclear how age impacts average `BP`, since Figure 5c shows that the effect of age varies significantly (though a smoother line shows slight positive increase as age increases). We will test this formally in Section 3.



(a) BP vs. Time Passed by Subjects

(b) BP vs. Time Passed Separate Lines

Figure 3: BP vs. Minutes since first measurement

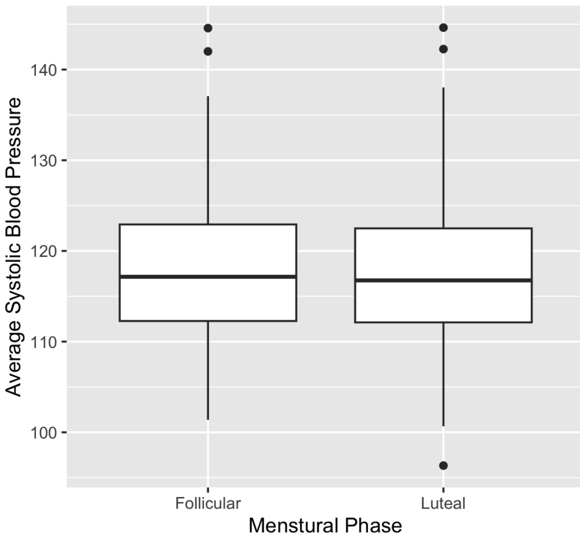


(a) BP vs. Activity Level by Subjects

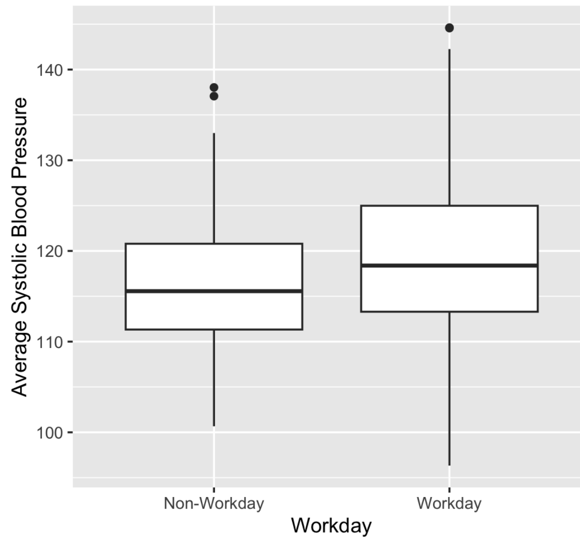
(b) BP vs. Mood Ratings by Subjects

(c) BP vs. Posture (Standing) by Subjects

Figure 4: BP vs. Level 1

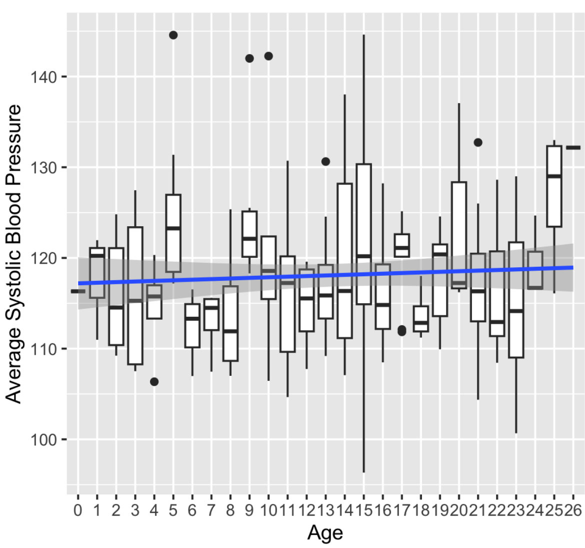


(a) Avg BP vs. Menstrual Phase

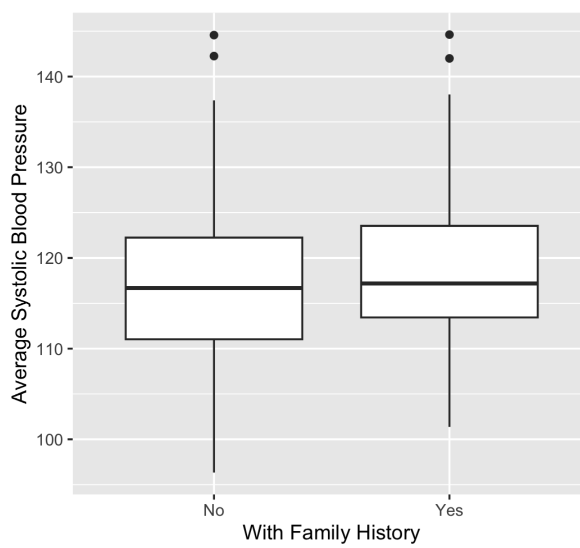


(b) Avg BP vs. Workday

Figure 5: Average BP vs. Level 2 Covariates



(c) Avg BP vs. Age



(d) Avg BP vs. Family History

Figure 5: Average BP vs. Level 2 Covariates

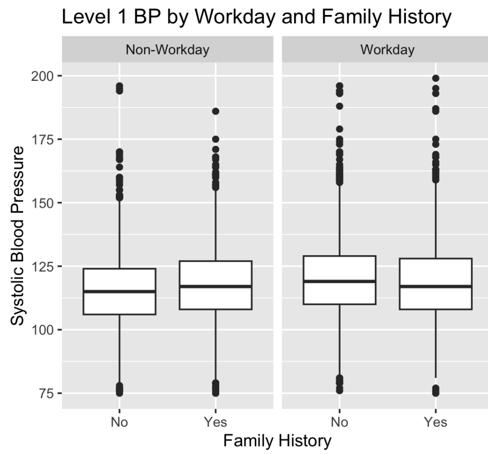


Figure 6: Level 1 BP by Workday and Family History

2.3 Level 1 by Level 2 Covariates

Next we investigate whether our level 1 covariates depend, in any way, on our level 2 covariates. There are several ways one could explore this.[2] In Section 2.2, we suspect that `BP` depends on `Workday` and `FH`. Instead of showing the average `BP`, we can also show simply the level 1 responses, faceted by workday and family history (see Figure 6).

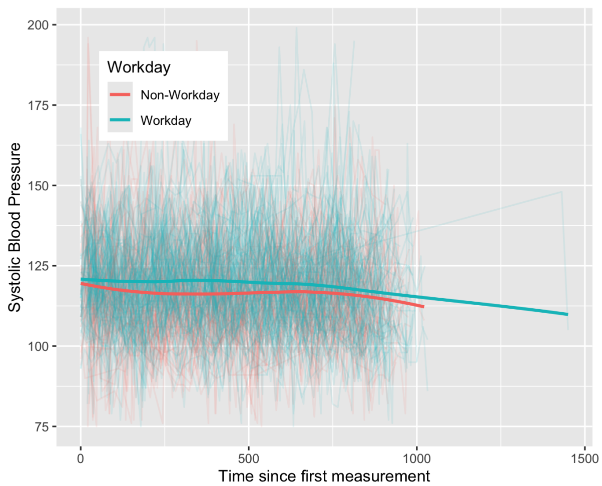
To better understand the interaction between level 2 covairates and the level 1 covairates, however, it is most clear to explore with spaghetti plots or separate LS models to see the difference in intercepts and slopes. We first explore with spaghetti plots. From Figure 7 alone, it is hard to say anything conclusive about whether one of the level one covariates depend on `Workday`; there all seem to be marginal difference in intercepts, but slopes seem largely similar. Figure 8 likewise shows inconclusive results in terms of `Family History` (though there appears to be some difference in slopes in Mood by `FH`). We proceed to analyze separate MLR models containing level 1 variables for each participant. This also allows us to better understand the continuous variable `Age24`.

After fitting separate MLR models for each participant, we see that the intercepts appears to depend on `Workday` and `Age` (see Figure 9b and Figure 9a), the slopes for `Time` appears to depend on `Family History` and `Workday` (see Figure 10c and Figure 10b), and the fixed effect for `Standing` appears to depend on `Age` and `Workday` (see Figure 12a and Figure 12b).

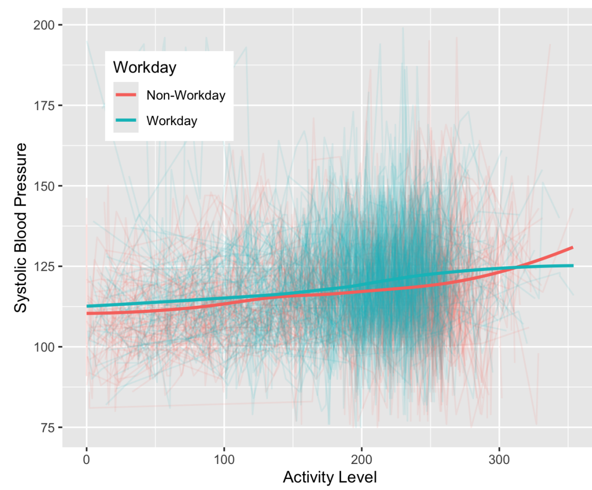
3 Model Selection

We will use a bottom-up approach to build a saturated model from scratch. Since, our data is ordered by time, we start with a base model (Model 1) with only the random intercept and level 1 covariate time: for the j^{th} measurement in the i^{th} participant, where $i = 1, \ldots, n$

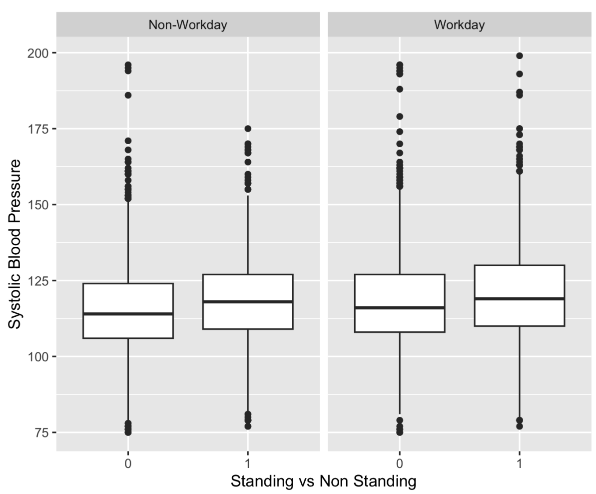
[2] Additional visualizations are shown in Appendix A.



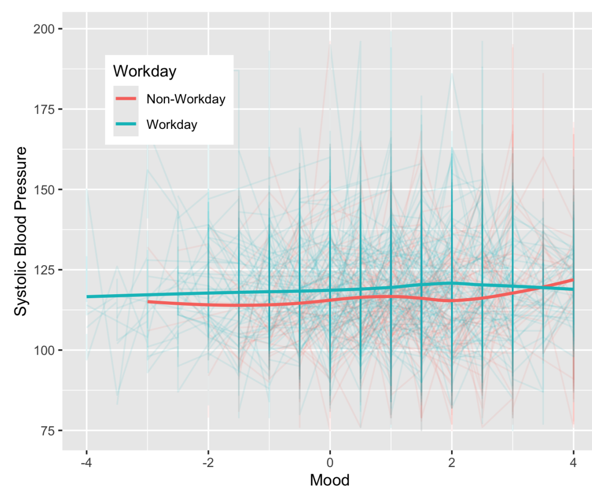
(a) L1 BP vs. Time by Workday



(b) L1 BP vs. Activity by Workday

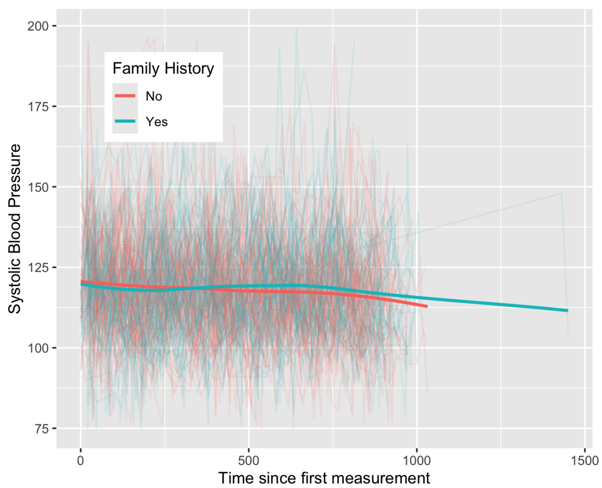


(c) L1 BP vs. Standing by Workday

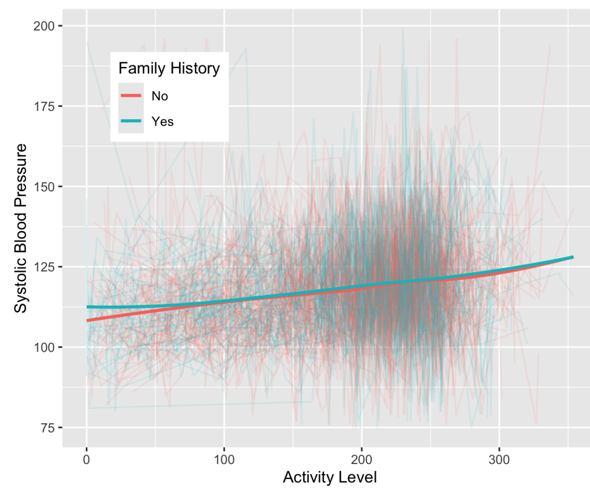


(d) L1 BP vs. Mood by Workday

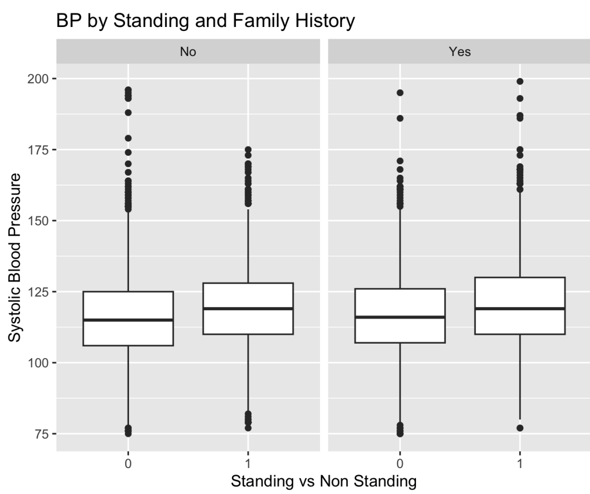
Figure 7: L1 BP vs. Level 1 Covariates by Workday



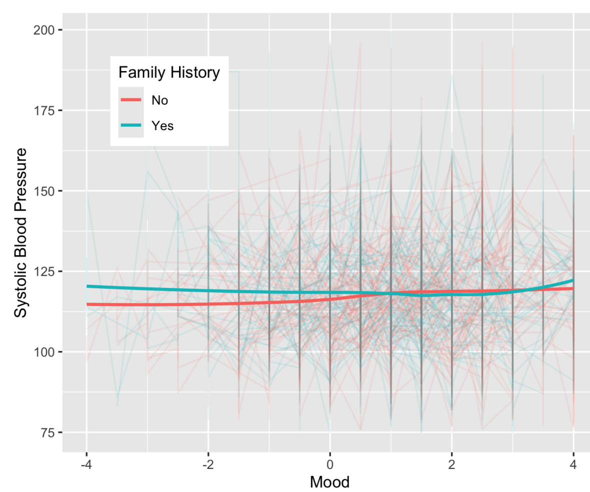
(a) L1 BP vs. Time by Family History



(b) L1 BP vs. Activity by Family History

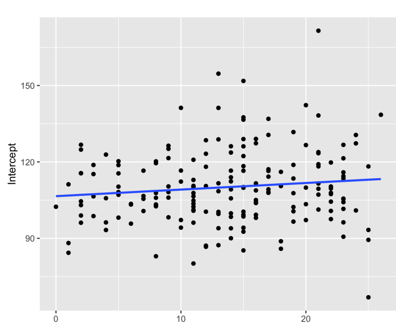


(c) L1 BP vs. Standing by Family History

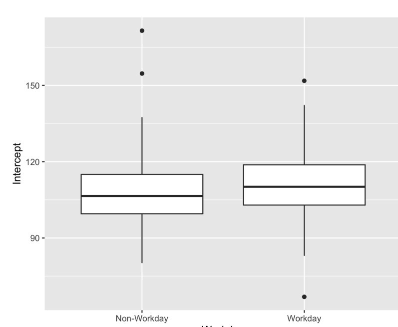


(d) L1 BP vs. Mood by Family History

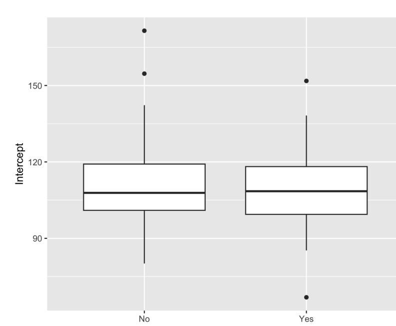
Figure 8: L1 BP vs. Level 1 Covariates by Family History



(a) MLR Intercepts by Age

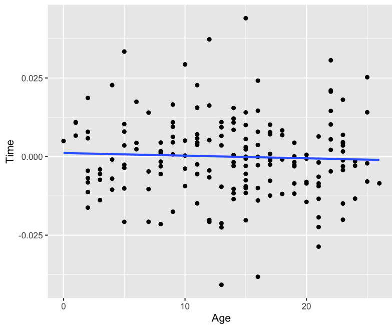


(b) MLR Intercepts by Workday

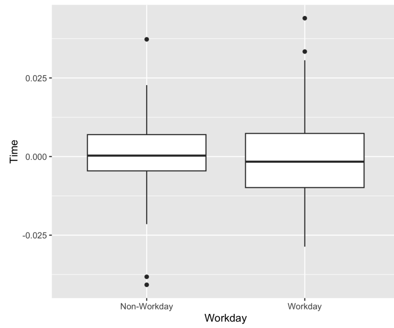


(c) Intercepts by Family History

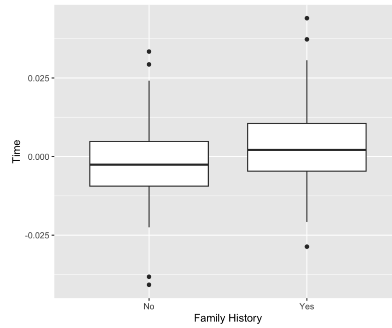
Figure 9: MLR Intercepts by Level 2 Covariates



(a) MLR Time by Age

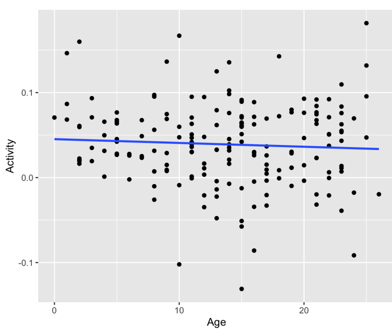


(b) MLR Time by Workday

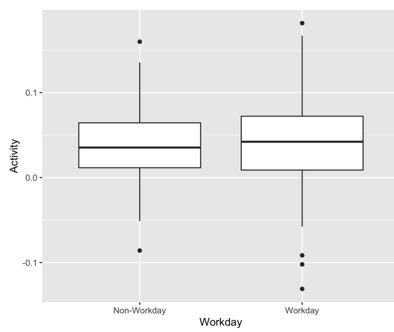


(c) Time by Family History

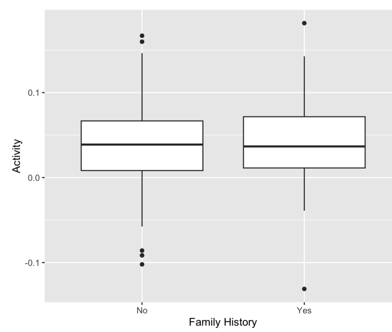
Figure 10: MLR Time Slopes by Level 2 Covariates



(a) MLR Activity by Age

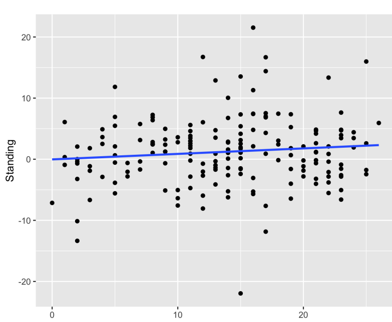


(b) MLR Activity by Workday

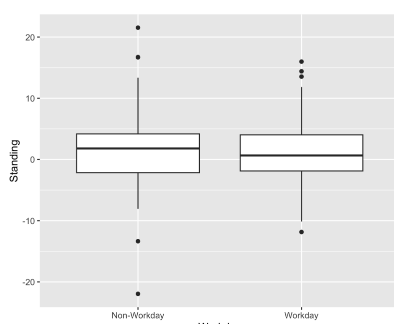


(c) Activity by Family History

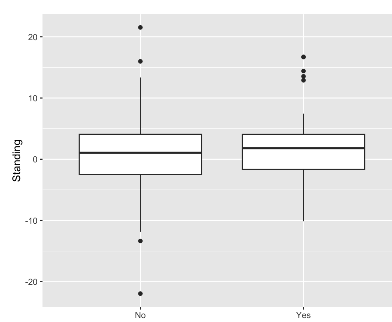
Figure 11: MLR Activity Slopes by Level 2 Covariates



(a) MLR Standing by Age

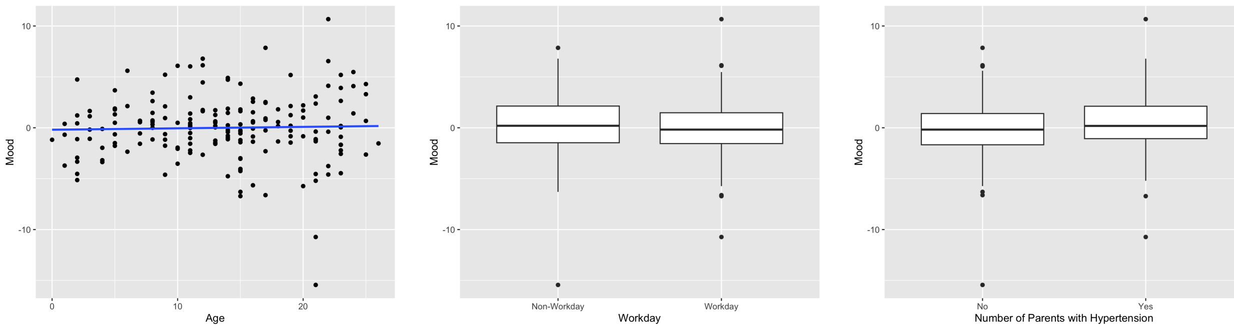


(b) MLR Standing by Workday



(c) Standing by Family History

Figure 12: MLR Standing Fixed Effects by Level 2 Covariates



(a) MLR Mood by Age (b) MLR Mood by Workday (c) Mood by Family History

Figure 13: MLR Mood Slopes by Level 2 Covariates

and $j = 1, \ldots, n_i$,

$$\begin{aligned} \text{Level 1}: \quad \text{BP}_{ij} &= a_i + \beta_1 \,\text{Time}_{ij} + \epsilon_{ij}, & \epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2) \\ \text{Level 2}: \quad a_i &= \alpha + u_i, & u_i &\overset{i.i.d.}{\sim} N(0, \sigma_u^2). \end{aligned} \tag{1}$$

A Likelihood Ration Test (LRT) reveals that compared to a Ordinary Least Squares (OLS) model with no random intercept, the random effect is significant ($\chi^2(0,\, 1) = 2044.1$, $p < 2.2 \times 10^{-16}$). Note that here the p-value is a mixture from a chi-square model with 0 degrees of freedom and 1 degree of freedom.

We then proceed to add in fixed effects for level 1 to investigate whether activity level, whether or not the participants were standing, and mood ratings were *associated* with elevating BP. We propose the following (Model 2) with a random intercept and all level 1 covariates:

$$\begin{aligned} \text{Level 1}: \quad \text{BP}_{ij} &= a_i + \beta_1 \,\text{Time}_{ij} + \beta_2 \,\text{Activity}_{ij} + \beta_3 \,\text{Standing}_{ij} + \\ & \quad\ \beta_4 \,\text{Mood}_{ij} + \epsilon_{ij}, \\ \epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2), \\ \text{Level 2}: \quad a_i &= \alpha + u_i, \\ u_i &\overset{i.i.d.}{\sim} N(0, \sigma_u^2). \end{aligned} \tag{2}$$

A LRT between Model 1 and 2 reveals that the three added fixed effects are statistically significant ($\chi^2(3) = 365.37$, $p < 2.2 \times 10^{-16}$). Naturally, one might wonder if we need all three fixed effect. Indeed, a LRT suggests that the following Model $2'$ without the covariate `Mood` ($\beta_4 = 0$) is just as good as Model 2 ($\chi^2(1) = 0.43$, $p = 0.51$). Because we are interested in *whether* something is associated with elevated `BP` at all, we proceed the analysis with the simpler Model $2'$.

$$\begin{aligned} \text{Level 1}: \quad \text{BP}_{ij} &= a_i' + \beta_1' \,\text{Time}_{ij} + \beta_2' \,\text{Activity}_{ij} + \beta_3' \,\text{Standing}_{ij} + \epsilon_{ij}, \\ \epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2), \\ \text{Level 2}: \quad a_i' &= \alpha' + u_i, \\ u_i &\overset{i.i.d.}{\sim} N(0, \sigma_u^2). \end{aligned} \tag{2'}$$

Interestingly, LRT tests also suggest that compared to Model 2′, a model without the `Time` term ($\beta_1' = 0$) is also just as good ($\chi^2(1) = 0.99$, $p = 0.32$). However, because time is such an important aspect of the longitudinal model, we think it is important that we keep `Time` as a level 1 covariate in our model.

Next we investigate what random effects are needed in our model. The most clear evidence of random slopes from our EDA comes from `Mood`, but since LRT tests suggest that the fixed effect of `Mood` is not significant, we will not consider its random effect. For the rest of the covariates, EDA do show marginal signs of varying slopes. However, only adding a random effect for `Standing` do not raise convergence issues. This leads us to the following model with a random effect for `Standing`:

$$\begin{aligned}
\text{Level 1}: \quad \text{BP}_{ij} &= a_i + \gamma_1 \,\text{Time}_{ij} + \gamma_2 \,\text{Activity}_{ij} + b_i \,\text{Standing}_{ij} + \epsilon_{ij}, \\
\epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2), \\
\text{Level 2}: \quad a_i &= \alpha + u_i, \\
b_i &= \beta + v_i, \\
\begin{bmatrix} u_i \\ v_i \end{bmatrix} &\overset{i.i.d.}{\sim} N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right).
\end{aligned} \tag{3}$$

A LRT suggests that the random effect for `Standing` is significant, based on a mixture from a chi-square model with 1 and 2 degrees of freedom ($\chi^2(1, 2) = 377.34.1$, $p < 2.2 \times 10^{-16}$). Interestingly, removing the correlation term (ρ_{uv}) between the random effect for the intercept and the random effect for `Standing` causes further convergence issues, so we preserve the it in our following analysis.

We move on to consider the interaction between level 2 and level 1 covaraites. Recall that our EDA identifies `Age`, `Workday`, and `Family History` as possibly interacting with the effect for `Standing`, the effects for `Time` and `Standing`, and the effect for `Time`, respectively. We will first consider whether the intercept and the effect of `Standing` depend on `Age` (recentered at 24):

$$\begin{aligned}
\text{Level 1}: \quad \text{BP}_{ij} &= a_i + \gamma_1 \,\text{Time}_{ij} + \gamma_2 \,\text{Activity}_{ij} + b_i \,\text{Standing}_{ij} + \epsilon_{ij}, \\
\epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2), \\
\text{Level 2}: \quad a_i &= \alpha + \alpha_1 \,\text{Age}_i + u_i, \\
b_i &= \beta + \beta_1 \,\text{Age}_i + v_i, \\
\begin{bmatrix} u_i \\ v_i \end{bmatrix} &\overset{i.i.d.}{\sim} N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right).
\end{aligned} \tag{4}$$

When compared with Model 3 ($\alpha_1 = \beta_1 = 0$), a LRT suggests that the smaller model without `Age` is just as good as the full model with the `Age` term ($\chi^2(2) = 0.97$, $p = 0.61$). Because this project concerns *whether* a covaraite is associated with elevated `BP`, we will proceed with the smaller model without `Age`.[3]

[3]Note that Goldstein and Shapiro (2000) does include `Age` as a predictor because they are interested in preventative measures *controlling for* age. This is not the aim of this project, even though `Age` is a meaningful and important predictor.

Next, we consider the interaction of `Workday` between `Time` and `Standing`.

$$
\begin{aligned}
\text{Level 1}: \quad \mathrm{BP}_{ij} &= a_i + b_i \,\mathrm{Time}_{ij} + \delta \,\mathrm{Activity}_{ij} + c_i \,\mathrm{Standing}_{ij} + \epsilon_{ij}, \\
\epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2), \\
\text{Level 2}: \quad a_i &= \alpha_0 + \alpha_1 \,\mathrm{Workday}_i + u_i, \\
b_i &= \beta_0 + \beta_1 \,\mathrm{Workday}_i \\
c_i &= \gamma_0 + \gamma_1 \,\mathrm{Workday}_i + v_i, \\
\begin{bmatrix} u_i \\ v_i \end{bmatrix} &\overset{i.i.d.}{\sim} N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right).
\end{aligned}
\tag{5}
$$

A LRT suggests that, when compared to Model 3 ($\alpha_1 = \beta_1 = \gamma_1 = 0$), at least one of coefficients is statistically significantly different from zero; we have some evidence to reject the Model 3 and prefer Model 5 for better model fit ($\chi^2(3) = 8.22$, $p = 0.04$). However, when one removes the interaction between `Workday` and `Standing` ($\gamma_1 = 0$), a LRT test suggests that the smaller model is just as good ($\chi^2(1) = 0.45$, $p = 0.50$); we drop the γ_1 for the sake of simplicity, which, when compared to the simpler Model 3 ($\alpha_1 = \beta_1 = 0$), still presents better explanatory power ($\chi^2(2) = 7.77$, $p = 0.02$). Nevertheless, if we were to remove the interaction between `Time` and `Workday` ($\beta_1 = \gamma_1 = 0$ vs. $\gamma_1 = 0$), a LRT suggests that we have some evidence to prefer having the interaction between `Time` and `Workday` for a better model fit ($\chi^2(1) = 4.46$, $p = 0.03$):

$$
\begin{aligned}
\text{Level 1}: \quad \mathrm{BP}_{ij} &= a_i + b_i \,\mathrm{Time}_{ij} + \delta \,\mathrm{Activity}_{ij} + c_i \,\mathrm{Standing}_{ij} + \epsilon_{ij}, \\
\epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2), \\
\text{Level 2}: \quad a_i &= \alpha_0 + \alpha_1 \,\mathrm{Workday}_i + u_i, \\
b_i &= \beta_0 + \beta_1 \,\mathrm{Workday}_i \\
c_i &= \gamma_0 + v_i, \\
\begin{bmatrix} u_i \\ v_i \end{bmatrix} &\overset{i.i.d.}{\sim} N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right).
\end{aligned}
\tag{5$'$}
$$

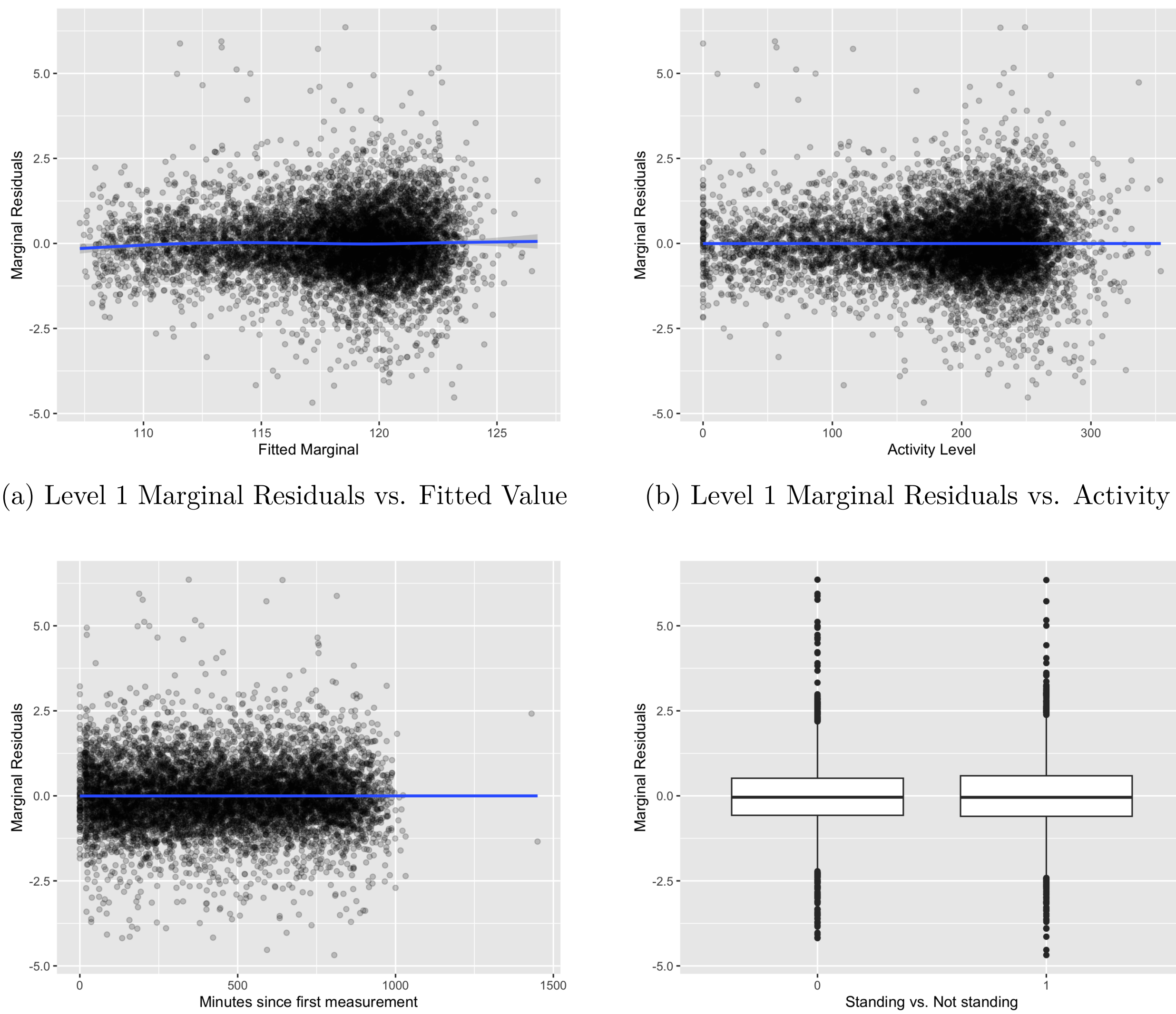
Lastly, we will consider the interaction between `Family History` and `Time`:

$$
\begin{aligned}
\text{Level 1}: \quad \mathrm{BP}_{ij} &= a_i + b_i \,\mathrm{Time}_{ij} + \delta \,\mathrm{Activity}_{ij} + c_i \,\mathrm{Standing}_{ij} + \epsilon_{ij}, \\
\epsilon_{ij} &\overset{i.i.d.}{\sim} N(0, \sigma^2), \\
\text{Level 2}: \quad a_i &= \alpha_0 + \alpha_1 \,\mathrm{Workday}_i + \alpha_2 \,\mathrm{Family\ History}_i + u_i, \\
b_i &= \beta_0 + \beta_1 \,\mathrm{Workday}_i + \beta_2 \,\mathrm{Family\ History}_i \\
c_i &= \gamma_0 + v_i, \\
\begin{bmatrix} u_i \\ v_i \end{bmatrix} &\overset{i.i.d.}{\sim} N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right).
\end{aligned}
\tag{6}
$$

When compared with 5$'$ ($\alpha_2 = \beta_2 = 0$), a LRT suggests that there is strong evidence to reject the smaller model and prefer having the interaction between `Time` and `Family History` for better model fit ($\chi^2(2) = 16.3$, $p = 0.00029$). As before, removing the correlation between random effects results in convergence issues. We will use Model 6 as our final model for interpretation.

3.1 Residual Analysis

We first consider the level 1 *marginal residuals* for each of the predictors as well as fitted values. From Figure 14a, there are some slight curvature in the fitted values, but not a huge concern. We see a similar pattern in `Activity Levels` (Figure 14b), but generally not a huge concern. `Time` and `Standing` looks fine (see Figure 14c and Figure 14d).



(a) Level 1 Marginal Residuals vs. Fitted Value

(b) Level 1 Marginal Residuals vs. Activity

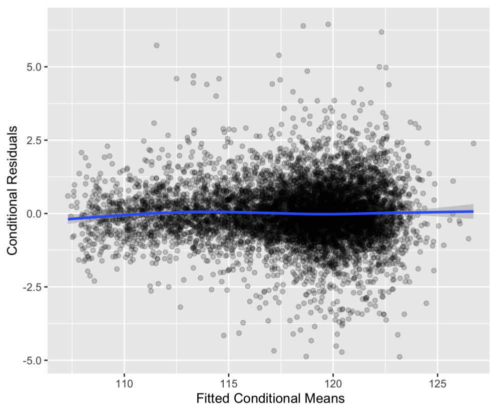
(c) Level 1 Marginal Residuals vs. Time

(d) Level 1 Marginal Residuals vs. Standing

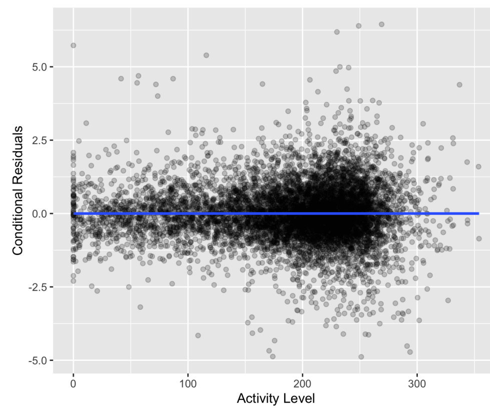
Figure 14: Level 1 Marginal Residuals vs. Predictors

Because we see some worry of constant variance in the fitted values, we proceed to check that assumption with the *conditional residuals*. As before, `Time` and `Standing` looks fine (see Figure 15c and Figure 15d). However, we see that for fitted values (Figure 15a) and `Activity`

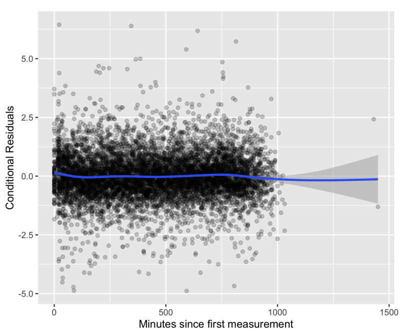
`Levels` (Figure 15b), we have higher variations for higher values of fitted values and `Activity Levels`, respectively. The conditional residuals are, however, very heavy-tailed.



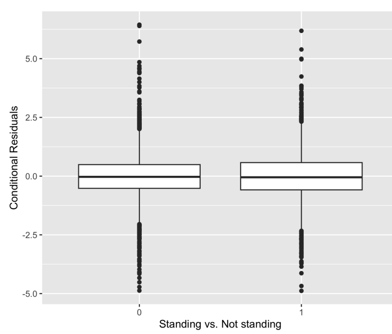
(a) Level 1 Conditional Residuals vs. Fitted Value



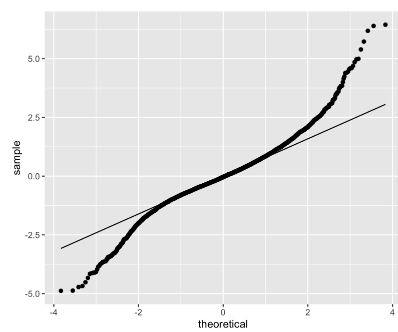
(b) Level 1 Conditional Residuals vs. Activity Level



(c) Level 1 Conditional Residuals vs. Time



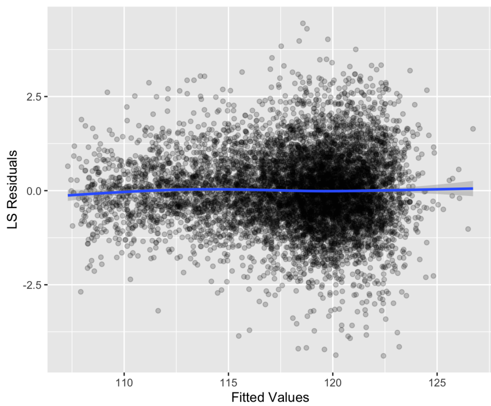
(d) Level 1 Conditional Residuals vs. Standing



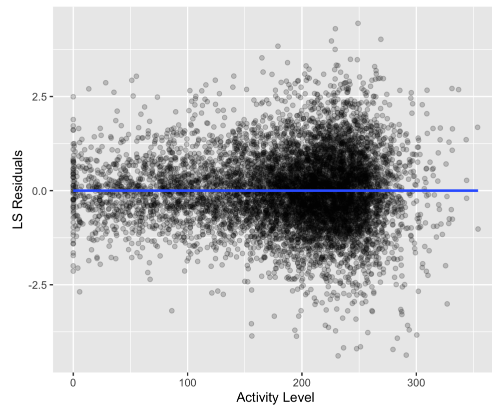
(e) Distribution of Level 1 Conditional Residuals

Figure 15: Level 1 Conditional Residuals vs. Predictors

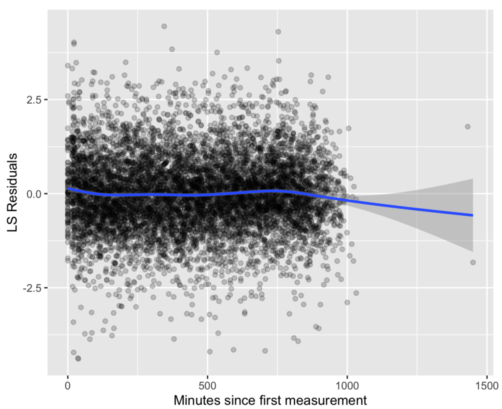
We reach similar conclusions for *least squares* (LS) residuals (Figure 16). There is still some slight curvature in the fitted values (Figure 16a), some worry of constant variance (see Figure 16a and Figure 16b), and violation of normality (Figure 16e, heavy-tailed). Nevertheless, when we look at LS residuals for each participant (Figure 16f), we see that residuals across participants have more or less similar within cluster variation, which suggests that our constant variance might not be as severe as we imagined.



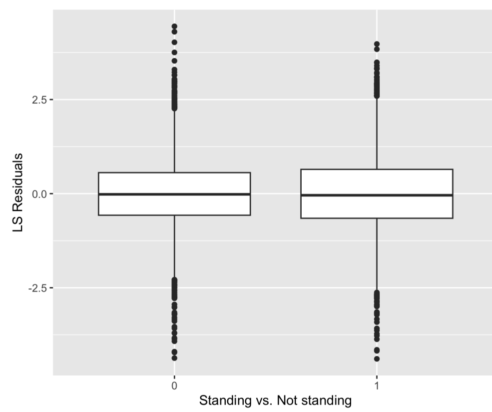
(a) Level 1 LS Residuals vs. Fitted Value



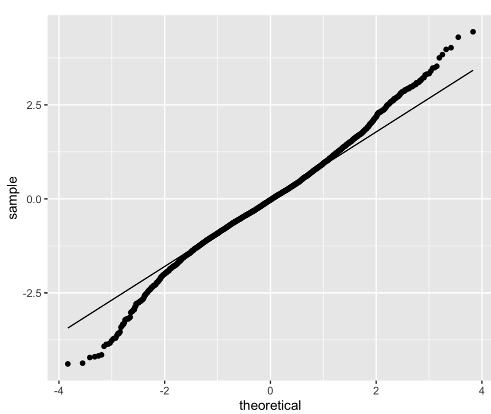
(b) Level 1 LS Residuals vs. Activity Level



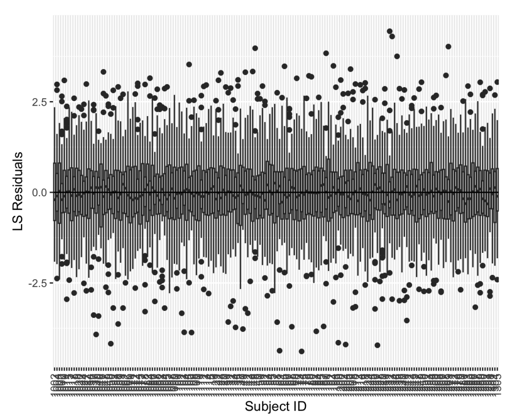
(c) Level 1 LS Residuals vs. Time



(d) Level 1 LS Residuals vs. Standing



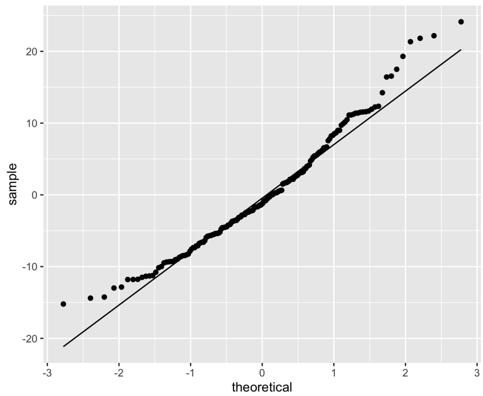
(e) Distribution of Level 1 LS Residuals



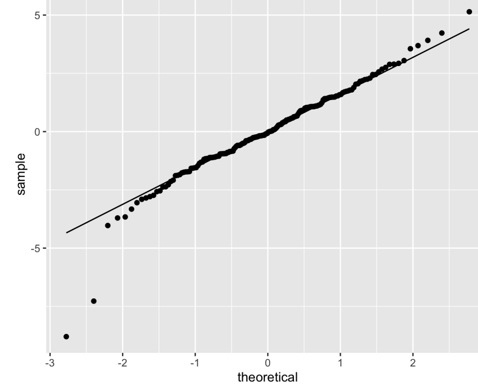
(f) Level 1 LS Residuals vs. Subject

Figure 16: Level 1 LS Residuals vs. Predictors

We move on to examine level 2 residuals. We first examine the normality of the intercepts and slopes (Figure 17). The intercepts are right-skewed (Figure 17a), whereas the slopes (for `Standing`) are heavy-tailed (Figure 17b). We also examine whether there is any missing variables (Figure 18). We chose not to include `Phase` in our model because our EDA did not suggest any relationship between `Phase` and `BP`; indeed, Figure 18a shows that the intercept random effect does not seem to be related to `Phase`.[4] Whether the intercept random effect depends on `Age` has been tested in Section 3, and Figure 18b is consistent with our analysis above. That is, the intercept random effect does not seem to be dependent on age (even though the smoother suggest a very slight increasing trend).

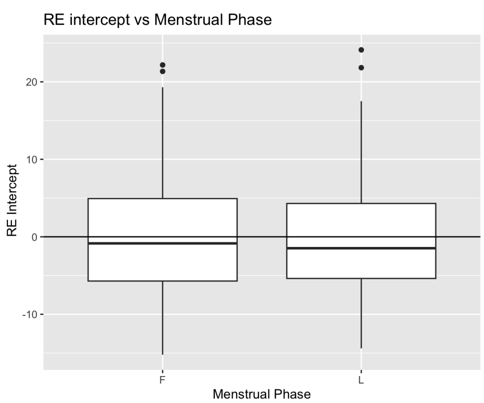


(a) Distribution of Level 2 Intercepts

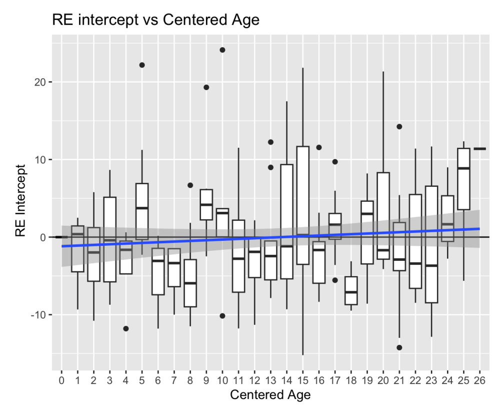


(b) Distribution of Level 2 Slopes

Figure 17: Distribution of Level 2 Residuals



(a) Level 2 Residuals by Phase



(b) Level 2 Residuals by Age

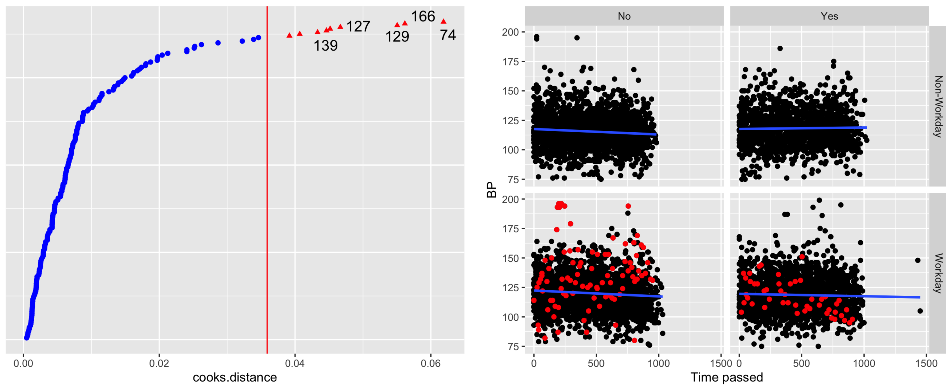
Figure 18: Missingness of Level 2 Covariates

3.2 Influential Statistics

We also examine if some of the measurements carry greater weights in our model. It is important to note first that while many cases have been flagged, there is no *a priori* reason that we should remove any of them.

From Figure 19, we see that most flagged cases (e.g., participant 1133, 1276, 1223) in terms of Cook's Distance come from observations during a workday as opposed to non-workday.

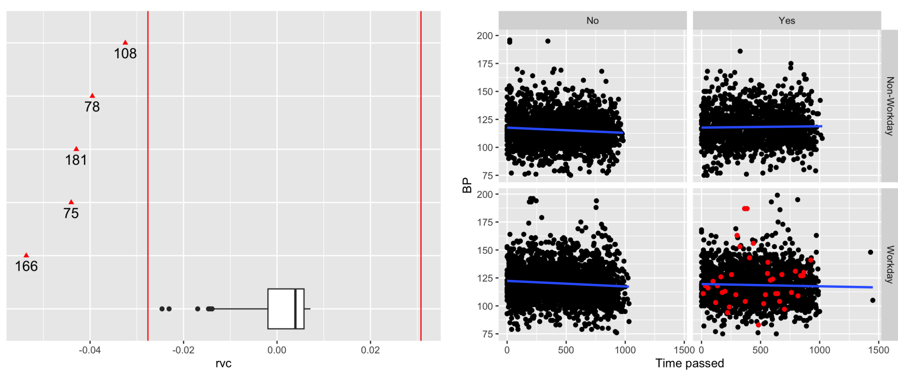
[4]Goldstein and Shapiro (2000; p. 231) reached the same conclusion.



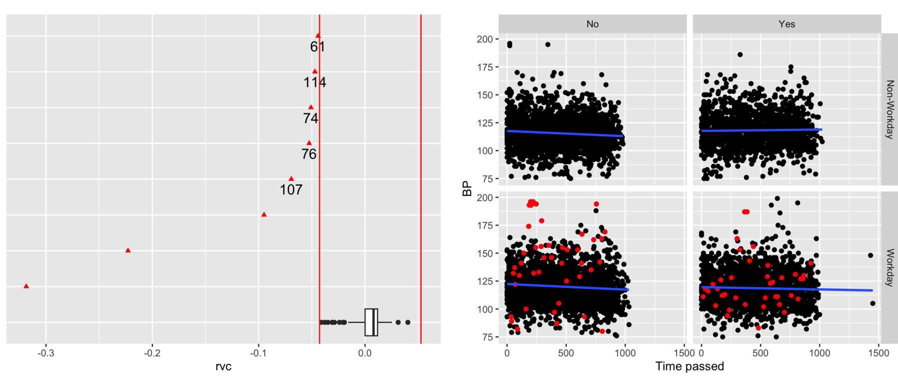
(a) Cook's Distance (b) Flagged (first 3)

Figure 19: Fixed Effects Influence

For random effects influence, we see that most influential cases resuot in lower RE variation when omitted (Figure 20). Specifically, most flagged cases (e.g., participant 1290) in terms of RVC for the intercept come observations during a workday from those *with* family history. Similarly for slopes, most flagged cases (e.g., participant 1276, 1290) in terms of RVC for the slope come from observations during a workday.



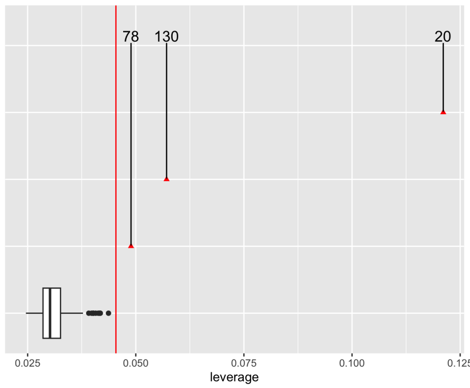
(a) RVC for intercepts (b) Flagged (first 3)



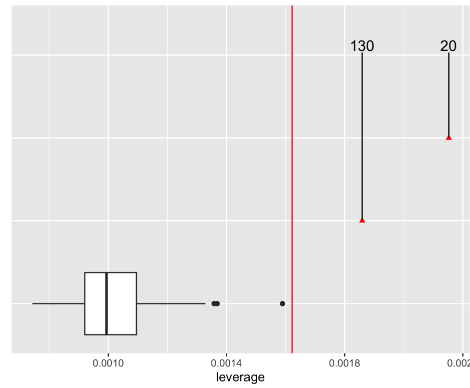
(c) RVC for slopes (d) Flagged (first 1)

Figure 20: Random Effects Influence

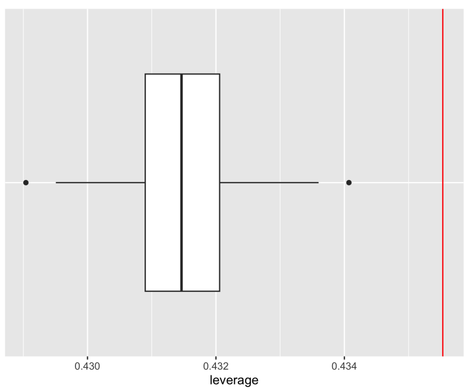
Participants 1046, 1224, and 1140 have the highest overall leverage. Participants 1046 and 1224 also have the highest fixed effects leverage. We see that these cases occur mostly during the workday and mostly for those with family history (except participant 1046, which also have the highest overall leverage).



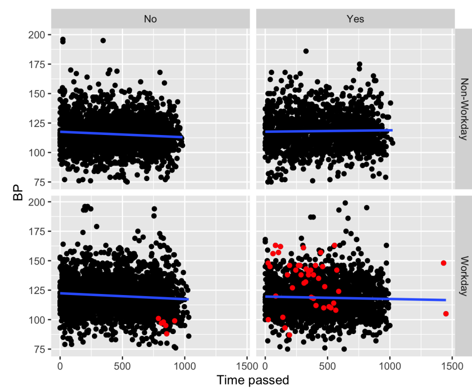
(a) Overall Leverage



(b) Fixed Leverage



(c) (Unconfounded) RE Leverage



(d) Flagged (first 1)

Figure 21: Leverage Influence

4 Results

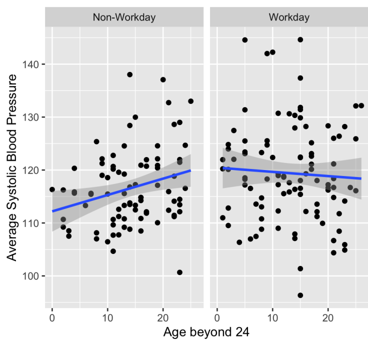
Appendix A Additional Visualizations

References

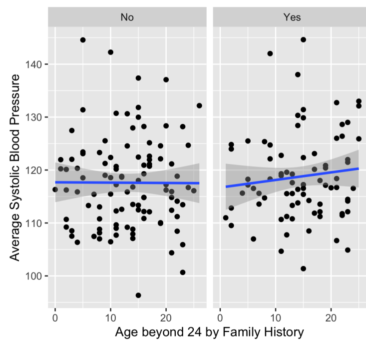
Goldstein, I. B., & Shapiro, D. (2000). Ambulatory blood pressure in women: Family history of hypertension and personality [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/713690197]. *Psychology, Health & Medicine*, *5*(3), 227–240. https://doi.org/10.1080/713690197

Roback, P., & Legler, J. (2021). *Beyond multiple linear regression* (1st ed.). Chapman; Hall/CRC. Retrieved February 10, 2025, from https://bookdown.org/roback/bookdown-BeyondMLR/

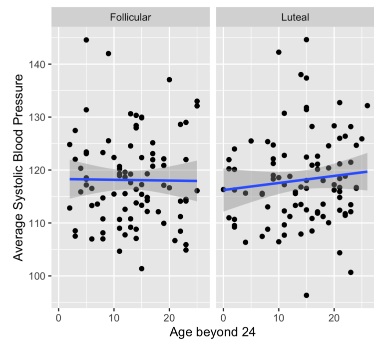
Weiss, R. E. (2005). *Modeling longitudinal data*. Springer. https://doi.org/10.1007/0-387-28314-5



(a) BP vs. Age Level by Workday

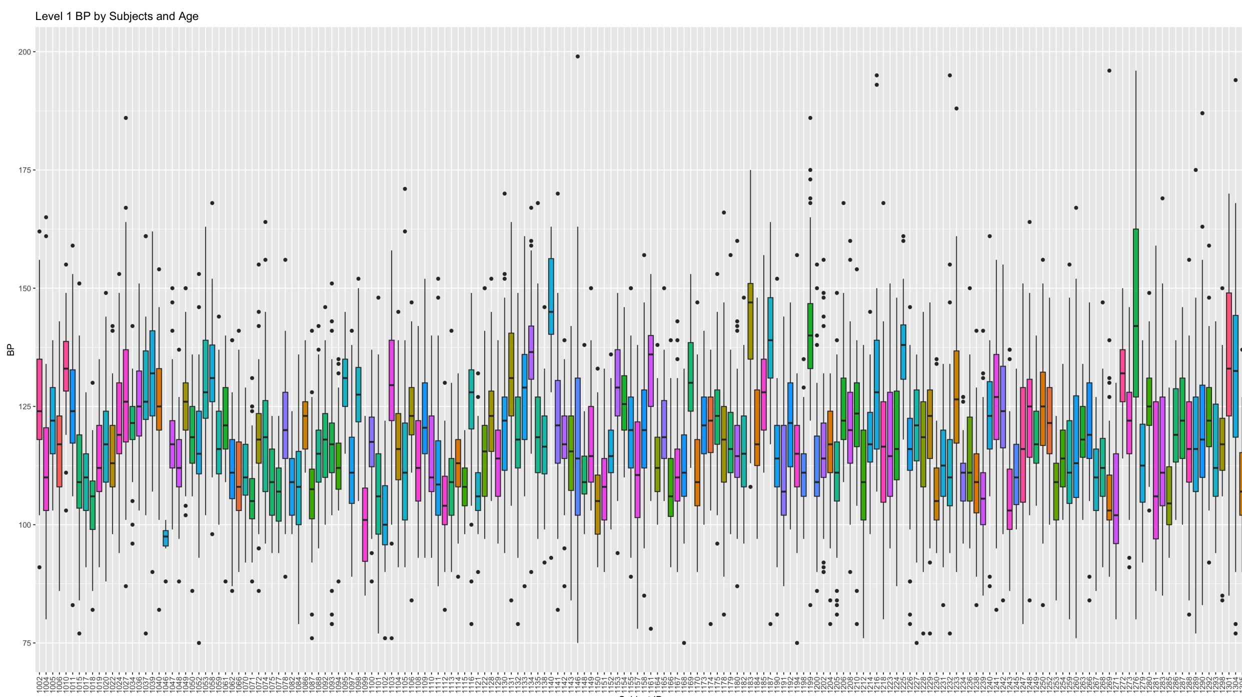


(b) BP vs. Age by Family History



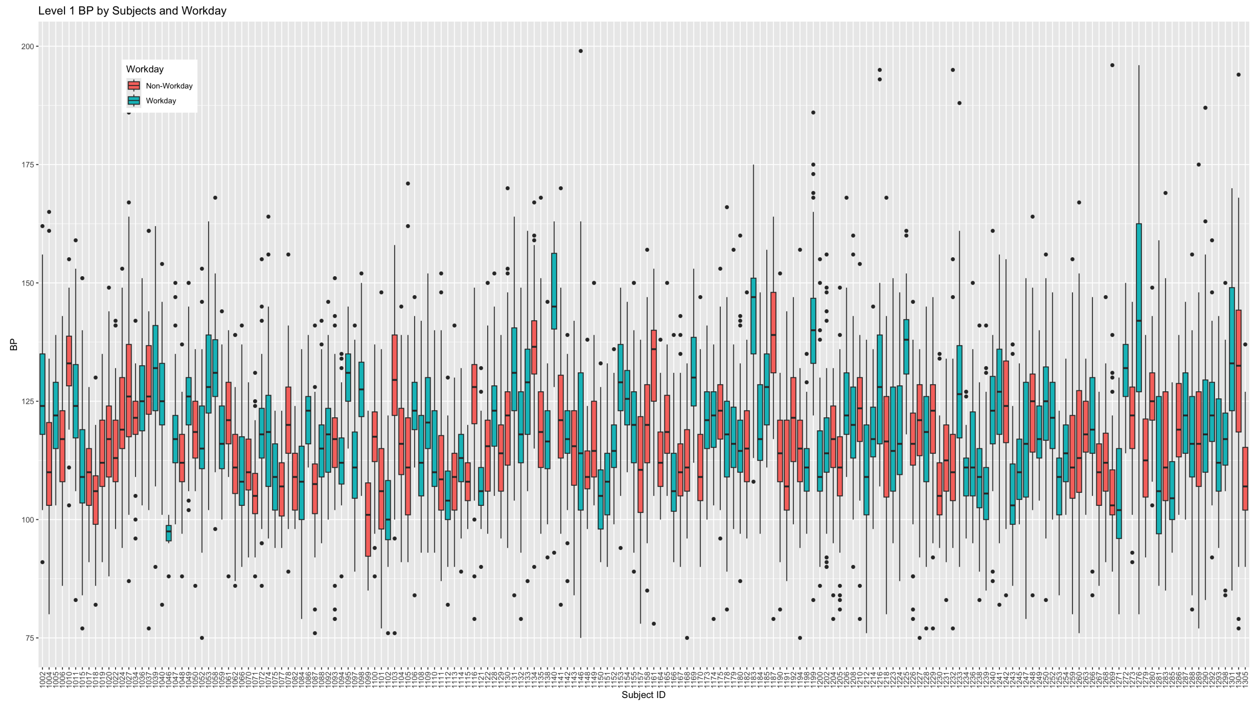
(c) BP vs. Age by Menstrual Phase

Figure 22: BP vs. Age



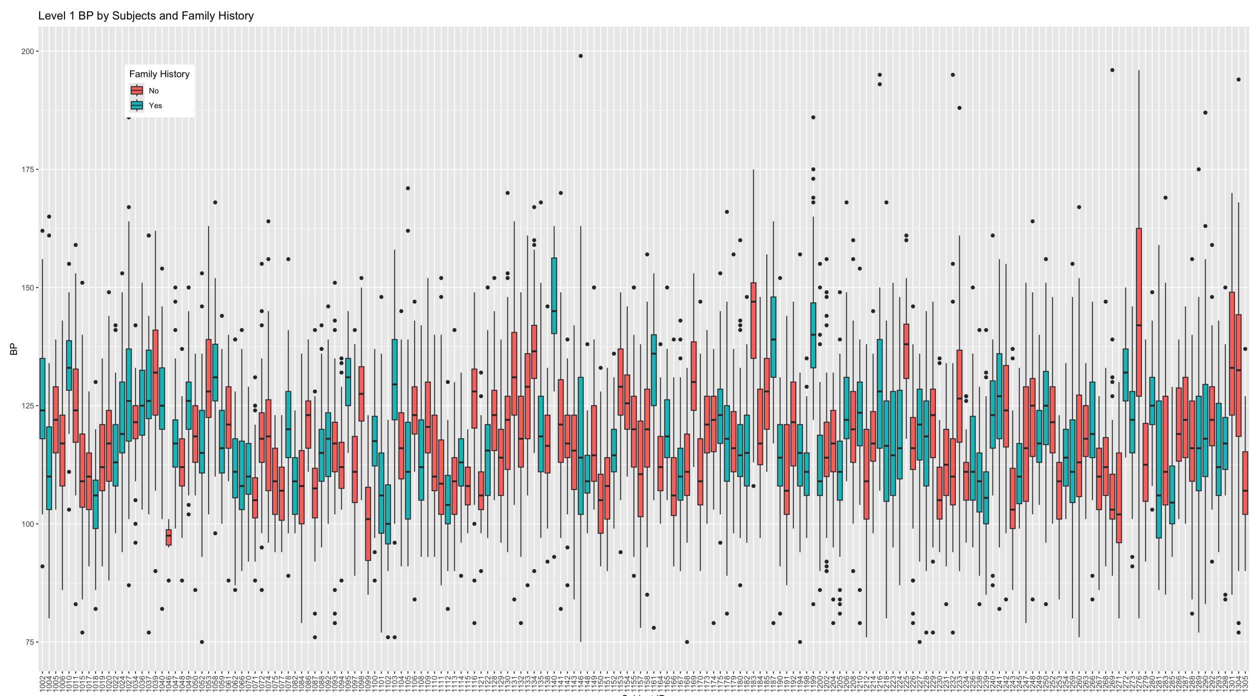
(a) BP vs. Age by Subjects

Figure 23: BP vs. Level 2



(b) BP vs. Workday by Subjects

Figure 23: BP vs. Level 2



(c) BP vs. Family History by Subjects

Figure 23: BP vs. Level 2