

# Quantile Regression Analysis for Statin Effects on Body Mass Index

September 03, 2021

## Abstract

An individual's body mass index (BMI) is an important predictor of a cardiovascular diseases, stroke, diabetes and many other medical conditions. In this article, we use quantile regression to study the effects of age, total cholesterol, and use of medication on the conditional distribution of BMI and compare these results to those obtained by ordinary least squares. We study age and total cholesterol effect on the BMI in two different ways: quadratic terms and spline basis expansions. The national health and nutrition examination survey (NHANES) data is used in this study. The differences between using quadratic terms and spline basis expansion for the predictors is also presented.

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Quantile Regression</b>	<b>5</b>
<b>3</b>	<b>Splines Verses Polynomials</b>	<b>7</b>
<b>4</b>	<b>Methods</b>	<b>8</b>

<b>5</b>	<b>Numerical Example</b>	<b>9</b>
5.1	Data . . . . .	9
<b>6</b>	<b>Results</b>	<b>11</b>
6.1	Age effects Analysis . . . . .	14
6.2	TC effects Analysis . . . . .	16
6.3	Analysing Race\ Ethnicity effect on BMI . . . . .	18
<b>7</b>	<b>Conclusion</b>	<b>21</b>
<b>8</b>	<b>References</b>	<b>22</b>

# 1 Introduction

High body mass index (BMI) has a negative impact on an individuals health (Nuttall 2015) because there is a positive relationship between the risk of death from all causes including cardiovascular disease, cancer, or other diseases and BMI for both genders (Calle et al. 1999). Due to high cost of obesity on an individuals health, it is crucial to get better understanding of the underlying risk factors for obesity. Understanding the causes of obesity can help decision makers create new recommendations that help to prevent and stop the dramatic increase in BMI.

BMI plays an important role in predicting heart disease risk (Katzmarzyk et al. 2012). Approximately 18 million deaths annually per year world wide are caused by cardiovascular diseases(CVD), and this number is similar to the number of nonfatal cardiovascular events (Hay et al. 2017). In 2011, annual costs for CVD and stroke were estimated at \$320.1 billion, which is more than what is spent on cancer. The total CVD cost includes \$195.6 billion in direct costs (health-care costs) and \$124.5 billion of future productivity loss (Mozaffarian et al. 2015). Important risk factors associated with CVD are high BMI and abnormal lipid ratio (Yusuf et al. 2004; Anderson et al. 1991). There are some drugs that are used reduce CVD risks by lowering low-density lipoprotein (LDL) including statins. Statin use reduces the risk of cardiovascular diseases, even with a population with no CVD risk factors (Yusuf et al. 2016). On the other hand, statin uses is associate with negative side effects on health conditions including increased BMI. The effect of BMI on the choice of lipid-lowering treatment has been studied by (Ferrières et al. 2018). Where they found that there is a positive correlation ( $\rho = 0.13$ ) between statin does levels and BMI levels. Additionally, it is been shown that statin users consume 192 additional calories per day which causes gaining a 6 lb to 11 lb. in a year. Statin users gain 1.3 units in the BMI measure while non-staining users gain  $0.4 \text{ kg/m}^2$ . Moreover, consumption of fat in statin users raised by 14.4%(Sugiyama et al. 2014). This study gives us insight about the association between cholesterol drug uses and condition means of the BMI.

An important question to ask is, are there associations between cholesterol drug uses and different BMI quantiles? This question helps us to understand cholesterol drug uses effect on the obese, overweight population, so that a decision maker take an optimal action about treatment policy.

Ordinary least squares (OLS) regression helps us to study the effects of predictors on the mean of the response. For example, estimating the association of different factors like a lipid-lowering drug with the conditional mean of BMI (Ferrières et al. 2018). However, OLS approaches does not give us insight into the predictor's effects on different quantiles of the responses.

Instead, quantile regression (QR) is used to investigate the conditional relationship heterogeneity of the  $\tau$ th quantile  $0 < \tau < 1$  of BMI given with a set of independent predictors. Estimation of underweight and overweight regressed on age could be hard because of the trend change and nonlinearity behavior of the BMI with respect to age, as it is shown in (Flegal 1999). Therefore, QR is an appropriate method used to estimate the differences in the association of the BMI with predictors.

QR has a wide range of applications. For example, it is used to study the association between BMI and the set of predictors; low childhood socioeconomic position, high maternal weight, low childhood general cognition is stronger in the upper end of BMI quantile in the UK population(Bann, Fitzsimons, and Johnson 2020). One cause of this heterogeneity is that risk factors may have stronger effect on patients with worse health, and these effects may diminishes when conditional distribution of the BMI is studied. Moreover, another application of using QR is in ecology where different factors interact in a complicated way that produces different variations of one factor for different levels of another variable (Cade and Noon 2003).

We compare two approaches to study the QR of BMI regressed on age and total cholesterol (TC) levels. One approach is called quantile polynomial regression. It is defined as a way to model the non-relationship between predictors and distribution of the response as an  $n$ th

degree polynomial in the predictors. And the other approach is using a spline basis expansion of the predictors.

Polynomial regression using the quadratic term of predictors forces the response to take convex or concave shape, see Figure 1.12 (Koenker 2005). This is because the limit of the response variable  $y$ , when modeled as a quadratic function of the predictor variable  $x$ , approach  $\pm\infty$  ( i.e., as  $x \rightarrow \pm\infty$ ,  $y \rightarrow \pm\infty$ ). However, when we model the response variable using a spline basis expansion of the predictor variable the response is not forced to take a specific shape. In the latter case, different polynomials are constructed for the different regions in the range of predictors that satisfy continuity and smoothness conditions at the knots. Because of the above issue in the polynomial regression, we found that spline regressions produce a different estimate than polynomial regressions.

## 2 Quantile Regression

QR is a tool used to regress the dependent variable with high variance over the independent variables. QR is developed to study the relationships between predictors and the response when characteristics other than conditional mean of the response are of interest. Also, QR is a valuable tool to study association between predictors and response when we have a weak or no-relationships between predictors and the conditional mean of the response. Moreover, one of the advantages of using QR over OLS is QR is robust for outliers. A brief description of quantile regression is introduced in the following paragraphs. For a continuous random variable  $X$ , the cumulative distribution function (CDF) is

$$F(X) = P(X \leq x),$$

and the  $\tau$ th quantile of  $X$  is defined by

$$F^{-1}(\tau) = \inf\{x : F(x) \geq \tau\},$$

where  $0 < \tau < 1$ . Let the loss function of interest be defined by

$$\rho_\tau(u) = u(\tau - I_{(u<0)}),$$

where  $I$  is the indicator function (Koenker 2005). The quantile estimator is the value that minimizes the expected loss function

$$E\rho_\tau(X - \hat{x}) = (\tau - 1) \int_{-\infty}^{\hat{x}} (x - \hat{x}) dF(x) + \tau \int_{\hat{x}}^{-\infty} (x - \hat{x}) dF(x).$$

Differentiating with respect to  $\hat{x}$ , we get

$$0 = (\tau - 1) \int_{-\infty}^{\hat{x}} dF(x) + \tau \int_{\hat{x}}^{-\infty} dF(x) = F(\hat{x}) - \tau.$$

Due to the monotonicity of the cumulative distribution function, any solution that satisfies  $\{x : F(x) = \tau\}$  is a minimizer for the expected loss function.

OLS method expresses the conditional mean of  $y$  given  $x$  as  $\mu(x) = x^T \beta$  and it solves

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n (y_i - x_i^T \beta)^2.$$

The above equation represent the loss function for the OLS regression. By solving this equation for the minimum values for  $\beta$ , we get regression equation, which is  $\hat{y} = x_i^T \beta$ . On the other hand, QR expresses the conditional quantile function  $Q_y(\tau|x) = x^T \beta(\tau)$  and solves

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(y_i - x_i^T \beta)^2.$$

The above This minimization problem can be reformulated to a linear programming problem and it can be solved using standard statistical software.

### 3 Splines Verses Polynomials

Most of the relations between the responses and predictors variables are complicated to the point that linear regressions are not suitable to model these relationships(Bruce, Bruce, and Gedeck 2020). For example, the response to different levels of drug doses is not a linear relationship. Linear regressions can be generalized to deal with nonlinear effects. One approach is through including polynomial terms in the regression equation. A continuous predictor can be modeled as a linear, say  $x$ , or nonlinear term, say  $x^2$  depending on the relationship with the response variable. The mathematical model for  $n$  degree polynomial regression is shown in the Eq(1)

$$y = \beta_0 + \beta_1 X + \beta_2 X^2 + \cdots + \beta_n X^n. \quad (1)$$

One of the limitations of using polynomial regression is the curvature that can be captured is limited with low order terms. However, including higher-order terms has a negative impact on the model by introducing undesirable “wigginess” in the regression equation especially near the boundaries of the data. Another robust approach to model nonlinear relationships is splines. It is similar to a technique used by draftsmen in plotting curves. The spline is a process of constructing a set of piece-wise continuous polynomials that are smoothly connected at a set of points in the range of the predictor variable. The connection points are called knots i.e. splines are used to smoothly interpolate between certain points. Let  $a, b$ , and  $c$  are the endpoints of the  $x$ -axis intervals. A transformation for some or all of predictors is needed to capture the nonlinearity in the model. The family of transformation of the predictors that can be fit together to built the model’s shape is known as a basis function. the basis functions are  $b_1(x), b_2(x), \dots, b_k(x)$ , and the estimation of  $y_i$  is computed as follows:

$$y_i = \beta_0 + \beta_1 b_1(x_i) + \beta_2 b_2(x_i) + \cdots + \beta_k b_k(x_i) \quad (2)$$

This concept of a family of transformations that can fit together to capture general shapes is called a basis function. In this case, our objects are functions:  $b_1(X), b_2(X), \dots, b_K(X)$ . In a more general way to represent a value of  $y = f(X)$  using a piecewise cubic polynomials with a single knot  $c$ .

Imposing a continuity condition and first, and second derivative at  $c$  are equal in the two sides, we get a cubic splines.

## 4 Methods

A multivariate quantile regression model is used to assess the characteristics of the association variability in different quantiles of the conditional distribution of the body mass index.

The independent variables in our model are gender, race, age, total cholesterol and cholesterol drug use (yes or no). All types of cholesterol drugs are used including statins. The included races are non-Hispanic white, non-Hispanic black, Hispanic, and others. The relation between total cholesterol and fasting glucose has been studied in (Tsaousis 2014; Chang et al. 2011). They found that there is a positive correlation between the two groups on average. Therefore, fasting glucose level is not included to avoid collinearity. The following formulas represents the different QR models that fit conditional distribution of BMI on :

- Model I: Age enters the model as spline basis expansion, Race, Gender, and cholesterol drug use (Yes or No).
- Model II: TC enters the model as spline basis expansion, Race, Gender, and cholesterol drug use.
- Model III: Age enters the model as a quadratic term, Race, Gender, and cholesterol drug use.
- Model IV: TC enters the model as a quadratic term, Race, Gender, and cholesterol drug use.

## 5 Numerical Example

### 5.1 Data

The data used in this study is the National Health and Nutrition Examination Survey data (NHANES)(Disease Control and (CDC) 2018). The survey examines a nationally representative sample of the U.S. population focusing on a variety of health and nutrition measurements released in two year cycle. In this study, we accumulated 6 cycles of NHANES data (2007–2018). We used two data files: One contains demographic variables, such as age, sex, race, income, etc. and the other contains data that are related to body measurements, such as BMI, head circumference, etc. These files are merged by using the respondent sequence number (SEQN) resulting in around 12,000 records. The missing values are deleted. We selected a population age between 20 and 80. BMI are classified into different categories according to underweight,  $18.5 \text{ kg}/\text{m}^2$ ; normal weight,  $18.5$  to  $25 \text{ kg}/\text{m}^2$ ; overweight,  $25$  to  $30 \text{ kg}/\text{m}^2$ ; obese,  $30$  to  $35 \text{ kg}/\text{m}^2$ ; and very obese more than  $35 \text{ kg}/\text{m}^2$ .

Syntax	Male	Female
count	5990	6416
Mean of Age	49.9	49.73
BMI	28.549	29.379
TC		
Statin use (ratio)	0.198	0.181

## BMI for Colesterol drug users and who are not

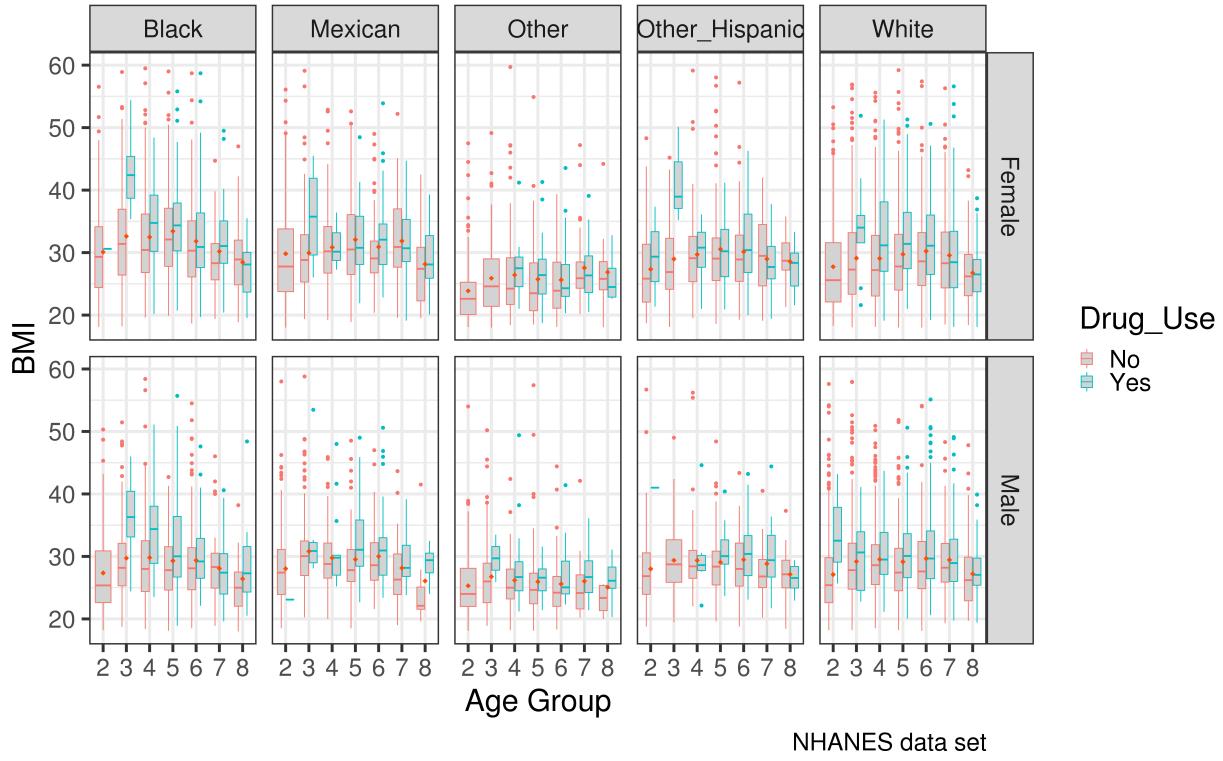


Figure 1: Illustration of data distribution using a box plot. The plot compares cholesterol drug users vs non-cholesterol drug users

Figure 1 shows the association between age and BMI with respect to cholesterol drug changes based on age. At the younger ages, the correlation is positive while at the middle age the correlation is almost flat, and at the old age the correlation is negative. Moreover, at low age cholesterol medications are described only for a population with high BMI, but at a higher age, the difference in term of the BMI for a population that takes cholesterol medication and without cholesterol medication decreases until it diminishes as in the white population. The heterogeneity can also be seen in gender ?, for the Hispanic female population, the cholesterol medication is taken by the population with lower BMI if compared to the population that does not take cholesterol medication.

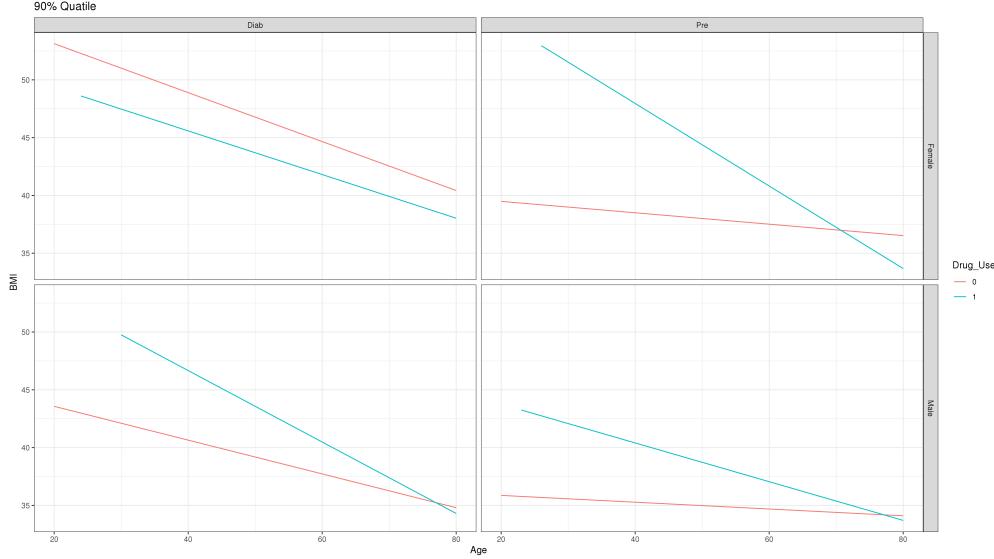


Figure 2: A 90th quantile of BMI plotted with respect to age. The population grouped with respect gender and fasting blood glucose level: prediabetes nad diabetes

## 6 Results

Figures 3 and 4 illustrate marginal effects of predictors on different quantiles, where in the first plot age and total cholesterol entered the model as quadratic terms while in the second plot, spline basis expansion is used to model age and total cholesterol. First, we look at the differences in the plot that is due to using OLS vs QR, then we look at the differences due to the using quadratic term to model predictors vs spline basis expansion.

From Figures 3, we can see the effects on conditional mean of BMI level may not reflect the size and nature of these effects on lower or upper quantiles. For example, the conditional mean effect of gender on BMI level is about  $-1 \text{ kg}/\text{m}^2$ , that is, on average a male BMI is less than female BMI by  $1 \text{ kg}/\text{m}^2$ . However, at the lower quantiles males have higher BMI by  $1 \text{ kg}/\text{m}^2$ . Then, the differences start to diminish up to 0 at 40th quantile. At higher quantiles, a male BMI is less than a female BMI. For example, at 60th quantile a male BMI is less than a female BMI by  $1 \text{ kg}/\text{m}^2$ , and it continue to increase up to  $3\text{kg}/\text{m}^2$  at 90th quantile.

From the OLS it is obvious cholesterol drug users have on average higher BMI levels if

compared to non-cholesterol drug users which are around 1.75. The disparity is almost regular through different BMI quantiles. For example, in the lower quantiles of the distribution the cholesterol drug users have higher BMI value by about  $1.5 \text{ kg/m}^2$ , but around at 50 th percentile of the conditional distribution the difference is  $1.8 \text{ kg/m}^2$ . Overall, cholesterol drug use seems to be associated with rather large effects on BMI levels somewhat between 1 to  $2 \text{ kg/m}^2$ .

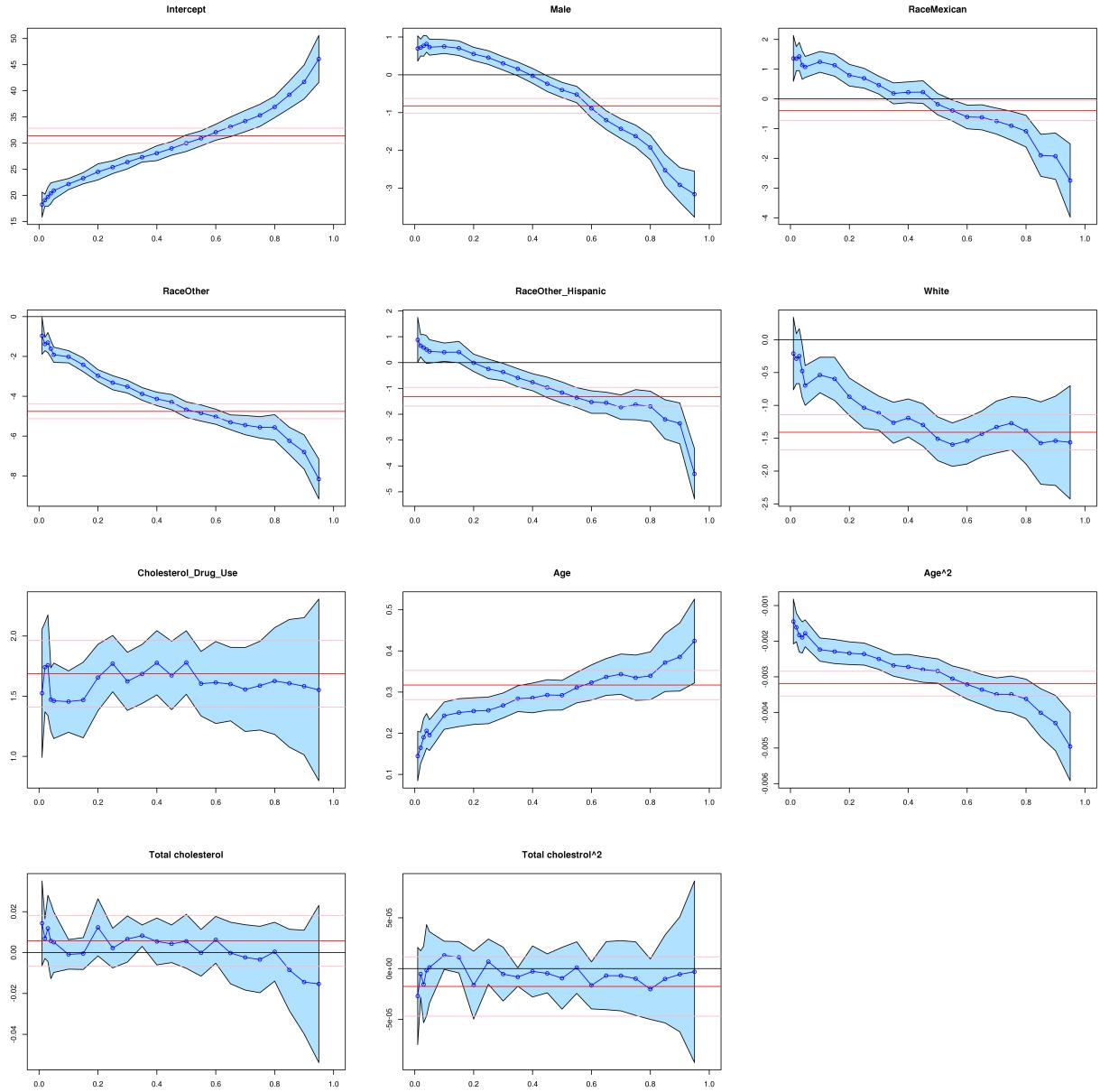


Figure 3: Quantile regression of BMI plotted for different marginal effects. The predictors modeled using quadratic terms.

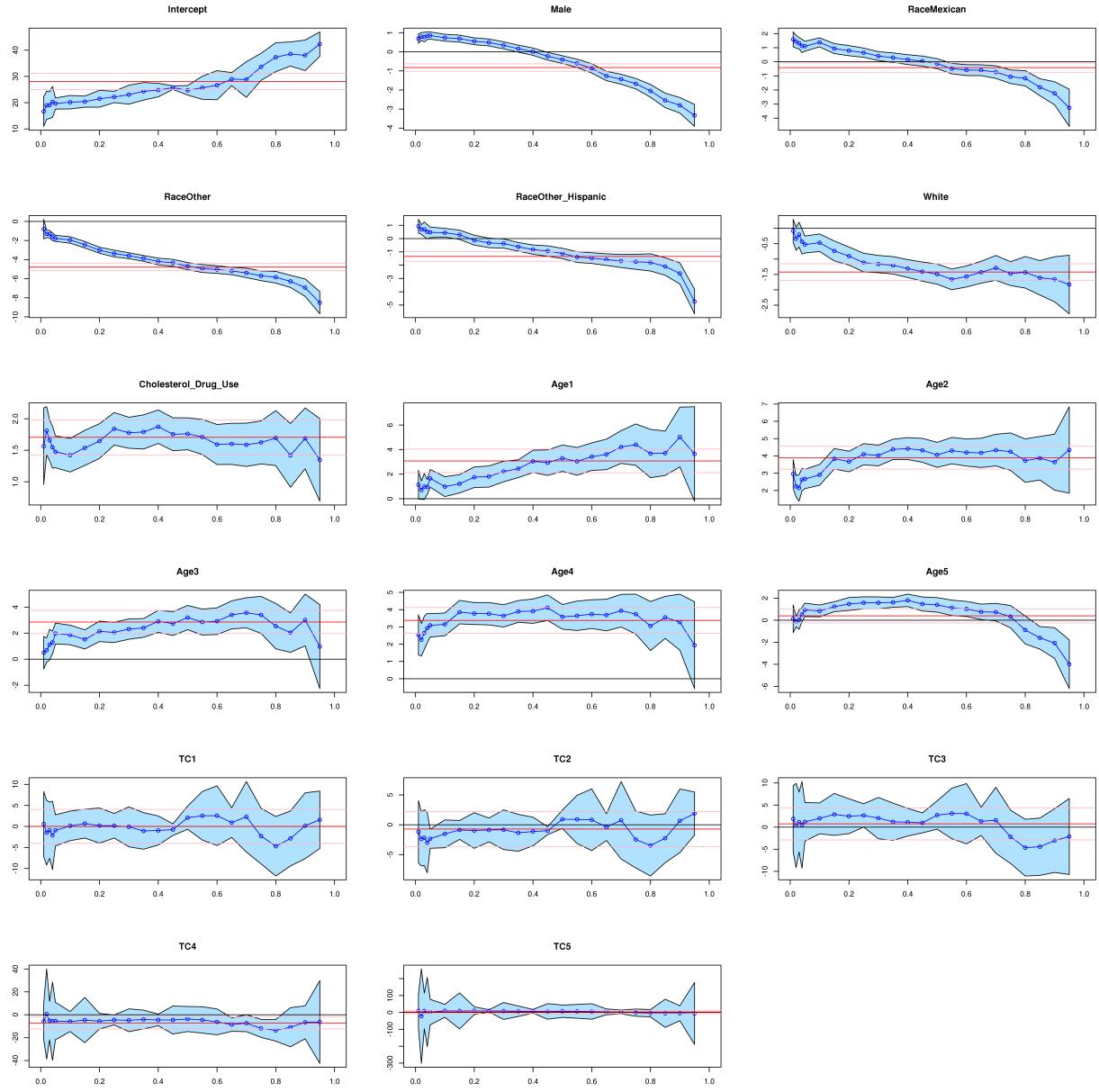


Figure 4: Quantile regression of BMI plotted for different marginal effects. The predictors modeled using splines.

## 6.1 Age effects Analysis

Figure 5 illustrate the differences in modeling Age in two different ways (Model I; spline basis expansion) and (Model III; quadratic term ).

Model I shows, at the lower quantile, the marginal age effects on BMI increases from age 20 to around age 42, and then it stays flat up to age 70 but it starts to decrease after that. While model III shows that there is positive relationship between Age and BMI in the age range between 20 and 50. Then age effects stay flat in the age range between 50 and 60 years old and decreases after that.

At the 50th quantile, model I shows the highest age effect on BMI occurs earlier by 10 years when compares to the lower quantile, but model III shows maximum effect occurs at around age 50, which is similar to the age effect at lower quantiles. The age effects start to decrease using Model I at around 70 years old while model III shows the decay starts at around age 60.

At higher quantiles  $\tau = 75$ , using Model I, the maximum age effects occurs earlier by three years, but Model III shows the highest age effects occurs at around age 50. The decay, according to Model I, starts at around age 70 years old, but Model III shows the decay starts earlier at around 70 years old. Convex behavior of the quadratic term has a clear pattern in formulating age effects on the BMI. The box plot (Fig. 1) supports Model I results because the highest mean and inter quantile range occurs at age 30.

Using Model I, at the 90th quantile of the conditional BMI distribution, we have a similar behavior as in 0.75th quantile except the relationship is stronger, i.e, for example, the age effects at age 20 is  $37 \text{ kg}/\text{m}^2$  while at the later the effect is  $30 \text{ kg}/\text{m}^2$ . The BMI trend in this modeling is close to the trend shown in (Chen 2005), which computed using a complicated polynomial and log transformation for the response. Our conclusion from Figure 5 is that the marginal age effects on BMI is different on different quantiles. Moreover, modeling age as quadratic term or as spline basis expansion produce different effect with respect to different quantiles of the conditional BMI distribution.

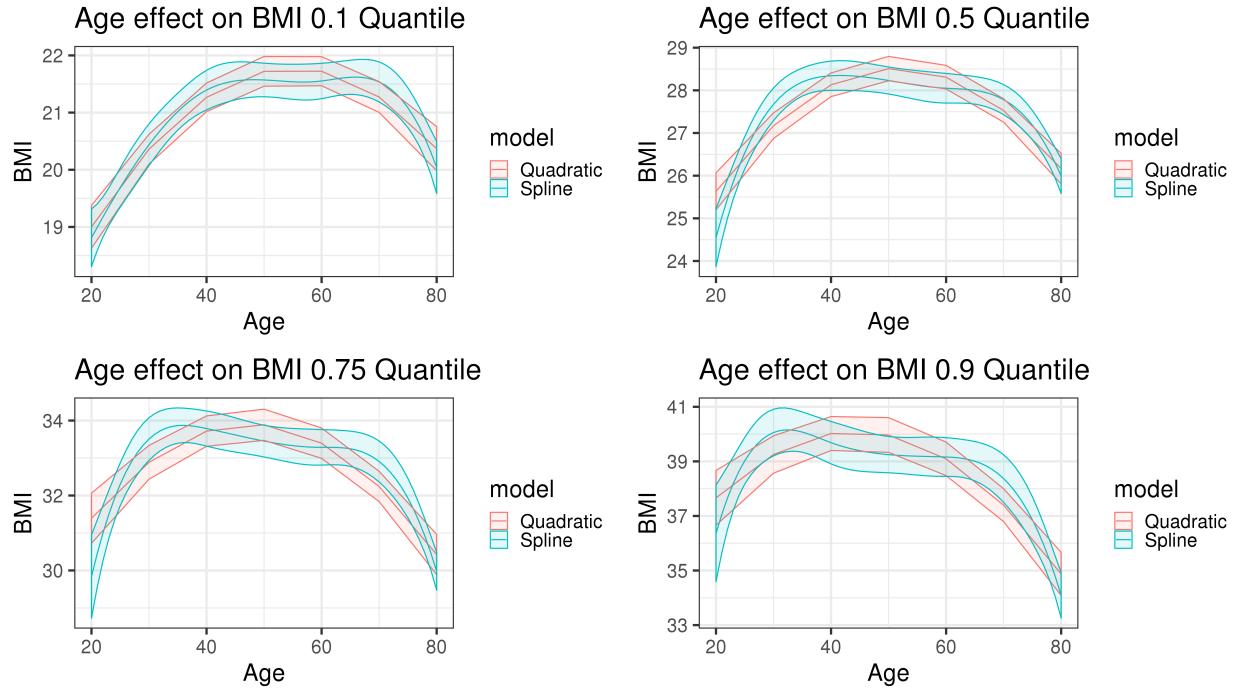


Figure 5: An illustration for the differences in modeling the age effects on different BMI quantiles. The predictors (Age) modeled using quadratic terms and basis splines expansion.

## 6.2 TC effects Analysis

Next, we investigate marginal effect of TC on four different quantiles of the conditional distribution of the BMI. The quadratic effect of total cholesterol on the conditional distribution of BMI is convex in general when age is modeled using quadratic term. However, age effect on the BMI is different when age is modeled using spline basis expansion. The disparity is due to the same reason mentioned earlier, which is as  $TC \rightarrow \pm\infty$ ,  $BMI \rightarrow 0$ ). At the lower tail ( $\tau = 0.1$ ), for TC less than 300 mg\dl, we have positive association, that is, when TC increases, the BMI increases also. Moreover, when TC passes this level, spline basis expansion modeling of TC presents negative correlation between TC and BMI while quadratic modeling of TC shows positive correlation.

At the 50th quantile of BMI, the two methods of modeling TC shows almost the same trend through whole TC range.

At higher quantiles ( $\tau = 75$ ), for TC values less than 350, the association is negative that is as TC cholesterol level increases the BMI decreases, for both ways of modeling TC. However, the two models differ after TC passes this value. The quadratic term force the TC effect to take convex shape. At a higher quantile ( $\tau = 90$ ), both ways of modeling TC shows BMI take a convex shape, but quadratic way of modeling BMI is more convex than the other. When TC passes 380 mg\|dL, where TC modeled using spline basis expansion, the BMI starts to decrease slow if compared to the quadratic way for modeling BMI, see Figure 6.

The population with low BMI, it is expected when TC increases (more than 200 mg\|dL), people start to reduce their BMI. However, this trend is captured by using spline basis expansion for the age but it was missed by using quadratic term for the age, see the upper left corner plot in Figure 6. In addition, he/she with high BMI reduces their weight to reduce him/her TC. The quadratic modeling of the TC shows a sharp decrease in the BMI if compared to spline basis expansion. This can be seen in the lower right corner of Figure 6. As it was shown in (Mc Auley 2020) the obese population has higher TC, this trend is not captured by quadratic term modeling for TC.

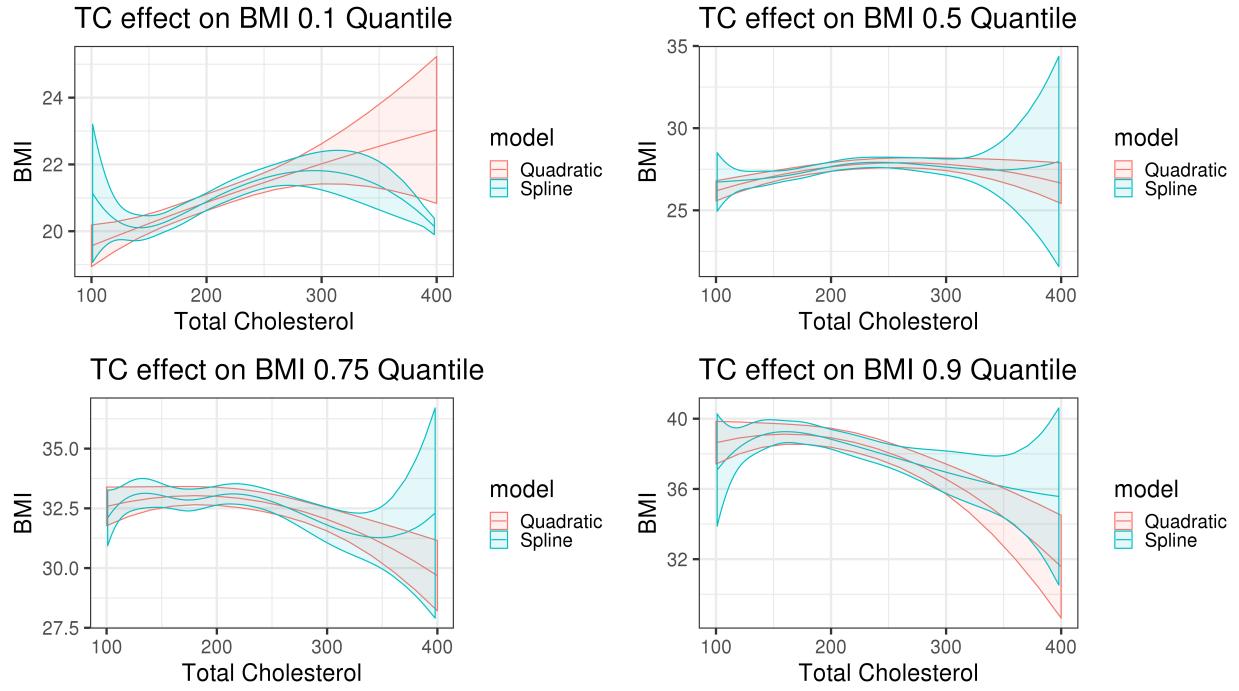


Figure 6: An illustration of marginal effects of total cholesterol on different BMI quantiles. The predictors modeled using quadratic terms and splines. The BMI quantiles  $au = 0.1, 0.5, 0.75, 0.90$  corresponds to  $21.5, 27.9, 32.4, 37.7 \text{ kg}/m^2$ , respectively.

### 6.3 Analysing Race\ Ethnicity effect on BMI

Figure 7 illustrates age effects on 90th BMI quantile, where age modeled as a quadratic term. There is a clear association between age and BMI. For example, at early ages for a non cholesterol drug users female Mexican, age effects is around  $40 \text{ kg}/m^2$ , but it decreases later up to 35 at age 80 years. The effects of age is higher for cholesterol drug users which is about  $2 \text{ kg}/m^2$ . Moreover, male Mexicans have higher age effect on BMI than females by about  $2 \text{ kg}/m^2$ . The lowest age effects on 90th BMI quantile is found in the white race. In general age effect on BMI for males are higher than for females. Figure 8 presents age effect on BMI through using spline basis expansion.

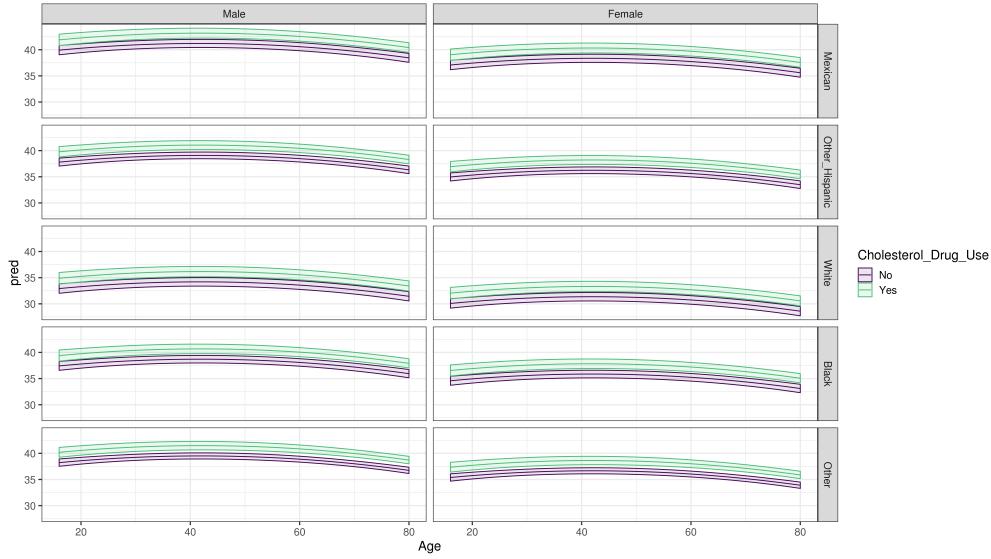


Figure 7: Illustration of the quadratic age effect on 90th BMI quantile for different race ethnicity groups.

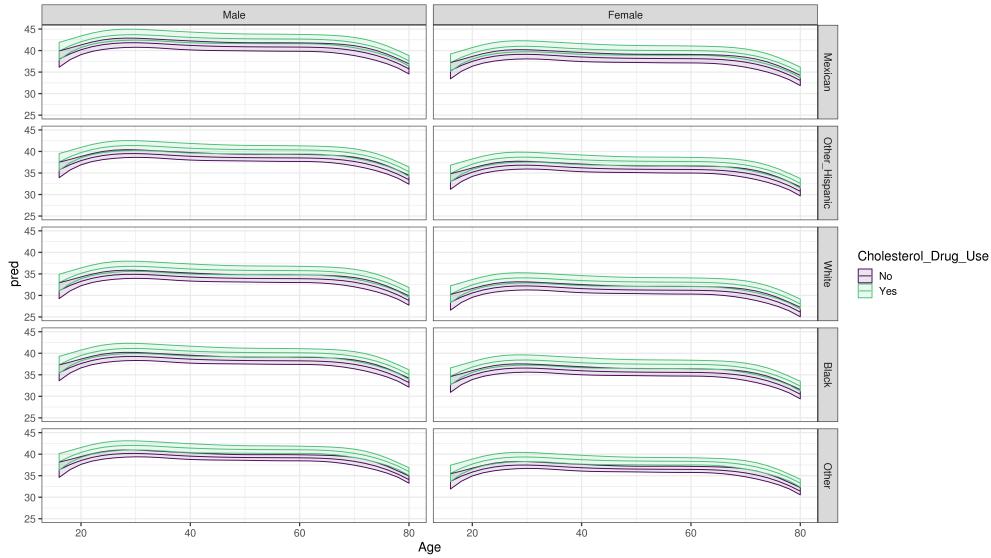


Figure 8: Illustration of the age effect on the 90th BMI quantile, where age modeled using spline basis expansion. The effect shown for different race/ethnicity groups and gender.

Figure 9 and Figure 10 present TC effect on BMI by modeling TC as a quadratic term and spline basis expansion, respectively. Male population have higher

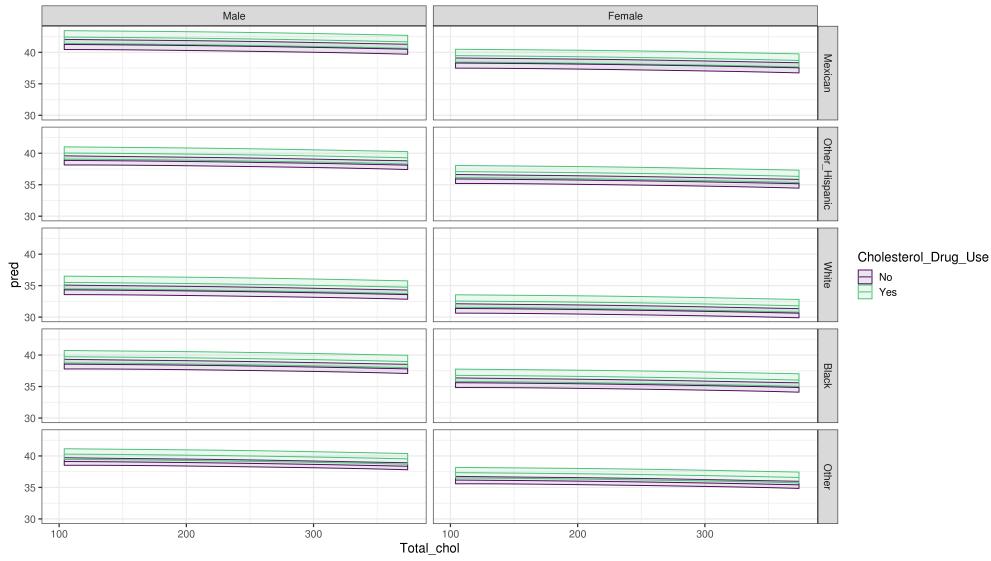


Figure 9: Illustration of a 90th quantile regression of the BMI. The cholesterol term modeled as quadratic.

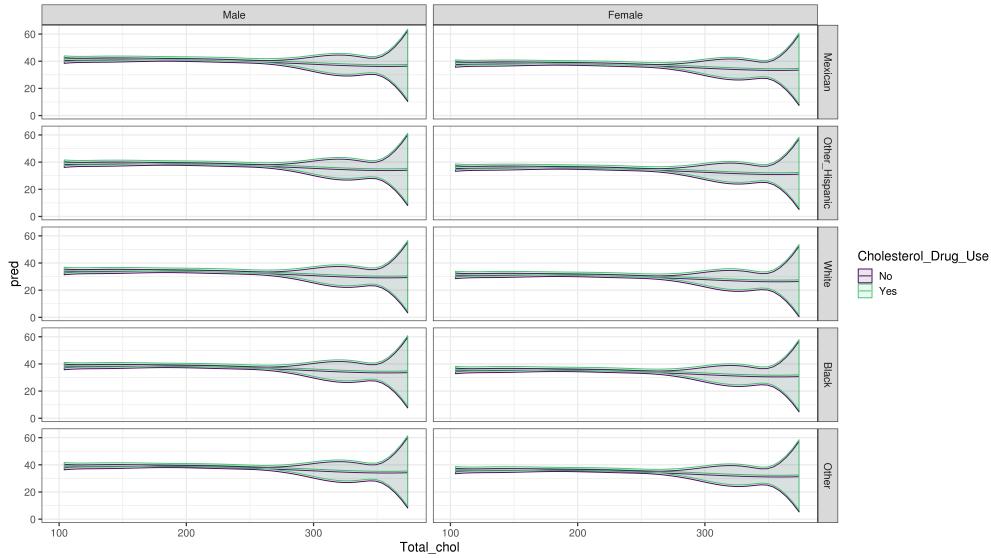


Figure 10: Illustration of a 90th quantile regression of the BMI. The cholesterol term modeled using splines.

## 7 Conclusion

Multivariate quantile regression is used to study the effects of different risk factors on different BMI levels. our study shows that different risk factors for higher BMI have different effects on different quantiles. Moreover, predictors effect on BMI is subject to the methods used to represent it, which are quadratic modeling spline basis expansion. Quadratic term for the response to take a convex shape. This is different from spline basis expansions in which different polynomials have been used to model the predictors. Our results shows the turning point in the association between age and BMI is 30 years old. This is the age when people start to consider reducing weight. This trend is shown through the descriptive statistics and QR. However, quadratic modeling for the age is not able to detect this trend.

The effects of TC on the BMI is studied with respect to two ways for modeling TC. The turning point for patients to start reducing their BMI is when TC hit around 200 mg\|dL. However, if TC hits 350, the association become positive. This is a sign where patients are losing their control on their TC and BMI levels.

It is recommended to investigate why the effect estimates are varying across different BMI quantiles.

## 8 References

- Anderson, Keaven M, Patricia M Odell, Peter WF Wilson, and William B Kannel. 1991. “Cardiovascular Disease Risk Profiles.” *American Heart Journal* 121 (1): 293–98.
- Balkau, Beverley, Gang Hu, Qing Qiao, Jaakko Tuomilehto, Knut Borch-Johnsen, K Pyorala, DECODE Study Group, European Diabetes Epidemiology Group, and others. 2004. “Prediction of the Risk of Cardiovascular Mortality Using a Score That Includes Glucose as a Risk Factor. The Decode Study.” *Diabetologia* 47 (12): 2118.
- Bann, David, Emla Fitzsimons, and William Johnson. 2020. “Determinants of the Population Health Distribution: An Illustration Examining Body Mass Index.” *International Journal of Epidemiology* 49 (3): 731–37.
- Bruce, Peter, Andrew Bruce, and Peter Gedeck. 2020. *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python*. O'Reilly Media.
- Cade, Brian S, and Barry R Noon. 2003. “A Gentle Introduction to Quantile Regression for Ecologists.” *Frontiers in Ecology and the Environment* 1 (8): 412–20.
- Calle, Eugenia E, Michael J Thun, Jennifer M Petrelli, Carmen Rodriguez, and Clark W Heath Jr. 1999. “Body-Mass Index and Mortality in a Prospective Cohort of Us Adults.” *New England Journal of Medicine* 341 (15): 1097–1105.
- Castro, M Regina, Gyorgy Simon, Stephen S Cha, Barbara P Yawn, L Joseph Melton, and Pedro J Caraballo. 2016. “Statin Use, Diabetes Incidence and Overall Mortality in Normoglycemic and Impaired Fasting Glucose Patients.” *Journal of General Internal Medicine* 31 (5): 502–8.
- Chang, Jin-Biou, Nain-Feng Chu, Jhu-Ting Syu, An-Tsz Hsieh, and Yi-Ren Hung. 2011. “Advanced Glycation End Products (Ages) in Relation to Atherosclerotic Lipid Profiles in Middle-Aged and Elderly Diabetic Patients.” *Lipids in Health and Disease* 10 (1): 228.
- Chen, Colin. 2005. “Growth Charts of Body Mass Index (Bmi) with Quantile Regression.” *AMCS* 5: 114–20.
- Disease Control, Centers for, and Prevention (CDC). 2018. “National Health and Nutri-

tion Examination Survey Data (Nhanes.”

Ferrières, Jean, Dominik Lautsch, Anselm K Gitt, Gaetano De Ferrari, Hermann Toplak, Moses Elisaf, Heinz Drexel, et al. 2018. “Body Mass Index Impacts the Choice of Lipid-Lowering Treatment with No Correlation to Blood Cholesterol—Findings from 52 916 Patients in the Dyslipidemia International Study (Dysis).” *Diabetes, Obesity and Metabolism* 20 (11): 2670–4.

Flegal, Katherine M. 1999. “The Obesity Epidemic in Children and Adults: Current Evidence and Research Issues.” *Medicine and Science in Sports and Exercise* 31 (11 Suppl): S509–14.

Hay, Simon I, Sudha P Jayaraman, Alejandra G Contreras Manzano, Anoushka Millear, Laura Kemmer, Brent Bell, Juan Jesus Carrero, et al. 2017. “GBD 2015 Risk Factors Collaborators. Global, Regional, and National Comparative Risk Assessment of 79 Behavioural, Environmental and Occupational, and Metabolic Risks or Clusters of Risks, 1990-2015: A Systematic Analysis for the Global Burden of Disease Study 2015 (Vol 388, Pg 1659, 2016).” *Lancet* 389 (10064): E1–E1.

Katzmarzyk, Peter T, Bruce A Reeder, Susan Elliott, Michel R Joffres, Punam Pahwa, Kim D Raine, Susan A Kirkland, and Gilles Paradis. 2012. “Body Mass Index and Risk of Cardiovascular Disease, Cancer and All-Cause Mortality.” *Canadian Journal of Public Health* 103 (2): 147–51.

Koenker, Roger. 2005. “Quantile Regression, Volume 38 of.” *Econometric Society Monographs*.

Mc Auley, Mark Tomás. 2020. “Effects of Obesity on Cholesterol Metabolism and Its Implications for Healthy Ageing.” *Nutrition Research Reviews* 33 (1): 121–33.

Mozaffarian, Dariush, Emelia J Benjamin, Alan S Go, Donna K Arnett, Michael J Blaha, Mary Cushman, Sarah De Ferranti, et al. 2015. “Executive Summary: Heart Disease and Stroke Statistics—2015 Update: A Report from the American Heart Association.” *Circulation* 131 (4): 434–41.

Nuttall, Frank Q. 2015. “Body Mass Index: Obesity, Bmi, and Health: A Critical Review.” *Nutrition Today* 50 (3): 117.

Pandya, Ankur, Stephen Sy, Sylvia Cho, Milton C Weinstein, and Thomas A Gaziano. 2015. “Cost-Effectiveness of 10-Year Risk Thresholds for Initiation of Statin Therapy for Primary Prevention of Cardiovascular Disease.” *Jama* 314 (2): 142–50.

Ridker, Paul M, Aruna Pradhan, Jean G MacFadyen, Peter Libby, and Robert J Glynn. 2012. “Cardiovascular Benefits and Diabetes Risks of Statin Therapy in Primary Prevention: An Analysis from the Jupiter Trial.” *The Lancet* 380 (9841): 565–71.

Sugiyama, Takehiro, Yusuke Tsugawa, Chi-Hong Tseng, Yasuki Kobayashi, and Martin F Shapiro. 2014. “Different Time Trends of Caloric and Fat Intake Between Statin Users and Nonusers Among Us Adults: Gluttony in the Time of Statins?” *JAMA Internal Medicine* 174 (7): 1038–45.

Tsaousis, Konstantinos T. 2014. “Blood Glucose and Cholesterol Concentrations in a Mediterranean Rural Population of Andros Island, Greece.” *International Journal of Preventive Medicine* 5 (11): 1464.

Van de Kassteele, Jan, RT Hoogenveen, PM Engelfriet, PHM Van Baal, and HC Boshuizen. 2012. “Estimating Net Transition Probabilities from Cross-Sectional Data with Application to Risk Factors in Chronic Disease Modeling.” *Statistics in Medicine* 31 (6): 533–43.

Yusuf, Salim, Jackie Bosch, Gilles Dagenais, Jun Zhu, Denis Xavier, Lisheng Liu, Prem Pais, et al. 2016. “Cholesterol Lowering in Intermediate-Risk Persons Without Cardiovascular Disease.” *New England Journal of Medicine* 374 (21): 2021–31.

Yusuf, Salim, Steven Hawken, Stephanie Ôunpuu, Tony Dans, Alvaro Avezum, Fernando Lanas, Matthew McQueen, et al. 2004. “Effect of Potentially Modifiable Risk Factors Associated with Myocardial Infarction in 52 Countries (the Interheart Study): Case-Control Study.” *The Lancet* 364 (9438): 937–52.