TITLE: How Well Do the Standard Body-Mass Index or Variations With A Different Exponent Predict Human Lifespan?

AUTHORS: Dean Foster, Howard Karloff, Kenneth E. Shirley.

AFFILIATIONS:

For Dean Foster: Amazon, Inc., New York, NY, and Department of Statistics, University of Pennsylvania, Philadelphia, PA.

For Howard Karloff: None.

For Kenneth E. Shirley: None.

KEYWORDS: BMI, Cox proportional hazards model, hazard rate, survival times, personalized medicine

RUNNING TITLE: Improved BMI

CONTACT INFO:

Corresponding author: Howard Karloff, howard@cc.gatech.edu.

Dean Foster: dean@foster.net.

Kenneth E. Shirley: shirley.kenny@gmail.com

WORD COUNT: 3500

FUNDING: No funding sources for any coauthor.

DISCLOSURE: No potential conflicts of interest for any coauthor.

AUTHOR CONTRIBUTIONS: All authors participated in the design of the study, the analysis of the data, and the writing of the manuscript.

What Was Known Before

(1) Despite being a continuous numeric trait, BMI was traditionally "binned" into categories.
(2) The WHO "normal" interval is [18.5, 24.99]. Earlier research suggested that the relative hazard rate for those with BMI's in [25, 29.99] was no higher than that for those in the [18.5, 24.99] bin.
(3) The exponent of 2 on height in the BMI formula was chosen to maximize correlation with body-fat level and minimize correlation with height, but not to best predict death.

Study Importance

By treating BMI as a continuous variable, unlike most existing papers which study BMI, this paper derives two main results:
(1) Personalized optimal BMI's for individuals, with mean value around 26 (which is in the [25, 29.99] range currently labeled "overweight"). For an individual, the "personalized optimal BMI" is the BMI that, based on his or her covariates(race, smoking status, educational status, etc.), is associated with the lowest relative risk of death. By personalizing the recommendations, advice can be tailored to individuals.
(2) Instead of using "mass/height^2", as in the traditional formula for BMI, an exponent close to 1 on height would give better risk-of-mortality predictions. The interesting possibility that the optimal exponent, separately for men and women is 1.0, which would give the formula "mass/height," cannot be excluded.

Abstract

- Objective: The objective is two-fold: (1) To estimate for each individual the BMI which is associated with the lowest risk of death, and (2) to study variants of the BMI formula to determine which gives the best predictions of death.
- Methods: Treating BMI=mass/height^2 as a continuous variable and estimating its interaction effects with several other variables, the authors analyze the NIH-AARP study data set of approximately 566,000 individuals and fit Cox proportional hazards models to the survival times.
- Results: For each individual a "personalized optimal BMI," the BMI for that individual which, according to the model, is associated with the lowest risk of death, is estimated. The average personalized optimal BMI is approximately 26, which is in the current "overweight" category. In fact, mass/height is a better predictor of death on the data set than BMI itself.
- Conclusions: The model suggests that an individual's "optimal" BMI depends on his or her features; "one-size-fits-all" recommendations may be not best.

Introduction

Body-mass index, mass in kilograms divided by the square of height in meters, was created in 1832 by Belgian polymath Adolphe Quetelet, who wrote, "the weight increases as the square of the height" [1]. Known at various times as the "Quetelet Index" [1], "index of weight relative to stature," "index of build," and "weight-height index" [2], BMI was originally intended to model human populations or to measure adiposity, not to predict mortality. The groundbreaking paper by Keys et al. [2] promoted BMI over alternate measures such as weight/height, weight/height^3, and the percentage of average weight for a given height, age, and sex because BMI overall had a lower correlation with height and a higher correlation with a sum of skin-fold thicknesses, a measure of body fat. Indeed, it was Keys et al. [2] who coined the term "body-mass index."

Standard American medical advice, which follows World Health Organization guidelines, categorizes BMI's under 18.5 as underweight; those in the range $[18.5, 24.99]$ as normal or of healthy weight; those in $[25, 29.99]$ as overweight; and those 30 and higher as obese [3]. (See [4] also.) This paper suggests to the contrary that a BMI of 25 should not be considered overweight, but instead close to optimal. That a BMI in $[25, 29.99]$ might be no worse than one in the "normal" range was also suggested by [5, 6, 7, 8, 9, 10, 11] and others.

Several people, e.g., mathematicians Nick Trefethen [12] of Oxford University and Keith Devlin [13] of Stanford University, have questioned the value of BMI as a measure of health. For example, Trefethen [14] criticized the current BMI formula and suggested using an exponent of 2.5 on height instead [12], but with no evidence supporting the choice. Devlin [13] emphasized the fact that BMI was designed as a measure of populations rather than a predictor of mortality.

In this paper human mortality is modeled as a function of various health-related characteristics of individuals, one of which is BMI, which is treated as a continuous variable. By contrast, most recent papers on BMI, e.g., [6, 7, 8, 9], "binned" BMI, treating it as an unordered categorical variable, a notable exception being [10].

Table 1: Variables used

Variable Name	Units or Number of Levels
age at entry	years
height	meters
BMI	kg/m^2
race	7
education	8
smoking status	32
physical activity frequency	7
alcohol consumption	9
self-reported health status	6
marital status	6
diabetes status	2
chronic disease status	2

Chronic disease status is true if and only if the respondent currently had, or had previously been diagnosed with, cancer, heart disease, renal disease, emphysema, or stroke.

Methods

Data Set

This paper uses the NIH-AARP data set [15] of approximately 566,000 members of the American Association for Retired People (AARP) living in California, Louisiana, Florida, Atlanta, North Carolina, New Jersey, Pennsylvania, and Detroit who responded to a diet and health survey in 1996-1997, when they were of age between 50 and 70, and whose status, living or dead, was determined at least 12.9 years later, the last available follow-up date having been December 31, 2009. (The data set was obtained in response to a proposal to the NIH-AARP Diet and Health Study governing committee, which provided anonymized data.)

At the end of the study, for each individual, either the date of his or her death or the fact that he or she was alive at the end, was recorded. Approximately 111,800 respondents died by December 31, 2009. Participants completed a detailed survey by mail giving, among other variables, the variables we used, which are given in Table 1.

Three and a half million people were sent surveys. There were 566,398 respondents. After the data were cleaned by removing respondents with extreme values of height, weight, caloric

Table 2: Data cleaning

Initial number of respondents:	566,398
Respondents removed because height not in [1.4m,2.1m] or weight not in [31.8kg,181.8kg]:	16,987
Respondents removed because daily caloric consumption not in [200,6000]:	3,180
Respondents removed because daily alcohol consumption exceeded 200g:	4,099
Respondents removed because BMI_2 not in [15,50]:	1,455
Respondents removed because of having a chronic illness:	133,178
Final number of respondents ("participants") remaining:	407,499

consumption, alcohol consumption, and BMI, 540,677 survey responses remained. (See Table 2 and the supplemental material.) Next, those respondents who reported currently having a chronic disease (defined as any one of cancer, heart disease, renal disease, emphysema, or stroke) at the time of the survey, were excluded, to avoid any issue of reverse causation, which further removed 133,178 respondents. (See the Discussion for a sensitivity analysis regarding the removal of the chronically ill.) This left us with $n = 407,499$ total respondents. Of these, 235,546 were men (38,425 deaths, or 16.3%), and 171,953 were women (20,820 deaths, or 12.1%). Hereafter "men" and "women" in the data refer to these two samples of non-chronically-ill men and women. Such respondents are called "participants."

Models

A series of Cox proportional hazards regression models were fit to the right-censored times until death for men and women separately, using age as the underlying time variable. (See [16] and [17] for a discussion of how to choose the time variable when fitting Cox proportional hazards models.)

The Cox proportional hazards model [18, 19, 20] is a standard way of estimating one's relative instantaneous risk of death, i.e., *hazard rate*, across individuals. Table 3 summarizes three of the models that were fit. Four other models that were fit for comparison's sake, namely, models M_0, M_1, M_4, and M_5, are discussed in the supplemental material.

Let BMI_α denote the formula "mass/height$^\alpha$", traditional BMI being BMI_2.

Model M_2 models the hazard function for participant i as $h_i(t) = h_0(t) \exp(\mathbf{X}_i \cdot \boldsymbol{\beta})$, where t is the subject's age in years, $h_0(t)$ is the (unknown and, for the purpose of using the Cox model,

Table 3: Summary of models

	Model	df	LL (men)	LL (women)
M_2	All variables	74	-241.6	-138.7
M_3	All interactions ($\alpha = 2.0$)	145	-0.8	-0.2
M_6	All interactions ($\alpha =$ optimal)	145	0.0	0.0

Here, the column "df" reports the model degrees of freedom, i.e., the number of predictors in the model, and "LL" is the log-likelihood of the parameters of the model given the data, with M_6 used as the comparison model.

irrelevant) baseline hazard function, $\mathbf{X}_i$ is the vector of predictor variables for participant i corresponding to the first 11 input variables listed in the "Data Set" section, and $\boldsymbol{\beta}$ is the vector, of the same length as $\mathbf{X}_i$, of unknown coefficients that must be estimated. In other words, the part of the hazard rate that, among people of age t, depends on a person's 11 covariates, is $\exp(\mathbf{X}_i \cdot \boldsymbol{\beta})$, an exponential of the dot product of the vector of the values of that person's 11 covariates with the vector $\boldsymbol{\beta}$ of unknown coefficients to be found. Algorithms for fitting the Cox model find the coefficient vector $\boldsymbol{\beta}$ so that the resulting predicted hazard rates $h_i(t)$, in aggregate over participants, most accurately agree with the actual death rates. (Because some covariates are multidimensional, $\boldsymbol{\beta}$ will have more than 11 entries.) The fit of the model to the data is measured by the natural logarithm of its (partial) likelihood [18, chapter 4]. Higher values of log-likelihood imply a better fit to the data. Increasing the log-likelihood by two or more units when increasing the number of degrees of freedom by one is considered statistically significant.

Additionally, a cubic function of the input variable BMI_2, rather than just a linear function, was used, to allow the model to capture the well-known "J-shaped" relationship between BMI_2 and mortality [10]. Thus, there are three model degrees of freedom associated with BMI_2 in model M_2, and counting $J - 1$ model degrees of freedom for each categorical input variable with J levels, and one degree of freedom for each continuous or binary input variable in the model, model M_2 contains 74 model degrees of freedom. That is, $\boldsymbol{\beta}$ for this model has 74 components. The maximum likelihood estimates of $\boldsymbol{\beta}$ for model M_2, for both men and women, are available in the supplemental material.

Model M_3 includes all of the predictor variables in model M_2, as well as the two-way interaction

effects between all pairs of input variables, and quadratic terms for all variables except for diabetes (which is binary). To fit such a model using the existing categorical variables would have resulted in a huge number of parameters in the model—$((7-1)\times(32-1)=186$ variables, for example, for the interaction between the 7-level input variable "physical activity frequency" and the 32-level input variable "smoking"). To alleviate this problem, the authors "tied together" the parameters for the interactions to reduce the number of resulting model degrees of freedom; see the supplemental material for details. The effects of each of the seven categorical variables in model M_2 (race, education, smoking status, physical activity frequency, alcohol consumption, self-reported health status, and marital status) were "tied together," and two-way interaction effects between all pairs were estimated.

Model M_6 is the same as model M_3, except that the exponent α on height in the denominator of the definition of BMI_α is chosen (separately for men and women) to be the number in $\{0.1, 0.2, 0.3, ..., 3.0\}$ which maximizes the log-likelihood. See the "A Better BMI Formula" section.

Results

Personalized Optimal BMI

The value of BMI_2 that minimizes one's relative risk under model M_3, holding all other variables constant, is defined as his or her *personalized optimal BMI*$_2$, or POB_2. Hence one's minimum relative risk depends on the values of his or her other demographic and health-related variables (such as race, education, smoking status, etc.). Since under model M_3 the logarithm of one's relative risk $h_i(t)$ of death is a cubic function of his or her BMI_2, it is simple to compute one's POB_2 using differentiation and the quadratic formula. (For technical reasons, POB_2 is not defined for approximately 0.1% of the population. See the supplemental material for details.) A confidence interval for one's POB_2, defined as the interval of values around the POB_2 for which a hypothesis test that the derivative of the relative risk curve at that BMI_2 value was equal to zero was not rejected, was also computed. See the supplemental material for details.

The authors plan to make a calculator for POB_2's and their associated confidence intervals

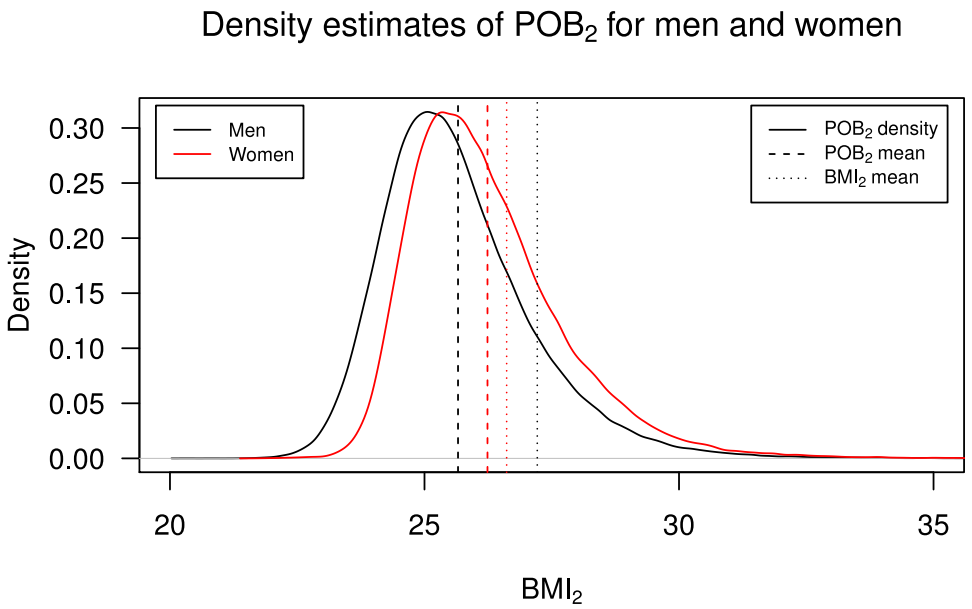

Figure 1: Density estimates of the POB_2's for (non-chronically ill) men and women in the study. The sample mean and standard deviation were 25.7 (1.51) among men ($n = 235,546$), and 26.3 (1.54) among women ($n = 171,953$). For comparison, note that the mean observed BMI values for these groups are 27.2 (men) and 26.6 (women).

available on the web.

The mean POB_2 for men is 25.7, whereas for women it is 26.3. The distributions of POB_2 are shown in Figure 1. The standard deviation for an individual man's POB_2 is 1.51 and that for an individual woman's POB_2 is 1.54. (By contrast, the average BMI_2's in the study for men and women are, respectively, 27.2 and 26.6, with standard deviations of, respectively, 3.99 and 5.30.) Approximately 60% of men's and 44% of women's POB_2 confidence intervals did not contain the mean POB_2 value for their respective sex, indicating that a single recommended BMI_2 value for each sex may not be appropriate. Figure 2 shows histograms of the differences between participants' BMI_2's and POB_2's.

To check the fit of model M_3, both men and women were stratified into subsets based on their POB_2 and their observed BMI_2, and then, in Figure 3, the death rate vs. observed BMI_2 for each subset was plotted. Each point represents about 4700 (men) or 3400 (women) participants. Quadratic curves were drawn through each set of death rates corresponding to a POB_2 group. The figure illustrates that the lowest death rate does, indeed, occur for participants whose observed

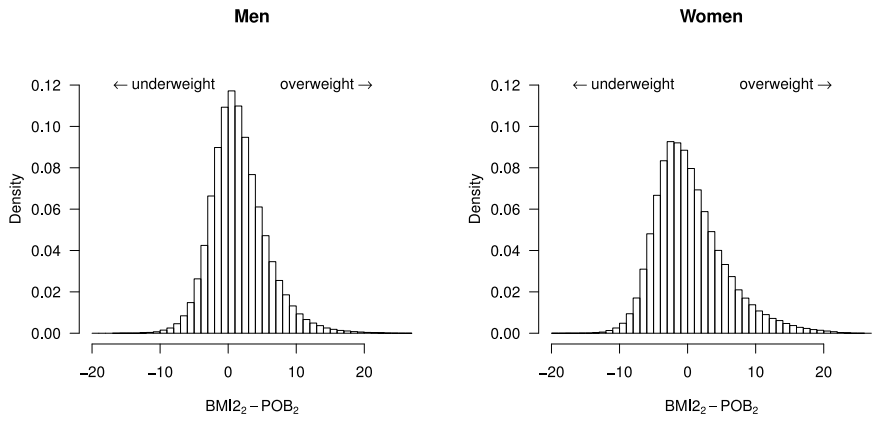

Figure 2: For each sex, the distribution of the difference between a participant's BMI_2 and his or her POB_2 is shown. By inspecting the absolute values of the differences, one finds that men are, on average, 3.2 units away from their POB_2's, while women are, on average, 3.9 units away.

BMI_2 is near their estimated POB_2 (for example, the nadirs of the orange curves occur at larger values of BMI_2 than do those of the blue or red curves, for both sexes).

It is important to note that one's POB_2 is the minimum point of one's relative risk curve when *only* BMI_2 (effectively, *weight) is changed*. Should one change one's BMI_2 by changing other covariates (e.g., by exercising more), that person's POB_2 will also change. From Figure 3, one sees that having a lower POB_2, *whatever one's actual* BMI_2, generally gives an individual a lower relative risk. To the extent that the model is causal, if at all, changing covariates other than BMI_2 so as to lower one's POB_2 may be more important than adjusting one's weight to match his or her POB_2.

A Better BMI Formula

Quetelet's claim that "the weight increases as the square of the height" is supported by the data set, in which the correlation between height and BMI is close to zero: it is -0.01 for men and -0.07 for women.

The goal of this section is a "BMI-like" formula that best predicts risk of mortality (rather than one that best models human populations). Specifically, the authors investigated whether using a different exponent α in the BMI formula could significantly improve the fit of an otherwise identical model that uses BMI_2. Naively, one might think that mass should grow as the cube of

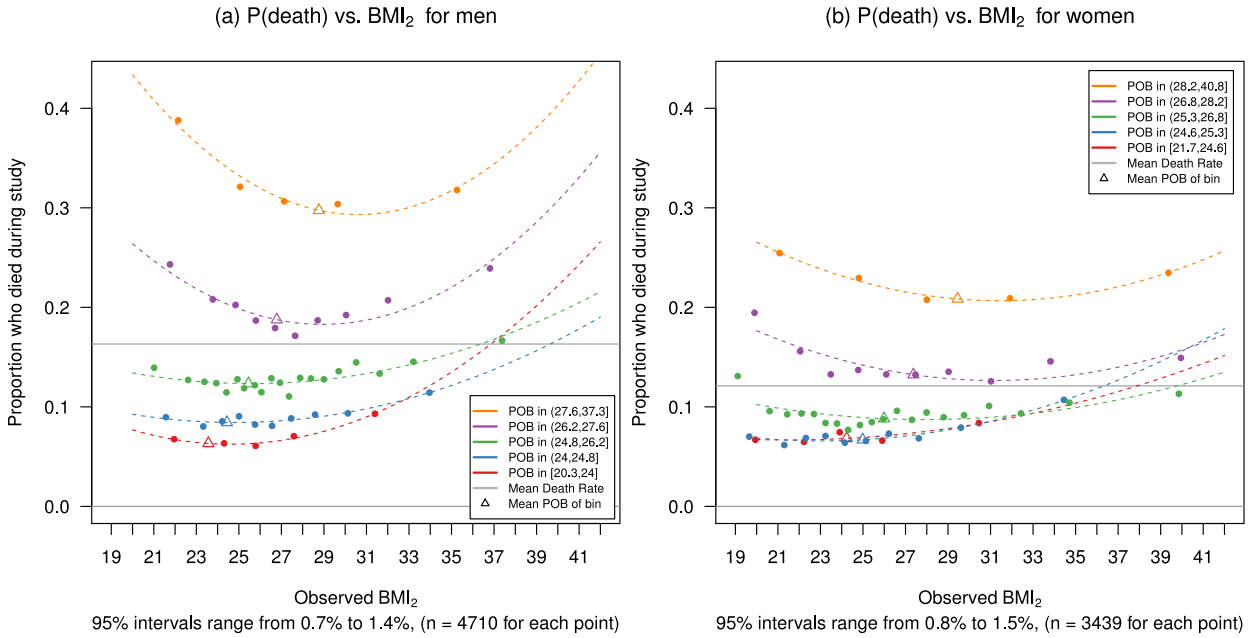

Figure 3: For each sex, the set of participants of that sex were divided into five groups based on their POB_2: lowest decile, 10th-30th percentile, 30th-70th percentile, 70th-90th percentile, and top decile. Next, for each sex, each of these five groups was subdivided into 5, 10, 20, 10, and 5 subgroups based on the participants' observed BMI_2 (using unequal splits at the group level to ensure equal subgroup sizes), and death rates within these subgroups were plotted. The nadirs of the curves do, in fact, occur near the mean POB_2 for each group.

height, volumes being cubic in length, but such a conclusion incorrectly assumes that a six-foot person is a scaled-up version of a four-foot person.

To find the best exponent α on height, first, a model, denoted M_4, identical to M_3 with the exception that it omits the interaction effect between BMI_2 and height, was fit. Then, for a grid of 30 values of $\alpha \in \{0.1, 0.2, 0.3, ..., 2.9, 3.0\}$, model M_4 was refit and the log-likelihood was recorded. Figure 4 shows the relationship between α (the exponent) and the log-likelihood of the resulting models. The optimal values of α for men and women were approximately 1.1 and 1.3, respectively, and they increased the log-likelihood of the model over using the traditional BMI_2 by 17.6 and 4.4 points, respectively, both of which are statistically significant margins ($p < 10^{-8}$ for men, and $p < 0.01$ for women). More precisely, for men, the log-likelihood is maximized at $\alpha = 1.1$, with a range of two units of log-likelihood (i.e., the confidence interval) extending from approximately 0.8 to 1.3. For women, the log-likelihood is maximized at approximately 1.3, with the range of two units of log-likelihood extending from 1.0 to 1.7; see Figure 4. It seems surprising that the optimal α is demonstrably less than 2. Note that $\alpha = 1.0$ is within the confidence interval for both men and women, which leads to the simple formula $\text{BMI}_1 = \text{weight/height}$.

Dependence of POB on Height

Interestingly, the POB_2 values for men and women are not independent of height. More specifically, the optimal BMI_2 for a given individual depends on the individual's height, and the optimal BMI_2 is higher for short people than for tall people (and the difference between optimal BMI_2 values for short and tall people is larger for men than for women).

When using $\alpha = 1.1$ for men and $\alpha = 1.3$ for women, however, the POB_α (defined as the BMI_α associated with the lowest risk of death for a given individual) no longer depends on height ($p = 0.67$ and $p = 0.98$ for men and women, respectively). Figure 5 illustrates the dependence, among women, of one's POB_2 on her height, and the lack of such dependence when using $\text{POB}_{1.3}$ instead.

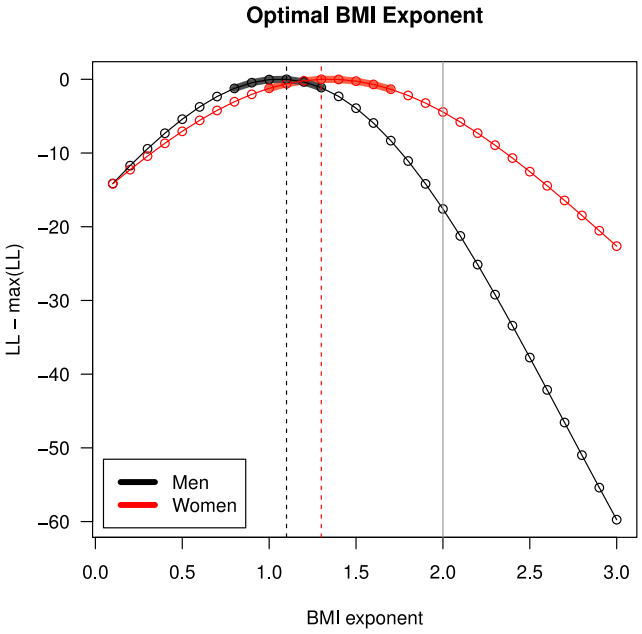

Figure 4: The relationship between the log-likelihood and the exponent of height in the BMI formula for men and women. The dotted lines indicate the values of α where the maximum likelihood was achieved for each sex, and the shaded regions of the curves indicate the set of values of α for which the likelihood was within 2 units of the maximum (i.e., the shaded regions can be thought of as confidence intervals for the optimal α values).

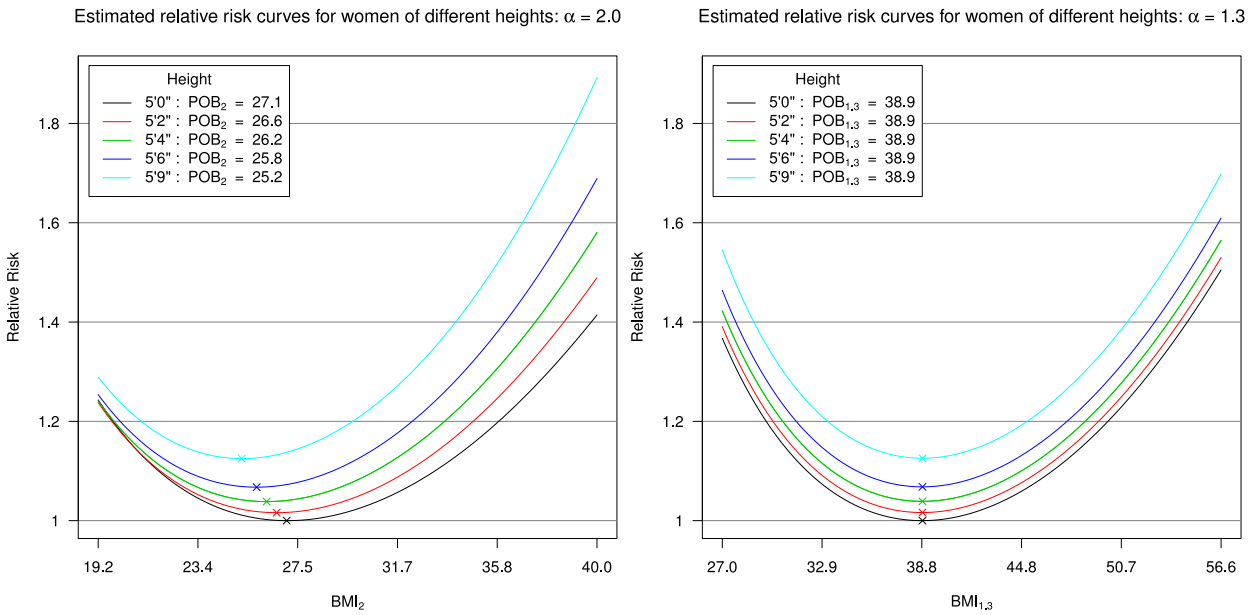

Figure 5: Under model M_3 (see the "Models" section), pictured on the left, the interaction effect between BMI_2 and height among women is significant, which means that a woman's relative risk curve as a function of her BMI_2 has a different minimum depending on her height. Under model M_6 (see Section "A Better BMI Formula"), pictured on the right, by using $\text{BMI}_{1.3} = \text{mass}/\text{height}^{1.3}$, the interaction effect between $\text{BMI}_{1.3}$ and height disappears, and a woman's minimum relative risk is a function solely of her $\text{BMI}_{1.3}$, rather than her height as well. The five curves in each plot correspond to the 2.5th, 25th, 50th, 75th, and 97.5th quantiles of women's height.

Discussion

Summary of Results

It was shown that treating BMI as a continuous variable and computing interaction effects between variables leads to a better estimate of the relationship between BMI and mortality. This paper's first result was a definition of POB_2 as the value of BMI_2 associated, in the model, with the lowest risk of death. The average POB_2's for men and women are in the "overweight" range. The second result was that among all positive α's, the BMI_α's which give the best predictions of mortality have α approximately 1.1 for men and 1.3 for women. One could not exclude the interesting possibility that the best α was 1.0 for both men and women, which would give rise to the simple formula $\mathrm{BMI}_1 = \mathrm{mass/height}$. That $\alpha = 2$ is the optimal α for either sex was rejected. The third result was that POB_2 is not independent of height, but that when using $\alpha = 1.1$ for men and $\alpha = 1.3$ for women, the POB_α no longer depends on height.

Practical Value of the Results

In the presence of all the covariates and interactions between them, whether one uses an exponent of 2 or the optimal exponent for one's sex makes no statistically significant difference.

One policy-related question is "How much longer, on average, could one expect to live if he or she changed his or her BMI_2 from its current value to the nearer endpoint of the WHO-recommended BMI_2 interval of [18.5, 24.99], if it is not already in this interval?" Then, for comparison, "How much longer, on average, could one expect to live by changing his or her BMI_2 to the nearer endpoint of his or her POB_2 confidence interval, if it is not already in this interval?" In other words, would it be beneficial, on average, to replace the current, WHO-recommended universal BMI_2 interval of [18.5, 24.99] by the personalized intervals computed with the model?

Because the data are observational, and no causal effects were estimated in the model, these questions cannot be answered directly. Proving causality would require doing, e.g., a study based on either traditional randomization, "Mendelian randomization" [21, 22], instrumental variables, or propensity score matching. Instead, for each person, the estimated hazard rate (according to

model M_3) at his or her current BMI_2 is compared to the estimated hazard rate of a person identical to him or her with the exception that the second person's BMI_2 is either (1) equal to the nearer endpoint of the [18.5, 24.99] interval (if it is outside the interval), or (2) equal to the nearer endpoint of the first person's POB_2 confidence interval (if it is outside the interval). In both cases, if the first person's BMI_2 lies within the interval, then the second person's BMI_2 is set equal to the first person's BMI_2. Then, assuming a Gompertz model [23] for the baseline hazard rate with a log-slope of 0.108 (men) or 0.115 (women), and a log-intercept of -12.1 (men) or -12.8 (women), where both the log-slopes and log-intercepts were estimated from the data in this study, the difference in mean residual life (using [24, equation (2.7)]) for each type of pair of individuals described above is computed.

Over all hypothetical pairs of men in the study, the mean difference in mean residual life between the man whose BMI_2 was set to the nearer endpoint of the WHO interval and the man with BMI_2 unchanged from the study is 4.6 months ($\sigma = 10.6$ months). When the second man in the pair had a BMI_2 set equal to the nearer endpoint of the POB_2 interval of the first man, however, the mean difference in mean residual life was 6.3 months ($\sigma = 10.6$ months).

For women, the corresponding means are 3.5 months ($\sigma = 9.8$ months) for the WHO interval and 6.6 months ($\sigma = 10.2$ months) for the POB_2 interval.

All averages here are over *all* participants for each sex. Note the large standard deviations for both sexes.

The increased mean residual lifetimes (of roughly 2-3 months) estimated for men and women whose BMI_2 values were moved in accordance with the POB_2 interval of their counterpart, rather than the WHO interval, suggest that a recommended universal BMI_2 is not optimal, and that it could be beneficial (especially in light of the high standard deviations) to tailor one's recommendation to individuals, for whom the differences could be much more pronounced, using a model such as ours.

Effect of Removing the Chronically Ill

The analysis was rerun without excluding the 133,178 chronically ill respondents, to see whether their removal had a large effect on the results. Let "new POB_2" refer to the POB_2 calculated based on the population including the 133,178 chronically ill and let "old POB_2" refer to the original POB_2. Among the non-chronically ill, the correlation between the new POB_2 and the old POB_2 is 0.96 and the standard deviation of the difference is 0.61. The new POB_2's on average (over the non-chronically ill) are 0.88 units higher than the old POB_2's.

The optimal exponent in the BMI formula, previously 1.1 (with a confidence interval of [0.8,1.3]) for men and 1.3 (with a confidence interval of [1.0,1.7]) for women, changes to 1.0 (with a confidence interval of [0.8,1.1]) for men and 1.4 (with a confidence interval of [1.1,1.8]) for women.

Acknowledgments

The authors gratefully acknowledge Lee Newberg for help in the early phase of the work, Prabhakar Chalise for help in fitting Cox models with R, to the referee for his or her thoughtful comments, and Alejandro Schäffer for tremendous help revising drafts of this paper in order to present our results more clearly.

References

[1] Eknoyan G (2008). Adolphe Quetelet (1796-1874)—the average man and indices of obesity. Nephrol Dial Transplant 23:47-51. Also available at `http://ndt.oxfordjournals.org/content/23/1/47.full.pdf+html`; accessed July 6, 2015.

[2] Keys A, Fidanza F, Karvonen M J, Kimura N, Taylor H L (1972). Indices of relative weight and obesity. J Chron Dis 25:329-343. Reprinted in 2014 Int J Epidemiol 43(3): 655-665. Also available from `http://ije.oxfordjournals.org/content/43/3/655.full.pdf+html`; accessed July 6, 2015.

[3] Centers for Disease Control (2012). Defining overweight and obesity. Available from `http://www.cdc.gov/obesity/adult/defining.html`; accessed May 21, 2015.

[4] Berrington de Gonzalez A, Hartge P, Cerhan J, et al. (2010). Body-mass index and mortality among 1.46 million white adults. N Engl J Med 363:2211-2219.

[5] Adams K, Schatzkin A, Harris T, et al. (2006). Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. N Engl J Med 355:763-778.

[6] 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. JAMA 309:71-82.

[7] McGee D L, Diverse Populations Collaboration (2005). Body-mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies. Ann Epidemiol 15:87-97.

[8] Janssen I, Mark A E (2006). Elevated body-mass index and mortality risk in the elderly. Obes Rev 8:41-59.

[9] Flegal K, Graubard B, Williamson D, Gail M (2005). Excess deaths associated with underweight, overweight, and obesity. JAMA 293:1861-1867.

[10] Gronniger J T (2006). A semiparametric analysis of the relationship of body-mass index to mortality. Am J Public Health 96:173-178.

[11] Orpana H, Berthelot J-M, Kaplan M S, Feeny D H, McFarland B, Ross N A (2009). BMI and mortality: results from a national longitudinal study of Canadian adults. Obesity 18:214-218.

[12] Trefethen N. BMI (body mass index). Available at `http://people.maths.ox.ac.uk/trefethen/bmi.html`; accessed May 21, 2015.

[13] Devlin K (2009). Do you believe in fairies, unicorns, or the BMI? Available at `https://www.maa.org/external_archive/devlin/devlin_05_09.html`; accessed May 21, 2015.

[14] Trefethen N (2013). Letter to the editor. The Economist.

[15] US National Institutes of Health. Resources for researchers. Available from `http://dietandhealth.cancer.gov/resource`; accessed May 21, 2015.

[16] Thiébaut A, Bénichou J (2004). Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. Stat Med 23:3803-3820.

[17] Chalise P, Chicken E, McGee D (2009). Time scales in epidemiological analysis: an empirical comparison. Available at `http://arxiv.org/abs/1502.02534`; accessed May 26, 2015.

[18] Kalbfleisch J D, Prentice R L (2002). The statistical analysis of failure time data (second edition). Wiley, Hoboken, NJ.

[19] Wikipedia (2014). Proportional hazards model. Available from `http://en.wikipedia.org/wiki/Proportional_hazards_model`; accessed May 21, 2015.

[20] Collett D (1994). Modelling survival data in medical research. Chapman and Hall, New York.

[21] Nordestgaard B G, Palmer T M, Benn M, Zacho J, Tybjaerg-Hansen A, Smith G D, Timpson N J (2012). The Effect of Elevated Body Mass Index on Ischemic Heart Disease Risk: Causal Estimates from a Mendelian Randomisation Approach. PLOS Med 10.1371/journal.pmed.1001212. Available at `http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001212`; accessed July 21, 2015.

[22] Smith G D, Ebrahim S (2013). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 32:1-22. Also available at `http://ije.oxfordjournals.org/content/32/1/1.full.pdf+html`; accessed July 21, 2015.

[23] Gompertz B (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos Trans R Soc Lond 115:513-583.

[24] Poynor, V A (2010). Bayesian inference for mean residual life functions in survival analysis. Available at `https://users.soe.ucsc.edu/~vpoynor/MastersRevised.pdf`; accessed May 21, 2015.