Shape of the BMI-Mortality Association by Cause of Death, Using Generalized Additive

Models: NHIS 1986-2002.

Anna Zajacova

University of Michigan

April 7, 2008

Contact: Anna Zajacova, Population Studies Center, ISR, University of Michigan, 426 Thompson St. #2048, Ann Arbor MI 48104; email: zajacova@umich.edu.

Abstract.

The relationship between body mass index (BMI) and mortality is curvilinear; with very high and low BMIs are associated with increased odds of dying. The precise shape of the association for the majority of adults in the center of the BMI distribution, however, has not been established. This paper employs generalized additive models (GAMs) to describe the relationship between BMI and cause-specific mortality. GAMs allow modeling the shape of the relationship without imposing a priori constraints on the functional form of the predictors, which helps avoid model misspecification and results in an accurate but still parsimonious description of the predictors' effects. We examine the association between BMI and all-cause and cause-specific mortality (cardiovascular, respiratory, cancer, and diabetes). Analyses are based on the National Health Interview Survey 1986-2000 linked to the National Death Index for NH white adults age 50-80. Results suggest that the BMI-mortality association is more V-shaped than U-shaped for both men and women.   The association differs substantially for different cause of death: from essentially flat for cancer, V-shaped for cardiovascular, inverted J-shape for respiratory, to monotonically increasing for diabetes.

The sharp increase in the prevalence of obesity among adults in the U.S. in recent decades constitutes one of the most prominent public health issues. About a third of American adults is overweight and another third obese (Hedley, Ogden et al. 2004). Health consequences of excess weight encompass a wide range of negative outcomes, from chronic conditions such as diabetes and osteoarthritis, to physical limitations, poor health-related quality of life, to higher mortality rates (Troiano, Frongillo et al. 1996; Durazo-Arvizu, McGee et al. 1997; Bender, Trautner et al. 1998; Ferraro and Booth 1999; Himes 2000; Ferraro, Su et al. 2002; Mokdad, Ford et al. 2003).

These population trends motivated the development of a large body of literature examining the association between BMI and mortality. There is a general consensus that both very low and very high BMIs are associated with increased mortality (Lee, Manson et al. 1993; Bender, Jockel et al. 1999; Greenberg 2001; Allison, Zhu et al. 2002; Zhu, Heo et al. 2003; Flegal, Graubard et al. 2005; Flegal, Graubard et al. 2007). The knowledge about the higher odds of dying for the very thin and very heavy has been documented as far back as the 19th century (Czerniawski 2007). In contrast, there is little agreement about the shape of the relationship for the bulk of the BMI distribution in the population (for instance, see Mokdad, Marks et al. 2004; Flegal, Graubard et al. 2005; Greenberg 2006; Greenberg, Fontaine et al. 2007). Across different studies and population segments, the nadir of the mortality curve varies from BMI below 22 to above 30. In studies using categorical BMI, there is similar lack of consensus about the mortality risks of overweight adults in comparison to their normal-weight counterparts (McGee 2005).

Moreover, there is a number of factors that have a strong impact on the association between body mass and mortality. Controlling for smoking, exclusion of early deaths, baseline health status, and other adjustments sometimes has a profound effect on the shape of the association (i.e., Manson, Willett et al. 1995; Allison, Gallagher et al. 1997; Adams, Schatzkin et al. 2006; Gelber, Kurth et al. 2007). In addition, the BMI-mortality relationship is context-specific: age, gender, and race groups all evidence somewhat different effects of body weight on mortality (Sobal and Stunkard 1989; Inelmen, Sergi et al. 2003; Yan, Daviglus et al. 2004; Flegal, Graubard et al. 2005). Thus, despite the wealth of research on BMI and mortality, we still don't have a sufficient understanding of the association between BMI and mortality for the majority of the adult population in the center of the weight distribution.

In this paper, we address two drawbacks ubiquitous in published studies: examination of all-cause rather than cause-specific mortality, and forcing an a priori functional form of BMI on mortality risk. The focus on all-cause mortality is driven in part by data constraints: studying cause-specific mortality in adults requires a fairly large dataset, which naturally has to include both body weight information and mortality followup. It is understood, however, that BMI has a different association with deaths from diabetes or cardiovascular disease compared to cancers or even external causes like accidents and homicides. Second, all published studies we are aware of constrain the shape of BMI to be either a step function (when BMI is included in models as a categorized predictor with 'normal' weight as reference), linear (usually for a restricted range of the weight distribution), or quadratic (to allow the increased mortality at both ends of the BMI distribution). This type of constraints can potentially bias findings and obscure the true shape of the relationship because they impose an apriori global solution to the functional form of the predictor, BMI, rather than allowing the data to reveal the best fitting form.

This is an exploratory analysis that examines the shape of the relationship between body mass index and all-cause as well as cause-specific mortality, using a large nationally representative data (NHIS) from 1986-2000 linked to the National Death Index. Our paper contributes to the literature in two main aspects. First, we study mortality by specific causes of death. Such disaggregation is crucial in order to provide a clearer picture of the effects of body weight on mortality processes that presumably differs for causes as varied as, say, cardiovascular, smoking-related respiratory, and diabetes. Second, we employ generalized additive models, a type of nonparametric approach, to examine the shape of the association. Many researchers who examine the effect of BMI on health outcomes note the need to more closely examine the shape of the association. We fill this serious gap by providing a picture of the shape that is driven entirely by the data structure, rather than by a priori modeling assumptions for the functional form of the BMI effect.

DATA AND METHODS

Data

The analyses are based on the National Health Interview Survey (NHIS) data for years 1986-2000, linked to multiple cause-of-death files from the National Death Index that includes deaths through 2002. The NHIS is a large annual cross-sectional health survey conducted through

face-to-face household interviews by the National Center for Health Statistics (NCHS). NHIS uses a complex multistage stratified sampling design to obtain a sample representative of the civilian non-institutionalized U.S. population. Every year, around 100,000 households are selected for interviews. The total household response rate was around 90%; the rates were over 95% in the earlier years (Massey, Moore et al. 1989) and declined to 89% in 2000 (CDC 2002). Additional information about the sampling design are available in Massey et al. (1989) for years 1985-1994 and in Botman et al. (2000) for 1995-2004. The data and accompanying documentation for all survey years are available on the NCHS website: http://www.cdc.gov/nchs/about/major/nhis/quest_data_related_doc.htm.

Vital status for adults who participated in 1986-2000 NHIS was determined by a probabilistic match to the National Death Index (Lochner, Hummer et al. 2007). The NDI includes information on the date of death and detailed cause of death. The NHIS-MCD linked data file includes only the underlying cause of death and select contributing causes, coded using the ICD-10 classification.

Sample

The analyses are based on non-Hispanic white men and women who were 50-80 years at the time of the interview and whose BMI was between 15 and 45 points. The upper age boundary was set because of indications that the matching of death information was less successful than at younger ages (Masters, Brown and Hayward, unpublished analyses), which could lead to biased findings. The lower age boundary and the BMI range were chosen to obtain a relatively small sample (N=109,345 for men and N=128,563 for women), without outliers in ages at death. The BMI range was chosen to exclude individuals (3.1%) with extreme values of the body mass index.

Measures

Body mass index (BMI) was calculated from self-reported height and weight using formula BMI=703*(weight in pounds/squared height in inches). From the NHIS respondent pool, about 12 percent did not report their height or weight, were not asked the height and weight questions, or their information was topcoded or bottomcoded and thus their BMI was not calculated. The top- and bottom-coded individuals were generally at the ends of the BMI distribution (though not necessarily—for instance, a very tall and heavy person may have been topcoded although they had a normal BMI). With a flexible modeling of the BMI as a predictor of mortality, this would

not bias the results – however, data sparseness at the boundaries of the distribution may present an estimation problem. In the first part of the analysis, BMI was used as a continuous predictor with a quadratic functional form centered around 25.

The mortality data file included the underlying cause of death and select contributing causes coded using the ICD-10 classification. We categorized causes into one of the following major categories: cardiovascular, cerebrovascular disease, respiratory including lung cancer, cancer excluding lung, diabetes, external, and other causes. Lung cancer was included among lower-respiratory diseases such as COPD or emphysema because all these causes are strongly affected by smoking behaviors.   In the analyses, we presented results for all-cause mortality, as well as for four select causes: cardiovascular, respiratory, cancer, and diabetes. Duration of followup was measured in years and ranged from 3 to 17 for survivors and 0-17 for respondents who died. Control variables included age, region of residence (Northeast as reference), and rural residence (not rural as reference). Auxiliary analyses also adjusted for marital status (married as reference) and education categorized as less than high school, high school (reference), and more than high school.   There were no missing values on demographic variables used in presented models. The proportion of observations with missing values on marital status and education was less than 1 percent. These were excluded from the auxiliary analyses. The inclusion of the control variables, on which the survey sampling was based, eliminated the necessity of using weights in order to obtain unbiased estimates for the effects of the predictors (Winship and Radbill 1994). However, to assess the possible issues with point estimates and their standard errors, we estimated parallel sets of parametric models, with and without adjusting for sampling design. We found no substantive differences in the values of the point estimates or their p-values.

Analysis.

In the first part of the analyses, we estimated models comparable to those used in existing literature: Cox proportional hazard models with a quadratic form for BMI. The results are presented in table 2 and figure 1. In the rest of analyses, we used Poisson models on a person-year data structure (Carstensen 2005), because GAMs have not been developed for use with proportional hazard models. In the Poisson analysis, the unit of observation is not time of observation like for Cox models but a person-year time interval. The response is a dichotomous 0/1 variable indicating dead or survive for each person-time unit. This approach has multiple advantages over proportional hazard models, including an easy incorporation of time-varying

predictors. The disadvantages include computational cost (expanding a person-level dataset to person-time dataset multiplies the number of observations in the dataset, in our case by a factor of 10). We first these estimated these models with the same set of predictors as the Cox models to examine the comparability of estimates from proportional hazard models and person-year Poisson models.

 In the main part of the analyses, we employed generalized additive models to flexibly estimate the shape of the BMI- mortality association. This approach is ideal for examining the precise shape of a predictor's effect in a way that parametric methods cannot capture (Xiang 2001). The GAMs allow fitting any predictor with an arbitrary nonparametric smoothing function, which avoids the need for a priori assumptions about the shape of the predictor's effect. The models assume only that the mean of the outcome variable is a function of additive predictors through a nonlinear link function (Stone 1985; Hastie and Tibshirani 1990). These models offer a great compromise between some nonparametric methods such as projection pursuit, recursive partitioning regression, or neural networks, which are very flexible but difficult to interpret, and traditional models, which are parsimonious and easy to read but may not capture the effect of a predictor if the effect varies across the range of its values (Beck and Jackman 1998). GAMs are more versatile that other nonparametric regression approaches, such as splines or local regression methods. Moreover, GAM can accommodate interaction terms (we used tensor product splines), which allows us to examine the shape of the BMI-mortality association by lag time from interview to death/censoring. R version 2.6.1 was used to estimate the models. Results are shown two- and three-dimensional plots (the latter to display how the BMI-mortality curve varies by duration).

The intuition for GAMs is straightforward. In simple linear regression, the dependent variable is a linear function of one predictor: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. Multiple linear regression extends this relationship to more than one predictor; the dependent variable y is a *linear* combination of the predictors $x_1, x_2, ... x_k$ or $y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki} + \varepsilon_i$. There are additional assumptions such as that the errors $\varepsilon \sim i.i.d. N(0, \sigma^2)$. Generalized linear regression is an extension for non-normally distributed response variable: $\eta_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki} + \varepsilon_i$ where some link function $g(y_i)$ transforms the mean of *y to* η_i.

Nonparametric regression relaxes the assumption of linearity. The strength of nonparametric methods is a more accurate modeling of the regression function. The drawbacks differ according to the specific method but generally include computation cost and interpretability of results. The outcome is assumed to be some smooth (but not necessarily linear or monotonic) function of the predictors. Simple nonparametric regression parallels simple linear regression in that the dependent variable is a function of one predictor, but their association is not necessarily linear: $y_i = \beta_0 + f(x_i) + \varepsilon_i$. This regression is also called scatterplot smoothing because it estimates a smooth curve in a scatterplot of y against x. It provides intuition for how higher-dimension nonparametric regression models work and it forms essential estimation blocks for additive nonparametric regression. The most widely used simple nonparametric regression approach is lowess (locally weighted scatterplot smoothing) , a local regression with tricube weights and bisquare robustness weights (Cleveland 1979).

Additive nonparametric regression is a multiple regression with the following constraint: the outcome is assumed to be a combination of one or more smooth functions of the predictors: $y_i = \beta_0 + f_1(x_{1i}) + \cdots + f_k(x_{ki}) + \varepsilon_i$. Its major strengths are in interpretability and estimation (avoids the curse of dimensionality that plagues most nonparametric methods, among them lowess/loess). In terms of both estimation and interpretation, the additive model reduces to two-dimensional partial-regression equations. The additive model can be semi-parametric, that is, some of the predictors can be linear predictors and others nonparametric: $y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + f_{p+1}(x_{(p+1)i}) + \cdots + f_k(x_{ki}) + \varepsilon_i$. Finally, despite its name, additive models can incorporate the 'interaction' of two or more predictors. Generalized additive models extend the linear additive regression to non-Gaussian responses including binary and count data: $\eta_i = \beta_0 + f_1(x_{1i}) + \cdots + f_k(x_{ki}) + \varepsilon_i$. We use a Poisson distributed response with a log link on a person-year file structure to model the odds of dying by BMI.

RESULTS.

Table 1 presents characteristics of the analysis sample that comprises non-Hispanic white men and women age 50-80 at interview with BMI between 15 and 45. The mean age at interview was 62.8 years for men and 63.9 years for women. On average, men contributed 9.7 years of exposure, women contributed 10.1 years. During the followup time, 29.4% of the male

respondents died at average age of 74 years, as did 22.4% of the female respondents, at 76.1 years. During the interview, men's mean BMI was 26.3 and women's BMI was 25.4 – in the low overweight range for both groups. More women than men had normal weight. More men than women were overweight (47.7% vs. 30.5%) but slightly more women were obese (16.3% vs. 15.6%). The modal cause of death was cardiovascular disease for both genders.

Table 2 shows results from parametric proportional hazard models of all-cause mortality for men and women, using quadratic and categorical classifications of BMI. In all models, BMI was a significant predictor of mortality hazard. Controlling for education and marital status had little impact on the BMI coefficients (results not shown, available on request). In the quadratic specification, a simple calculation shows that the nadir of the BMI-mortality curve was at BMI=28.2 for men and 27.1 for women. The models with categorical BMI show that obese individuals had a significantly higher risk of dying compared to their normal-weight counterparts, although the effect was weaker for men. Obese men had about 9% higher odds of dying compared to normal-weight men while obese women had 26% higher odds. Overweight men had substantially lower mortality hazard as compared to normal-weight adults; their odds of dying were about 15% lower. The overweight vs. normal weight difference was significant for women as well but substantively smaller, about 3% lower odds of dying for overweight women. Being underweight was associated with the highest mortality hazard: compared to their normal-weight counterparts, the odds of dying were 240% higher for underweight men and about 80% higher for underweight women.

Figure 1 shows the results from model 1 for men and women. The plot suggests that BMI was a somewhat stronger predictor of all-cause mortality for men than for women, with men's mortality increasing more steeply from the nadir of the curve. The figure also shows that in this quadratic specification of BMI, there was a wide range of BMIs, between BMIs of about 25 to the low obese range around 32 where mortality was roughly flat for both men and women.

Next, results from nonparametric GAMs are presented. Because the GAMs have not been developed for proportional hazard modeling network, we switched to Poisson models of death indicator on a person-year data. We first estimated the Poisson models parametrically, with the same set of predictors as the models above (table not shown). We found the effects of BMI to be essentially identical: for instance, the effect of BMI was .958 and BMI squared 1.007 in the

Poisson model, compared to Cox results of .957 and 1.007. The standard errors were identical to the fourth decimal place as well. This comparison ensured that the results described below are comparable to results obtained from Cox models.

Figures 2a and 2b display the shape of the association between BMI and all-cause mortality, adjusted for basic demographics – comparable to those shown in table 2 and figure 1. The association retains the curvilinear shape but it's narrower than the quadratic form would suggest. For men, the shape is closer to a V, with a nadir just slightly above BMI of 25. For women, the nadir of the mortality curve is below 25, in contrast to the BMI=27.1 calculated from the parametric models. There is also a leveling of the mortality curve at high BMIs (above 40).

Duration of followup time has been shown an important mediating factor in the BMI-mortality relationship. We first incorporated the lag time between interview and a given person-year of observation into the models additively, and allowed both the lag variable and age to be modeled flexibly. The results are presented in figures 3a for men and b3 for women. Adjusting for lag time does not have a strong effect on the shape of the BMI predictor. Age is linear, except for the lower mortality at the highest ages for men. Mortality is also essentially flat across the duration of the survey except for the first year or two. This pattern is probably due to the fact that respondents who were likely to die immediately after the survey were less likely to participate in it. Despite the absence of a strong effect of lag time on the mortality risk, it may moderate the effect of BMI throughout the followup time.

The results from models that include the BMI by lag time interaction (using tensor product splines) are presented in figures 4a and 4b. The patterns for both men and women look similar: for deaths that occur shortly after the interview where BMI is reported, low BMIs are associated with much higher mortality while mortality appears to be increasing more slowly with higher BMIs. With increased followup time, this patterns reverses: low BMIs are associated with relatively low mortality, while individuals who reported high BMIs are interviews have much higher chances of dying. This patterns likely represents reverse causality whereby disease causes weight loss – hence, individuals who report low BMIs at interviews may be those who have low weight because they are seriously ill and are therefore likely to die shortly after the interview.

Finally, in figure 5 we show the association of BMI with cause-specific mortality, focusing on cardiovascular, respiratory, cancer, and diabetes deaths. The main point from these results is the substantial differences among the causes in the functional form of the BMI predictor. For cardiovascular deaths, the association is U- or V-shaped for men and women. The nadir of the relationship is around 25 for men and 22-23 for women. For respiratory deaths, there is a strong increase in mortality at low BMIs and little or no increase at high BMIs. This pattern is likely due to smoking whereby smokers tend to have lower body weights (Flegal, Troiano et al. 1995) and suffer increased mortality from respiratory illnesses such as lung cancer or emphysema. The relationship between BMI and cancer is relatively flat, compared to other causes. For men, there appears to be slightly increased risk of dying from cancer at low and high BMIs, while for women the mortality curve dips slightly at the extremes of the BMI range. This difference may be due to the gender difference in the distribution of various types of cancers, each of which may have a different association with body mass. Finally, diabetes shows a clear, strong, monotonically increasing association with BMI for both men and women.

DISCUSSION.

Despite the wealth of research addressing the effects of body weight on mortality, the shape of the association has not been established beyond the general curvilinear pattern where high and low BMIs are associated with increased mortality. Parametric models that have been used for the analyses, with BMI specified as quadratic or step function by definition cannot describe the precise shape because they impose an apriori global shape on the full range of data. Nonparametric models are better suited to allowing the data to reveal the shape of the association.

In this paper, we used generalized additive models, a type of nonparametric models, to examine the shape of the BMI-mortality association for all-cause and cause-specific mortality in non-Hispanic white male and female older adults. The analyses revealed several interesting results. First, the effect of BMI on all-cause mortality tends toward a V-shaped one, rather than the U-shape under the quadratic specification. The corollary of this finding is that researchers may use splines to model the association rather than a quadratic form. The placement of the knot for the spline function could be determined using a nonparametric exploratory analysis like those presented here. Second, the association between BMI and mortality is somewhat stronger for

men than for women but the general shape is similar for both genders in all-cause, as well as cause-specific mortality. Third, duration is an important factor influencing the BMI-mortality association. For short-duration followup, low BMIs are associated with increased mortality more so than high BMIs. A reversal occurs with increased lag time whereby for deaths that occur after a long duration since interview, high BMIs are strongly associated with mortality while low BMIs are not. Finally, there are large differences between BMI and mortality from different causes of death – BMI is only weakly related to cancer deaths, low BMI is strongly associated with respiratory deaths, and high BMI with diabetes deaths.

The goal of this study was to begin analyzing the nature of the shape of the BMI-mortality association. There are limitations in the data and analyses that make the finding only a first step in a progress toward a better modeling of their association. The NHIS dataset was used because of its size, which permitted cause-specific analyses. However, smoking information was collected only in periodic supplements. Since smoking is a crucial mediating variable, its absence for the full dataset is a strong drawback of the presented results. Another limitation is not controlling for baseline health status -- this problem, however, can be addressed relatively easily in future studies. A third limitation concerns the BMI measure. BMI changes across the lifecourse, sometimes dramatically. The NHIS collected BMI at one time point only and we had to model BMI as if it remained static during the followup.

A next step in this line of inquiry will involve extending this analysis to control for smoking status, physical activity, baseline health, and socioeconomic status. Another desirable feature of the data to be used is measured, rather than self-reported, body mass index. A dataset such as the National Health and Nutrition Examination Survey III linked to mortality files includes all these data, although is much smaller size (N=20,050 adults) would preclude cause-of-death analyses.

Table 1. Characteristics of the analysis sample at interview and followup, by gender.

	Men	Women
N	109,345	128,563
Age at interview (mean, s.d.)	62.8 (8.4)	63.9 (8.7)
BMI (mean, s.d.)	26.3 (3.9)	25.4 (4.8)
Underweight	1.0%	3.2%
Normal weight	35.7%	49.9%
Overweight	47.7%	30.5%
Obese	15.6%	16.3%
Education (mean, s.d.)	12.7 (2.5)	12.6 (2.2)
Rural	28.3%	28.1%
Region of residence		
Northeast	22.3%	22.9%
Midwest	27.2%	27.2%
South	31.9%	32.0%
West	18.5%	17.9%
Followup information		
Proportion died	29.4%	22.4%
Years of followup (mean, s.d.)	9.7 (4.1)	10.1 (4.0)
Age at death (mean, s.d.)	74.0 (8.3)	76.1 (8.5)
Cause of death (as proportion of total deaths)		
Cardiovascular	36.3%	33.8%
Cerebrovascular	5.5%	8.0%
Respiratory[1]	17.3%	14.2%
Cancer[1]	19.4%	20.2%
Diabetes[2]	2.4%	3.0%
External	3.0%	2.0%
Other	16.1%	18.7%

Note: the analysis sample is restricted to non-Hispanic white men and women, age 50-80 at interview, with BMI between 15 and 45.
[1] Lung cancer is excluded from the cancer category and included in the respiratory category.
[2]Additional deaths had diabetes as an underlying cause: 2,816 observations (8.8% of all deaths) for men and 2,774 observations (9.6% of all deaths) for women.

Table 2. Effect of BMI on all-cause mortality, by gender.

	Men		Women	
	1	2	1	2
Age	1.093***	1.093***	1.095***	1.094***
	(0.000782)	(0.000780)	(0.000835)	(0.000836)
Rural	1.038***	1.037***	0.993	0.990
	(0.0129)	(0.0129)	(0.0132)	(0.0131)
Midwest	0.995	0.997	1.051***	1.050***
	(0.0161)	(0.0161)	(0.0177)	(0.0177)
South	1.083***	1.093***	1.063***	1.067***
	(0.0167)	(0.0168)	(0.0173)	(0.0174)
West	0.921***	0.926***	1.000	1.008
	(0.0164)	(0.0165)	(0.0188)	(0.0189)
Categorical BMI, normal weight as ref.				
Underweight		2.399***		1.805***
		(0.0924)		(0.0457)
Overweight		0.847***		0.971**
		(0.0104)		(0.0135)
Obese		1.087***		1.258***
		(0.0187)		(0.0210)
Continuous BMI				
BMI	0.957***		0.982***	
	(0.00165)		(0.00140)	
BMI squared	1.007***		1.004***	
	(0.000189)		(0.000154)	
N	109342		128562	

*** $p<0.01$, ** $p<0.05$.
Note: Mortality hazard rations, standard errors in parentheses.
Results are from all-cause proportional hazard models for non-Hispanic white men and women age 50-80 at interview with BMI 15-45. Model 1 specifies a continuous quadratic shape of the BMI predictor. BMI is centered around 25. Model 2 includes BMI as a categorical predictor with normal weight (BMI 18.5-24.9) as reference.

Fig. 1. Mortality hazard by BMI, quadratic specification
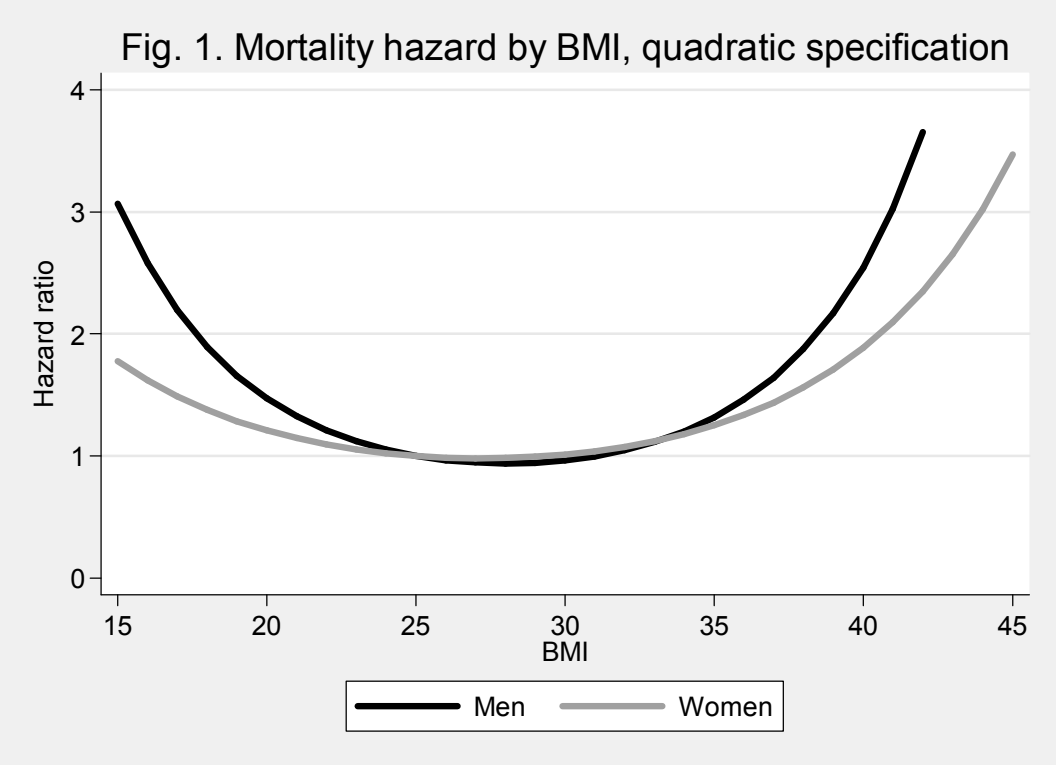
Fig. 2a. All-cause mortality risk by BMI, for men
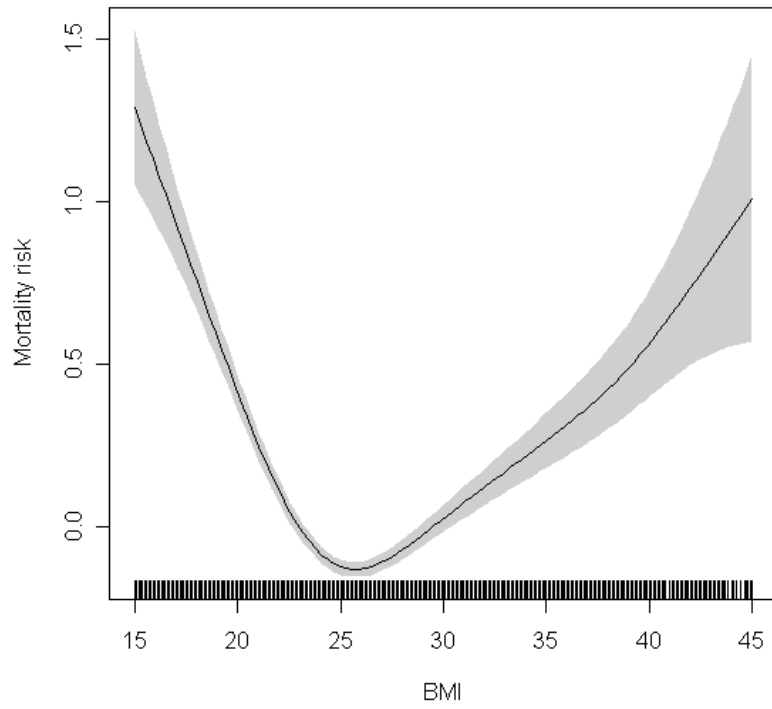

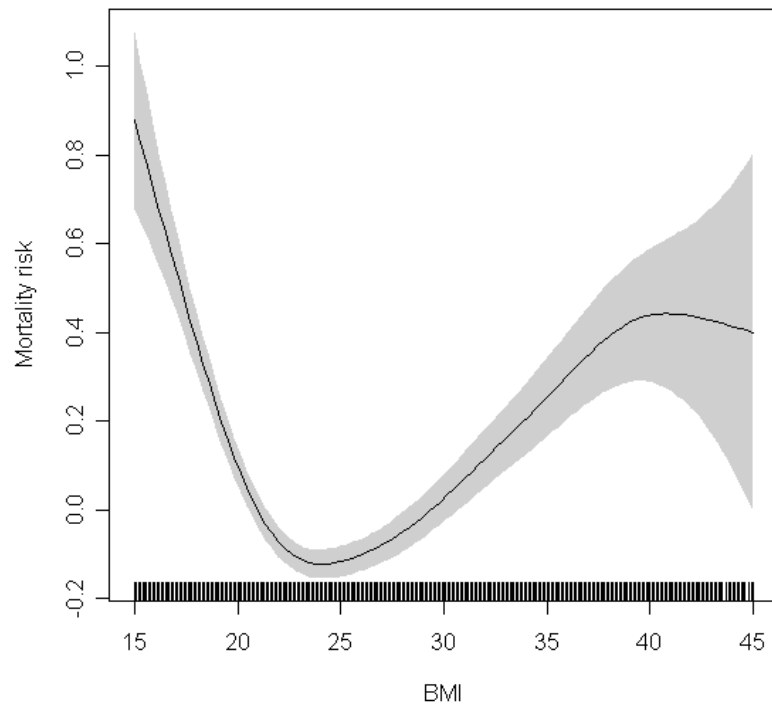**Fig. 2b. All-cause mortality risk by BMI, for women**

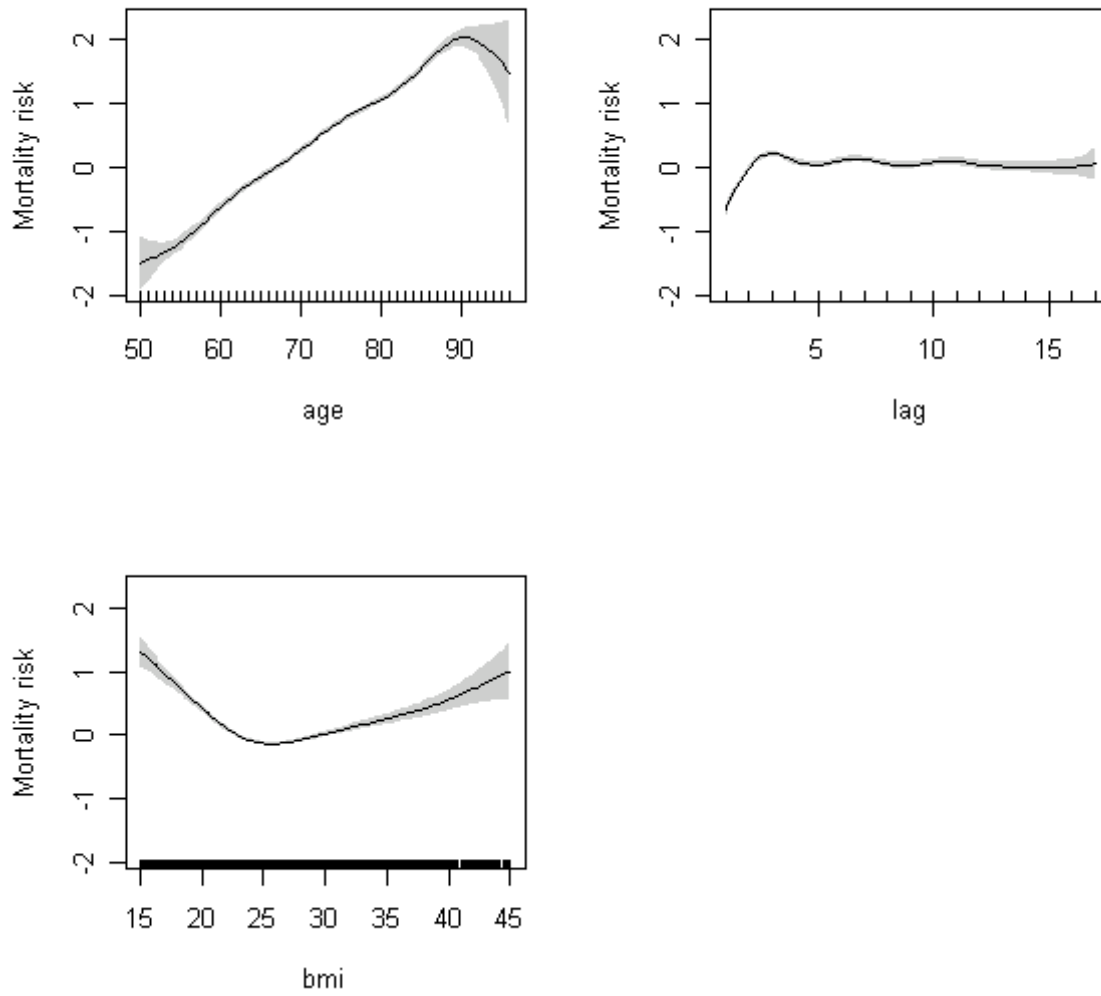Fig.3a. Mortality Risk by BMI, age, and lag time, for men

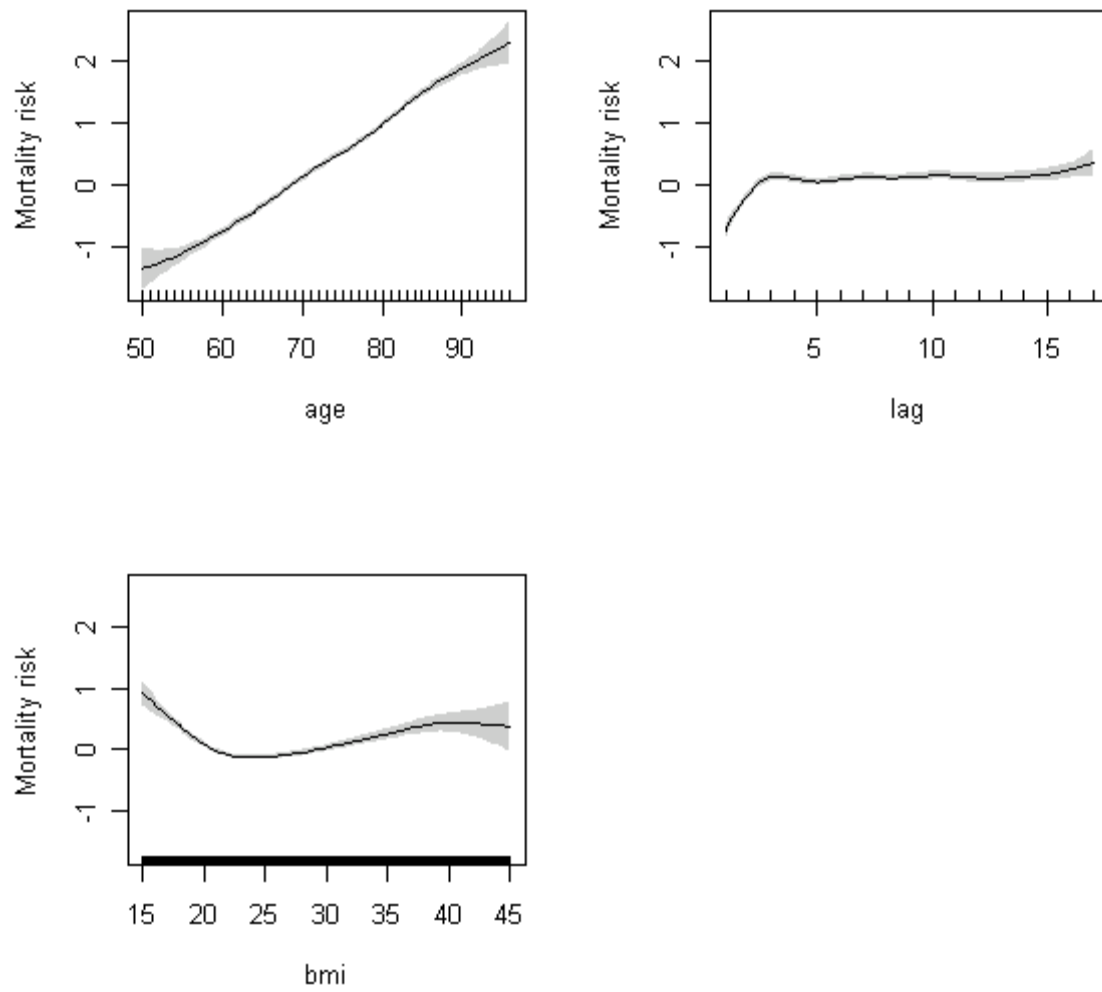Fig.3b. Mortality Risk by BMI, age, and lag time, for women

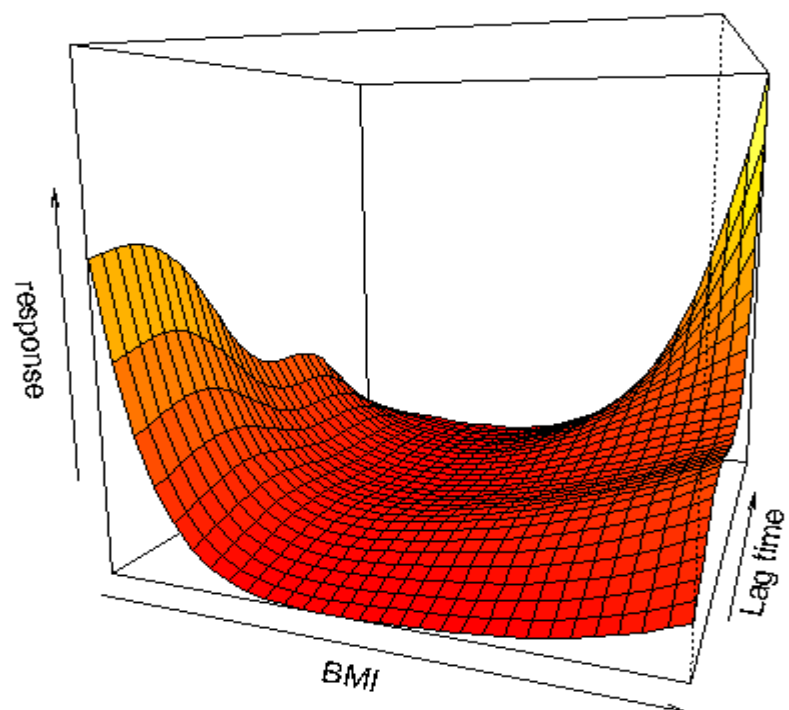Fig. 4a. Mortality by BMI and lag time, for men

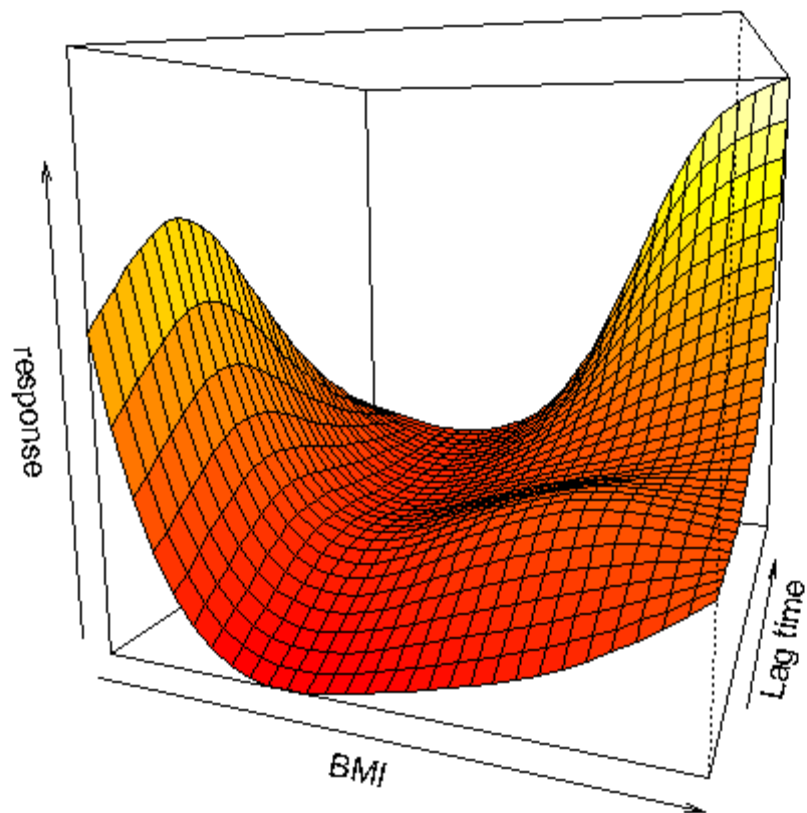**Fig. 4b. Mortality by BMI and lag time, for women**

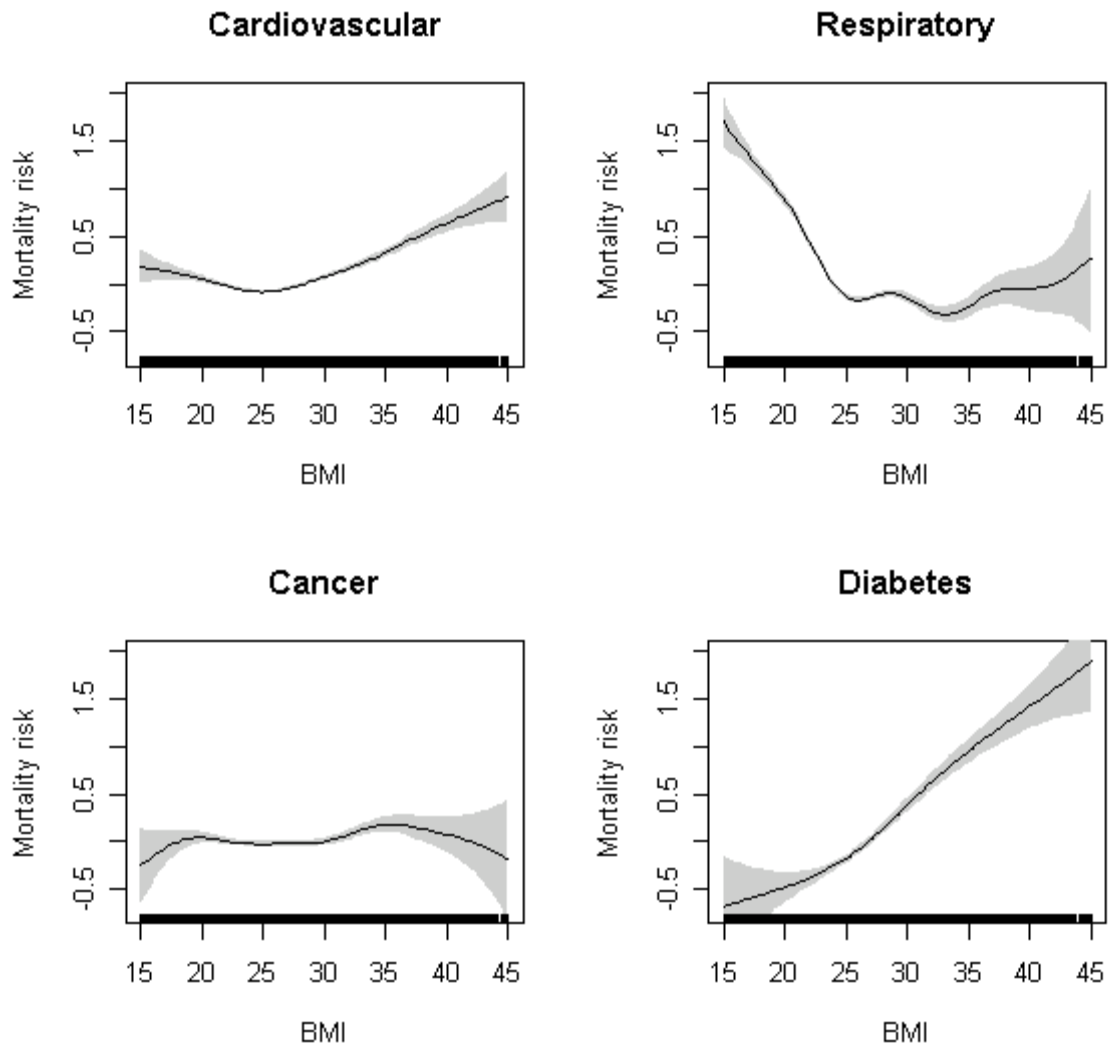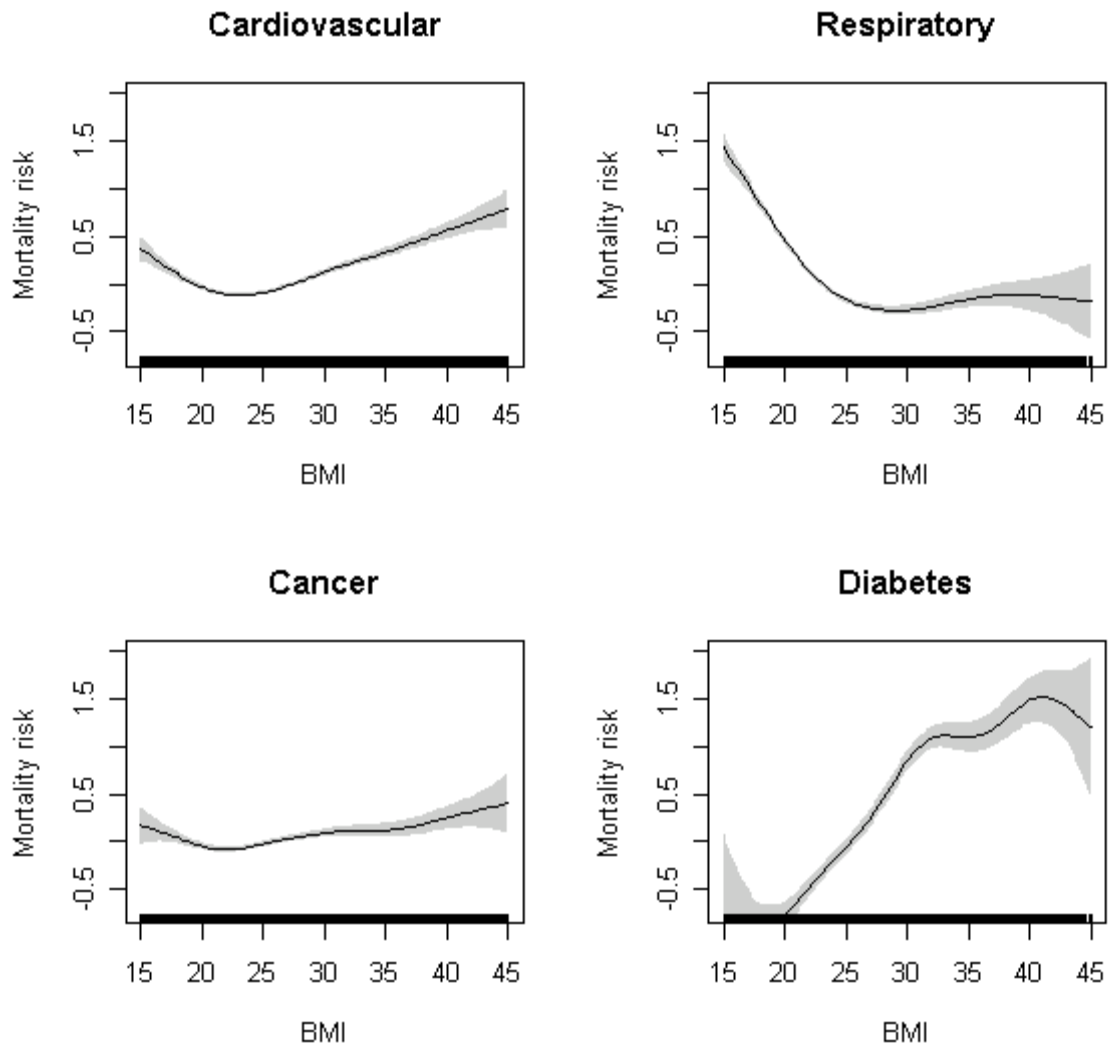## Fig. 5a. Cause-specific mortality risk by BMI, for men

Fig. 5b. Cause-specific mortality risk by BMI, for women

REFERENCES

Adams, K. F., A. Schatzkin, et al. (2006). "Overweight, Obesity, and Mortality in a Large Prospective Cohort of Persons 50 to 71 Years Old." New England Journal of Medicine **355**(8): 763-778.

Allison, D. B., D. Gallagher, et al. (1997). "Body Mass Index and All-Cause Mortality among People Age 70 and over: The Longitudinal Study of Aging." International Journal of Obesity **21**(6): 424-31.

Allison, D. B., S. Zhu, et al. (2002). "Differential Associations of Body Mass Index and Adiposity with All-Cause Mortality among Men in the First and Second National Health and Nutrition Examination Surveys (NHANES I and NHANES II) Follow-up Studies." International Journal of Obesity **26**(3): 410-6.

Beck, N. and S. Jackman (1998). "Beyond Linearity by Default: Generalized Additive Models." American Journal of Political Science **42**(2): 596-627.

Bender, R., K. H. Jockel, et al. (1999). "Effect of Age on Excess Mortality in Obesity." JAMA **281**(16): 1498-504.

Bender, R., C. Trautner, et al. (1998). "Assessment of Excess Mortality in Obesity." American Journal of Epidemiology **147**(1): 42-8.

Botman, S. L., T. F. Moore, et al. (2000). Design and Estimation for the National Health Interview Survey, 1995-2004. , National Center for Health Statistics. Vital Health Statistics 2(130).

Carstensen, B. (2005). Demography and Epidemiology: Practical Use of the Lexis Diagram in the Computer Age., Paper presented at the Annual Meeting of the Finnish Statistical Society, Oulu, Finland, May 2005.

CDC (2002). 2000 National Health Interview Survey (NHIS) Public Use Data Release, Division of Health Interview Statistics, National Center for Health Statistics Hyattsville, MD.

Cleveland, W. S. (1979). "Robust Locally Weighted Regression and Smoothing Scatterplots." Journal of the American Statistical Association **74**: 829-36.

Czerniawski, A. M. (2007). "From Average to Ideal: The Evolution of the Height and Weight Table in the United states, 1836-1943." Social Science History **31**(2): 273-96.

Durazo-Arvizu, R., D. McGee, et al. (1997). "Establishing the Nadir of the Body Mass Index-Mortality Relationship: a Case Study." Journal of American Statistical Association 92(440): 312-9.

Ferraro, K. F. and T. L. Booth (1999). "Age, Body Mass Index, and Functional Illness." Journal of Gerontology: Social Sciences 54(6): S339-48.

Ferraro, K. F., Y.-P. Su, et al. (2002). "Body Mass Index and Disability in Adulthood: a 20-Year Panel Study." American Journal of Public Health 92(5): 834-40.

Flegal, K. M., B. I. Graubard, et al. (2007). "Cause-Specific Excess Deaths Associated With Underweight, Overweight, and Obesity." JAMA 298(17): 2028-2037.

Flegal, K. M., B. L. Graubard, et al. (2005). "Excess Deaths Associated with Underweight, Overweight, and Obesity." JAMA 293(15): 1861-7.

Flegal, K. M., R. P. Troiano, et al. (1995). "The Iinfluence of Smoking Cessation on the Prevalence of Overweight in the United States." New England Journal of Medicine 333(18): 1165-70.

Gelber, R. P., T. Kurth, et al. (2007). "Body Mass Index and Mortality in Men: Evaluation the Shape of the Association." International Journal of Obesity 31: 1240-1247.

Greenberg, J. A. (2001). "Biases in the Mortality Risk Versus Body Mass Index Relationship in the NHANES-1 Epidemiologic Follow-Up Study." International Journal of Obesity 25(7): 1071-8.

Greenberg, J. A. (2006). "Correcting Biases in Estimates of Mortality Attributable to Obesity." Obesity 14(11): 2071-2079.

Greenberg, J. A., K. R. Fontaine, et al. (2007). "Putative Biases in Estimating Mortality Attributable to Obesity in the US Population." International Journal of Obesity 31: 1449-1455.

Hastie, T. and R. Tibshirani (1990). Generalized Additive Models. London, Chapman and Hall.

Hedley, A. A., C. L. Ogden, et al. (2004). "Prevalence of Overweight and Obesity Among US Children, Adolescents, and Adults, 1999-2002." JAMA 291(23): 2847-50.

Himes, C. L. (2000). "Obesity, Disease, and Functional Limitation in Later Life." Demography 37(1): 73-82.

Inelmen, E. M., G. Sergi, et al. (2003). "Can Obesity Be a Risk Factor in Elderly People?" Obesity Reviews 4(3): 147-55.

Lee, I.-M., J. E. Manson, et al. (1993). "Body Weight and Mortality. A 27-Year Follow-up of Middle-Aged Men." JAMA **270**(23): 2823-8.

Lochner, K. A., R. A. Hummer, et al. (2007). Comparative Analysis of the Public-Use and Restricted-Use NHIS Linked Mortality Files. Hyattsville, MD (Available at the following address: http://www.cdc.gov/nchs/data/datalinkage/nhis_mort_compare_2007.pdf). National Center for Health Statistics**:** 34.

Manson, J. E., W. C. Willett, et al. (1995). "Body Weight and Mortality among Women." New England Journal of Medicine **333**(11): 677-85.

Massey, J. T., T. F. Moore, et al. (1989). Design and Estimation for the National Health Interview Survey, 1985-94. , National Center for Health Statistics. Vital Health Statistics 2(110).

McGee, D. L. (2005). "Body Mass Index and Mortality: A Meta-Analysis Based on Person-Level Data from Twenty-Six Observational Studies." Annals of Epidemiology **15**(2): 87-97.

Mokdad, A. H., E. S. Ford, et al. (2003). "Prevalence of Obesity, Diabetes, and Obesity-Related Health Risk Factors, 2001." JAMA **289**(1): 76-9.

Mokdad, A. H., J. S. Marks, et al. (2004). "Actual Causes of Death in the United States, 2000." JAMA **291**(10): 1238-1245.

Sobal, J. and A. J. Stunkard (1989). "Socioeconomic Status and Obesity: a Review of the Literature." Psychological Bulletin **105**(2): 260-275.

Stone, C. J. (1985). "Additive Regression and Other Nonparametric Methods." Annals of Statistics **12**: 689-705.

Troiano, R. P., E. A. Frongillo, et al. (1996). "The Relationship Between Body Weight and Mortality: a Quantitative Analysis of Combined Information from Existing Studies." International Journal of Obesity **20**(1): 63-75.

Winship, C. and L. Radbill (1994). "Sampling Weights and Regression Analysis." Sociological Methods and Research **23**(2): 230-57.

Xiang, D. (2001). Fitting Generalized Additive Models with the GAM Procedure., SAS Institute Paper P256-26.

Yan, L. L., M. L. Daviglus, et al. (2004). "BMI and Health-Related Quality of Life in Adults 65 Years and Older." Obesity Research **12**(1): 69-76.

Zhu, S., M. Heo, et al. (2003). "Associations of Body Mass Index and Anthropometric Indicators of Fat Mass and Fat Free Mass with All-Cause Mortality among Women in the First and Second National Health and Nutrition Examination Surveys Follow-up Studies." <u>Annals of Epidemiology</u> **13**(4): 286-93.