Overweight Adults May Have the Lowest Mortality -- Do They Have the Best Health?

Anna Zajacova,[1,2] Jennifer Beam Dowd,[3] Sarah A. Burgard[4]

February 28, 2010

Running head: BMI and health indicators

Key words: Body mass index, overweight, biomarkers, health, generalized additive models, nonparametric models

[1] Department of Sociology, University of Wyoming
[2] Direct correspondence to Anna Zajacova, Department of Sociology, Dept. 3293, 1000 E University Ave., University of Wyoming, Laramie WY 82071; email: zajacova@uwyo.edu
[3] Department of Epidemiology and Biostatistics, Hunter College, School of Public Health, City University of New York (CUNY)
[4] Department of Sociology and Population Studies Center, Institute for Social Research, University of Michigan

Abstract

Several recent studies have reported that overweight adults experience lower overall mortality than those who are underweight, normal-weight, or obese. These widely publicized findings carry critical implications for public health and policy because they suggest that overweight may be the optimal weight category for adult health and longevity. In this study, we test this assumption using nationally representative NHANES surveys (2005-2008) with adults age 20-80. We employ generalized additive models, a type of semiparametric model, to examine the relationship between body mass and key biological risk measures, including inflammatory markers and indicators of respiratory, cardiovascular, and metabolic function. The key finding is that the association between BMI and biological risk factors is generally monotonically increasing rather than U-shaped as mortality analyses suggest. We document the modifying effects of age and sex on the BMI-risk association and attempt to reconcile our findings with those from the mortality literature.

This paper examines the shape of the association between BMI and health. What body size is associated with the lowest biological risk? How does the optimal body weight vary with age, by sex, and across different health indicators? Employing a non-parametric regression to analyze a nationally-representative sample of US adults, we find that the profile of health risks tends to worsen monotonically from the low-normal to very high body weights.

The motivation for this study comes from the literature on the relationship between BMI and mortality. These studies almost uniformly report a U-shaped association between body weight and mortality, such that obesity (BMI above 30 kg/m^2) and underweight (BMI below 18.5 kg/m^2) are associated with a significantly increased risk of dying, relative to normal weight (BMI between 18.5 and 25 kg/m^2). The controversial result in this literature is that overweight adults (BMI between 25 and 30 kg/m^2) experience mortality no higher, and often lower, than their normal-weight counterparts [1-9]. This finding has attracted a lot of attention from the scientific community and the general public because it contradicts the predominant message from clinical and public health research that body weight above the normal range is detrimental to health. Moreover, the *mortality* results have been used to draw conclusions about *health* outcomes, implying that having body weight in the overweight range does not impair health, especially among older people. Currently, over 34% of U.S. adults are overweight range and an additional 34% are obese [10-11]. Therefore, understanding to what extent these categories of body weight are associated with increased health risks, relative to normal weight, is critical to public-health policies and interventions.

In response to the reports about the low mortality of overweight adults, some researchers have argued that this pattern is an artifact of confounding and reverse causation [12-13]. The two most frequent adjustments to the survival models included controlling for smoking status or eliminating smokers from the sample and discarding observations in which death occurred within the first several years after the measurement of body weight. These adjustment often succeed in changing the results to a point where the mortality for overweight becomes significantly higher than that of normal-weight adults [9, 12-16]. However, these new results come at an expense of excluding a large portion of the study samples, a questionable step in any analysis aiming to obtain unbiased and generalizable findings.

Apart from improving the mortality models, a more direct way to understand the impact of excess body weight on health is to examine measures of health directly. Compared to the literature on mortality, there are relatively few studies focused on the BMI-health relationship. The health consequences of *obesity* are known to include a range of negative outcomes, from potentially life-

threatening conditions such as diabetes or cardiovascular disease [17-19], nonfatal chronic illness such as osteoarthritis [20-21], to physical limitations and poor general health [8, 22-27]. The findings for overweight, relative to normal-weight adults, are mixed. Some studies report a higher prevalence of chronic conditions among the overweight [28-29], while others find overweight on par with normal-weight adults [4, 25, 30].

A key limitation of these studies is their analytic approach. The majority of researchers, both in the mortality and health-outcomes line of research, categorize BMI along the standard cutpoints and compare outcomes among the resulting groups. Alternatively, researchers impose a quadratic shape on the BMI predictor to take into account the U-shape association between body mass and outcomes. Both of these approaches force an apriori shape on the relationship that may not be optimal for the data, and may bias the findings. For the categorical specification of BMI, the models assume homogeneity within each category, although the range of BMIs within each group is fairly wide. This might obscure meaningful variation within categories, a problem particularly acute for the reference 'normal weight' category. For the continuous specification, the quadratic shape may similarly hide the true shape of the BMI-health associations. Only a few studies have attempted to overcome this problem by employing nonparametric methods to examine mortality among US adults [5, 16]. These two studies, however, could not adequately model the mortality process -- survival was a binary outcome variable and time to death was not modeled. We are not aware of any published studies on that used non-parametric modeling to determine the shape of the association between BMI and health outcomes.

In addition to inflexible modeling of the BMI-health association, the studies may have failed to take age into account adequately. All studies controlled for age, but did not evaluate the possibility that age moderates the association between weight and health outcomes. This is a potentially important omission. Body weight changes systematically across age [31-32], and its effect on health may change as well [14, 33].

The present study offers several contributions to the BMI-health literature. We show how BMI and different health indicators are associated among adults in the US. Using generalized additive models overcomes the need for any apriori constraints on the shape of the association, letting the data drive the findings. With the exception of self-rated health, all health-related variables are measured rather than self-reported. BMI is measured during a clinical examination and health is captured by multiple risk indicators, assessed using laboratory tests and through examination by a trained health technician. Second, we use a range of biomarkers to assess health risks, from

indicators of cardiovascular fitness to systemic inflammation. This allows us to model health status differences in all age groups, including among younger people, for whom clinical symptoms of disease are relatively rare. Third, we use nationally-representative data and the full range of adult ages, from 20 to 80, and we explicitly examine how age modifies the association between BMI and health outcomes. This approach gives us a lifecourse view of the impact of excess body weight on health.

DATA AND METHODS

Data

The analyses are based on the two most recent waves of data from the National Health and Nutrition Examination Surveys, NHANES 2005-2006 and NHANES 2007-2008 [34]. The NHANES is a series of studies initiated in 1960s, designed to measure the health status of US adults and children. The study collect demographic, socioeconomic, lifestyle, and health-related information. Since 1999, the NHANES design changed to a cross-sectional continuous data collection conducted in two-year cycles, with over 5,000 individuals interviewed every year. We combined two 2-year cycles of data in order to obtain sufficient statistical power for all analyses, as recommended by the NCHS [35]. NHANES uses a stratified, multistage probability sampling design with an oversample of African Americans, Hispanics, low-income persons, and older adults. The sample is representative of non-institutionalized civilian US population. The data collection procedure combines a household interview with a medical examination at a mobile examination unit by a team of 16 people, including a physician, trained interviewers, and health technicians. During the examination, blood and urine are collected for laboratory analysis. The response rate for the interview was approximately 80% in 2005-2006 and 78% in 2007-2008; about 96% of interviewed individuals also participated in the subsequent medical examination.

Analytic sample

We defined the analytic sample as adults age 20 to 80 who had a valid sociodemographic measures, BMI ranging from 15 to 45, at least one biological risk marker, and were not pregnant. Out of the 10,566 adults in the desired age range, this definition required the exclusion of 12 adults with missing education and 393 women who were pregnant; additionally, 612 adults did not participate in the medical examination and hence were missing health outcomes and/or BMI information. Finally, restricting the BMI range to 15-45 excluded 247 adults with extremely low or high body mass. The final sample size is 9,302, or 88% of those who underwent the household interview.

Measures

Body mass index (BMI) was calculated from height and weight, using the formula BMI=703*(weight in pounds/squared height in inches). Height and weight were measured by trained technicians during the medical examination. We restricted the BMI range to 15-45 because of low-density problems at the extremes of the body mass distribution. The restriction excluded only 2.5% of the sample.

The outcomes include markers of cardiovascular function (systolic and diastolic BP, HDL and total cholesterol), glycosylated hemoglobin as a marker of metabolic risk, and inflammatory biomarkers C-reactive protein and fibrinogen. All biomarkers are modeled as continuous outcomes, with appropriate transformation or appropriate model link function for non-normally distributed variables. Self-reported health outcomes include self-rated health, measures on a 5-point scale from excellent to poor.

Control variables include demographic, socioeconomic, and lifestyle measures. Demographic controls comprise age, centered on the sample mean; race/ethnicity coded as non-Hispanic white (reference), non-Hispanic black, Hispanic, and other race/ethnicity; region of residence (Northeast as reference); and type of place of residence (urban as reference). SES is measured by educational attainment and poverty-income ratio. Controlling for SES matters because this factor is strongly linked to both BMI and health outcomes [36]. Finally, control for smoking is important because of the strong inverse relationship between smoking and body weight. Smoking is trichotomized as current smoker, past smoker, and never smoker.

Analysis

We employ generalized additive models (GAMs) to examine the association between BMI and each risk marker [37]. The GAM is an extension of the generalized linear model in that one or more predictors may be specified using a smooth function $f(x)$. The general structure of GAM is $\eta_i = \beta_0 + f_1(x_{1i}) + \cdots + f_k(x_{ki}) + \varepsilon_i$. where a smooth monotonic link function $[g(y)]_i$ transforms the expected value of Y $to\ \eta_i$. Y can follow any distribution from the exponential family. The key strength of this approach over GLMs is the flexibility from the data-driven shape of the $f(x)$ functions, avoiding the need for a priori assumptions about the shape of a predictor's effect, for instance, a linear or quadratic functional form for the effect of BMI on some outcome [38-39]. The functions $f(x)$ can be specified by various spline smoothing functions but typically thin plate regression splines are used [40]. This smoother is efficient to compute, the knots do not need to be placed by the researcher, and they can be constructed for two or more predictors jointly. The

degree of smoothness for the functions $f(x)$ is also estimated from the data, with the aim of optimal balance between the fit to the data against a penalty for excessive 'wiggliness' of the functions. GAMs can accommodate the 'interaction' of two or more predictors, in a way that is conceptually comparable to interactions in GLMs. The joint smooth function can be specified different ways; we will use tensor product smooths. The GAM is estimated using a penalized ML procedure; typically iteratively re-weighted least squares [41], although alternative fitting approaches can be used [41-42]. After the basis for the function $f(x)$ is chosen, the GAM becomes essentially a GLM, which makes it possible to conduct standard model building and checking procedures. Model fit is estimated using generalized cross-validation (GCV) based on the prediction mean square error, un-biased risk estimator (UBRE)m and Akaike's information criterion (AIC).

Confidence intervals for parameter estimates are calculated using the posterior distribution of the model coefficients. The p-values associated with model coefficients are estimated from the covariance matrix of $\hat{\beta}$'s, but are known to be marginally smaller than expected under then null hypothesis. Models can be compared using an approximation to the LR test for nested models or ANOVA. In addition to model summaries, results are presented graphically in line graphs showing the smooth function for a predictor as a line with 95% Bayesian credible (confidence) intervals, or as a surface describing the interaction of two predictors on some outcome [41]. The software package R 2.9.2 is used to estimate the models.

PRELIMINARY RESULTS

Figures 1 to 4 below summarize a subset of results from preliminary models of log-transformed C-reactive protein. These figures illustrate how final results will be presented and offer some preliminary answers to the key research question: is the association between BMI and biological risk factors U-shaped and thus comparable to results from the literature on the BMI-mortality association? Figure 1 shows the sex-specific association of BMI and log-CRP for the full sample -- depicted is the smoother estimated by an age-adjusted GAM model (95% CI). The results show a monotonically increasing level of CRP with BMI, whereby lower BMI is associated with a healthier profile of the CRP levels.

We use two approaches to examine whether age modifies the association between BMI and biomarkers. The first approach stratifies the sample into three age groups: 18-39, 40-64, and 65-85. The second approach uses the full age spectrum but incorporates a tensor-product interaction between age and BMI. Figure 2 depicts findings from age-stratified models. At the younger and older ages, the association is relatively linear. In mid-adulthood, there is a suggestion toward a flat

slope of BMI through most of the normal BMI range. Finally, figure 3 presents the age by BMI interactive effects on C-reactive protein in a continuous form using a 3-dimensional plot. This figure illustrates the steeper gradient in CRP by BMI at younger ages, as well as the monotonic relationship across the full age range. The age-focused analyses are intended to inform the question of whether the shape of the BMI-biomarker associations may become U shaped at older ages, in way that would corroborate the pattern observed in mortality studies. However, the preliminary set of results suggests that it will not be the case -- at least not uniformly for all biomarkers.

Additionally, we conducted a series of models stratified by race. Figure 4 shows the association between BMI and (ln) CRP for men -- there is a substantial variation across the four race/ethnic groups in the overall slope and the shape of the relationship. The *existence* of race differentials in the association between BMI and biological risk profiles is in accordance with literature [43-45]. The key point is that no race/ethnic group evidenced a pronounced U- or J-shaped association between the examined biomarkers and BMI.

These preliminary findings offer evidence that the U-shaped pattern described for the relationship between BMI and mortality may not generalize to other health outcomes. Our analyses show that overweight adults generally have worse profiles of biological risk than normal-weight adults -- there appears to be a dose-response (monotonic) relationship for most biomarkers from very low levels of BMI. This pattern holds for a range of biological risk indicators, for both sexes, all major race/ethnic groups, and across the full adult age spectrum. These findings suggest that public health messages should emphasize efforts to lower BMI for overweight adults.

Figure 1. Association between BMI and CRP

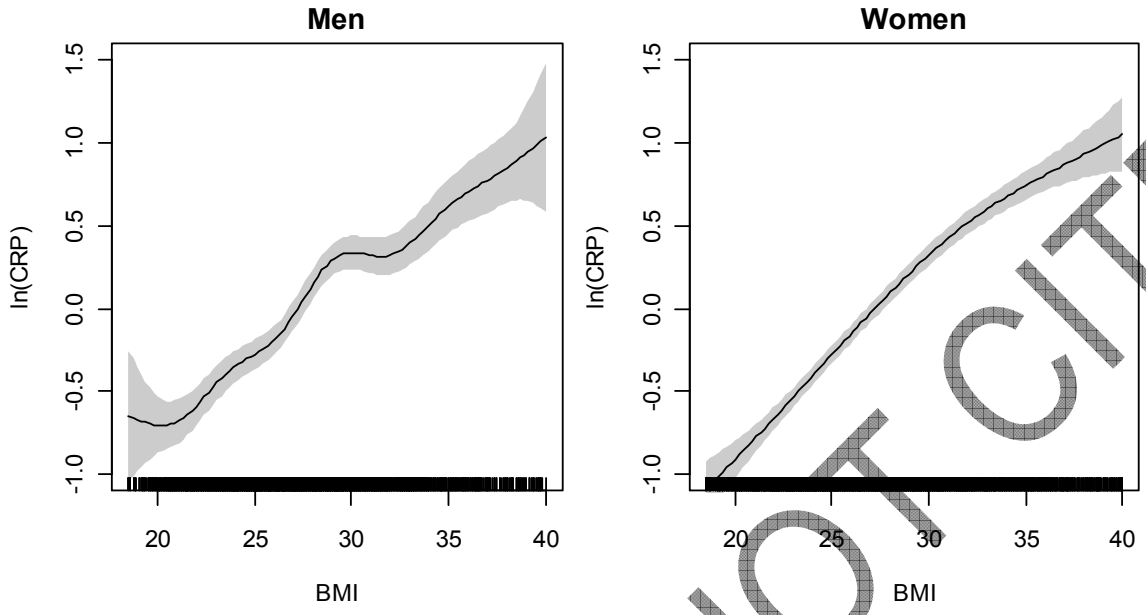


Figure 2. Association between BMI and CRP by age, for men

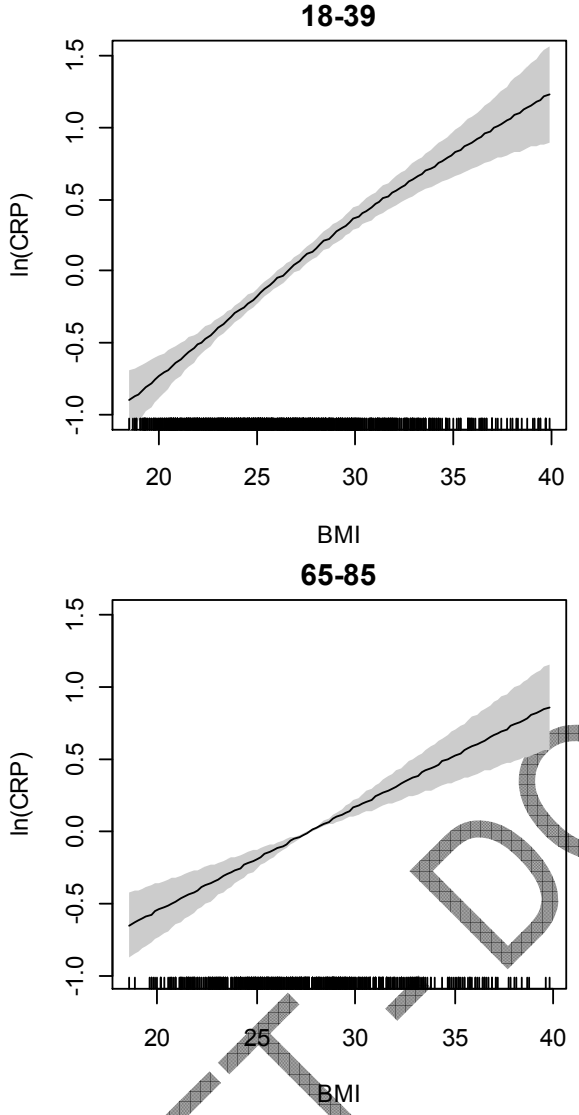


Figure 3. Association between CRP and BMI across age, for men

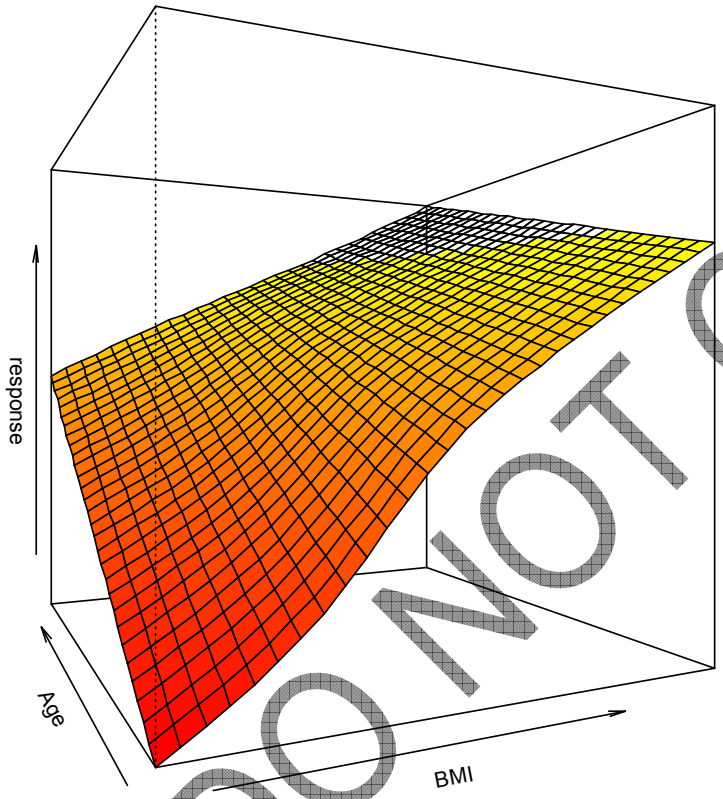
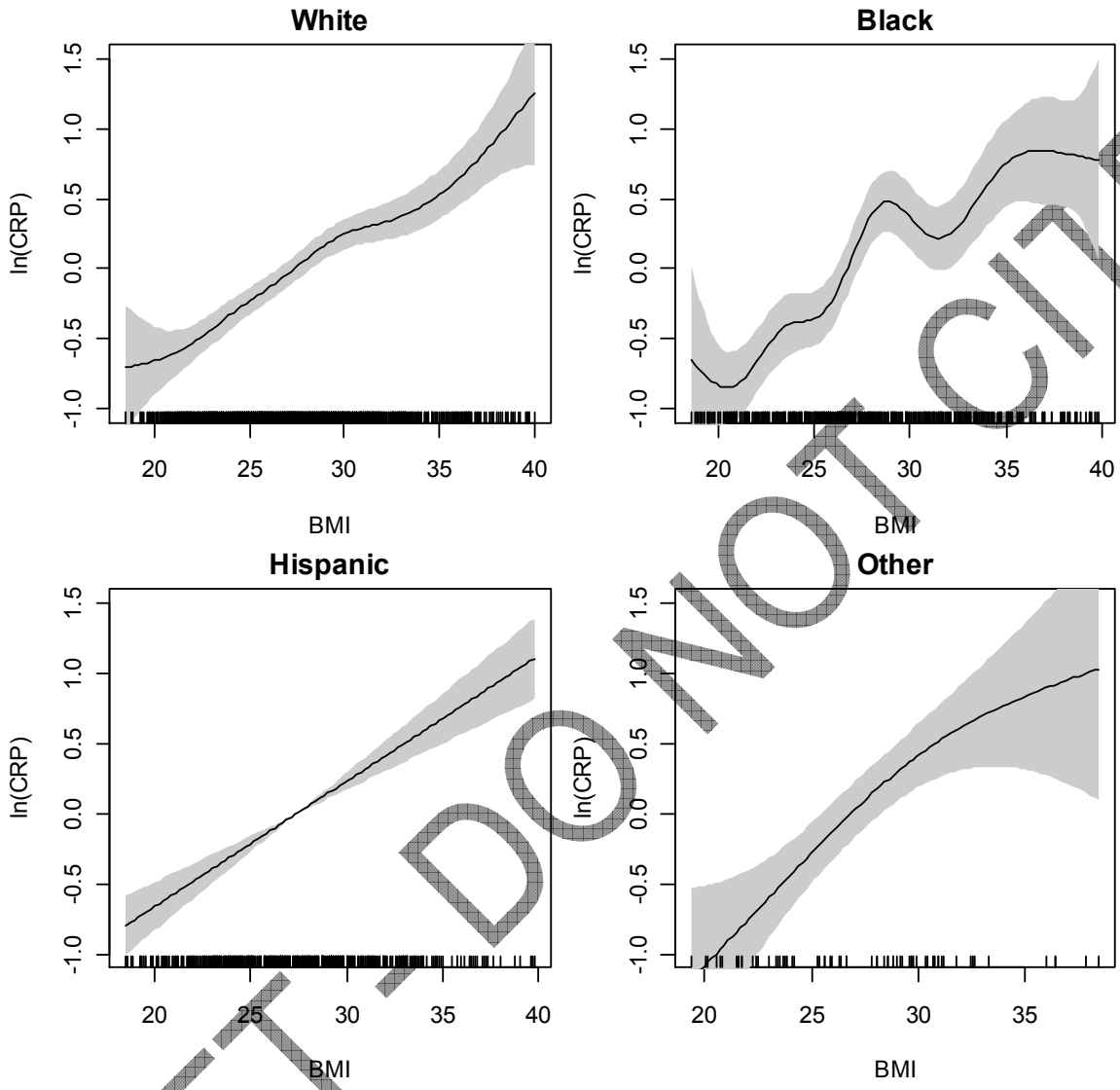


Figure 4. Association between BMI and CRP by race, for men



References

1. Flicker, L., et al., *Body Mass Index and Survival in Men and Women Aged 70 to 75.* Journal of the American Geriatrics Society, 2010. **58**(2): p. 234-241.
2. Flegal, K.M., et al., *Excess Deaths Associated with Underweight, Overweight, and Obesity.* JAMA, 2005. **293**(15): p. 1861-7.
3. Flegal, K.M., et al., *Cause-Specific Excess Deaths Associated With Underweight, Overweight, and Obesity.* JAMA, 2007. **298**(17): p. 2028-2037.
4. Janssen, I., *Morbidity and Mortality Risk Associated With an Overweight BMI in Older Men and Women.* Obesity, 2007. **15**(7): p. 1827-40.
5. Mehta, N.K. and V.W. Chang, *Mortality Attributable to Obesity Among Middle-Aged Adults in the United States.* Demography, 2009. **46**(4): p. 851-872.
6. Berraho, M., et al., *Body Mass Index, Disability, and 13-Year Mortality in Older French Adults.* Journal of Aging and Health, 2010. **22**(1): p. 68-83.
7. McGee, D.L., *Body Mass Index and Mortality: A Meta-Analysis Based on Person-Level Data from Twenty-Six Observational Studies.* Annals of Epidemiology, 2005. **15**(2): p. 87-97.
8. Troiano, R.P., et al., *The Relationship Between Body Weight and Mortality: a Quantitative Analysis of Combined Information from Existing Studies.* International Journal of Obesity, 1996. **20**(1): p. 63-75.
9. Adams, K.F., et al., *Overweight, Obesity, and Mortality in a Large Prospective Cohort of Persons 50 to 71 Years Old.* New England Journal of Medicine, 2006. **355**(8): p. 763-778.
10. Hedley, A.A., et al., *Prevalence of Overweight and Obesity Among US Children, Adolescents, and Adults, 1999-2002.* JAMA, 2004. **291**(23): p. 2847-50.
11. Flegal, K.M.P., et al., *Prevalence and Trends in Obesity Among US Adults, 1999-2008.* . JAMA, 2010. **303**(3): p. 235-241.
12. Greenberg, J.A., K.R. Fontaine, and D.B. Allison, *Putative Biases in Estimating Mortality Attributable to Obesity in the US Population.* International Journal of Obesity, 2007. **31**: p. 1449-1455.
13. Greenberg, J.A., *Correcting Biases in Estimates of Mortality Attributable to Obesity.* Obesity, 2006. **14**(11): p. 2071-2079.
14. Freedman, D.S., et al., *Body Mass Index and All-cause Mortality in a Nationwide US Cohort.* International Journal of Obesity, 2006. **30**: p. 822-829.
15. Zamboni, M.M., G.; Zoico, E.; Harris, T.B.; et al., *Health Consequences of Obesity in the Elderly: A Review of Four Unresolved Questions.* International Journal of Obesity, 2005. **29**(9): p. 1011-29.
16. Gronniger, J.T., *A Semiparametric Analysis of the Relationship of Body Mass Index to Mortality* American Journal of Public Health, 2006. **96**(1): p. 173-178.
17. Wannamethee, S.G., A.G. Shaper, and M. Walker, *Overweight and Obesity and Weight Change in Middle Aged Men: Impact on Cardiovascular Disease and Diabetes.* Journal of Epidemiology and Community Health, 2005. **59**: p. 134-139.
18. Gregg, E.W., et al., *Secular Trends in Cardiovascular Disease Risk Factors According to Body Mass Index in US Adults.* JAMA, 2005. **293**(15): p. 1868-1874.
19. Mokdad, A.H., et al., *Prevalence of Obesity, Diabetes, and Obesity-Related Health Risk Factors, 2001.* JAMA, 2003. **289**(1): p. 76-9.
20. Himes, C.L., *Obesity, Disease, and Functional Limitation in Later Life.* Demography, 2000. **37**(1): p. 73-82.
21. Felson, D.T., et al., *Obesity and Knee Osteo-Arthritis -- The Framingham Study.* Annals of Internal Medicine, 1988. **109**(1): p. 18-24.
22. Bender, R., et al., *Assessment of Excess Mortality in Obesity.* American Journal of Epidemiology, 1998. **147**(1): p. 42-8.

23. Durazo-Arvizu, R., et al., *Establishing the Nadir of the Body Mass Index-Mortality Relationship: a Case Study.* Journal of American Statistical Association, 1997. **92**(440): p. 312-9.
24. Ferraro, K.F. and T.L. Booth, *Age, Body Mass Index, and Functional Illness.* Journal of Gerontology: Social Sciences, 1999. **54**(6): p. S339-48.
25. Ferraro, K.F., et al., *Body Mass Index and Disability in Adulthood: a 20-Year Panel Study.* American Journal of Public Health, 2002. **92**(5): p. 834-40.
26. Alley, D.E. and V.W. Chang, *The Changing Relationship of Obesity and Disability, 1988-2004.* JAMA, 2007. **298**(17): p. 2020-2027.
27. Vincent, H.K., K.R. Vincent, and K.M. Lamb, *Obesity and Mobility Disability in the Older Adult.* Obesity Reviews, 2010. **Published online 01/06/2010**(9999).
28. Must, A., et al., *The Disease Burden Associated with Overweight and Obesity.* JAMA, 1999. **282**(16): p. 1523-9.
29. Katz, D.A., C.A. McHorney, and R.L. Atkinson, *Impact of Obesity on Health-Related Quality of Life in Patients with Chronic Illness.* Journal of General Internal Medicine, 2000. **15**(11): p. 789-96.
30. Diehr, P., et al., *Weight, Mortality, Years of Healthy Life, and Active Life Expectancy in Older Adults.* Journal of the American Geriatrics Society, 2008. **56**(1): p. 76-83.
31. Munoz, A.M., et al., *Basic Anthropometry and Health Status of Elderly: Findings of the Maracaibo Aging Study.* Journal of Aging and Health, 2010. **22**(2): p. 242-261.
32. Heo, M., et al., *Hierarchical Linear Models for the Development of Growth Curves: An Example with Body Mass Index in Overweight/Obese Adults.* Statistics in Medicine, 2003. **22**(11): p. 1911-42.
33. Kulminski, A.M., et al., *Body Mass Index and Nine-Year Mortality in Disabled and Nondisabled Older U.S. Individuals.* Journal of the American Geriatrics Society, 2008. **56**(1): p. 105-110.
34. CDC, *National Health and Nutrition Examination Survey Data 2007-2008*. 2010, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics: Hyattsville, MD.
35. CDC, *Analytic and Reporting Guidelines: The National Health and Nutrition Examination Survey (NHANES)*. 2006, National Center for Health Statistics, CDC: Hyattsville, MD.
36. Brunner, E.J., et al., *Social Inequality in Coronary Risk: Central Obesity and the Metabolic Syndrome. Evidence from the Whitehall II Study.* Diabetologia, 1997. **40**(11): p. 1341-1349.
37. Hastie, T. and R. Tibshirani, *Generalized Additive Models (with Discussion).* Statistical Science, 1986. **1**(3): p. 297-318.
38. Stone, C.J., *Additive Regression and Other Nonparametric Methods.* Annals of Statistics, 1985. **12**: p. 689-705.
39. Hastie, T. and R. Tibshirani, *Generalized Additive Models*. 1990, London: Chapman and Hall.
40. Wood, S.N., *Thin Plate Regression Splines.* Journal of the Royal Statistical Society. Series B (Statistical Methodology), 2003. **65**(1): p. 95-114.
41. Wood, S.N., *Generalized Additive Models: An Introduction with R*. 2006, Boca Raton, FL: Chapman & Hall.
42. Andersen, R., *Nonparametric Methods for Modeling Nonlinearity in Regression Analysis.* Annual Review of Sociology, 2009. **35**(1): p. 67-85.
43. Carroll, J.F., et al., *Impact of Race/Ethnicity on the Relationship Between Visceral Fat and Inflammatory Biomarkers.* Obesity, 2009.
44. Wee, C.C., et al., *Obesity and C-reactive Protein Levels Among White, Black, and Hispanic US Adults.* Epidemiology, 2008. **16**(4): p. 875-880.
45. Nelson, T.L., et al., *Obesity and Associated Coronary Heart Disease Risk Factors in a Population of Low-Income African-American and White Women: the North Carolina WISEWOMAN project.* Preventive Medicine, 2002. **35**(1): p. 1'6.