Eric Buehler

**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). Unfortunately, the measurement of blood pressure is not as straightforward as one might expect due to its inherent variability and inconsistencies in measurement techniques. New approaches have emerged to meet this problem. However, many of them are largely tailored towards privileged groups, ignoring the frequently cited observations that those in rural communities, low-income communities and of minority status have some of the highest risk of hypertension. Thus, the primary goal of this study is to investigate how machine learning might be used in a straightforward and accessible manner to assist healthcare workers in measuring blood pressure. To do this, the study compares newer modeling techniques against traditional OLS regression and leverages these techniques to design a simple tool that can determine if an observed blood pressure is anomalous or not.

**2 Literature Review**

The Lancet reports that 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). There are many reasons why blood pressure remains difficult to control, however this study chooses to focus on the problems of measurement, and limitations of information and technology. 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.

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-Zolnierek, Kelly, et al., 2015). The effect of these problems is very 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).

Beyond human based errors, a litany of instrumental errors also affects in office and out of office blood pressure measurements. 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 are a common alternative to the mercury sphygmomanometer. Oscillometric 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 (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). 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 it’s 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). HBPM is often a more practical approach to ABPM as it can be reported at a patient’s convenience and has 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. Since HBPM relies on several measurement, it avoids many of the accuracy issues associated with oscillometric devices (Whelton, Paul K., et al., 2017). Nonetheless, the devices are still subject to miscalibration, which are difficult to detect in home. Overall, the application of either ABPM and HBPM is somewhat situational, but HBPM perhaps provides the best tradeoff between practicality and accuracy.

Despite the promise of ABPM and HBPM, their implementation in low‐resource settings would requires a change in the training of healthcare professionals, regulatory changes regarding the production and sale of BP devices, and technological advances (Seravalle, G., et al., 2018). The irony is that the individuals at the highest risk of hypertension are those, who have limited access to healthcare professionals, technology, and the latest public health evidence and training opportunities (Harris, Jenine K., et al., 2016). No studies have robustly evaluated the cost of implementing HBPM among these high-risk communities (Pickering, Thomas G., 2008). Thus, it seems that one of the most supported new techniques for taking accurate blood pressure measurements, HBPM, generally assumes a privileged audience, and pays little mind to rural and low-income communities who perhaps have the most to gain from more accurate blood pressure readings. In general, the cacophony of instruments and standards used in blood pressure measurements presents the need for useful baseline measurements.

The issue of a privilege audience in new approaches to hypertension control also extends to other novel approaches, specifically those that leverage information technologies. 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 uses a unified electronic health record that stores data and makes it accessible across the entire health system, providing consistent information to professional (Merai, Rikita, et al., 2016). In addition, algorithms can be deployed on NorthShore’s health records to accurately identify those at risk of hypertension (Rakotz, M. K., et al., 2014). Telehealth is another common 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 as they have shown variability and inconsistency (Whelton, Paul K., et al., 2017). Additionally, communication between developers and health professionals on these telehealth technologies has been less than ideal (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.

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 professionals (Whelton, Paul K., et al., 2017; Whelton, Paul K., 2015; Harris, Jenine K., et al., 2016). While some of these solutions may prove scalable, little of the literature has investigated a detailed approach to implementing these solutions in disadvantaged communities.

Some research has already been done on how machine learning might be used in a straightforward and accessible manner to assist healthcare workers in measuring blood pressure. 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; Latifoğlu, Fatma, et al., 2018). These models provide an accessible way to assist in the measurement of blood pressure if trained on reliable data from appropriate populations. 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 NorthShore University Health System and KPNC. These simpler technology-based approaches provide the motivating foundation for this studies goal of proposing a simple tool that helps workers in the detection of anomalies in blood pressure measurements.

**3 Methods**

The primary needs of the proposed study are to predict an expected blood pressure for a given individual based on predictors of readily available medical data, using modern machine learning methods. The prediction will be used as a baseline to determine if an observed blood pressure measurement is erroneous or not based on Mahalanobis distance. Key limitations of this methodological goal lie in the data and prediction process. Data limitations are largely a result of lack of access to longitudinal or 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 limitations of the model being applied. The study also limits the scope of predictors used to common medical data for practical purpose, which further limits the model’s predictive power. The implantation of this methodology and limitations are discussed explicitly below.

**3.1 Data**

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 and children 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. While some of the data were taken by medical professional according to strict guidelines, they are still subject to measurement errors.

The primary dependent variables of this study are diastolic and systolic blood pressure. NHANES provides blood pressure 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 of the problems in 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 measurement 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 most desirable for decreasing 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.

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 (categorical and 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 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 does not require IRB approval (Denison, 2019). Therefore, the use of NHANES data for this study is in line with the standards of National Center of Health Statistics and Denison IRB.

**3.3 Analytical Approach**

The primary methods for this study will include a variety of predictive models, and outlier detection via Mahalanobis distance measure.

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 and KNN will be considered as well.

**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 values (quote). Since blood pressures vary 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 provides a more holistic method of stratification that may more appropriately control for the unequal variance in blood pressure. 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.

**Decision Tree**

Decision trees are common alternative to linear regression that are easier to understand and blood pressure’s issue of correlated predictors. The basic idea is that algorithm divides a predictor space into distinct, nonoverlapping regions. It does this by partitioning regions in a way that minimizes the sum of the square of the residuals (RSS). This partitioning allows for interactions of variables as each variable is forced to interact with every variable in subsequent partitions. 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 simple tree graph that is even easier to understand than a regression equation, and thus particular desirable for medical predictions. Despite these benefits, decision trees typically overfit data and have high variance. 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, this weakness may still result in even linear regression outperforming the decision tree.

**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*. 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. However, these benefits come at a cost to interpretability. There is no simple way to visualize a random forest and thus 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.

**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.

K-nearest Neighbors is another effective model that predicts a points value based on its distance to some number of neighbors. The benefit of this approach is that K-nearest can predict based on a point’s non-Euclidean distance to others, which is useful considering this study is predicting a multivariate distribution, constituted by systolic and diastolic blood pressure. However, K-nearest Neighbors has the drawback of not producing an easy to understand equation or graph to explain predictions. These methods performance will be compared by RMSE across the clusters determined in 3.1.

The models will be compared on the basis of mallahonbis

**3.3 Outlier Detection**

After a blood pressure curve is fitted to each cluster by the best performing model, the mahalanobis distances for the distributions will be calculated. Then, an undetermined heuristic will be used to determine an appropriate cut-off for the mahalanobis distance for each cluster. The cut-off values can then be used to determine if an observed blood pressure is within a reasonable distance in terms of both systolic and diastolic of a predicted blood pressure (McLachlan, 1999). If it is under the cut-off, it will be considered valid, else it will be considered erroneous. If time allows, different distance metrics/approaches may also be considered.

**4 Technology**

**4.1 Tools**

All algorithmic approaches described above will be implemented using R (R Core Team, 2018). Specific packages are still to be determined. Some Python may be used for scraping the necessary NHANES data (Python Core Team, 2019). All code will be made available on github (Buehler, 2019).

**4.2 App-Interface**

The system described in the analytical approach section will be implemented in 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 erroneous. If it is under, it will be considered valid.

**5 Conclusion**

In concluding, the methodology aims to update techniques used to stratify blood pressure groups via clustering, compare newer techniques of blood pressure prediction, and design a simple tool for assisting the measurement of blood pressure that is simple in comparison to techniques such as ABPM and HBPM, using NHANES data.

**Results**

The predictive baseline for blood pressure was established by fitting an OLS regression model to both systolic and diastolic blood pressure. Predictors were mainly selected drawing on support from the literature on blood pressure prediction (Whelton, P. K., et al., 2018). 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.) and interaction terms. Assumptions can be found in the appendix of the article. Ultimately the following equations for systolic and diastolic blood pressure were produced:

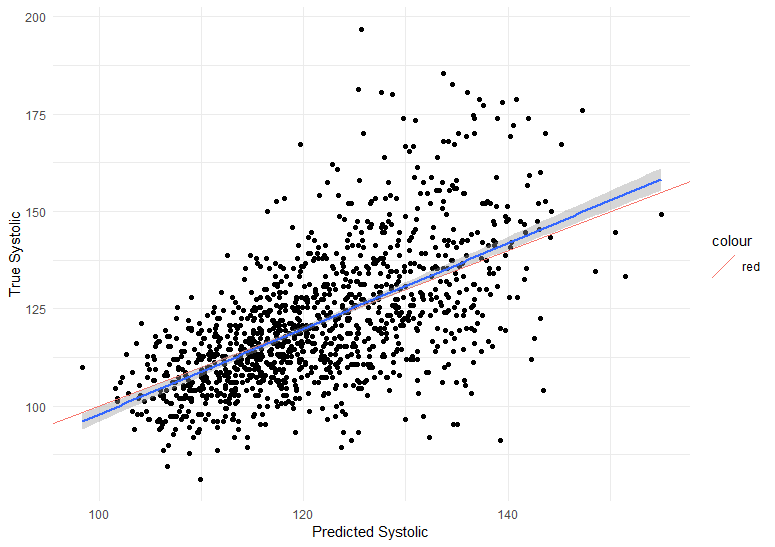
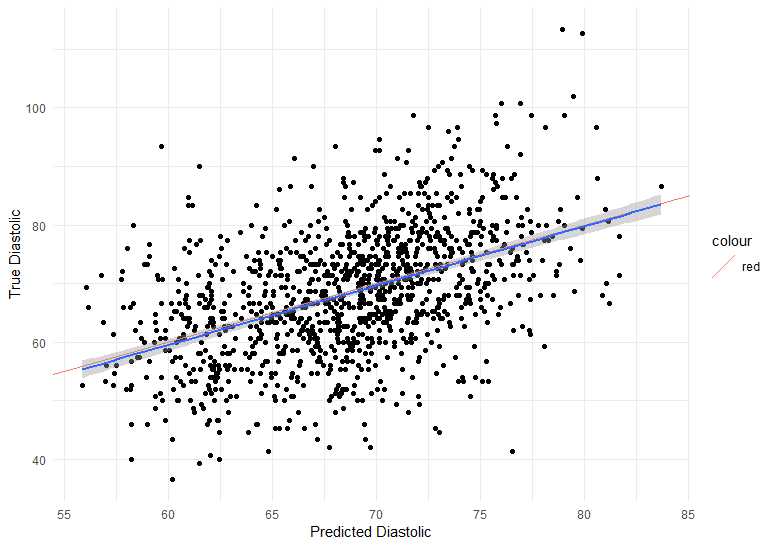
**Model 1:**

|  |  |  |  |
| --- | --- | --- | --- |
| names | Systolic Model |  |  |
|  |  |  |  |
| (Intercept) | 66.22 \*\*\* | | |
|  | -0.33 |  |  |
| arm circumference | -1.66 \*\*\* | | |
|  | -0.35 |  |  |
| abdominal circumference | 1.16 \*\*\* | | |
|  | -0.34 |  |  |
| log(weight) | 1.46 \*\* | | |
|  | -0.44 |  |  |
| gender | 1.99 \*\*\* | | |
|  | -0.31 |  |  |
| age | 15.22 \*\*\* | | |
|  | -0.91 |  |  |
| age^2 | -15.17 \*\*\* | | |
|  | -0.96 |  |  |
| white | 0.64 |  |  |
|  | -0.35 |  |  |
| black | 1.28 \*\* | | |
|  | -0.41 |  |  |
| asian | 2.98 \*\*\* | | |
|  | -0.49 |  |  |
| diabetes | 0.33 |  |  |
|  | -0.46 |  |  |
| `log(ldl cholesterol)` | -1.00 \*\* | | |
|  | -0.34 |  |  |
| weight | 1.73 \*\*\* | | |
|  | -0.33 |  |  |
| hypertensive | 3.83 \*\*\* | | |
|  | -0.68 |  |  |
| hypertension medication | 0.91 |  |  |
|  | -0.78 |  |  |
| age:white | -0.59\* | | |
|  | -0.29 |  |  |
| abdominal diameter:ridageyr | -1.01 \*\*\* | | |
|  | -0.16 |  |  |
| age:hypertension medication | -2.88 \*\*\* | | |
|  | -0.48 |  |  |
| gender:ldl cholesterol | 0.56 \* | | |
|  | -0.28 |  |  |
| N | 4950 |  |  |
| R2 | 0.204539 |  |  |
| \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05. | | | |

**Model 2:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| names | Diastolic Model |  |  | |
|  |  |  |  | |
| (Intercept) | 66.22 \*\*\* | | |
|  | -0.33 |  |  | |
| arm circumference | -1.66 \*\*\* | | |
|  | -0.35411 |  |  | |
| abdominal diameter | 1.16 \*\*\* | | |
|  | -0.34 |  |  | |
| `log(weight)` | 1.46 \*\* | | |
|  | -0.44 |  |  | |
| Gender | 1.99 \*\*\* | | |
|  | -0.31667 |  |  | |
| Age | 15.22 \*\*\* | | |
|  | -0.91204 |  |  | |
| age^2 | -15.17 \*\*\* | | |
|  | -0.96 |  |  | |
| white | 0.64 |  |  | |
|  | -0.35 |  |  | |
| black | 1.28 \*\* | | |
|  | -0.41 |  |  | |
| asian | 2.98 \*\*\* | | |
|  | -0.49 |  |  | |
| diabetes | 0.33 |  |  | |
|  | -0.46 |  |  | |
| log(ldl cholesterol) | -1.00 \*\* | | |
|  | -0.34 |  |  | |
| total cholesterol | 1.73 \*\*\* | | |
|  | -0.33 |  |  | |
| hypertensive | 3.83 \*\*\* | | |
|  | -0.68 |  |  | |
| hypertension medication | 0.91 |  |  | |
|  | -0.78 |  |  | |
| age:White | -0.59 \* | | |
|  | -0.29 |  |  | |
| abdominal:age | -1.01 \*\*\* | | |
|  | -0.16 |  |  | |
| age:hypertensive medication | -2.88 \*\*\* | | |
|  | -0.48 |  |  | |
| gender:ldl cholesterol | 0.56 \* | | |
|  | -0.28 |  |  | |
| N | 4950 |  |  | |
| R2 | 0.204539 |  |  | |
| \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05. | | | |

**Figure 1: Predicted Blood Pressures Compared to True Blood Pressures**



**Discussion**