Assessing Impacts on Mortality of Lifestyle Factors: Allowing for Model Uncertainty

Peter Congdon[1]

[1] School of Geography, Queen Mary University of London, London, UK

Correspondence: Peter Congdon, School of Geography, Queen Mary University of London, Mile End Rd, London E1 4NS, UK. E-mail: p.congdon@qmul.ac.uk

Abstract

Ongoing debate concerns how many deaths or disease cases are linked to lifestyle-linked (modifiable) factors such as adiposity, smoking and physical activity. This paper considers mortality attributable to modifiable risks using US cohort data from the NHANES III survey, focussing on mortality in a cohort of adults under age 65 at baseline. A piecewise-exponential regression analysis is adopted, with predictor selection, so acknowledging model uncertainty. Attributable risks are estimated using a Bayesian approach, with risks estimated by gender and population sub-groups (ethnic, income). For the cohort considered, smoking has the highest attributable risk, especially for males.

Keywords: modifiable risks, health behaviours, attributable risk, mortality, piecewise exponential, predictor selection.

1. Introduction

1.1 Assessing Impacts on Mortality of Modifiable Risk Factors

Recent studies have considered how far mortality and disease incidence are linked to modifiable risk factors, namely those reflecting health behaviours (Flegal et al, 2005; Mehta and Chang, 2009; Ekelund et al, 2015). Examples of such risk factors are physical activity, smoking, and adiposity, encompassing general and central obesity (e.g. Goh et al, 2014). Impacts of behavioural risk factors can be measured by attributable risks (also called population attributable fractions), which incorporate both associations between the outcome and exposure, and the proportion of subjects exposed (Flegal et al, 2015; Benichou, 2001). The attributable risk due to a particular exposure measures the reduction in mortality, prevalence or incidence due to a disease if the exposure were to be eliminated.

Issues raised by attributable risk studies of mortality include impacts of adiposity as against smoking and physical activity; potential biases in assessing adiposity impacts; and impacts of fat mass as against central obesity. Thus some investigations question the potential extent of rising obesity levels in shortening lifespans as compared to smoking (Flegal et al. 2005). Regarding adiposity and physical inactivity, some evidence indicates they independently contribute to mortality risk (Hainer et al, 2009; Hu et la, 2004). Possible biases in assessing adiposity impacts include reverse causation in samples including older subjects, since they are more likely to have chronic diseases leading to weight loss (Manson et al, 2007; Richman and Stampfer, 2010). Regarding the relative importance of general and central adiposity, evidence is inconclusive, some studies reporting stronger impacts of central adiposity on future mortality (Price et al, 2006), others finding no difference in predictive strength (Taylor et al, 2010), or concluding that both types of adiposity measure be used in tandem (Pischon et al, 2008).

1.2 Focus of This Paper and Data Used

This paper considers attributable mortality risks due to excess adiposity, smoking and physical inactivity, classed as *primary* risk factors in the terminology of Basu and Landis (2004), with a focus on US all-cause mortality. While many studies only consider population wide attributable risks, attributable risks are here obtained separately for males and females, and by income and ethnic group within genders.

The analysis is based on data from the National Health and Nutrition Examination Survey (NHANES), specifically the NHANES III health survey (1988-1994), and a mortality follow up to 2011. NHANES is a series of surveys assessing health and health behaviours, unique in combining interviews and physical examinations. The interview component includes demographic, socioeconomic, and behaviour questions, while the examination component consists of medical and physiological readings by trained personnel. The mortality follow up involves linkage of survey participants with the US National Death Index, facilitating investigation of links between health factors and mortality. The analysis considers

adults aged 35-64 at baseline to alleviate biases due to weight loss among older patients with chronic disease. Exclusion on the basis of disease status is not considered as this involves sample depletion, and possible alternative biases as thresholds or disease definitions are varied.

1.3 Methodological Aspects of the Analysis

A survival regression approach, with a piecewise-exponential representation of the hazard function, is adopted. Attributable risks are based on survival regression on the primary risk factors (adiposity, exercise, smoking), adjusting for confounders (Morgenstern, 2008, p 60). A Bayesian estimation approach using Markov Chain Monte Carlo (MCMC) methods is used, with predictor selection to allow for regression model uncertainty. Specifically, a stochastic search variable selection (SSVS) strategy is used for predictor selection (George and McCulloch, 1993).

Bayesian estimation also simplifies inferences on attributable risks, whereas confidence intervals are complex to obtain using classical approaches. Data missingness is explicitly considered via a multiple imputation strategy, avoiding biases present in complete cases analysis. Differential survey weights are also incorporated (Heeringa et al, 2015).

2. Methods

2.1 Attributable Risk Definitions

The attributable risk (AR), also sometimes called the attributable fraction, or population attributable fraction, is a measure of the importance of a risk factor (Samuelsen and Eide, 2008). Let D denote mortality, disease incidence or presence of disease (=1 for death or disease, = 0 otherwise), then one has

$$AR = \frac{P(D) - P(D|Intervention)}{P(D)} ,$$

where P denotes probability, and a generic intervention corresponds to the (possibly counterfactual) total or partial removal of one or more risk factors.

Consider mortality or disease status D_i, and risk factor data X_i for subjects i=1,...,n. Let X_i^* be the risk factor under intervention, with selected levels of the risk factor modified or removed. Let C_i denote other risk factors or confounders, and let E_i denote any postulated interactions $E_i = C_i X_i$ (i.e. allowing effect modifications) between these remaining confounders and X_i. Under exposure modification the latter would become $E_i^* = C_i X_i^*$.

Denote model based probability estimates of the outcome under observed and counterfactual scenarios as $r_i(X_i, C_i, E_i | \beta)$ and $r_i(X_i^*, C_i, E_i^* | \beta)$, respectively, where β measures impacts of X, C, and E on D. Let population wide totals be denoted

$$R = \sum_{i=1}^{n} r_i(X_i, C_i, E_i | \beta)$$

and

$$R^* = \sum_{i=1}^{n} r_i(X_i^*, C_i, E_i^* | \beta)$$

respectively. Then the attributable risk may be written (Greenland and Drescher, 1993) as

$$AR = \frac{R - R^*}{R} = 1 - \frac{R^*}{R}. \tag{1}$$

2.2 Population Sub-Groups

Often policy interest is on differentials in health impacts or burdens according to socio-demographic category, such as income or ethnic group G_i. One may estimate attributable risks AR_g for subgroups g by aggregating probability estimates within such groups, namely

$$R_g = \sum_{i=1, G_i=g}^{n} r_i(X_i, C_i, E_i | \beta)$$

and

$$R_g^* = \sum_{i=1, G_i=g}^{n} r_i(X_i^*, C_i, E_i^* | \beta).$$

Then

$$AR_g = 1 - \frac{R_g^*}{R_g}, \tag{2}$$

with the population wide AR obtained as

$$AR = \sum_g AR_g W_g,$$

with weights

$$W_g = R_g/R.$$

Differences in attributable risks between subgroups may be due to differences in risk factor prevalence, for example, physical activity and/or adiposity differences between income or ethnic groups (Powell et al, 2006). Differences in ARs between groups may also be due to risk factor modification by group, such as effects of obesity on mortality varying by ethnic group (Flegal et al, 2004).

2.3 Attributable Risks in Survival Analysis

More specifically consider survival data, with observed death times or right censored times t_i and predictors $Z_i = (X_i, C_i, E_i)$ Then estimation of attributable risks may involve the survival or hazard function (Samuelsen and Eide, 2008). For example, let $S(t_i|Z_i)$ denote the probability of surviving till time t, then the probability of mortality between t and $t+\Delta t$ is

$$r_i(t, t+\Delta t \,|Z_i) = P(t < t_i \leq t+\Delta t) = S(t|Z_i) - S(t+\Delta t|Z_i).$$

The corresponding attributable risks, based on totals $R(t, t+\Delta t)$ and $R^*(t, t+\Delta t)$ obtained as above, are specific to intervals $(t, t+\Delta t)$. A suitable parametric form for survival times may be difficult to obtain, and a flexible semi-parametric model involves a piecewise exponential assumption (Ibrahim et al, 2001) (see Appendix 1).

3. Case Study: Regression Specification

3.1 Dataset

We consider data from the National Health and Nutrition Examination Survey (NHANES) which assesses health and nutritional status, using both interviews and physical examinations. Specifically data are from the NHANES III survey of 1988-1994, and linked mortality follow-up on the respondents of that survey through to the end of 2011. The analysis involves subjects aged 35-64 at baseline: 3694 male subjects, and 4204 female subjects, with 1081 and 855 deaths respectively. Following Korn et al (1997), age at death (or censoring) is used as the time unit. J=10 intervals are used in the piecewise exponential regression, with $a_0=0$, interior cut-points a_j (j=1,..,9) based on the 10th, 20th,...,90th percentiles of observed ages at death, and with a_J exceeding the maximum age at death.

3.2 Defining Primary Risk Predictors

Predictors of mortality first consist of indicators of the primary risk factors. As a measure of central obesity, the waist-hip ratio category (Z_1) has three categories, using gender specific cut-points among subjects aged 35-64 in NHANES III ($Z_1=1$ for WHR under 75th percentile; $Z_1=2$ for WHR between 75th and 90th percentile; $Z_1=3$ for WHR over 90th percentile). As a measure of general obesity, BMI category (Z_2) has values 1, 2, or 3, according as BMI is under 30, 30-34.99, or over 35. Extreme BMI (i.e. over 35) has been found to have a strong impact on mortality outcomes (Kitahara et al, 2014). Both measures of adiposity are included as potential risk factors: for example, high BMI and abdominal obesity combined lead to particularly high cardiovascular risk (Snijder et al, 2006; Coutinho et al, 2013).

The other two primary risk factors are physical activity status and smoking status. Physical activity is represented using metabolic equivalent (MET) intensity levels (Ainsworth et al, 1993), based on leisure time physical activity in exercises, sports, or physically active hobbies. A fourfold categorisation was used: $Z_3=1$ for subjects reporting no activity (around 17% of males and 25% of females); $Z_3=2$ for those reporting some activity but below the median MET (by gender); $Z_3=3$ for those between the median and upper quartile MET; and $Z_3=4$ for those in the upper MET quartile.

Smoking status, Z_4, is an eight-fold category as in Ho and Elo (2013). This categorisation distinguishes between never-smokers, former smokers and current smokers, taking into daily cigarettes smoked for current smokers and time since quitting among former smokers. The categories are thus: never-smoker; former smoker who quit over 30, 20–29, 10–19, 5–9, or 0–4 years ago; and current smoker, distinguishing between smoking less than one pack, or one or more packs of cigarettes per day.

3.3 Other Predictors

Other predictors are behavioural and socio-demographic confounders. Behavioral confounders are first Z_5, a binary indicator of healthy eating. This indicator is based on the NHANES healthy eating index (Kappeler et al, 2013), with $Z_5=1$ for HEI scores in the upper quartile, and value 0 otherwise. A second behavioral confounder is alcohol consumption category, Z_6 (1=non-drinker as reference, 2=moderate, 3=heavy) defined as in Freiberg et al (2004). Socio-demographic confounders are poverty status Z_7, with value 1 for income to poverty ratios under 1, and value 0 otherwise; and

ethnic/race category, Z_8 (1=white non-Hispanic as reference, 2=black non-Hispanic, 3=Hispanic, 4=other). Poverty status is included because of evidence of confounding by socio-economic status (Mehta and Chang, 2009), and because of well-established effects of income-defined poverty on health and mortality, whether poverty is defined for individuals or communities (Marmot, 2002; Raphael, 2011). Ethnic group is included due to evidence of ethnic mortality differentials, controlling for other risk factors (Elo and Preston, 1997).

Four forms of interaction are considered, based on evidence from the literature. These are an interaction between adiposity and current smoking, Z_9 (binary, with value 1 for subjects who are both adipose and currently smoking); an interaction between adiposity and physical inactivity, Z_{10} (binary, with value 1 for adipose subjects reporting low or no activity); an interaction between adiposity and higher age, Z_{11} (binary, with value 1 for adipose subjects aged over 55); and interactions between adiposity and ethnicity (Z_{12}-Z_{14}). In defining interactions, adiposity consists of any one of the four excess BMI or excess WHR categories. The first interaction is included because associations between adiposity and mortality among current smokers may be weaker than among non-smokers (Manson et al, 2007; Koster et al, 2008). The second interaction is included based on recent evidence (Ekelund et al, 2015), though studies are inconsistent on this effect (Brown et al, 2012). The latter two types of interaction express possible effect modification by age and ethnicity (Mehta and Chang, 2009). For example, Crimmins et al (2011) state that "BMI has its largest effect on the risk of mortality for adults under 50, and the correlation between BMI and mortality decreases beyond that age".

3.4 Counterfactual Settings

The counterfactual scenario to derive attributable risks due to adiposity involves setting $Z_{1i}^*=Z_{2i}^*=1$ for all subjects (i.e. no excess central obesity and no high BMI levels), and also setting to zero all interactions (Z_9 to Z_{14}) involving adiposity. For the attributable risk due to physical inactivity, the setting $Z_{3i}^*=4$ is adopted. For the attributable risk due to smoking, subjects currently or formerly smoking are assigned to the never smoked category, and the interaction Z_9 is set to zero.

3.5 Predictor Selection

One benefit of a Bayesian approach is that model uncertainty can be allowed for in estimating attributable risks, since not all the above predictors may be relevant to varying mortality risks. Collinearity may produce unexpectedly insignificant or diminished effects, or effects contrary to established knowledge. Predictor selection is therefore applied to all regression parameters except for the interval intercepts β_{0k} in the piecewise exponential (Appendix 1).

For predictor j, let $I_j \sim \text{Bern}(\pi)$ be binary indicators, with $I_j=1$ corresponding to retention of coefficient β_j, and $I_j=0$ corresponding to exclusion. A stochastic search variable selection strategy is adopted whereby β_j has a conventional prior when $I_j=1$, but for $I_j=0$ the prior is centred at zero with high precision, so that while Z_j is still in the regression, it is essentially irrelevant to that regression. Thus

$$P(\beta_j|I_j)= I_j \, N(0,V_j)+(1-I_j) \, N(0,c_j V_j), \tag{3}$$

where V_j is the prior variance under retention, while c_j is small to ensure the range of β_j under $P(\beta_j|I_j=0)$ is confined to substantively insignificant values. The relevance of a risk factor is summarised in the marginal inclusion probability, namely the estimated probability that β_j is selected as relevant to explaining the outcome (equivalently the probability that $I_j=1$).

The predictors in the NHANES application are discrete valued (binary or categorical). The setting $V_j=1$ is adopted in line with a prior 95% expectation that relative risks will be between 0.14 (for protective factors) and 7.1 (for adverse factors).

Prior evidence in fact suggests less extreme relative risks, even for risk factors established to have strong adverse impacts on mortality. For example, Friedman et al (1997) report mortality relative risks under 2 for current smokers vs. never-smokers. The setting $c_j=1/10000$ is adopted, as in George and McCulloch (1993), with π set to 0.5 so that retention and rejection are a priori equally likely.

4. Case Study: Implementation

4.1 Survey Weights

Two complications are present in the analysis of these data. The first is the presence of survey weights (Heeringa et al, 2014). Let such weights be denoted w_i (rescaled to average 1). Then one defines

$$R=\sum_{i=1}^{n} w_i r_i(Z_i, \beta),$$
$$R^*=\sum_{i=1}^{n} w_i r_i(Z_i^*, \beta),$$

in deriving the attributable risk via formula (1). Estimation of the piecewise exponential survival model also involves a

weighted product over subjects i and intervals k of Poisson likelihoods L_{ik}, namely $\prod_i \{\prod_k L_{ik}\}^{w_i}$.

4.2 Missing Covariate Data

The second complication is the presence of missing covariate data. Missingness rates are highest for the healthy eating index (11%), waist-hip ratio (5%), BMI (8%), and income-poverty ratio (IPR, 9%). Otherwise missingness is under 0.25% (1 in 400) for smoking status, physical activity status, and alcohol use. To avoid the potential biases present in complete cases analysis, a multiple imputation strategy is adopted (see Appendix 2). To reflect differential survey weights (He et al, 2010), weighted likelihoods (discrete outcomes), or differentially weighted precisions (continuous outcomes) are adopted in the imputation.

The main survival analysis then involves MCMC estimation applied to K=5 multiply imputed datasets (Little et al, 2014) containing imputed covariate values where these values are missing. Variances of parameter estimates take account of within and between imputation variances. Inferences are from iterations 5000-10000 of two chain runs of 10000 iterations for each multiply imputed dataset, with convergence assessed using Brooks-Gelman-Rubin (BGR) statistics (Brooks and Gelman, 1998). Model checks are discussed in Appendix 3.

5. Results

Table 1 shows satisfactory posterior predictive checks, for both male and female mortality analyses. The predictive check probabilities are between 0.1 and 0.9. Figures 1.1 and 1.2 show that the 95% intervals of Kaplan-Meier estimates of survival probabilities over the J intervals contain the posterior means for the modelled survival probabilities.

Tables 2 and 3 show, for males and females respectively, parameter summaries of relative risks, calculated over imputed datasets (Marshall et al, 2009). Also shown is the percent relative efficiency of estimation using a finite number, M, of imputed datasets rather than an infinite number (Dong and Peng, 2013).

5.1 Significance and Retention of Risk Factors

As discussed above, the model selection framework allows that some risk factors, or possibly some sub-categories of categorical risk factors, be redundant in terms of predicting mortality risk. The results of such predictor selection are represented in Tables 2 and 3 by posterior retention probabilities $Pr(I_j=1|Y)$, where Y denotes observed data.

Consider first the three primary risk factors for which attributable risks are to be obtained (adiposity, inactivity, smoking). Table 3 shows that all four elevated WHR and BMI categories are significant factors that enhance risks of female mortality. By contrast, for males, of the waist-hip ratio indicators, only highly elevated WHR is a significant influence.

For both genders, both class 1 (moderate) obesity and extreme obesity are mortality risk factors, albeit with extreme obesity having a higher relative risk. Significance for moderate obesity (BMI 30-34.99) as a mortality risk factor, among both men and women, contrasts with the finding of some studies reporting an "obesity paradox" (e.g. Flegal et al, 2013)

Current smoking is significantly associated with higher mortality as compared to the reference never smoked category. The posterior mean relative risks of current smoking (1 or more packs per day) for all-cause mortality are 2.12 for males and 2.54 for females, similar to the estimates of Friedman et al (1997). Current moderate smoking (under 1 pack a day) is also associated with elevated risk. Among males, risks are elevated for former smokers who stopped relatively recently (within 0-4 years and within 5-9 years).

High physical activity (with $Z_3=4$) is a protective influence for both genders, with similar benefit (around 15% reduction

in mortality risks) across both genders, though for males, the 95% credible interval just straddles 1. This confirms other studies (e.g. Hainer et al, 2009) that activity is independently related to mortality after accounting for the effect of adiposity. Moderate activity levels (i.e. $Z_3=2$ or $Z_3=3$) are associated with lower male mortality, but the 95% credible intervals straddle 1, and the inclusion probabilities are inconclusive.

Effects of activity are weaker than those of adiposity and smoking, and this may be partly due to intercorrelation with other health behaviours, although predictor selection will control for such intercorrelation. For example, among males in the highest activity quartile, around 29% are also in the highest quartile on the healthy eating index, compared to a rate of 16% among males in the lowest activity quartile.

With regard to effects of disease or lifestyle confounders, it can be seen that moderate alcohol consumption is a protective factor for both sexes, in line with other evidence (e.g. Rimm et al, 1991; de Labry et al, 1992; Gaziano et al, 2000). The reduction in mortality is slightly greater for males. The indicator of healthy eating is also associated with lower mortality, an effect apparent even after controlling for adiposity (cf. McCullough et al, 2000).

As to socio-demographic factors, poverty status emerges as a significant adverse influence on mortality (cf. Hahn et al, 1996), especially among males. Black ethnicity is associated with elevated mortality risk for both genders, even after allowing for poverty status (cf. Hummer and Chinn, 2011).

As to risk factor interactions, no interaction emerges as significant except that between adiposity and older age at baseline, namely a reduced mortality impact for adiposity at older ages (cf. Crimmins et al, 2011).

5.2 Estimates of Attributable Risk

Estimates of attributable risk are obtained by MCMC sampling while allowing for predictor selection among primary risk factors and confounders. A full posterior density estimate of the attributable risk is available, without techniques such as delta approximation. AR estimates are also obtained, as in (2), for population sub-groups, namely by income group (high vs low income, defined by income to poverty ratios above and below 2 respectively), and by ethnic group (white N-H, black N-H, Hispanic) within gender. Table 4 and Figures 2.1-2.3 show all-cause mortality AR estimates by sub-group.

The overall AR estimates due to adiposity are around 7% for males and 13% for females. These estimates are similar to percentages of mortality attributable to obesity reported by Flegal et al (2015). The higher AR estimate for females reflects significance of all adiposity variables for females, and also the mostly higher relative risk estimates as compared to males. For example, the mean relative risk associated with extreme BMI is 1.92 for females compared to 1.79 for males.

Figure 2.1 also shows a higher AR due to adiposity among low income females. This reflects different prevalence of extreme WHR and BMI by income: lower income groups have higher prevalence of extreme adiposity. For example, among women aged 35-64, and using the first imputed dataset, among low income women, 9.3% have an extreme WHR (over 0.98), as against 6.1% among higher income women, and 11.8% have an extreme BMI (over 35), as against 8.4% among higher income women.

A higher AR due to adiposity also applies to black and Hispanic females, with associated cardiovascular disease and mortality risk (Agyemang and Powell-Wiley, 2013). This reflects higher adiposity levels for these ethnic categories (e.g. 47% obesity among black women compared to 37% across all females) (Ogden et al, 2013). AR estimates for physical inactivity are also higher among women than men, reflecting a higher prevalence of inactivity among women (25%, as against 17% of men).

AR estimates due to smoking show a slightly higher attributable risk for men than women. Higher current smoking rates among men than women (31% vs 25%) are offset by higher relative mortality risks for current smoking among women. Higher AR estimates for smoking occur among lower income groups. Among females, lower ARs show for black and Hispanic women, reflecting lower smoking prevalence (CDC, 2001).

6. Discussion

Attributable risk estimates may be sensitive to a number of methodological choices. Existing reviews mention factors such as choice of formula (Flegal et al, 2015), choice of counterfactual scenario (Flegal et al, 2015), whether to allow effect modification (Flegal et al, 2004), how to account for mediating effects (Mason and Tu, 2008), and choice of exposure threshold (Rockhill et al, 1998). Another issue is whether a single population is used, as against obtaining estimates of relative risks and risk factor prevalence from different sources (van Dam et al, 2008; Flegal et al, 2015).

The present paper follows a regression model based strategy for estimating attributable risks from a single population (e.g. Greenland and Drescher, 1993), adapted to cohort survival data. However, the approach adopted is distinct from existing studies in incorporating impacts of regression model uncertainty on attributable risks, specifically regarding the relevance or not of predictors. A realistic model for mortality will include a range of risk factors, confounders and interaction effects. This raises issues of potential collinearity between predictors, and hence the need for predictor selection.

Often a set of regression models (differing in subsets of retained predictors) differs little in terms of fit. Unlike classical predictor selection (e.g. using forward selection with PHREG in SAS), Bayesian methods do not select a single best model in such a situation, but allow for model uncertainty (Piironen and Vehtari, 2016). An additional benefit is an index of relevance for each risk factor on a single scale, the marginal inclusion probability, namely the probability that a risk factor should be included in the regression as a predictor of mortality. This principle can be extended to arbitrary combinations of risk factors, so that one could (for example) consider the joint inclusion probability for the four adiposity indicators in the above analysis (Ghosh and Ghattas, 2015).

Bayesian estimation also simplifies inferences on attributable risks, with credible intervals or highest posterior density intervals readily obtained from MCMC output. By contrast, standard errors and hence confidence intervals are complex to obtain using classical approximations, especially where observations have varying survey weights.

Limitations of the analysis may also be mentioned. Although the NHANES survey has the benefit of being nationally representative, a limitation, becoming apparent for cause-specific mortality, is the relatively small sample sizes for deaths in the NHANES mortality follow up, especially when attention is confined to ages under 65 at baseline.

Important substantive findings of the paper are obtained regarding risk factors for mortality and population sub-group differences. Regarding mortality risk, we find evidence for

(a) the enduring importance of smoking, the impacts of which outweigh those of adiposity (cf. Mehta and Chang, 2009), though relative impacts of these two factors vary between population groups;

(b) the relevance of both central obesity and BMI to the total effect attributable to adiposity, with no evidence for an obesity paradox.

Leading examples of sub-group contrasts, apparent from attributable risk estimates, are that

(a) adiposity-linked mortality risk is higher for females as compared to males, for Hispanics as compared to other ethnic categories, and for black females;

(b) smoking-linked mortality risk is higher for males, and for lower-income groups, in both cases reflecting differences in smoking prevalence.

A final point to emphasize is that estimates (e.g. of attributable risk) in the analysis here condition on the prevalence of risk factor patterns around 1990. However, for the current generation of adults under 65, there have been significant changes in risk factor levels. Since 1990, adiposity levels have increased, both in terms of BMI and central obesity (Walls et al, 2011; Agyemang and Powell-Wiley, 2013), while smoking rates have fallen. Therefore the excess attributable risk for smoking as compared to adiposity (apparent in Table 4) will be expected to diminish.

References

Agyemang, P., & Powell-Wiley, T. (2013). Obesity and black women: special considerations related to genesis and therapeutic approaches. *Curr Cardiovasc Risk Rep. 7*(5), 378-386. https://doi.org/10.1007/s12170-013-0328-7

Ainsworth, B., Haskell, W., Leon, A., Jacobs, D., Montoye, H., Sallis, J., & Paffenbarger, R. (1993). Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise*, *25*(1), 71-80. https://doi.org/10.1249/00005768-199301000-00011

Basu, S., & Landis, J. (2004). Model-based estimates of population attributable risks for ordinal data. *Bulletin of Informatics and Cybernetics*, *36*, 73-90

Beddhu, S., Baird, B., Zitterkoph, J., Neilson, J., & Greene, T. (2009). Physical activity and mortality in chronic kidney disease (NHANES III). *Clin J Am Soc Nephrol. 4*(12), 1901-6. https://doi.org/10.2215/CJN.01970309

Benichou, J. (2001). A review of adjusted estimators of attributable risk. *Stat Methods Med Res. 10*(3), 195-216. https://doi.org/10.1177/096228020101000303

Berkhof, J., van Mechelen, I., & Hoijtink, H. (2000). Posterior predictive checks: principles and discussion. *Computational Statistics*, *15*(3), 337–354. https://doi.org/10.1007/s001800000038

Brooks, S., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. J. *Computational and Graphical Stat*, *7,* 434-45

Brown, W., McLaughlin, D., Leung, J., McCaul, K., Flicker, L., Almeida, O., Hankey, G, Lopez, D., & Dobson, A. (2012). Physical activity and all-cause mortality in older women and men. *Br J Sports Med*, *46*(9), 664-8. https://doi.org/10.1136/bjsports-2011-090529

Burbidge, J., Magee, L., & Robb, A. (1988). Alternative Transformations to Handle Extreme Values of the Dependent Variable. *J American Statistical Association, 83*(401), 123-127. https://doi.org/10.1080/01621459.1988.10478575

Center for Disease Control. (2001). Women and Smoking: A Report of the Surgeon General. Office on Smoking and Health and Centers for Disease Control and Prevention.

Coutinho, T., Goel, K., Corrêa de, SáD., & et al. (2013). Combining body mass index with measures of central obesity in the assessment of mortality in subjects with coronary disease: role of "normal weight central obesity". *J Am Coll Cardiol. 61*(5), 553-60. https://doi.org/10.1016/j.jacc.2012.10.035

Cowan, C., Hauser, R., & Kominski, R. (2013). Improving the measurement of socioeconomic status for the national assessment of educational progress: a theoretical foundation. *National Centre for Educational Statistics*.

Crimmins, E., Preston, S., & Cohen, B. (2011). Understanding Divergent Trends in Longevity in High-Income Countries. National Research Council; Washington, DC.

de Labry, L., Glynn, R., Levenson, M., Hermos, J., LoCastro, J., & Vokonas, P. (1992). Alcohol consumption and mortality in an American male population: recovering the U-shaped curve: Findings from the normative Aging Study. *J Stud Alcohol. 53*(1), 25-32. https://doi.org/10.15288/jsa.1992.53.25

Dong, Y., & Peng, C. (2013). Principled missing data methods for researchers. *Springerplus., 2*(1), 222. https://doi.org/10.1186/2193-1801-2-222

Ekelund, U., Ward, H., Norat, T., & et al (2015). Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC). *Am J Clin Nutr. 101*(3), 613-21. https://doi.org/10.3945/ajcn.114.100065

Elo, I., & Preston, S. (1997). Racial and ethnic differences in mortality at older ages. In "Racial and Ethnic Differences in the Health of Older Americans", Martin L, Soldo B (eds), *Washington: National Academies Press*, 10-42.

Flegal, K., Graubard, B., & Williamson, D. (2004). Methods of calculating deaths attributable to obesity. *Am J Epidemiol, 160*(4), 331-8. https://doi.org/10.1093/aje/kwh222

Flegal, K., Graubard, B., Williamson, D., & Gail, M. (2005). Excess deaths associated with underweight, overweight, and obesity. *J Amer Med Assoc, 293*(15), 1861-7. https://doi.org/10.1001/jama.293.15.1861

Flegal, K., Kit, B., Orpana, H., & Graubard, B. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *J Amer Med Assoc. 309*(1), 71-82. https://doi.org/10.1001/jama.2012.113905

Flegal, K., Panagiotou, O., & Graubard, B. (2015). Estimating population attributable fractions to quantify the health burden of obesity. *Ann Epidemiol. 25*(3), 201-7. https://doi.org/10.1016/j.annepidem.2014.11.010

Freiberg, M., Cabral, H., Heeren, T., Vasan, R., & Ellison, R. (2004). Alcohol consumption and the prevalence of the metabolic syndrome in the US: a cross-sectional analysis of data from the Third National Health and Nutrition Examination Survey. *Diabetes Care, 27*(12), 2954-2959. https://doi.org/10.2337/diacare.27.12.2954

Friedman, M. (1982). Piecewise exponential models for survival data with covariates. *The Annals of Statistics,* 101-113. https://doi.org/10.1214/aos/1176345693

Friedman, G., Tekawa, I., Sadler, M., & Sidney, S. (1997). Smoking and mortality: the Kaiser Permanente experience. In "Changes in Cigarette-Related Disease Risks and Their Implications for Prevention and Control" Rockville, MD, *National Cancer Institute*, 477-499.

Gaziano, J., Gaziano, T., Glynn, R., Sesso, H., Ajani, U., Stampfer, M., Manson, J., Hennekens, C., & Buring, J. (2000). Light-to-moderate alcohol consumption and mortality in the Physicians' Health Study enrollment cohort. *J Am Coll Cardiol. 35*(1), 96-105. https://doi.org/10.1016/S0735-1097(99)00531-8

George, E., & McCulloch, R. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association, 88*, 881-889. https://doi.org/10.1080/01621459.1993.10476353

Goh, L., Dhaliwal, S., Welborn, T., Lee, A., & Della, P. (2014). Anthropometric measurements of general and central obesity and the prediction of cardiovascular disease risk in women: a cross-sectional study. *BMJ Open. 6;4*(2), e004138. https://doi.org/10.1136/bmjopen-2013-004138

Greenland, S., & Drescher, K. (1993). Maximum likelihood estimation of the attributable fraction from logistic models. *Biometrics, 49*(3), 865-72. https://doi.org/10.2307/2532206

Hahn, R., Eaker, E., Barker, N., Teutsch, S., Sosniak, W., & Krieger, N. (1996). Poverty and death in the United States. *Int J Health Serv. 26*(4), 673-90. https://doi.org/10.2190/967K-LC4F-DU66-W5P9

Hainer, V., Toplak, H., & Stich, V. (2009). Fat or fit: what is more important? *Diabetes Care, 32*(suppl 2), S392-S397. https://doi.org/10.2337/dc09-S346

He, Y., Zaslavsky, A., Landrum, M., Harrington, D., & Catalano, P. (2010). Multiple imputation in a large-scale complex survey: a practical guide. *Stat Methods Med Res. 19*(6), 653-70. https://doi.org/10.1177/0962280208101273

Heeringa, S., Berglund, P., West, B., Mellipilán, E., & Portier, K. (2015). Attributable fraction estimation from complex sample survey data. *Ann Epidemiol. 25*(3), 174-8. https://doi.org/10.1016/j.annepidem.2014.11.007

Ho, J., & Elo, I. (2013). The contribution of smoking to black-white differences in U.S. mortality. *Demography, 50*(2), 545–568. https://doi.org/10.1007/s13524-012-0159-z

Hu, F., Willett, W., Li, T., Stampfer, M., Colditz, G., & Manson, J. (2004). Adiposity as compared with physical activity in predicting mortality among women. *N Engl J Med. 351*(26), 2694-703. https://doi.org/10.1056/NEJMoa042135

Hummer, R., & Chinn J. (2011). Race/Ethnicity And U.S. Adult Mortality: Progress, Prospects, and New Analyses. *Du Bois Rev., 8*(1), 5-24. https://doi.org/10.1017/S1742058X11000051

Ibrahim, J., Chen, M-H, D., & Sinha, D. (2001). Bayesian Survival Analysis. New York: Springer Verlag

Kappeler, R., Eichholzer, M., & Rohrmann, S. (2013). Meat consumption and diet quality and mortality in NHANES III. *Eur J Clin Nutr. 67*(6), 598-606. https://doi.org/10.1038/ejcn.2013.59

Katzmarzyk, P., & Staiano, A, (2012). New race and ethnicity standards: elucidating health disparities in diabetes. *BMC Med. Apr*, *30*(10), 42. https://doi.org/10.1186/1741-7015-10-42

Kitahara, C., Flint, A., Berrington de Gonzalez, A., & et al (2014). Association between class III obesity (BMI of 40-59 kg/m2) and mortality: a pooled analysis of 20 prospective studies. *PLoS Med. 11*(7), e1001673. https://doi.org/10.1371/journal.pmed.1001673

Korn, E., Graubard, B., & Midthune, D. (1997). Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *American Journal of Epidemiology, 145,* 72-80. https://doi.org/10.1093/oxfordjournals.aje.a009034

Koster, A., Leitzmann, M., Schatzkin, A., Adams, K., van Eijk, J., Hollenbeck, A., & Harris, T. (2008). The combined relations of adiposity and smoking on mortality. *Am J Clin Nutr. 88*(5), 1206-12.

Little, T., Jorgensen, T., Lang, K., & Moore, E. (2014). On the joys of missing data. *J Pediatr Psychol. 39*(2), 151-62. https://doi.org/10.1093/jpepsy/jst048

McCullough, M., Feskanich, D., Rimm, E., & et al. (2000). Adherence to the Dietary Guidelines for Americans and risk of major chronic disease in men. *Amer J Clinical Nutrition, 72*(5), 1223-1231.

Manson, J., Bassuk, S., Hu, F., Stampfer, M., Colditz, G., & Willett, W. (2007). Estimating the number of deaths due to obesity: can the divergent findings be reconciled? *J Womens Health., 16*(2),168-76. https://doi.org/10.1089/jwh.2006.0080

Marmot, M. (2002) The influence of income on health: views of an epidemiologist. *Health Affairs. 21*(2), 31-46. https://doi.org/10.1377/hlthaff.21.2.31

Marshall, A., Altman, D., Holder, R., & Royston, P. (2009). Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Medical Research Methodology, 9,* 57. https://doi.org/10.1186/1471-2288-9-57

Mason, C., & Tu, S. (2008) Partitioning the population attributable fraction for a sequential chain of effects. *Epidemiol Perspect Innov., Oct 2*(5), 5. https://doi.org/10.1186/1742-5573-5-5

Mehta, N., & Chang, V. (2009). Mortality attributable to obesity among middle-aged adults in the United States. *Demography, 46*(4), 851-872. https://doi.org/10.1353/dem.0.0077

Meng, X-L. (9994). Posterior Predictive p-Values. *The Annals of Statistics, 22,*1142-1160 https://doi.org/10.1214/aos/1176325622

Morgenstern, H. (2008). Attributable fractions. In Boslaugh S (Ed.) Encyclopedia of Epidemiology. Thousand Oaks, CA: Sage Publications, 55-63.

Ogden, C., Carroll, M., Kit, B., & Flegal, K. (2013). Prevalence of obesity among adults: United States, 2011-2012. NCHS data brief, no. 131. Hyattsville, MD: National Center for Health Statistics.

Piironen, J., & Vehtari, A. (2016) Comparison of Bayesian predictive methods for model selection, Statistics and Computing (online pre-publication). https://doi.org/10.1007/s11222-016-9649-y

Pischon, T., Boeing, H., Hoffmann, K., & et al. (2008) General and abdominal adiposity and risk of death in Europe. *N

*Engl J Med. 359(*20), 2105-20. https://doi.org/10.1056/NEJMoa0801891

Pollack, C., Chideya, S., Cubbin, C., Williams, B., Dekker, M., & Braveman, P. (2007) Should health studies measure wealth? A systematic review. *American Journal of Preventive Medicine, 33*, 250–64. https://doi.org/10.1016/j.amepre.2007.04.033

Powell, L., Slater, S., Chaloupka, F., & Harper, D. (2006). Availability of physical activity-related facilities and neighborhood demographic and socioeconomic characteristics: a national study. Am J Public Health. 96(9), 1676-80.

Price, G., Uauy, R., Breeze, E., Bulpitt, C., & Fletcher, A. (2006). Weight, shape, and mortality risk in older persons: elevated waist-hip ratio, not high body mass index, is associated with a greater risk of death. *Am J Clin Nutr., 84*(2), 449-60. https://doi.org/10.2105/AJPH.2005.065573

Raphael, D. (2011). Poverty in Canada: Implications for Health and Quality of Life. Canadian Scholars' Press.

Richman, E., & Stampfer, M. (2010). Editorial comment: weight loss and mortality in the elderly: Separating cause and effect. *Journal of Internal Medicine, 268*(2), 103–105. https://doi.org/10.1111/j.1365-2796.2010.02227.x

Rimm, E., Giovannucci, E., Willett, W., Colditz, G., Ascherio, A., & Rosner, B. Stampfer M Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet. 338*(8765), 464-8.

Rockhill, B., Newman, B., & Weinberg, C. (1998). Use and misuse of population attributable fractions. *American Journal of Public Health, 88*(1), 15-19. https://doi.org/10.2105/AJPH.88.1.15

Rubin, D. (1987). Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316696

Samuelsen, S., & Eide, G. (2008). Attributable fractions with survival data. *Statistics in Medicine, 27*(9), 1447-1467. https://doi.org/10.1002/sim.3022

Schafer, J. (2001). Analyzing the NHANES III Multiply Imputed Data Set: Methods and Examples. https://wwwn.cdc.gov/nchs/data/nhanes3/7a/midocs.pdf

Snijder, M., van Dam, R., Visser, M., & Seidell, J. (2006). What aspects of body fat are particularly hazardous and how do we measure them? *Int J Epidemiol. 35*(1), 83-92. https://doi.org/10.1093/ije/dyi253

Stolt, P., Bengtsson, C., Nordmark, B., Lindblad, S., Lundberg, I., Klareskog, L., & Alfredsson, L. (2003). Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. *Annals of Rheumatic Diseases, 62*(9), 835-841. https://doi.org/10.1136/ard.62.9.835

Taylor, A., Ebrahim, S., & Ben-Shlomo, Y. (2010). Comparison of the associations of body mass index and measures of central adiposity and fat mass with coronary heart disease, diabetes, and all-cause mortality: a study using data from 4 UK cohorts. *Am J Clin Nutr. 91*(3), 547-56. https://doi.org/10.3945/ajcn.2009.28757

Thomas, A., Best, N., Lunn, D., Arnold, R., & Spiegelhalter, D. (2004). GeoBUGS User Manual. MRC Biostatistics Unit, University of Cambridge.

van Dam, R. M., Li, T., Spiegelman, D., Franco, O. H., & Hu, F. B. (2008). Combined impact of lifestyle factors on mortality: prospective cohort study in US women. *BMJ, 337*, a1440. https://doi.org/10.1136/bmj.a1440

Walls, H., Stevenson, C., Mannan, H., Abdullah, A., Reid, C., McNeil, J., & Peeters, A. (2011). Comparing trends in BMI and waist circumference. *Obesity, 19*(1), 216-9. https://doi.org/10.1038/oby.2010.149

Ward, E., Jemal, A., Cokkinides, V., Singh, G., Cardinez, C., Ghafoor, A., & Thun, M. (2004). Cancer disparities by race/ethnicity and socioeconomic status. *CA: a Cancer Journal for Clinicians, 54*(2), 78-93. https://doi.org/10.3322/canjclin.54.2.78

Appendix 1 Piecewise Exponential Models.

Under piecewise exponential regression (e.g. Friedman, 1982) with predictors Xi, the conditional hazard is

$$h(t_i|X_i) = h_0(t_i) \exp(X_i\beta) = \lambda_j \exp(X_i\beta),$$

for times $t_i \in (a_{j-1}, a_j]$, j=1,...,J, with baseline hazard λ_j constant within each interval $(a_{j-1}, a_j]$, and predictors X_i excluding an intercept. If model checks indicate departure from proportional hazards, one may specify time varying predictor effects via $h(t_i|X_i) = \lambda_j \exp(X_i\beta_j)$. With death indicators δ_i (=1 for death, 0 otherwise), the likelihood contribution in the jth interval is

$$[\lambda_j \exp(X_i\beta)]^{\delta_i} \exp[-\sum_{k=1}^{j} \lambda_k d_{ik} \exp(X_i\beta)],$$

where $d_{ik} = \min(t_i, a_k) - a_{k-1}$ is the time spent in the kth interval for subjects still under observation, and $d_{ik}=0$ otherwise. Letting $z_{ik}=1$ for a subject dying in interval k, and $z_{ik}=0$ otherwise, likelihood contributions are equivalently

$$\prod_{k=1}^{j} [\lambda_k \exp(X_i\beta)]^{z_{ik}} \exp[-\lambda_k \exp(X_i\beta) d_{ik}], \qquad (A1.1)$$

these being proportional to a Poisson likelihood for responses z_{ik} with means $\mu_{ik} = \lambda_k \exp(X_i\beta_k) d_{ik}$, and offsets d_{ik}. The hazard terms $\lambda_k = \log(\beta_{0k})$ may be included in an expanded covariate term $B_{ik} = \exp(\beta_{0k} + X_i\beta_k)$.

Under the piecewise exponential approach, the probability of death between t and t+Δt is

$$r_i(t, t+\Delta t | X_i) = S(t|X_i) - S(t+\Delta t|X_i) = \exp[-\sum_{k=1}^{J} d_k(t) B_{ik}] - \exp[-\sum_{k=1}^{J} d_k(t+\Delta t) B_{ik}]$$

where $d_k(t) = 0$ if $t \le a_{k-1}$, $d_k(t) = (t-a_{k-1})$ if $a_{k-1} < t \le a_k$, and $d_k(t) = (a_k - a_{k-1})$ if $t > a_k$. Attributable risks are obtained using estimated mortality probabilities $r_i(t, t+\Delta t | X_i)$ for observed covariates, X_i, and under the counterfactual intervention, where X_i is replaced by X_i^*.

Appendix 2 Missing Data Imputation

Missing data imputation is carried out across the combined sample (males and females) before the survival regression analysis. Missingness rates over 5% are confined to the healthy eating index (HEI, 11%), waist-hip ratio (WHR, 5%), BMI (8%), and income-poverty ratio (IPR, 9%). Otherwise missingness is low, under 0.25% (1 in 400), for smoking status, alcohol use, and MET level.

Missing data imputation packages in R and STATA, such as the MICE and AMELIA packages, use general purpose methods, using all observed data to predict missing data. Here we use an imputation method taking account of substantive epidemiological linkages. Thus for HEI and IPR (in continuous form), predictors are education level (<12 years, 12 years, some college, completed college), sex and age. For WHR (in continuous form), predictors are BMI, HEI, activity status, and sex. For BMI (in continuous form), predictors are WHR, HEI, activity status, and sex. Bayesian methods are used with a joint likelihood model defined over all variables subject to missingness.

Weighted lognormal regression is used for HEI, WHR and BMI (to ensure imputations are positive). Weights are NHANES III survey weights, scaled to average 1. Normal linear regression is applied to inverse hyperbolic sine transformed values of IPR, rather than a log transformed IPR, as IPR can have zero values (Burbidge et al, 1988). For variables with low missingness, binary or multinomial sampling is used to impute missing values, with weighted forms

of likelihood (Heeringa et al, 2015).

Samples of imputed values are taken at intervals of 100 iterations after MCMC convergence. K=5 imputed datasets are obtained (as in Schafer, 2001). Denote posterior means and variances of the K estimates of a parameter θ as $Q_1,..,Q_K$ (with average $\bar{Q} = \sum_{k=1}^{K} Q_k/K$) and $V_1,..,V_K$ respectively. The within imputation variance of θ is estimated as $\bar{V}= \sum_{k=1}^{K} V_k/K$, the between imputation variance as $B= \sum_{k=1}^{K} (Q_k-\bar{Q})^2/(K-1)$, and the total variance of the pooled estimate $\bar{Q}$ of θ as $T=B(1+1/K)+\bar{V}$. The statistic $(Q-\bar{Q})T^{-0.5}$ is approximately (Rubin, 1987) t-distributed with $\nu=(K-1)[1+\bar{V}/(B+B/K)]$ degrees of freedom, with the ratio $\rho=(B+B/K)/\bar{V}$ called the relative increase in variance due to nonresponse. This defines another summary statistic regarding missingness, the fraction of missing information $\phi=[\rho+2/(\nu+3)]/(\rho+1)$. The relative efficiency of using the finite number K of imputations, rather than using an infinite number, is then approximately $(1+\phi/K)^{-1}$.

Appendix 3 Model Checks

Model checking involves posterior predictive assessment (Meng, 1994; Berkhof et al, 2000), based on actual and predicted deaths in each of the J intervals, and on sum of squares differences between observation and model based Kaplan Meier (K-M) curves.

For the first check, replicates $z_{ik,new}$ are sampled (see equation A1.1), and replicate deaths $m_{k,new}=\sum_i z_{ik,new}$ within intervals k=1,...,J obtained. These are compared with predicted deaths $\mu_k=\sum_i \mu_{ik}$ using a chi-square statistic, X^2_{new}. The analogous statistic X^2_{obs} is also obtained by comparing observed death totals m_k with μ_k. A posterior predictive p-value is estimated by the proportion of MCMC iterations where X^2_{new} exceeds X^2_{obs}. Extreme p-values, namely under 0.05 or over 0.95, indicate model discrepancies. Under the second check, modelled survival probabilities are compared with Kaplan-Meier observation-based estimates $S(t)=\prod_{t \le a_k}(1 - \frac{m_k}{R_k})$, where R_k are numbers at risk at the start of the kth interval. Let $\hat{S}(t)=\prod_{t \le a_k}(1 - \frac{\mu_k}{R_k})$ denote the corresponding model based survival curve, with m_k replaced by μ_k. Then error sum of squares over the J intervals compare (a) S(t) with $\hat{S}(t)$, and (b) $S_{new}(t)$ with $\hat{S}(t)$, where $S_{new}(t)$ is the K-M curve obtained by replacing m_k by $m_{k,new}$. With the respective sum of squares denoted SSKM and $SSKM_{new}$, the posterior predictive p-value is estimated by the proportion of MCMC iterations where $SSKM_{new}$ exceeds SSKM.
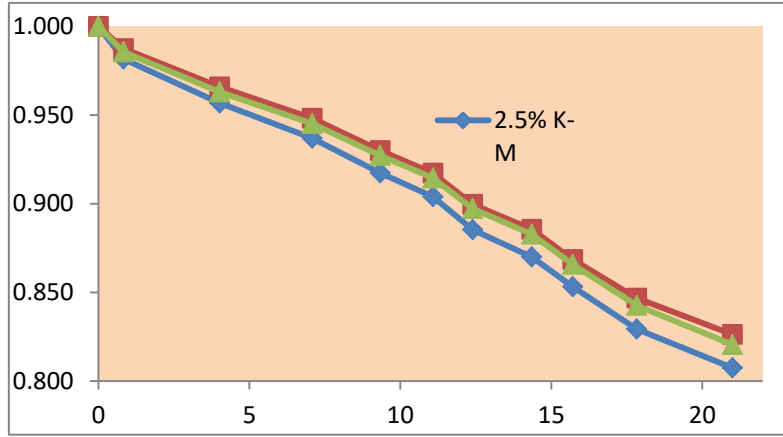
Figure 1.1 Male Survival Curve, Model and 95% Kaplan-Meier Limits
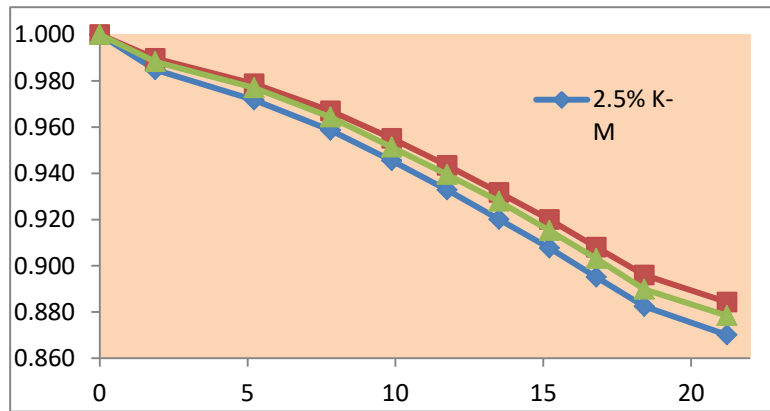

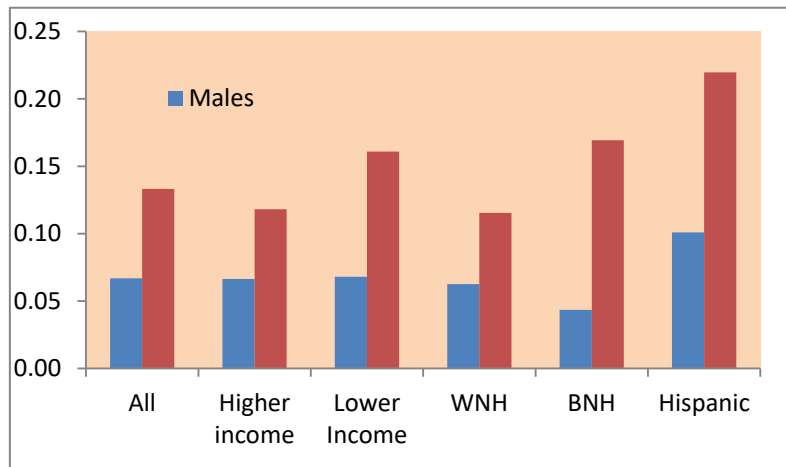Figure 1.2 Female Survival Curve, Model and 95% Kaplan-Meier Limits
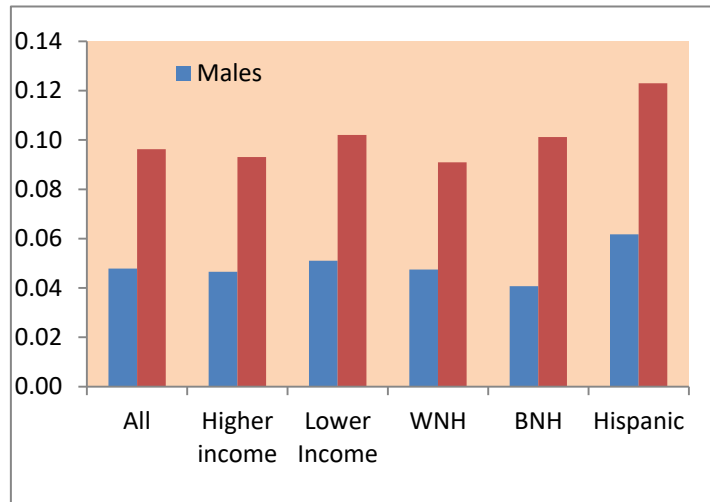

Figure 2.1 Adiposity Attributable Risks
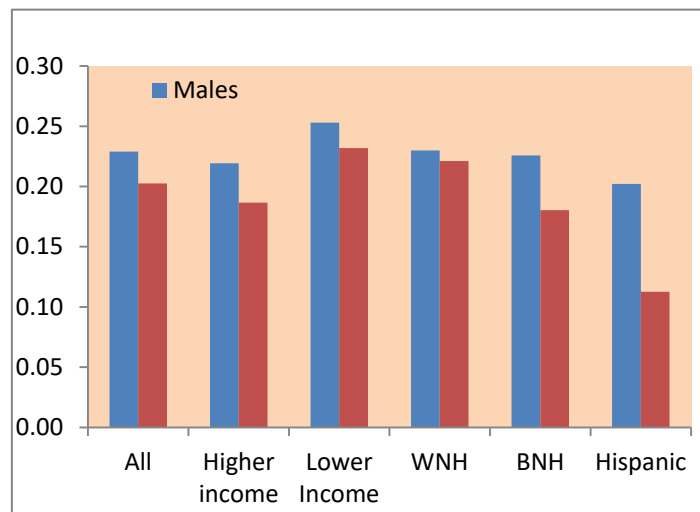
Figure 2.2 Inactivity Attributable Risks

Figure 2.3 Smoking Attributable Risks

Table 1. Posterior Predictive Checks

	Males	Females
Chi-square against predicted deaths by interval	67.3%	69.0%
Sum Squared Errors, K-M plot	69.4%	71.1%

Table 2 All Cause Mortality, Males, Risk Factor Effects and Regression Retention						
		Posterior Mean and 95% Credible Interval			% Relative Efficiency	Inclusion Probability
Risk Factors	Variable or Category	Mean	Lower	Upper		
Adiposity	High waist-hip ratio	1.16	0.86	1.47	88	0.60
	V high waist hip ratio	1.52	1.15	1.89	89	1.00
	High BMI (Class 1)	1.39	1.18	1.60	92	1.00
	Extreme BMI	1.79	1.42	2.16	93	1.00
Activity	Activity Category 2	0.87	0.69	1.06	91	0.67
	Activity Category 3	0.91	0.73	1.08	91	0.51
	Activity Category 4 (highest MET)	0.82	0.63	1.01	92	0.81
Smoking	Former smoker, quit 30+ years	0.92	0.62	1.22	84	0.30
	Former smoker, quit 20-29 years	0.98	0.85	1.11	93	0.18
	Former smoker, quit 10-19 years	1.03	0.93	1.12	98	0.15
	Former smoker, quit 5-9 years	1.64	1.25	2.04	91	1.00
	Former smoker, quit 0-4 years	1.62	1.36	1.88	96	1.00
	Current smoker, < 1 pack	2.03	1.72	2.35	95	1.00
	Current smoker, 1+ pack	2.12	1.82	2.43	93	1.00
Confounders	Healthy Eating Indicator	0.81	0.65	0.98	88	0.89
	Moderate alcohol consumption	0.78	0.70	0.85	97	1.00
	Heavy alcohol consumption	0.99	0.94	1.05	100	0.06
	Poverty	1.52	1.34	1.69	99	1.00
	Black non-Hispanic	1.62	1.43	1.82	98	1.00
	Hispanic	1.01	0.95	1.07	99	0.07
	Other	0.82	0.46	1.18	89	0.60
	Adipose-Current Smoker	1.04	0.89	1.20	96	0.21
	Adipose- age over 55	0.55	0.46	0.64	91	1.00
	Adipose-BNH	0.97	0.85	1.10	97	0.19
	Adipose-Hispanic	1.03	0.97	1.08	100	0.09
	Adipose-Other	1.59	0.44	2.74	90	0.74
	Adipose-Inactive	0.99	0.87	1.11	98	0.16

Table 3 All Cause Mortality, Females, Risk Factor Effects and Regression Retention						
		Posterior Mean, 95% Credible Interval				
Risk Factors		Mean	Lower	Upper	% Relative Efficiency	Inclusion Probability
Adiposity	High waist-hip ratio	1.35	1.16	1.54	97	1.00
	V high waist hip ratio	1.40	1.19	1.62	93	1.00
	High BMI (Class 1)	1.51	1.29	1.73	94	1.00
	Extreme BMI	1.92	1.62	2.22	96	1.00
Activity	Activity Category 2	1.00	0.93	1.08	96	0.09
	Activity Category 3	0.98	0.91	1.06	97	0.11
	Activity Category 4 (highest MET)	0.85	0.71	0.99	95	0.82
Smoking	Former smoker, quit 30+ years	0.85	0.48	1.22	88	0.50
	Former smoker, quit 20-29 years	1.07	0.77	1.37	90	0.25
	Former smoker, quit 10-19 years	1.19	0.82	1.57	92	0.61
	Former smoker, quit 5-9 years	1.54	1.11	1.97	92	1.00
	Former smoker, quit 0-4 years	1.05	0.87	1.23	94	0.23
	Current smoker, < 1 pack	1.80	1.43	2.17	93	1.00
	Current smoker, 1+ pack	2.54	2.06	3.01	89	1.00
Confounders	Healthy Eating Indicator	0.82	0.67	0.98	90	0.87
	Moderate alcohol consumption	0.85	0.72	0.98	98	0.83
	Heavy alcohol consumption	1.01	0.94	1.08	99	0.08
	Poverty	1.39	1.23	1.56	100	1.00
	Black non-Hispanic	1.36	1.17	1.54	98	1.00
	Hispanic	1.00	0.93	1.07	98	0.08
	Other	0.97	0.88	1.07	97	0.14
	Adipose-Current Smoker	1.10	0.86	1.35	91	0.42
	Adipose- age over 55	0.50	0.43	0.57	94	1.00
	Adipose-BNH	0.95	0.81	1.09	96	0.30
	Adipose-Hispanic	0.99	0.89	1.10	96	0.14
	Adipose-Other Ethnicity	1.03	0.80	1.26	92	0.25
	Adipose-Inactive	1.21	0.88	1.54	89	0.70

Table 4 Attributable Fractions by Gender and Population Category					
Males					
		Posterior Mean	**Lower 5%**	**Upper 95%**	**Relative Efficiency**
All Males	Adiposity	0.07	0.05	0.09	93.6
	Inactivity	0.05	0.00	0.10	97.8
	Smoking	0.23	0.19	0.27	89.7
Higher Income Males	Adiposity	0.07	0.05	0.09	94.0
	Inactivity	0.05	0.00	0.10	98.1
	Smoking	0.22	0.18	0.26	89.6
Lower Income Males	Adiposity	0.07	0.05	0.09	92.3
	Inactivity	0.05	0.00	0.10	96.8
	Smoking	0.25	0.21	0.29	90.3
WNH Males	Adiposity	0.06	0.04	0.08	94.0
	Inactivity	0.05	0.00	0.10	98.0
	Smoking	0.23	0.19	0.27	89.3
BNH Males	Adiposity	0.04	0.03	0.06	94.7
	Inactivity	0.04	0.00	0.08	97.3
	Smoking	0.23	0.19	0.26	93.2
Hispanic Males	Adiposity	0.10	0.07	0.13	97.0
	Inactivity	0.06	0.00	0.12	96.7
	Smoking	0.20	0.16	0.24	90.7
Females					
		Posterior Mean	**Lower 5%**	**Upper 95%**	**Relative Efficiency**
All Females	Adiposity	0.13	0.11	0.16	93.8
	Inactivity	0.10	0.04	0.15	97.8
	Smoking	0.20	0.16	0.24	87.1
Higher Income Females	Adiposity	0.12	0.10	0.14	95.8
	Inactivity	0.09	0.03	0.15	97.7
	Smoking	0.19	0.15	0.23	86.8
Lower Income Females	Adiposity	0.16	0.12	0.20	90.8
	Inactivity	0.10	0.04	0.16	97.8
	Smoking	0.23	0.19	0.27	88.0
WNH Females	Adiposity	0.12	0.09	0.14	94.9
	Inactivity	0.09	0.03	0.15	97.7
	Smoking	0.22	0.18	0.26	86.8
BNH Females	Adiposity	0.17	0.11	0.22	93.1
	Inactivity	0.10	0.04	0.16	97.7
	Smoking	0.18	0.15	0.22	89.3
Hispanic Females	Adiposity	0.22	0.17	0.26	93.9
	Inactivity	0.12	0.06	0.19	97.8
	Smoking	0.11	0.08	0.14	89.6