Objective:

Chang Chen:

Pulse pressure is an important indicator of heart health, thus, it is meaningful to explore the relationships between the pulse pressure and different exposures and mediators. The goal of this analysis is: 1) to explore how the mean pulse pressure differs by age, sex and race; 2) to investigate whether the relationship between age and mean pulse pressure is mediated through weight or height or body mass index after adjusting for sex and race.

Scott Mu:

Blood pulse pressure, or the difference between systolic blood pressure and diastolic blood pressure, is a physiologic measure that could be an indicator of cardiovascular disease. The relationship between pulse pressure and demographic characteristics such as age, sex and race is incompletely characterized. Additionally, other physiologic measures such as height, weight or body mass index (BMI) could mediate the relationship between age and pulse pressure. Our first hypothesis is that age, sex, race and their mutual interactions are all statistically signficiant predictors of pulse pressure. Our second hypothesis is that the relationship between age and pulse pressure among individuals of the same sex and race is mediated through weight, height or BMI.

Data:

Scott Mu:

The National Health and Nurition Examination Survey is a U.S. based cross-sectional study of ambulatory adults and children, using a sophisticated sampling strategy. The available dataset includes a subset of the 2015-2016 data, with 5395 participants aged 20 or greater. The recorded measurements included sex, age in years, race, systolic blood pressure, diastolic blood pressure, weight, height, and body mass index. To analyze the first hypothesis, sex, age and race will be exposure variables and calculated pulse pressure the outcome variable. For the second objective, height, weight, and BMI will be considered potential mediators in the relationship between exposure of age and outcome of pulse pressure.

Methods:

Mena Ru:

Descriptive statistics were summarized for PP by sex and race. Kruskal-Wallis tests were performed to compare PP between sex, race and simple linear regression models were built to evaluate the effect of age on PP With significant associations for all three variables, we included them in the proposed model. Next, assumptions of the mean model, independence, constant variance and interactions between covariates were evaluated using diagnostic plots. Since a non-linear relationship was identified between age and PP as well as signs of interactions and heteroscedasticity, a generalized least square model was proposed with a spline knot at 60 years old and centering at 40 years old to adjust for the nonlinearity, non-constant variance (as a function of age), as well as interactions between age and sex/race. Mediation analysis was conducted as two-folds: first, check if weight, height and BMI met the definitions of mediators through simple linear regression models; second, use the difference method assess the scale of mediation effects qualitatively. All analyses were conducted in R. Hypothesis tests were conducted as two-sided at an alpha level of 0.05.

Xiangji Ying:

We fit a linear regression model where the mean pulse pressure was modeled as a smooth function of age (natural cubic spline with 3 degrees of freedom) separately for males and females and adjusted for race. The inclusion of race, age, and sex in the model was determined by the statistical significance of the univariate

linear models and likelihood ratio tests for the effect of adding a variable. We modeled two-way and three-way interactions among race, age, and sex and examined if adding the interaction term significantly contributed to the model using likelihood ratio tests. To examine if the mean model was correctly specified, we used fitted line graphs and the Akaike information criterion (AIC) statistic to compare models with different functional forms of age, including linear form, linear splines, natural cubic splines (with 2 and 3 degrees of freedom), and polynomial terms. The normality and constant variance assumption of the residuals were checked by QQ plot and residual plots, respectively. Due to the potential violation of the normality and constant variance assumptions, 95% confidence intervals for the coefficients were derived using the bootstrap method, where 1000 bootstrap samples were drawn with replacement from the original sample, and the percentile bootstrap method was used to compute the confidence interval. Data points with high influence and/or leverage were checked with dbetas, dfits, and leverage statistics. Sensitivity analysis was conducted by re-running the final model after excluding data points with high influence and/or leverage. We performed mediation analyses using the difference method by comparing the estimated coefficients before and after the inclusion of the potential mediators in the model. Indirect effect, which is the difference between the coefficients before and after adjustment, was estimated.

Scott Mu:

We performed initial data inspection to ensure data quality and to quantify the degree of missing or spurious data. Comparisons between individuals with missing data and those with complete data were made to find potential correlations between covariates and missing data.

Using only data from participants with no missing or spurious data, we used a manual, iterative model selection approach to decide on an optimal final model. We started with ordinary least squares regression of age on pulse pressure, and increased model complexity by adding natural cubic spline terms for age with different degrees of freedom. We also investigated relaxing the constant variance assumption and fit generalized least squares models that allowed the pulse pressure variance to vary with age. At each stage, we checked our model assumptions by plotting the residuals against age and predicted pulse pressure. After we settled on our final model which included natural cubic splines of age with 3 degrees of freedom, we conducted nested comparisons with the partial F-test, sequentially adding sex, race and their permuted interactions. The full model and simplified models were used to report specific age, sex and race outcomes.

For our mediation analysis, we sought to determine whether the relationship between age and average pulse pressure among individuals with the same sex and race was mediated through weight, height or body mass index. With a ordinary least squares model of age, sex, race, and all mutual interactions as exposures and pulse pressure as the outcome, we calculated for each potential mediator (height, weight, BMI) the average direct effect, average causal mediated effect, and proportion mediated, using a non-parametric bootstrap method.

Statistical analysis was performed in the R statistical programming language, version 4.0.4. Key statistical packages included mediation, tidyverse, splines, and nlme.

Results:

Scott Mu:

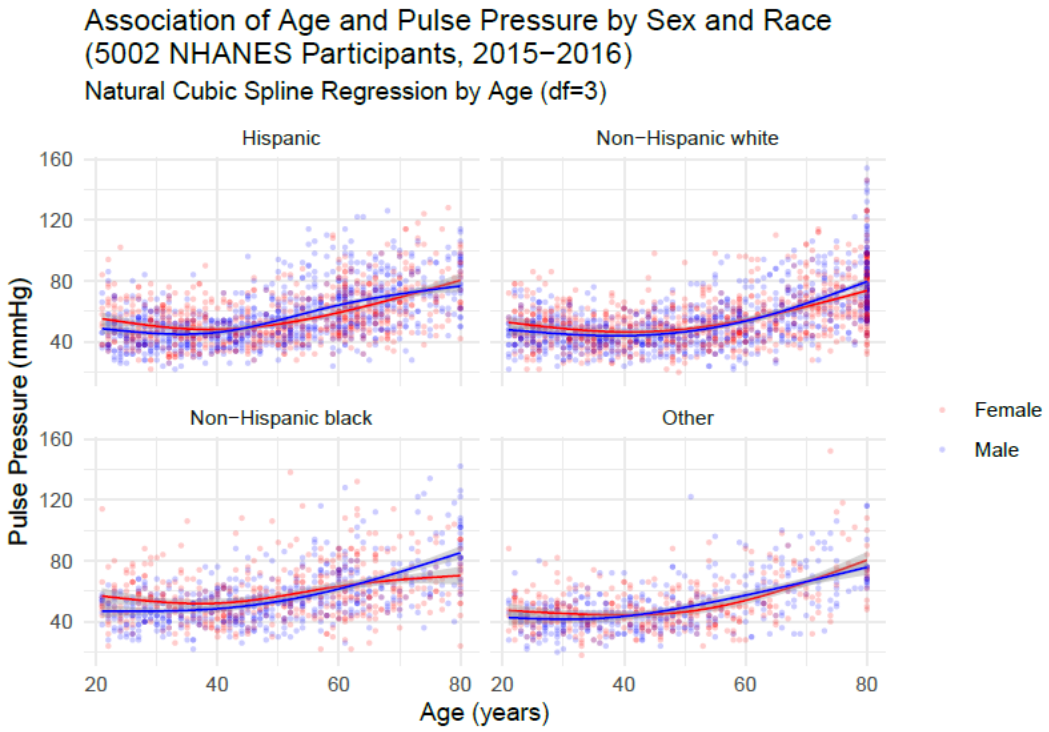
Effect of age, sex, race on pulse pressure: In total, 7 nested models were compared. The first model was a null model with intercept only, and each successive model added: cubic natural splines of age with 3 degrees of freedom, female sex, interaction of sex and natural splines of age, race, interaction of race and female sex, and full interactions of age splines, race and female sex. Adding natural splines of age was statistically significant (p=2.2e-16), as was the interaction of female sex with age (p=4.696e-12), the addition of race to the interaction model (p=2.2e-16) , and the full 3-way interaction model (p=1.384e-07). Because of the complexity of the full interaction model, selected comparisons will be drawn from simpler models. Figure

1 shows the natural cubic spline regressions by sex, stratified by race.

There was a complex non-linear association between age and pulse pressure, at first decreasing, reaching a nadir at age 36, then increasing with age. At age 20, not accounting for sex or race, the average pulse pressure was 50.3 mmHg, (95% CI 48.7 to 51.9). At age 36, the average pulse pressure was 46.2 mmHg (95% CI: 45.4 to 46.9) and age 80, 77.0 mmHg (95% CI 75.7 to 78.3).

Younger females tended to have higher pulse pressures than their male counterparts, but this relationship was reversed by age 50, whereupon males had higher pulse pressures. The average pulse pressure for 20 year old females, not accounting for race, was 53.8 mmHg (95% CI 51.6 to 56.1), and for males, 47.3 mmHg (95.% CI 45.2 to 49.4). By age 80, the average female pulse pressure was 74.8 (95% CI 72.9 to 76.6), and for males, 79.2 (95% CI 77.4 to 81).

Race and its interactions were significant predictors of pulse pressure. A simplified model without race interactions allows estimation of overall effects of race on pulse pressure, which showed that compared to Hispanics, non-Hispanic whites had 3.9 mmHg decreased pulse pressure (95% CI 5.0 to 2.9), non-Hispanic blacks had 1.4 mmHg increased pulse pressure (95% CI 0.2 to 2.5), and those with listed race of Other had 4.7 mmHg decreased pulse pressure (95% CI 6.0 to 3.4). However, as the figure demonstrates, there were additional interactions of pulse pressure by age and sex within each race strata.
Overall, there were complex non-linear relationships between covariates, but age, sex, race, and the mutual interactions were all statistically significant predictors of pulse pressure.



Mena Ru:

A non-linear relationship between age and PP was identified during model assumption checking. PP remained the same for people aged 20-40 years old and started to increase as people age. Independence of samples can be assumed with population sampling, but non-constant variance was detected thus variance was weighted as a function of age in the proposed model. Interactions between age and sex/race on PP were found significant

by F tests. Based on the extended model with both interactions (Table 1), within the same race, the annual growth rate difference (AGRD) for a female aged >40 years old to aged 20-40 years old compared to that of males is 0.38mm Hg, and the AGRD for a female aged >60 years old to aged 40-60 years old compared to that of males is 0.75mm Hg. Similarly within the same sex, the AGRD for a White person aged >40 years old to aged 20-40 years old is 3.49mm Hg less than that of Blacks, and the AGRD for a White person aged >60 years old to aged 40-60 years old is 2.93mm Hg less than that of Blacks.

Chang Chen:

2) Table 1 provided the bootstrap direct and indirect effect estimates along with 95% credible intervals of age on the outcome mediated by BMI, height and weight. The results showed that there existed mediation effects of weight and BMI on the relationship between age and mean pulse pressure when controlling sex and race, since the corresponding credible intervals did not include 0. For example, the mediated effect of BMI on the relationships between mean pulse pressure and age, age spline term were 0.028, -0.040 mmHg/year, respectively.

Table 1: Direct and indirect effects of age on mean pulse pressure meidated by BMI, height, weight

	Direct effect	Indirect effect
BMI-age	-0.289 (-0.367, -0.212)	0.028 (0.018, 0.04)
BMI-agesp42	1.051 (0.948, 1.156)	-0.040 (-0.057, -0.026)
Height-age	-0.263 (-0.343, -0.185)	0.001 (-0.001, 0.005)
Height-agesp42	1.004 (0.902, 1.106)	0.007 (-0.001, 0.016)
Weight-age	-0.283 (-0.36, -0.206)	0.021 (0.012, 0.032)
Weight-agesp42	1.049 (0.946, 1.153)	-0.038 (-0.055, -0.022)

Xiangji Ying:

A total of 5002 adults were included in the analysis. Figure 1 displays the observed and estimated mean pulse pressure as a smooth function of age, separately for males and females based on the least-squares model and adjusted for race. For both genders, pulse pressure was slightly decreasing before age 40 and started to increase rapidly afterward. The natural spline terms of age from the least-squares model were significant, indicating a non-linear relationship between age and pulse pressure (Table 1). Among Hispanic males, we estimated that the average pulse pressure for ages 21, 40, and 60 was 55.1 mmHg, 49.4 mmHg, and 59.4 mmHg, respectively, and that 95% of their pulse pressure will fall within [52.9, 57.3],[48.1, 50.8] and [58.1, 60.6]. We found significant evidence in the data to suggest that the change rate of pulse pressure differed by gender group (likelihood ratio test p-value<0.001; all p-value<0.05 in the model, Table 1). The average pulse pressure among females was 6.4 mmHg (model 95%confidence interval: [-9.2, -3.5]) lower than that of males of the same age and race (Table 1). Within the same age and gender, compared to Hispanics, the average pulse pressure was 3.9 mmHg (model 95%confidence interval: [-5, -2.9]) lower among Non-Hispanic White, 1.4 mmHg (model 95%confidence interval: [0.2, 2.5]) higher among Non-Hispanic Black, and 4.7 mmHg (model 95%confidence interval: [-6, -3.4]) lower among other races (Table 1). The confidence intervals generated by bootstrapping were close to those from the model. The model coefficients remained very similar after excluding the data points with high influence and/or leverage. There were no large differences in the coefficients of the main effect and interaction terms of age between the original model coefficients and the mediated model coefficients (Table 1). The estimated indirect effects via all three potential mediators were close to 0.

Discussion/Limitations:

Chang Chen:

The weighted multiple linear regression model we fitted had some limitations. First, we only added one age linear spline, so the model only allowed two constant change rates of mean pulse pressure in the two age intervals for the same sex and race group. Applying natural cubic splines may be more flexible to characterize the trends. Second, this model did not allow the change rates of mean pulse pressure for different race groups to vary after adjusting for sex, since the model including race-age interactions was not significantly different in the likelihood ratio test (p-value>0.05). However, if we used natural splines to more accurately characterize the data, the effect of race-age interactions may become significant. As for the mediation analysis, one limitation was that the suboptimal mean model may lead to biasedness and/or inprecision of the direct and indirect effect estimates. Furthermore, we did not consider interactions between mediators and age. For example, the mean BMI and weight for the participants in the data set may differ by age, thus the mediated effect estimates may be biased. Models allowing age-mediator interactions can address this problem and further validation is needed.

Xiangji Ying:

The current study has a few limitations. The complete case analyses we used relied on the assumption that data were missing completely at random or that the model was correctly specified. However, extensive model selection and checking were undertaken. Thus we believe the current model is an excellent fit to the data. Model checking indicated violations of the normality and constant variance assumptions. However, the similar confidence intervals generated by the bootstrapping process and the large sample size of the current study suggested confidence in the inferences. There may still be omitted variable bias as the current model only considers three variables. Future studies may want to explore the health disparities of high blood pressure by other factors such as socioeconomic status.