

Eric Buehler

DA 401

Rethinking the Application of Technology in Hypertension Control

- 1 Introduction
- 2 Literature Review
 - 2.1 Measurement Issues
 - 2.1.1 Human errors
 - 2.1.2 Instrument Error
 - 2.2 Information Technologies
 - 2.3 Motivating Work
- 3 Methods
 - 3.1 Data Overview
 - 3.1.1 Ethical Concerns
 - 3.1.2 Dependent Variables
 - 3.1.3 Independent Variables
 - 3.2 Predictive Modeling
 - 3.2.1 Linear Regression
 - 3.2.2 Decision Tree
 - 3.2.3 Random Forest
 - 3.2.4 Model Evaluation
 - 3.3 Outlier Detection
 - 3.4 Implementation and App Interface
 - 3.5 Pre-Processing
- 4 Results
- 5 Discussion and Conclusion
- Work Cited
- Appendix

1 Introduction

High blood pressure, or hypertension is one of the most important risk factors for morbidity and mortality (Paul K. Whelton, et al, 2017). According to the Lancet, detecting uncontrolled hypertension, including those untreated and those inadequately treated, would avert 10 million cardiovascular events worldwide over 10 years (Angell, Sonia Y, et al., 2015). Unfortunately, the measurement of blood pressure is not as straightforward as one might expect due to inherent issues related to human error, device calibration and patient physiology.

New techniques and technologies such as at home blood pressure monitoring have emerged to meet these challenges (Whelton, Paul K., et al., 2017). However, many of them are largely tailored towards privileged groups, ignoring the observation that those living in poor, rural communities have some of the highest risk of hypertension (Harris, J. K., Beatty, 2016). Thus, it appears that there is motivation to investigate how technologies might be used in a more equitable way to assist in hypertension control.

To do this, this study investigates if machine learning can improve upon traditional modeling techniques enough to provide baseline estimates for individuals in the US, using limited data from the National Health and Nutrition Examination Survey (NHANES) (CDC, 2019). Subsequently, the study proposes a tool that uses model predictions to assess whether an observed blood pressure is likely valid or a result of device miscalibration. The proposed system can predict blood pressure within 6 mmHg

of the true systolic and diastolic pressure and finds that observed blood pressures over 9 mmHg away from these predictions are likely to be the result of calibration issues.

2 Literature Review

There are many reasons why blood pressure remains difficult to control, however this study chooses to focus on the problems of measurement, and exclusivity of information technologies. These topics most clearly demonstrate the need for a simple measurement assistant tool that can be used for the detection of anomalies in blood pressure measurements.

2.1 Measurement Issues

2.1.1 Human Error

While taking BP measurements in office settings is easy, errors are common and can result in a misleading estimation of an individual's true blood pressure (Paul K. Whelton, et al., 2017). For example, a recent study demonstrated the inability of medical school students to follow blood pressure measurement guidelines and recommended that changes in medical school curriculum be made to emphasize blood pressure measurement (Rakotz, Michael K., et al., 2017). In addition to training issues, in office blood pressure measurement are also subject to white-coat hypertension which is when patients have high blood pressure levels in a physician's office and normal blood pressure levels at home. White-coat hypertension is thought to be explained by a patient's anxiety within a physician's office and in the presence of the physician (Haskard-Zolnieriek, Kelly, et al., 2015). The effect of these problems is real as the

NHANES survey found that 18.2%-33.5% individuals who were classified as hypertensive on the first reading were reclassified to lower BP categories (Handler, Joel, et al., 2012). In general, the impact of human error in blood pressure measurement is a well-documented problem, including additional issues such as inaccurate cuff selection and application, incorrect cuff positioning, inadequate rest period, rapid cuff deflation rate, poor observer concentration, digit bias, and lack of repeated measurements (Jones, Daniel W., et al., 2003). Unfortunately, these kinds of error are difficult to control for beyond drawing from measure that were taken in a consistent manner.

2.1.2 Instrument Error

Beyond human based errors, a litany of instrumental errors also affects blood pressure measurements in and out of office. In terms of in office measurement, the gold standard for in office blood pressure readings has traditionally been the mercury sphygmomanometer paired with the Korotkoff sound technique. However, more recently there has been decreased reliance on this method as mercury is being banned in countries due to environmental concerns (Pickering, Thomas G., et al., 2005). Additionally, systematic errors can occur in sphygmomanometer due to calibration issues and poor maintenance (Parati, Gianfranco, et al., 2006). Aneroid and oscillometric devices present a common alternative to the mercury sphygmomanometer. These devices mechanically measure the systolic blood pressure and predict diastolic blood pressure via a company designed algorithm (Whelton, Paul K., et al., 2017). As a result, only devices with a validated measurement protocol are recommended. In general, these devices have been found to be not as accurate as the

sphygmomanometer and are still subject to white-coat hypertension if used in office (Pickering, Thomas G., et al., 2005).

As an alternative to traditional in office sphygmomanometer and oscillometric readings, more cutting-edge techniques such as ambulatory blood pressure monitoring (ABPM) and at home monitoring (HBPM), have been proposed and implemented. ABPM uses a device to measure blood pressure over a 24 to 48-hour period in or out of office (Seravalle, G., et al., 2018). ABPM is generally accepted as the best measurement of blood pressure, but its 24 to 48 measurement period is time consuming and expensive. Additionally, guidelines over best practice have also yet to be established (Whelton, Paul K., et al., 2017). On the other hand, HBPM has a patient intermittently self-report their blood pressure electronically using an oscillometric device (Seravalle, G., et al., 2018). Since HBPM relies on several measurement, it avoids many of the accuracy issues associated with oscillometric devices (Whelton, Paul K., et al., 2017). HBPM is often a more practical approach to ABPM as it can be reported at a patient's convenience and has been shown to be similarly effective in measuring blood pressure as ABPM (Whelton, Paul K., et al., 2017; Seravalle, G., et al., 2018). Unfortunately, HBPM self-report style can also be a weakness as patients may fail to report their blood pressure, resulting in readings that are no more reliable than those taken by an oscillometric device in office. The devices are also still subject to miscalibration, which are difficult to detect in home. Overall, the application of either ABPM or HBPM is somewhat situational.

Despite the promise of ABPM and HBPM, their implementation in low-resource settings would require a change in the training of healthcare professionals, regulatory changes regarding the production and sale of BP devices and technological advances, and few studies have robustly evaluated the cost associated with these changes (Pickering, Thomas G., 2008; Seravalle, G., et al., 2018). The irony is that these are exactly the communities that suffer the most from uncontrolled hypertension (Harris, Jenine K., et al., 2016). Thus, it seems that some of the most cutting-edge techniques for addressing measurement issues in blood pressure pay little attention to the rural and low-income communities who perhaps have the most to gain from more accurate blood pressure readings.

2.2 Information Technologies

The issue of a privileged audience in hypertension control applies not only to new measurement techniques, but also to new detection techniques that leverage information technology. For example, The Kaiser Permanente Northern California (KPNC) has shown that structured, goal-oriented approaches backed up by a large data registry, capable of providing electronic monitoring and target achievements, can greatly improve the control of hypertension (Whelton, Paul K., 2015). NorthShore University Health System has deployed algorithms on their extensive health records to accurately identify those at risk of hypertension (Merai, Rikita, et al., 2016; Rakotz, M. K., et al., 2014). Telehealth is another emerging technology-based approach to hypertension management. These approaches use mobile computing and communication technologies to improve the detection and control of hypertension. However, while telehealth strategies demonstrate great promise, they are still a work in progress and

have shown variability and inconsistency (Whelton, Paul K., et al., 2017). Unfortunately, telehealth approaches also suffer from poor communication between developers and health professionals (Burke, Lora E., et al., 2015). The commonality between all these approaches is that they rely on vast data registries, expertise and cater to an affluent audience, like ABPM and HBPM.

2.3 Motivating Work

Despite appeals from the literature to find hypertension control methods that emphasize high risks groups, these are the common measurement and technology-based approaches being researched and recommended by the literature (Whelton, Paul K., et al., 2017; Whelton, Paul K., 2015; Harris, Jenine K., et al., 2016). However, some research has been done on more accessible approaches to hypertension control. For example, studies have shown that machine learning can be used to predict the presence of hypertension at a high degree of accuracy using a minimal number of inputs (Golino, Hudson Fernandes, et al., 2014; Ijaz, Muhammad, et al, 2018). Additionally, MedStar Health has found modest success using a simple system that inserts a default blood pressure goal for all their patients that can be used for providing structured feedback and a reference for in office blood pressure measurements (Merai, Rikita, et al, 2016). Thus, MedStar demonstrates a simpler alternative to the more robust approaches used at KPNC. These simpler technology-based approaches provide the motivating foundation for this study's goal of investigating how technologies might be used in a more equitable way to assist in hypertension control.

3 Methods

The methodological goal of this study is to predict an expected blood pressure for a given individual based on predictors of readily available medical data, using modern machine learning methods. And then, use these predictions as a baseline to determine if an observed blood pressure measurement is valid or likely an anomaly due to a miscalibration error. Key limitations of this methodological goal lie in the data and prediction process. Data limitations are largely a result of the use of non-experimental data. On the other side, the statistical prediction of any measure introduces a degree of uncertainty that stems from either assumption made about data or the limitations of the model being applied. The study also limits the scope of predictors to common medical data, which further limits the model's predictive power. The implantation of this methodology and its trade-offs are explicitly discussed below.

3.1 Data Overview

Data for this study comes from National Health and Nutrition Examination Survey (NHANES), which is a program of studies designed to assess the health and nutritional status of adults in the United States (CDC, 2019). NHANES is an ongoing cohort study that has been conducted every other year since 1999 and data are publicly available up to 2015. The individual units of analysis are individuals eighteen years or older living in the United States. NHANES breaks down into several parts, however this study is only concerned with the Demographic, Examination, Laboratory and Questionnaire portion of the survey. Since NHANES is a cohort study, the data cannot show causality as they come from a cohort study, not experimental or longitudinal. As a result, the proposed tool will likely fall short of true medical standards. However, by using NHANES, the

study meets its methodological goal of using machine learning in a way that does not rely on vast private health records. Thus, the research may still serve as a proof of concept.

3.1.1 Ethical Concerns

Since the data are produced by a Federal agency, they are in the public domain and may be reproduced without permission (NCHS, 2019b). While some NHANES data related to geolocation, STDs and youth participants are limited access and require approval from the NCHS, none of this data is required for this study. According to Denison's IRB guidelines, studies that use anonymous survey such as NHANES qualify as a category 4 exemption, and thus do not require IRB approval (Denison, 2019). Therefore, the use of NHANES data for this study is in line with the standards of the National Center of Health Statistics and Denison.

Other ethical concerns for this study include its overarching aim to development a tool that could theoretically be used to assist in detecting calibration issues in blood pressure devices. This is a particular source of concern, since NHANES was not designed to be used in medical support systems. The study also has no way of testing the effectiveness of its proposed design. Thus, in any condition, this study will fall short of medical standards and as a result should be taken as a proof of concept for how machine learning might be applied to hypertension control in a more democratizing way.

3.1.2 Dependent Variable

The primary dependent variables of this study are diastolic and systolic blood pressure. NHANES provides blood pressure measurements using a standardized procedure with sphygmomanometer with five minutes of rest between each

measurement (NCHS, 2019a). The standardized procedures for blood pressure measurement is a benefit of the NHANES survey as it controls for some problems associated with blood pressure measurement such as inconsistencies in device use and improper technique (Whelton, Paul K., et al., 2017). However, the measurements still fail to address the problems of white coat hypertension and shortcomings of traditional in-office techniques noted in the literature review (Whelton, Paul K., et al., 2017). While up to five measurements are available, in most cases, only three measurements are available. Thus, the “true” blood pressure can only be estimated from the mean of three single visit blood pressure measurements. This is problematic as more measurements from separate visits is the best practice for reducing variability and approximating the “true” blood pressure of an individual (Hughes, Michael D., and Stuart J. Pocock., 1992). Unfortunately, with the current resources available, there is little to correct for these issues. Keeping these limitations in mind, NHANES still presents a large amount of standardized blood pressure data that decently approximates levels in the United States and is reasonably suited to the aims of this study.

3.1.3 Independent Variable

Independent variables of interest include those supported by the literature that general practitioner might readily have available in the United States. These variables include age (continuous), race (categorical), gender (binomial), body mass index information (continuous), cholesterol (continuous), diagnosed diseases such as diabetes (binomial) and some basic blood measurements such fasting glucose levels (continuous) (Whelton, Paul K., et al., 2017; Stamler, Jeremiah, et al., 1975; Golino,

Hudson Fernandes, et al., 2014;). It is important to note that all diagnosis data is self-report. However, body measurement and other biological factors such as cholesterol were measured and recorded by medical professionals. The full documentation for these variables is publicly available by the CDC (NCHS, 2019a). For now, it will suffice to say that they were collected with similar standards as blood pressure.

By design the study relies only on common place medical data such as BMI and cholesterol and thus excludes more complicated biometric data, which has been shown valuable for precise predictions of blood pressures (Zhang, 2018). Therefore, the study is inherently limited in how much variation it will be able to explain.

3.2 Predictive Modeling

In terms of predictions methods, the study calls for approaches whose predictions are easily interpreted. Thus, methods such as linear regression, decision trees are prime candidates. However, it is possible that while these methods are desirable for their ease of interpretation that they may not have enough predictive power to be realistically applied for the proposed tool. Thus, methods such as random forest will be considered as well. Typical metrics of model assessment such as RMSE do not properly capture the two-dimension of blood pressure (systolic and diastolic), so mahalanobis distance is used instead.

3.2.1 Linear Regression

Linear regression performance is well documented in the blood pressure literature (Stamler, Jeremiah, et al., 1975.) While perhaps the most simplistic of the methods, it provides an easy to understand equation and has well documented properties. Thus, it serves as a solid baseline model to compare to others. The short

comings of linear regression are largely introduced in its rigorous assumptions of multi-normality, linear independence, no auto-correlation, and homoscedasticity. In terms of blood pressure, linear independence is a particularly troublesome assumption as it is nearly impossible to assume that biological and even demographic variables are completely independent of each other. Homoscedasticity also presents a sizable issue as blood pressure varies unequally across a range of variables (Hughes, Michael D., and Stuart J. Pocock., 1992; Rosner, Bernard, and B.frank Polk., 1979).

Since blood pressure varies unequally across variables, the stratification of individual's into different blood pressure groups has been shown to improve model quality (Stamler, Jeremiah, et al., 1975). However, these stratifications typically only rely on a few variables that are known to vary unequally such as race, gender and sex. Alternatively, K-prototype clustering can be used to define blood pressure group based on an entire subset of predictors. Thus, clustering on demographic variables may provide a more holistic method of stratification to account for unequal variance. K-prototype is limited by the fact that the "best" number of clusters is determined by a heuristic that examines within cluster distances. Due to this limitation, it is possible that the clustering offers little improvement over traditional stratification methods.

3.2.2 Decision Tree

Decision trees are common alternative to linear regression that are easier to understand and better account for blood pressure's correlated predictors. The basic idea of decision trees is that the algorithm divides a predictor space into distinct, nonoverlapping regions (Müller, K. R., et al., 2001). It does this by partitioning regions in a way that minimizes the sum of the square of the residuals (RSS). This partitioning

allows for complex interactions between variables as each variable is forced to interact with every variable in subsequent partitions (Müller, K. R., et al., 2001). This is a particularly desirable feature for the prediction of blood pressure, which relies upon dependent predictors. Additionally, decision trees partitioning can be understood in a simple to read graph that is easier to understand than a regression equation, a desirable feature in medical predictions. Despite these benefits, decision trees typically overfit data and have high variance (Müller, K. R., et al., 2001). To counteract this, decision trees can be “pruned”, or limited in how many splits they make by setting a threshold for the minimum decrease in RSS, *cp*. However, due to this weakness, a decision tree may underperform when compared to linear regression.

3.2.3 Random Forest

Random forest provides an overall improvement to the decision tree. It does this by aggregating the predictions of many different decision trees. Additionally, at each subspace split the tree is only allowed to consider some number of random predictors, m (Müller, K. R., et al., 2001). Thus, one strong predictor need not dictate the entire structure of the tree and new relations of predictors can be considered. The result is a model that decreases the variance of decision trees and that allows for interaction of predictors (Müller, K. R., et al., 2001). However, these benefits come at a cost to interpretability. There is no simple way to visualize a random forest and thus it is not ideal in the context of blood pressure predictions. Random forest does provide a way to examine variable importance by looking at the reduction in sum of squared errors whenever a variable is chosen for a split across trees, but this kind of analysis leaves something to be desired considering the linear regression and decision trees.

3.2.4 Model Evaluation

Typically, regression models are assessed based on root mean squared error (RMSE), which in this case is the square of the difference between the true blood pressure values and the predicted blood pressure values. Unfortunately, blood pressure isn't a one-dimensional measure. Rather, it consists of two dimensions: systolic and diastolic blood pressure. A simplistic adaption of RMSE would be to simply take the RMSE of systolic and diastolic blood pressure and compare the model's performance based on both measures. But what if one model performs better in systolic and worse in diastolic, it is unclear which we would choose as "better"? As a result, comparison by RMSE leaves something to be desired. This problem of comparison can be resolved by Mahalanobis distance, which considers the distance between two multi-dimensional distribution, or in this case the distance between the true systolic and diastolic, and the predicted systolic and diastolic values (McLachlan, Geoffrey J., 1999).

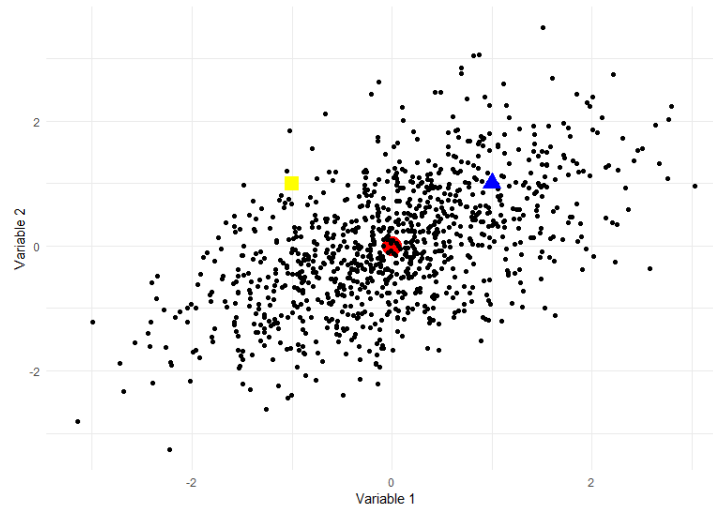
$$\Delta^2 = (\bar{\mu}_1 - \mu_2)^T \varepsilon^{-1} (\bar{\mu}_1 - \mu_2) \quad (1)$$

Equation one shows the calculation of the mahalanobis distance between true and predicted blood pressures with $(\bar{\mu}_1 - \mu_2)$ representing the matrix of differences between true and predicted systolic and diastolic blood pressure and ε^{-1} represents the inverse of the co-variance matrix between the true and predicted blood pressures. Thus, just as with RMSE, a smaller mahalanobis distance between true and predicted blood pressures represents a more accurate model. In short, mahalanobis distance allows for the assessment blood pressure predictions in terms of both its dimensions.

3.3 Outlier Detection

The primary goal of this methodology is not to evaluate model performance of models but to leverage the best model for detecting anomalous blood pressures. To do this, an observed blood pressure for an individual can be compared to the expected blood pressure generated by the model. If the observed blood pressure is over a certain distance from the blood pressure, it might be considered an anomalous reading as a result of calibration issues with a blood pressure instrument. Otherwise, it might be considered valid.

Mahalanobis distance again comes in handy as it allows for the observed blood pressure for an individual to be compared on the basis of its two dimensions. This is particularly handy in outlier detection as an observed blood pressure might differ greatly from a predicted blood pressure, but still be in line with the general distribution of blood pressures and still be considered valid (Clark, L. A., et al., 1987). Take for example Figure 1, both the triangle and square are roughly the same distance from the circle in the center of the distribution. However, considering these distances in the terms of Mahalanobis distance, the square may be thought to be further from the circle than the triangle as it is more in line with the general distribution.

Figure 1:

The only question that remains is how to determine the appropriate threshold for whether an observed point is anomalous or not. It turns out that the distribution of Mahalanobis distances for a multivariate normal distribution is a chi-square distribution (Garrett, R. G., 1989). Thus, a cumulative probability plot can be made using mahalanobis distances plotted against their corresponding chi-square values to determine which distances may be anomalous (Garrett, R. G., 1989). Again, the main concern here is finding an acceptable distance between an observed blood pressure and predicted, so it can be determined whether it is anomalous or not.

3.4 Implementation and App Interface

The system described in this methods section will eventually be implemented into a Shiny app (Winston Chang et al., 2019). The app will consist of a simple interface that allows users to enter their medical information, which in the background will be used to predict a blood pressure for the individual. The user will then enter their observed blood

pressure, which will be compared to the predicted value, using mahalanobis distance. If the blood pressure is over the predetermined cut-off value, it will be considered anomalous. If it is under, it will be considered valid. The overall practicality of this system is subject to critique but may demonstrate a novel application of machine learning to blood pressure measurement.

All algorithmic approaches described above will be implemented using R (R Core Team, 2019). All code will be made available on github (Buehler, 2019).

3.5 Pre-Processing

Each variable of interest in the NHANES survey was downloaded manually from the CDC's website from years 1999-2015 in .XPT format (CDC, 2019). The files were easily read into R and combined into a single data frame. All files included in the study can be found in the NHANES_DATA folder in this study's github repository (Buehler, 2019). Fortunately, due to the uniformity of NHANES, there were few discrepancies in the variables from year to year with one exception. From years 1999-2003, total cholesterol and HDL cholesterol were combined in the same .XPT file but were in separate .XPT files for years 2005-2015. This discrepancy was easily resolved by making sure rows bound correctly in the pre-processing stage.

Missing values were then removed, and the data was split into 75% training data and 25% testing data. The resulting training data set contained 4950 observations of 20 variables, while the test set contained 1204 observations of 20. Table 1 shows a snapshot of the resulting training set:

Table 1:

Estimated Diastolic (mmHg)	Estimated Systolic (mmHg)	Weight (kg)	BMI (kg/m ²)	Gender (male =1 / female = 0)	Age (yrs.)	Hypertensive (yes/no)	Hypertension Medication (yes/no)
55.33333333	93.33333333	74.4	25.4	1	22	0	0
80	128	62.9	24.3	0	42	0	0
59.33333333	124	63.1	24.8	0	56	0	0
80	122	79.5	28.4	1	24	0	0
75.33333333	126	72.8	26.2	1	51	0	0
44.66666667	110.6666667	53.8	17.9	1	22	0	0

As noted in the data overview, the estimated diastolic and systolic blood pressure were calculated from three observed readings using the mean of three single visit blood pressure measurements from a mercury sphygmomanometer, which introduces a degree of variance that cannot fully be accounted for.

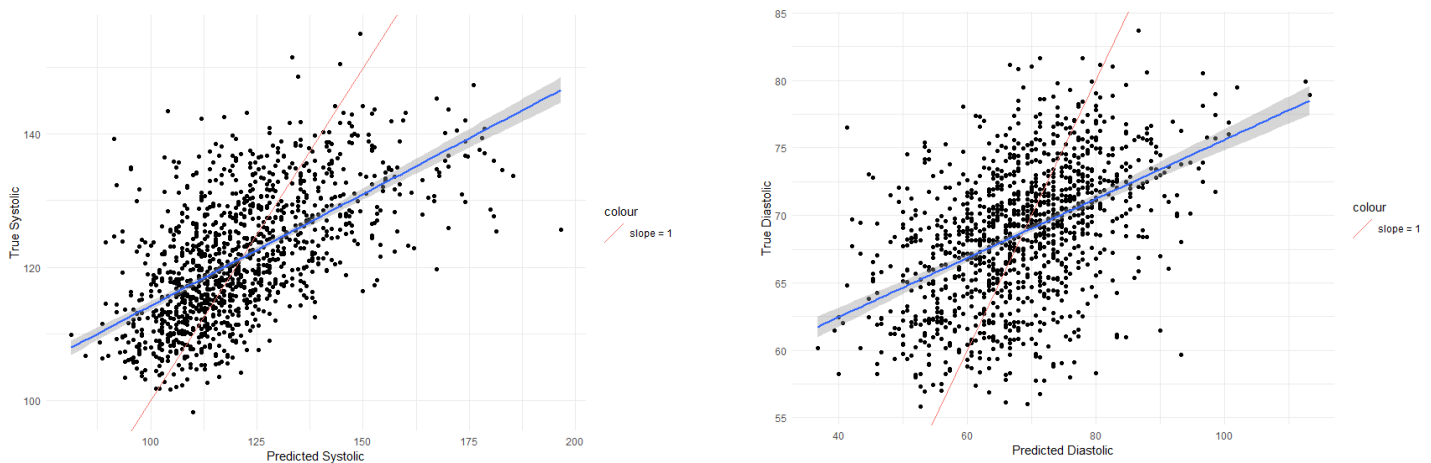
4 Results

The predictive baseline for blood pressure was established by fitting an OLS regression model to both systolic and diastolic blood pressure. Interaction terms were added to the model based on support from the literature and looking at two-way interactions using ANOVA. Assumptions of homoscedasticity, normality and multicollinearity were checked and corrected for as best as possible using transformations (log, squares, etc.). Overall, these assumptions were met rigorously (see appendix Figure A1 and A2 assumption checks).

The systolic and diastolic linear models achieved R-squares of 33% and 21%, respectively (see appendix Table A1 for full model). Figure 1 shows the true systolic blood and diastolic blood pressures compared to the linear predicted diastolic and

systolic blood pressures. The blue line with error bars represents the predicted fit of the models, while the red line without error represents a line with slope equal to one. As demonstrated by their divergent nature, there is considerable room for improvement over the linear model, which perhaps is not surprising considering the R-square values. For example, a predicted systolic value of 125 may have an underlying true value ranging anywhere from 115-138 mmHg, roughly speaking.

Figure 2: True Blood Pressure compared to Linear Predictions

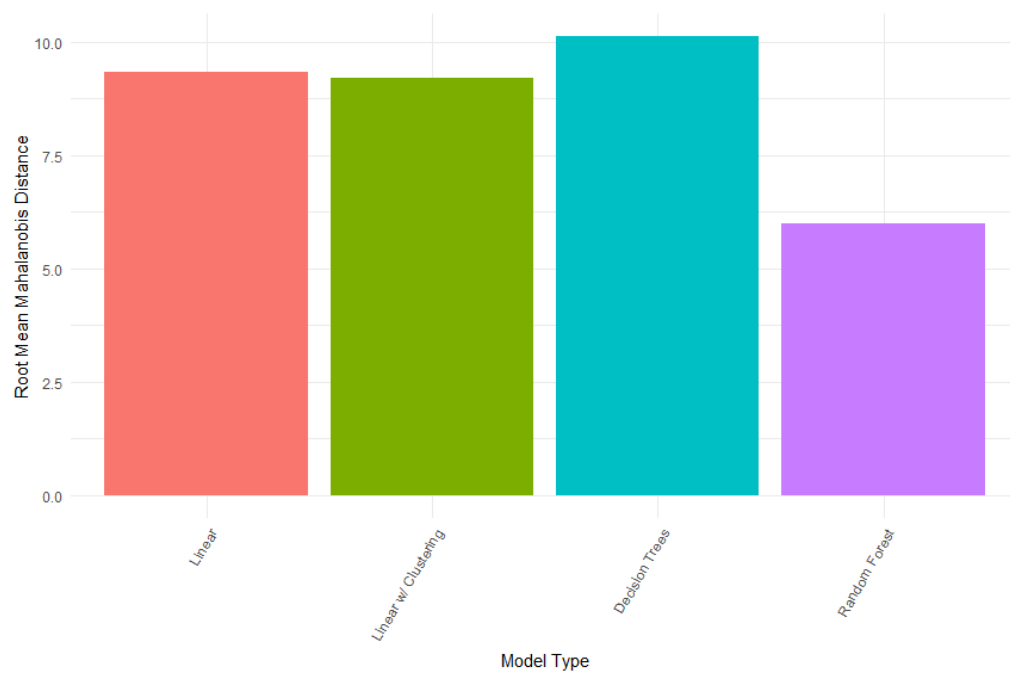


The average Mahalanobis distance is 9, which means on average the linear model is off by 9 mmHg in terms of both types of blood pressure. In terms of RMSE, systolic was off 13.8 mmHg on average and diastolic, 9.5 mmHg on average. The mahalanobis distance and RMSE are not ideal but demonstrate that on average the model does possess some predictive power and at least maintains a semblance of a moderate linear relation between the true and predicted values. Many of the model's shortcomings are likely due

to the unequal variance in blood pressure across age and other variables (see appendix Figures A3 and A4).

Fortunately, the baseline model can be considerably improved upon as shown in Figure 3. Random Forest outperforms the other models by a significant margin with a mahalanobis distance of 5.9, which means that on average the random forest model's blood pressure predictions were only off by about 6 mmHg. In terms of RMSE, systolic was off 7.8 mmHg on average and diastolic, 4.3 mmHg on average. Thus, random forest not only outperforms the other models in terms of joint blood pressure, but also in terms of systolic and diastolic pressure when considered separately (see Appendix 5a for true vs predicted values).

It's worth noting here that clustering on demographic variables did little to nothing to improve the linear model's performance, which is the opposite of the expected. This may be a result of poor separation between the clusters in terms of blood pressure. Thus, it may be possible for a better linear model to be constructed using more traditional approaches to stratification.

Figure 3: Comparison of Models in Predicting Blood Pressure

The final random forest models considered a subset of two predictors and were averaged over one-thousand trees (see Appendix Figures A6 & A7). Figure 4 and 5 on the following page show the variable importance of predictors for systolic and diastolic predictions in the models, respectively.

Figure 4: Variables Importance in Predicting Diastolic Pressure for Random Forest

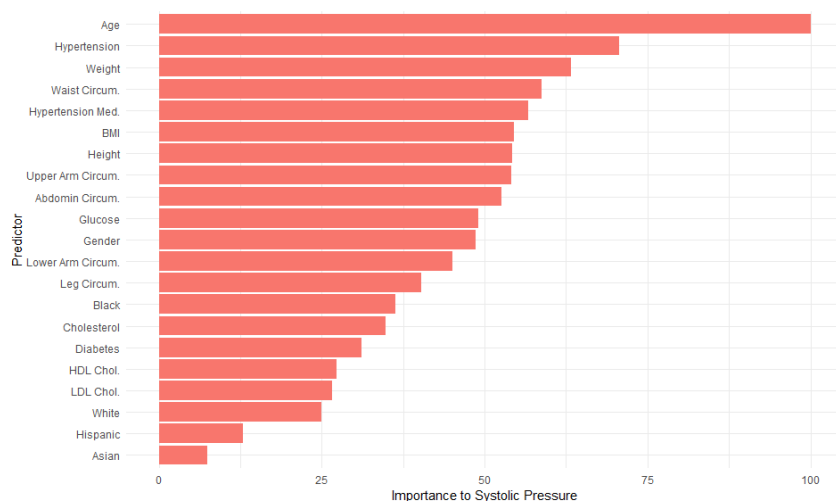
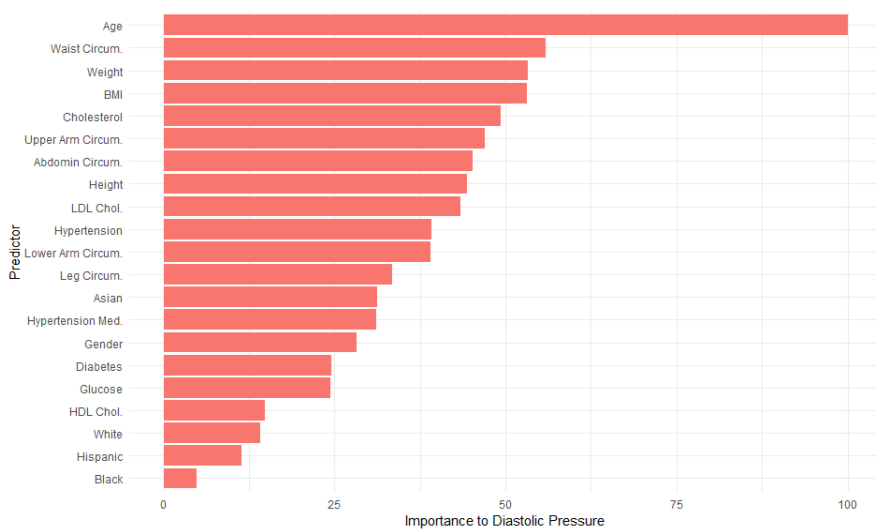


Figure 5: Variables Importance in Predicting Systolic Pressure for Random Forest

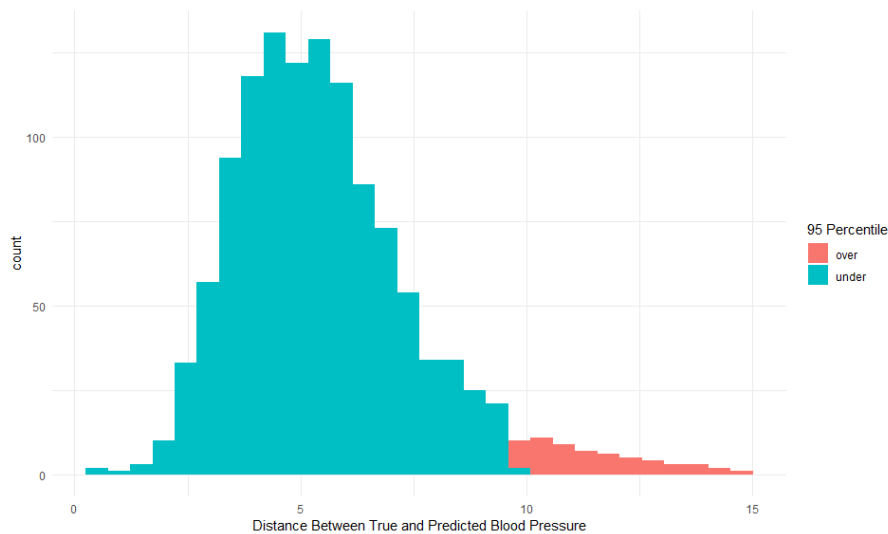


The graphs make intuitive sense in terms of the literature with variables such as age, weight and waist circumference being among the top five variables for both systolic and diastolic pressures. Additionally, the graphs give a good picture for how the two dimension of blood pressure differ. For example, hypertension condition seems more important in the prediction of systolic blood pressure, as both the hypertension

diagnosis and medication variables appear in the top five most important predictors. On the other hand, cholesterol seems more important for diastolic pressure with total cholesterol appearing in the top five predictors for diastolic pressure but appears in the bottom ten for systolic. However, here in lies one of the major drawbacks of these random forest model. Variable can only be discussed in terms of their relative importance. There is no equation that tells us the variables coefficients, or visual that explains how each decision is made. Thus, while the model's performance is impressive, it lacks what is often needed in medical support tools: explanatory power. Overall, the random forest's vast improvement in predictions is most likely due its ability to decrease the variance in predictions and to account for many interactions.

After comparing the distance between true and predicted blood pressure values to the chi-square distribution, a cut-off mahalanobis distance of 9 was considered reasonable for anomaly detection (see Appendix Figure A8). This translates to the top 2.5% of distances as shown in Figure 6. While the approach for selecting a cut-off value is reasonable, it is difficult to have full confidence in its validity without testing it in actual practice, highlighting a key limitation in this study.

Figure 6: Determining an Outlier Cut-off



5 Conclusion and Discussion

The goal of this study was to identify weaknesses in the literature concerning the measurement and prediction of blood pressure and design a suggested solution to the issue. This study finds that there are various competing standards and devices in the measurement of blood pressure, which complicates the problem of detecting hypertension. And, the technology driven solutions designed to resolve these measurement issues have largely ignored low-income and rural communities, who suffer from some of the highest levels of uncontrolled hypertension. The presented research outlines a suggested design for a tool that could possibly help bring technology driven solutions for measurement issue to low-income communities. The design works by providing baseline estimates for individuals in the US that can be used to detect if an individual's measured blood pressure is due to a calibration issue. The results show that this design may be feasible using a random forest model as it addresses inherent issues in the prediction of blood pressure such as the dependent

predictors and produces a fit that is only off by 6 mmHg on average. The results also yielded a cut-off value of 9 mmHg or greater for detecting outliers in observed blood pressures.

This being said, the question remains if such a tool would be feasible or useful in practice? To address this question, the study's data and medical limitations must be re-examined. For example, the data used for the study, NHANES, was not designed to be used in medical support systems. However, the proposed tool does not aim to supplant medical professional authorities, but rather detect calibrations errors in blood pressure devices. Thus, the use of NHANES for this task may not be completely inappropriate for the task, and this is at most the strongest assertion the study can make. Despite the limitations of NHANES, the study demonstrates that a machine learning algorithm trained on limited data has the power to produce powerful medical predictions. And thus emphasizes, the need for medical information to serve more than just the resource rich. This is a colossal task due to technical hurdles and the sensitivity of medical information, but nonetheless is an issue that needs more attention.

Beyond data limitations, the study is also limited in the sense that it is completely unclear if the proposed tool would work in practice. For example, what if blood pressure predictions within 6 mmHg of the true value is not good enough for detecting outliers? What if 9 mmHg is not a suitable cut-off value? What if tool predicts only miscalibrations or only miscalibrations? As reasonable as the number and methods might sounds, the fact remains that this study, at this time, has no way of answering these important questions. Even if the study assumes the tool does work, it is still entirely unclear if it would be useful to medical professional or even individuals. It very well may be the case

that the system adds too much complexity for its provided benefit. For example, a simple in-office chart based on a few predictors of high blood pressure may prove more useful in practice than this device. However, even if all the negatives prove true, this study still demonstrates how technologies in public health are largely being developed to cater to the resource rich and provides an example for how we might re-think the application of machine learning to these issues.

Work Cited

1. Whelton, Paul K., et al., 2017 “ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults, *Journal of the American College of Cardiology*”, May 2018, 71 (19) e127-e248; DOI: 10.1016/j.jacc.2017.11.006
2. Olsen, M. H., Angell, S. Y., Asma, S., Boutouyrie, P., Burger, D., Chirinos, J. A., ... & López-Jaramillo, P. (2016). A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet Commission on hypertension. *The Lancet*, 388(10060), 2665-2712.
3. Harris, J. K., Beatty, K., Leider, J. P., Knudson, A., Anderson, B. L., & Meit, M. (2016). The double disparity facing rural local health departments. *Annual review of public health*, 37, 167-184.
4. Golino, H. F., Amaral, L. S. D. B., Duarte, S. F. P., Gomes, C. M. A., Soares, T. D. J., Reis, L. A. D., & Santos, J. (2014). Predicting increased blood pressure using machine learning. *Journal of obesity*, 2014.
5. Centers for Disease Control and Prevention (CDC). (2019). National Health and Nutrition Examination Survey Questionnaire (NHANES). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. <https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>.
6. Rakotz, M. K., et al. “A Technology-Based Quality Innovation to Identify Undiagnosed Hypertension Among Active Primary Care Patients.” *The Annals of Family Medicine*, vol. 12, no. 4, 2014, pp. 352–358., doi:10.1370/afm.1665.
7. Whelton, Paul K. “The Elusiveness of Population-Wide High Blood Pressure Control.” *Annual Review of Public Health*, vol. 36, no. 1, 2015, pp. 109–130., doi:10.1146/annurev-publhealth-031914-122949.
8. Haskard-Zolnieriek, Kelly, et al. “White Coat Hypertension: Improving the Patient–Health Care Practitioner Relationship.” *Psychology Research and Behavior Management*, 2015, p. 133., doi:10.2147/prbm.s61192.
9. Jones, Daniel W., et al. “Measuring Blood Pressure Accurately.” *Jama*, vol. 289, no. 8, 2003, p. 1027., doi:10.1001/jama.289.8.1027.
10. Pickering, Thomas G., et al. “Recommendations for Blood Pressure Measurement in Humans and Experimental Animals.” *Hypertension*, vol. 45, no. 1, 2005, pp. 142–161., doi:10.1161/01.hyp.0000150859.47929.8e.
11. Parati, Gianfranco, et al. “Accuracy of Blood Pressure Measurement: Sphygmomanometer Calibration and Beyond.” *Journal of Hypertension*, vol. 24, no. 10, 2006, pp. 1915–1918., doi:10.1097/01.hjh.0000244935.19299.f5.
12. Seravalle, G., et al. “Blood Pressure Measurement, White-Coat and Masked Hypertension.” *Updates in Hypertension and Cardiovascular Protection Prehypertension and Cardiometabolic Syndrome*, 2018, pp. 383–394., doi:10.1007/978-3-319-75310-2_27.
13. Pickering, Thomas G., et al. “ ‘Call To Action on Use and Reimbursement for Home Blood Pressure Monitoring: Executive Summary.’” *The Journal of Clinical Hypertension*, vol. 10, no. 11, 2008, pp. 885–886., doi:10.1111/j.1751-7176.2008.0033_2.x.
14. Merai, Rikita, et al. “CDC Grand Rounds: A Public Health Approach to Detect and Control Hypertension.” *MMWR. Morbidity and Mortality Weekly Report*, vol. 65, no. 45, 2016, pp. 1261–1264., doi:10.15585/mmwr.mm6545a3.
15. Rakotz, M. K., et al. “A Technology-Based Quality Innovation to Identify Undiagnosed Hypertension Among Active Primary Care Patients.” *The Annals of Family Medicine*, vol. 12, no. 4, 2014, pp. 352–358., doi:10.1370/afm.1665.

16. Burke, Lora E., et al. "Current Science on Consumer Use of Mobile Health for Cardiovascular Disease Prevention." *Circulation*, vol. 132, no. 12, 2015, pp. 1157–1213., doi:10.1161/cir.0000000000000232.
17. Ijaz, Muhammad, et al. "Hybrid Prediction Model for Type 2 Diabetes and Hypertension Using DBSCAN-Based Outlier Detection, Synthetic Minority Over Sampling Technique (SMOTE), and Random Forest." *Applied Sciences*, vol. 8, no. 8, 2018, p. 1325., doi:10.3390/app8081325.
18. National Center for Health Statistics (NCHS). (2019, March 06-a). Retrieved from <https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx>.
19. National Center for Health Statistics (NCHS). (2019, March 06-b). Retrieved from <https://wwwn.cdc.gov/nchs/nhanes/NhanesCitation.aspx>.
20. Hughes, Michael D., and Stuart J. Pocock. "Within-Subject Diastolic Blood Pressure Variability: Implications for Risk Assessment and Screening." *Journal of Clinical Epidemiology*, vol. 45, no. 9, 1992, pp. 985–998., doi:10.1016/0895-4356(92)90114-3.
21. Rosner, Bernard, and B.frank Polk. "The Implications of Blood Pressure Variability for Clinical and Screening Purposes." *Journal of Chronic Diseases*, vol. 32, no. 6, 1979, pp. 451–461., doi:10.1016/0021-9681(79)90105-x.
22. Müller, K. R., Mika, S., Rätsch, G., Tsuda, K., & Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE transactions on neural networks*, 12(2).
23. McLachlan, Geoffrey J. "Mahalanobis distance." *Resonance* 4.6 (1999): 20-26.
24. Clark, L. A., Denby, L., Pregibon, D., Harshfield, G. A., Pickering, T. G., Blank, S., & Laragh, J. H. (1987). A data-based method for bivariate outlier detection: Application to automatic blood pressure recording devices. *Psychophysiology*, 24(1), 119-125.
25. Garrett, R. G. (1989). The chi-square plot: a tool for multivariate outlier recognition. *Journal of Geochemical Exploration*, 32(1-3), 319-341.
26. Team, R. C. (2019). R: A language and environment for statistical computing.
27. Buehler, Eric (2019). Senior-research. <https://github.com/buehlere/senior-research>.
28. Chang,W., Cheng, J., Allarie, J., Xie, Y., & McPherson, J. (2019). Shiny: Web Application Framework for R. R package version 1.2.0. <https://CRAN.R-project.org/package=shiny>.

Appendix

Figure A1: Quantiles of Residuals for Systolic Blood Pressure

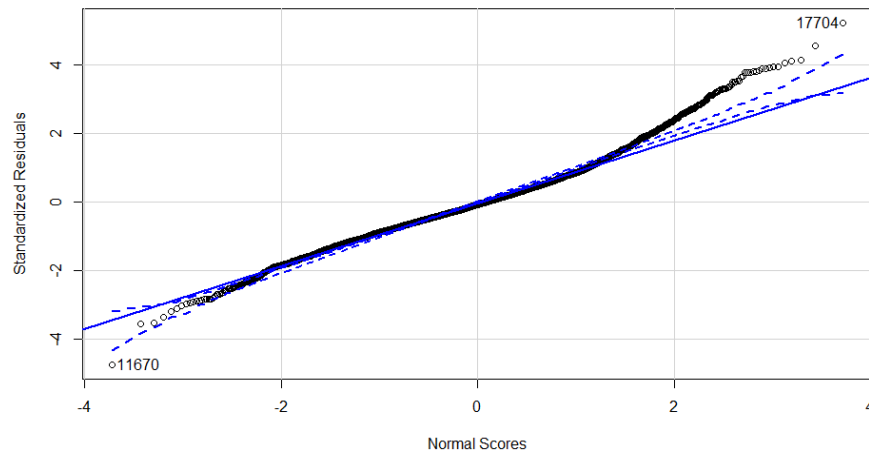


Figure A2: Quantiles of Residuals for Diastolic Blood Pressure

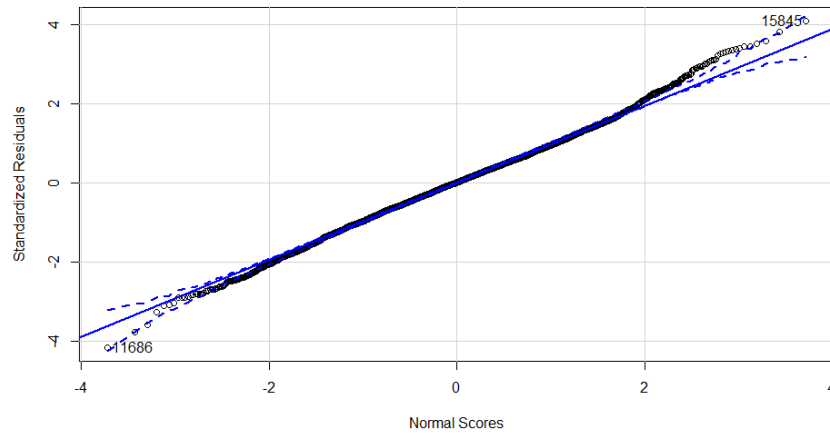


Table A1:

Variable	Systolic Model	Diastolic Model
----------	----------------	-----------------

(Intercept)	116.81 ***	66.22 ***
	-0.65	-0.33
Arm Circumference (cm)	`-1.97***	`-1.66 ***
	-0.52	-0.35
Sagittal Abdominal Diameter (cm)	5.48 ***	1.16 ***
	`-0.70	`-0.34
Standing Height (cm)	`-1.51 ***	
	-0.4	
Waist Circumference (cm)	`-2.95 ***	
	-0.82	
Weight (kg)	2.28 *	
	-0.93	
Gender (0=female/1=male)	5.52 ***	1.99 ***
	-0.59	-0.31
Age (yrs.)	-0.42	15.22 ***
	-1.23	-0.91
Age^2	6.81 ***	`-15.17 ***
	-1.23	-0.96
Hispanic	-0.87	
	-0.64	
White	`-1.85 **	0.64
	-0.62	-0.35
Black	3.34 ***	1.28 **
	-0.7	-0.41
Diabetes (0=no/1=yes)	-0.34	0.33
	-1.06	-0.46
LDL-Cholesterol (mg/dL)	`-2.63 ***	
	-0.51	
log(Total Cholesterol) (mg/dL)	3.12 ***	
	-0.5	
log(Fasting Glucose) (mg/dL))	1.61 ***	
	-0.33	
Hypertensive (yes/no)	7.26 ***	3.83 ***
	-0.95	-0.68
Hypertension Medication (yes/no)	1.06	0.91
	-1.05	-0.78
Arm Circumference:Age	`-0.98 ***	
	-0.23	
Gender:LDL	1.35 ***	0.56 *
	-0.4	-0.28
Gender:Fasting Glucose	-0.64	
	-0.39	
SAD:Diabetes	`-1.86 **	
	-0.644337149	
Gender:Age	`-2.48 ***	
	-0.42	
SAD:Hypertension Medication	`-1.97 ***	
	-0.53	
Age:White	`-2.31 ***	`-0.59 *
	-0.41	-0.29

Age:Diabetes	1.55	
	-0.86	
Age:LDL-Cholesterol	1.11 ***	
	-0.21	
`log(Weight)`		1.46 **
		-0.44
Asian		2.98 ***
		-0.49
`log(LDL-Cholesterol)` (mg/dL)		`-1.00 **
		-0.34
Total Cholesterol (mg/dL)		1.73 ***
		-0.33
SAD:Age		`-1.01 ***
		-0.16
Age:Hypertension Medication		`-2.88 ***
		-0.48
N	4950	4950
R2	0.33	0.21
*** p < 0.001; ** p < 0.01; * p < 0.05.		

Figure 3A: Variance in Systolic Blood Pressure by Age

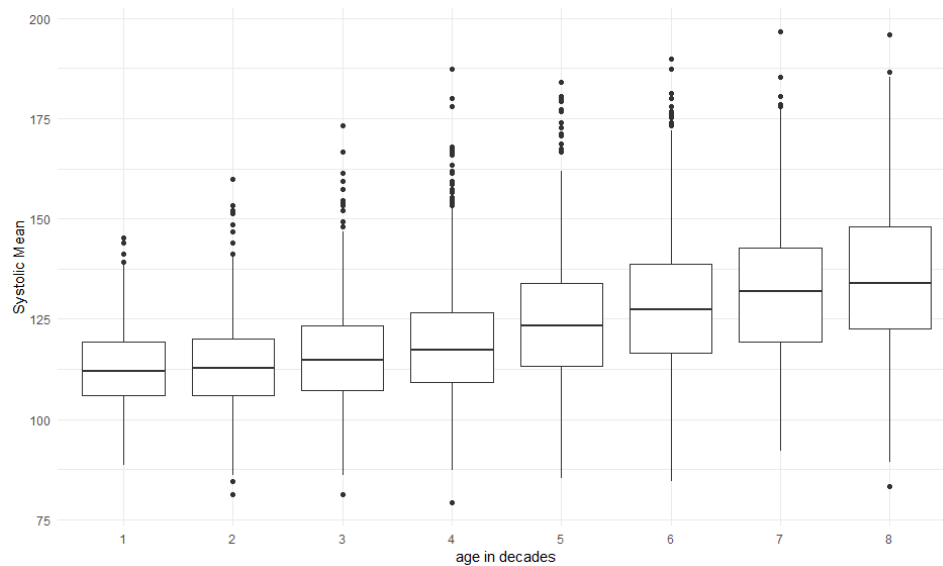


Figure 3A: Variance in Diastolic Blood Pressure by Age

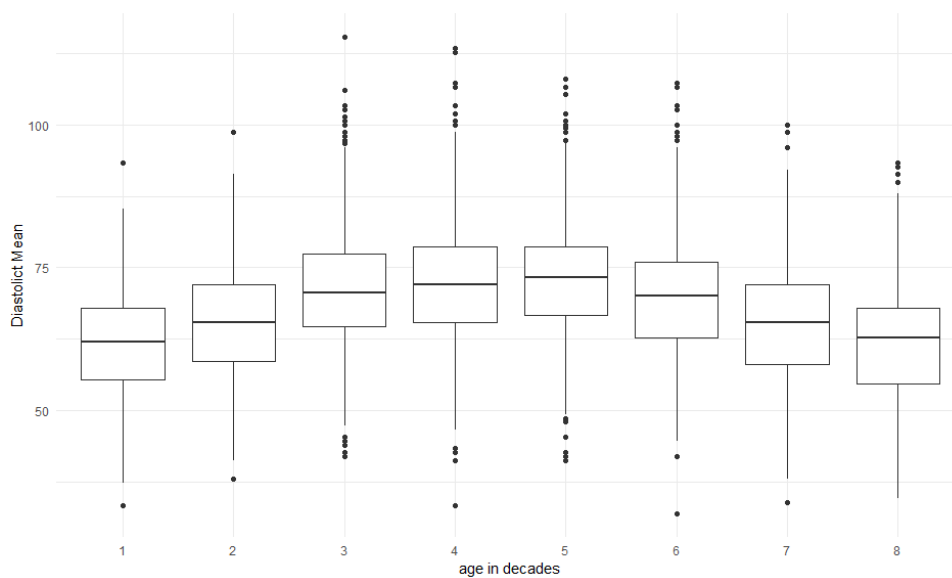


Figure 5A: True Blood Pressure compered to Random Forest Predictions

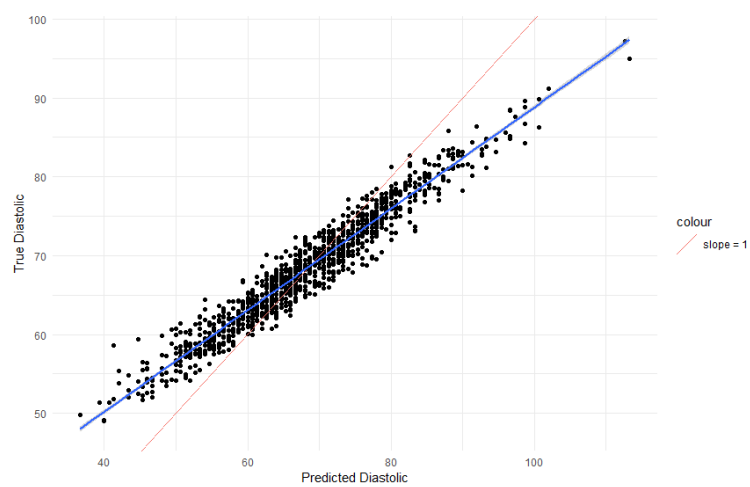
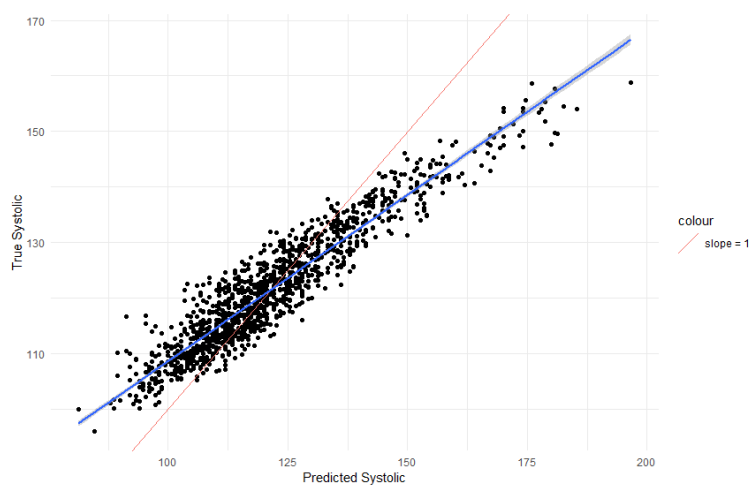


Figure Diastolic :

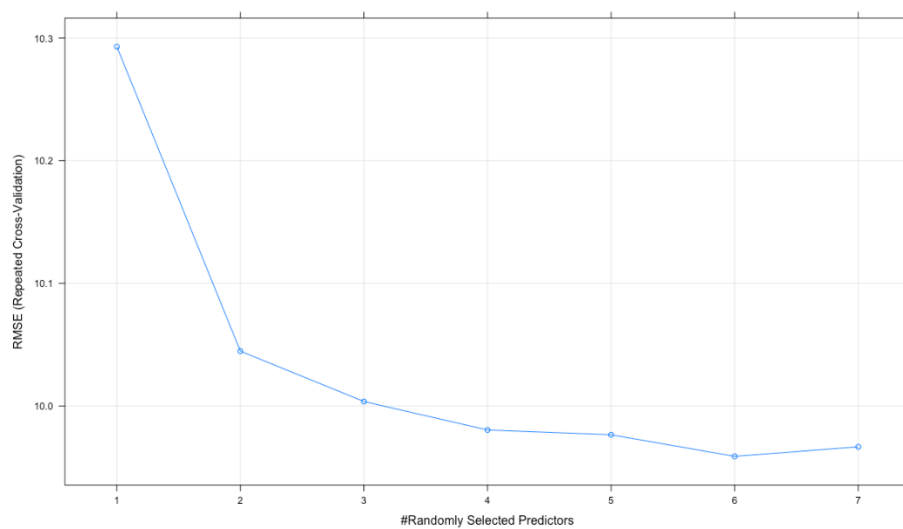
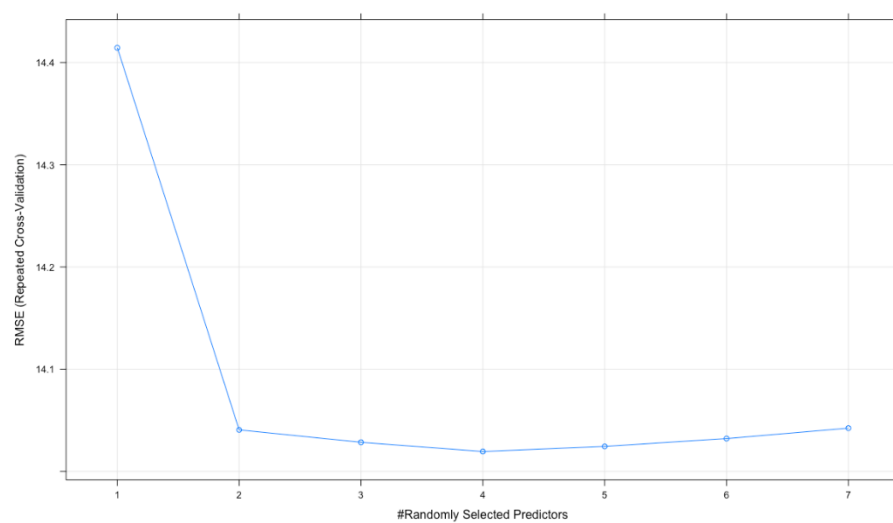
Figure A6: # of Variables considered for Systolic**Figure A7: # of Variables considered for Diastolic**

Figure A8: Mahalanobis distance against Chi-Square