From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

Doctoral Dissertations, 2011-2026 Population Studies

2013

Body Weight, Weight Change, and Mortality Risk

Blue, Laura

Princeton, NJ : Princeton University

http://arks.princeton.edu/ark:/88435/dsp01j6731390g

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

BODY WEIGHT, WEIGHT CHANGE, AND MORTALITY RISK

Laura Blue

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE PROGRAM IN

POPULATION STUDIES

Adviser: Noreen Goldman

November 2013

© Copyright by Laura Blue, 2013. All rights reserved.

ABSTRACT

Many studies find that overweight people have lower mortality than people with clinically normal body weight, despite higher rates of disease onset. One popular hypothesis to explain this is “reverse causation”: the idea that while weight affects health, health also affects weight. Specifically, the hypothesis states that death risk typically increases with body weight, all else being equal, but that sickness-induced weight loss creates a spurious association between low (though clinically normal) body weight and high mortality. The hypothesis remains largely untested, however. This dissertation assesses its plausibility.

Chapter 1 tests whether past weight loss can account for elevated death rates among people with clinically normal body mass index (BMI). Data come from five longitudinal surveys, each based in a different country: Australia, Costa Rica, England, Taiwan, and the USA. The study uses logistic regression with *P*-splines and finds that, in all five surveys, lowest mortality occurs among the overweight, adjusting for age (range: 50-79), age squared, sex, and smoking. However, people losing weight have higher mortality than weight-gainers or weight-maintainers. Adjusting the BMI–mortality curve for weight change thus attenuates the overweight longevity advantage. This provides modest support for the reverse causation hypothesis.

Chapter 2 explores possible reasons for high mortality among weight-losers, testing whether diagnosed disease predicts subsequent or concurrent weight loss in two population-based surveys: one in Costa Rica and one in England. The study uses logistic regression to estimate associations of two weight loss outcomes with six (self-reported) diagnoses – cancer, diabetes, heart attack, stroke, arthritis, and lung disease – adjusting for age (range: 52-79), sex, smoking, and initial BMI. Associations between disease and weight loss are more apparent in England than in Costa Rica. This suggests the impact of reverse causation could vary across populations.

Finally, Chapter 3 uses simulation to demonstrate that any of three competing hypotheses could explain a longevity advantage among the overweight. Where death risk increases with BMI, even modest sickness-induced weight loss is sufficient to produce the phenomenon. However, the simplest models of reverse causation are still surprisingly complex. Together, the three chapters provide tentative support for the reverse causation hypothesis.

TABLE OF CONTENTS

Abstract iii

Acknowledgments vii

List of tables xii

List of figures xiv

Introduction 1

1. Body weight, weight change, and mortality risk: Evidence from five countries 19

2. Disease and weight loss: A prospective study of middle-aged and older adults in Costa Rica and in England *[Joint work with Noreen Goldman and Luis Rosero-Bixby]* 49

3. Nutritional reserve or reverse causation? Comparing rival

hypotheses of the curvilinear relationship between body mass index and mortality 75

Conclusion 113

Appendices

1A 122

1B 128

1C 137

1D 146

2A 149

2B 158

2C 163

3A 164

ACKNOWLEDGMENTS

I am grateful to many people who have advised and encouraged me while at Princeton. First and foremost, I would like to thank my dissertation committee. To Noreen Goldman, my committee chair: Thank you for three years of meticulous reading and thoughtful critique, while still trusting me to find my own path. You have been an incredible adviser and a great source of support. To Scott Lynch: Thank you for your formidable statistical expertise and, equally, your conceptual guidance, both on this dissertation and during my general-exam preparation. To Germán Rodríguez: Thank you for insightful comments on methodology and my written work (you were always right), and, especially, for writing the custom Stata code to implement *P*-splines. I cannot imagine a more helpful, active, and congenial committee than you three.

I owe special thanks, in addition, to a number of other Princeton faculty: Marta Tienda, director of graduate studies, for wise counsel on conference presentations and on course selection, and for encouraging research travel; Georges Reniers, for helping to supervise my general exams; James Trussell, for meeting with me six years ago, in London, to convince me to come to Princeton in the first place; and, last but by no means least, Thomas Espenshade, for being a Reader on this dissertation, and for being a tremendous mentor, co-author, and friend over the past few years. I am grateful to you all.

I would like to thank Luis Rosero-Bixby at the University of California, Berkeley, as a co-author on Chapter 2 of this dissertation, and principal investigator of the Costa Rican Longevity and Healthy Aging Study (CRELES): Your advice on methodology and paper organization has been invaluable, and I am thrilled we were able to work together. Thank you for making the Costa Rican component of this study possible.

I am continually impressed by, and grateful to, the phenomenal library, IT, and administrative staff at Princeton's Office of Population Research. To Wayne Appleton, Elana Broch, Nancy Cannuli, Mary Lou Delaney, Joann Donatiello, Val Fitzpatrick, Jennifer Flath, Joyce Lopuh, Robin Pispecky, and Diana Sacke: Thank you for your tireless work. You make everything look effortless. I am especially indebted, however, to Lynne Johnson, the graduate student administrator: Thank you for handling my countless queries and concerns over the past few years, and for always finding the perfect solution.

To my fellow students: I have learned more from you than I can say. You are too numerous to call out by name, but if we ever chatted by the water cooler or studied together for an exam; if you sat through a practice talk or job interview; if we played a rock show together at the Ivy, or you took me cross-country skiing in New Hampshire when I thought I wanted to go to Mexico: thank you. You not only made my education richer, you did something I once feared impossible. You made Princeton cool.

To my boyfriend, Tom Essinger-Hileman: You have made me the happiest demographer I could be. I love you!

And, finally, to my family: To my parents, thank you for teaching me that questions are healthy, even if it meant no family dinner could pass without someone consulting the dictionary or an atlas; and to my brother, Pieter, for every conversation that began, “Don’t tell the parents, but I’m not sure how to put this back together,” or some similar gem. Together, you taught me that formal credentials are valuable, but that curiosity is its own reward.

***

Data for this dissertation come from five longitudinal surveys. The Australian Longitudinal Study of Ageing (ALSA) was supported by grants from the South Australian Health Commission, the Australian Rotary Health Research Fund, the U.S. National Institute on Aging (grant #AG 08523–02), the Office for the Ageing in South Australia, Elderly Citizens Homes in South Australia, Australia’s National Health and Medical Research Council (NH&MRC – ID:22922), the Premiers Science Research Fund in South Australia, and the Australian Research Council (ID:DP0879152; DP130100428). I gratefully acknowledge the work of the project team at the Flinders Centre for Ageing Studies, Flinders University, who carried out the ALSA and provided data for this dissertation.

The Costa Rican Longevity and Healthy Aging Study (CRELES) was conducted by the University of Costa Rica's Centro Centroamericano de Población in collaboration with the Instituto de Investigaciones en Salud, with funding support from the Wellcome Trust (grant #072406). Data for the 2005 wave are available through the Inter-university Consortium for Political and Social Research, at http://dx.doi.org/10.3886/ICPSR26681.v2. I am very grateful to principal investigator Luis Rosero-Bixby for providing data access to subsequent CRELES waves.

The English Longitudinal Study of Ageing was developed by a team of researchers based at Britain's National Centre for Social Research, University College London, and the Institute for Fiscal Studies. It was funded by the U.S. National Institute on Aging and a consortium of U.K. government departments. Data are available through the UK Data Service, at http://dx.doi.org/10.5255/UKDA-SN-5050-1.

The Social Environment and Biomarkers of Aging Study (SEBAS) in Taiwan is an extension of the Taiwan Longitudinal Study of Aging. It was funded by the U.S. National Institute on Aging (grant #R01 AG16790 and #R01 AG16661), the Taiwan Bureau of Health Promotion, the National Health Research Institutes in Taiwan (grant #DD01-861x-GR601S), and the Taiwan Provincial Government. I gratefully acknowledge the hard work and dedication of the staff at the Center for Population and Health Survey Research (BHP), who were instrumental in the design and implementation of the SEBAS and supervised all aspects of the fieldwork and data processing. Data from the SEBAS waves in 2000 and in 2006 are available through the Inter-university Consortium for Political and Social Research, at http://dx.doi.org/10.3886/ICPSR03792.v5. I am

grateful to Noreen Goldman, Maxine Weinstein, and Dana Glei for providing access to additional merged data files and to mortality data through 2010.

The U.S. National Health and Nutrition Examination Survey (NHANES) I Epidemiologic Follow-Up Study (NHEFS) was jointly initiated by the U.S. National Center for Health Statistics and the National Institute on Aging, in collaboration with other agencies of the U.S. Public Health Service. Data up to and including the year 1992 are publicly available online from the U.S. National Center for Health Statistics.

I gratefully acknowledge the time and effort invested in these five surveys by their participants, data collectors, and administrators. My work would not be possible without it. Additional support for this research comes from Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant #5R24HD047879), the National Institutes of Health (grant #5T32HD007163), and the National Science Foundation (grant #DGE-0646086). The analyses, conclusions, and opinions included in this work are my own (or, in Chapter 2, are mine and my coauthors'); they may not represent the positions or policies of the ALSA, CRELES, ELSA, SEBAS, or NHANES/NHEFS teams, or of any of the funding agencies. The Institutional Review Board at Princeton University has reviewed a proposal for this research, and deemed the work "exempt" from human subjects oversight.

LIST OF TABLES

Table I.1. Weight Values Corresponding to the Standard BMI Categories (by Select Values of Height) 4

Table 1.1. Participant Characteristics at Study Baseline: Five Surveys ... 29

Table 1.2. The Impact of BMI-Change Adjustment on the Relationship Observed Between BMI and Mortality 38

Table 2.1. Survey Measurement in Costa Rica (CRELES) and in England (ELSA) 53

Table 2.2. BMI Changes Between Adjacent Survey Waves 61

Table 3.1. Mean Hazard Ratios Associated with Weight Categories II and III, for Three Simulated Populations 100

Table 3.2. Selected Parameter Values: Three Models 101

Table 1A.1. Orderings of Events, with Mortality Outcomes..... 127

Table 1C.1. AIC Values for the Analyses Shown in Figures 1.3, 1.4, and 1.5, by Select Values of λ 144

Table 1D.1. Odds Ratios, Predicting Death Within *n* Years, for Five Possible Effect-Modification Scenarios 148

Table 2A.1. Descriptive Statistics, Costa Rica 153

Table 2A.2. Descriptive Statistics, England 154

Table 3A.1. Values Considered in the Parameter-Value Search 169

LIST OF FIGURES

Figure I.1. Mortality Hazard Ratios by Standard BMI Categories: Results from a Recent Meta-Analysis 3

Figure 1.1. Presumed Causal Pathways 22

Figure 1.2. Timing of Data Collection in Five Population Surveys 28

Figure 1.3. BMI and Mortality in Five Populations (with Controls for Age, Sex, and Smoking) 34

Figure 1.4. BMI Change and Mortality in Five Populations (with Controls for Age, Sex, and Smoking) 36

Figure 1.5. BMI and Mortality in Five Populations, with and Without Adjustment for BMI Change (with Controls for Age, Sex, and Smoking) 37

Figure 2.1. BMI Changes Observed in Costa Rica and in England 59

Figure 2.2. Self-Reported Disease Diagnoses and BMI Change < -1.0 BMI Point 62

Figure 2.3. Self-Reported Disease Diagnoses and BMI Change < -2.0 BMI Points 63

Figure 3.1. Transition Matrix: Using Rates of Population Processes to Calculate Annual Changes in Population Composition 93

Figure 3.2. Death Rates, by Age, in Three Simulated Populations 96

Figure 3.3. Weight Category Distribution, by Age, in Three Simulated Populations 97

Figure 3.4. Mortality Hazard Ratios Associated with Weight Categories II and III, by Age, in Three Simulated Populations 98

Figure 1B.1. Mortality Hazards in Three Countries 132

Figure 1B.2. Predicted Probabilities of Death Within Eight Years, Without Smoothing, in Three Countries 133

Figure 1B.3. Predicted Probabilities of Death Within Eight Years, with Smoothing, in Three Countries 135

Figure 1C.1. Smoothing the Body Mass Index–Mortality Relationship with *P*-Splines: Two Examples 138

Figure 1C.2. Examples of B-Splines. Reproduced from Eilers and Marx . 140

Figure 1C.3. Analysis of BMI and Mortality Net of BMI Change (over 11 Years), USA, 1982-84 143

Figure 2B.1. Self-Reported Disease Diagnoses and BMI Change (Continuous) 161

Figure 3A.1. Observed Death Rates by Age, with Target Age Pattern of Mortality for the Simulation Populations to Reproduce 165

Figure 3A.2. Observed BMI Category by Age, with Target Weight Category Distributions for the Simulation Populations 167

Figure 3A.3. Target Relationship Between Weight and Mortality, by Age, for the Simulation Populations 168

Introduction

In January 2013, four scholars at the U.S. National Center for Health Statistics (NCHS) published a review of nearly 100 previously written papers on body weight and mortality (1). Each paper included in the review was based on standard weight categories adopted by the World Health Organization and by the National Heart, Lung, and Blood Institute in the USA. Those standard weight categories, in turn, are based on body mass index, or BMI: a popular measure of weight (in kilograms) divided by the square of height (in meters). Yet despite these routine beginnings, the published report was immediately controversial. Overweight people have lower death rates, not higher ones, the study found, than people with clinically normal body weight.

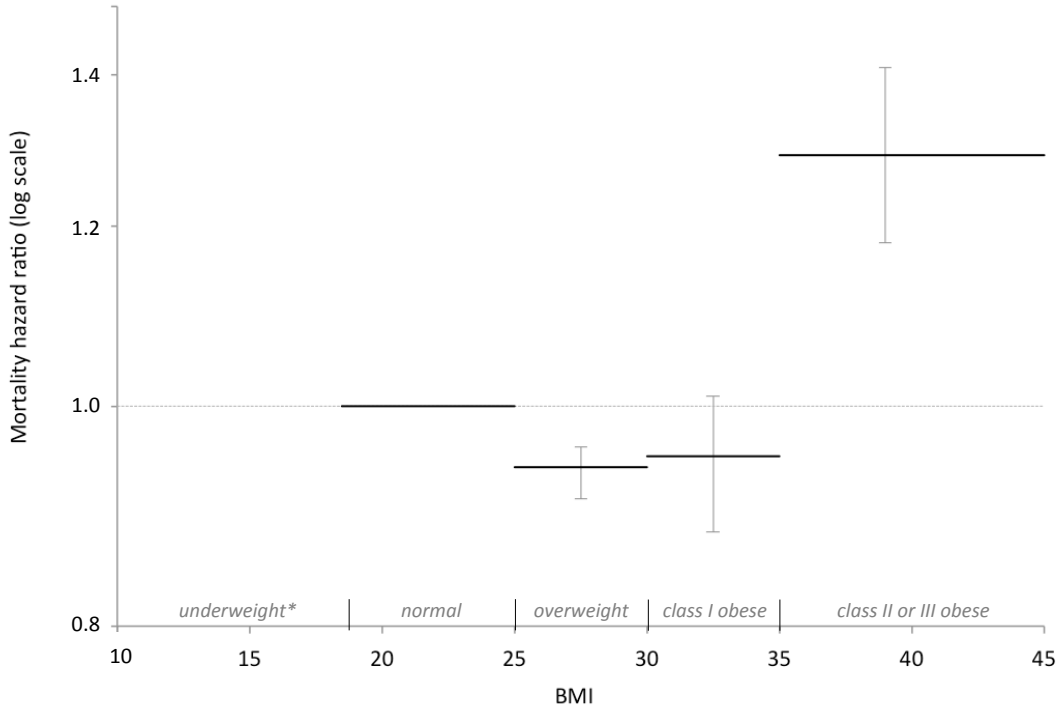
The results were contentious – and they made headlines worldwide (2-5) – because they would seem to contradict popular perceptions of healthy body weight (6) and extensive biomedical evidence that links overweight status to increased disease risk (7). The report found that overweight people (defined for both men and women as those with BMI ≥ 25, but < 30)

experience death rates that are 6% lower, on average, than people with clinically normal weight (defined as BMI ≥ 18.5 and < 25) (1). Furthermore, even people with mild obesity (BMI ≥ 30 and < 35) face a mortality risk that is statistically indistinguishable from the risk among people with normal weight, the NCHS authors concluded (1). (These results are illustrated in Figure I.1, below. See Table I.1 for weight values corresponding to BMI values.) Many scholars who had not worked on the new review urged caution in interpreting the results (2); other scholars, however, were more hostile still. “The new statistics are completely misleading for anyone interested in knowing about their optimal weight,” Harvard nutrition professor Walter Willett wrote in a public statement, covered in the popular press (8). “Stated politely,” Willett continued, “the paper is a pile of rubbish” (8).

The debate around body weight and mortality does not, in fact, center on the numbers themselves. To the contrary, the reported associations seem very real, and the latest systematic review, led by NCHS senior scientist Katherine Flegal, was published in the prominent and well-vetted *Journal of the American Medical Association* (*JAMA*). Instead, the debate centers on how we can meaningfully interpret studies of BMI and mortality, and whether a positive association between overweight status and longevity should be reported to the public at all.

At issue is whether the observed population-level associations reflect true, causal relationships. Nobody doubts that malnourished people may suffer poor health due to their low energy intake or to micronutrient deficiency (9, 10). But a large number of scholars argue that,

Figure I.1. Mortality Hazard Ratios by Standard BMI Categories: Results from a Recent Meta-Analysis. Original figure to show results from Flegal et al. (1), Table 1, for both sexes and all ages combined. Reference category is "normal weight" (BMI ≥ 18.5 and < 25). Grey vertical lines represent 95% confidence intervals.



* Data are insufficient to calculate a hazard ratio for "underweight."

Table I.1. Weight Values Corresponding to the Standard BMI Categories (by Select Values of Height).

BMI	*Standard category*	*Height*							
		5’2” (1.57 m)	5’4” (1.63 m)	5’6” (1.68 m)	5’8” (1.73 m)	5’10” (1.78 m)	6’0” (1.83 m)	6’2” (1.88 m)	6’4” (1.93 m)
< 18.5	Underweight	< 101 lbs (< 45.9 kg)	< 108 lbs (< 48.9 kg)	< 115 lbs (< 52.0 kg)	< 122 lbs (< 55.2 kg)	< 129 lbs (< 58.5 kg)	< 136 lbs (< 61.9 kg)	< 144 lbs (< 65.4 kg)	< 152 lbs (< 68.9 kg)
18.5 – 24.9	Normal	101 - 136 lbs (45.9 - 61.9 kg)	108 - 145 lbs (48.9 - 66.0 kg)	115 - 154 lbs (52.0 - 70.2 kg)	122 - 163 lbs (55.2 - 74.5 kg)	129 - 173 lbs (58.5 - 78.9 kg)	136 - 183 lbs (61.9 - 83.5 kg)	144 - 194 lbs (65.4 - 88.2 kg)	152 - 204 lbs (68.9 - 93.1 kg)
25 – 29.9	Overweight	137 - 163 lbs (62.0 - 74.3 kg)	146 - 174 lbs (66.1 - 79.2 kg)	155 - 185 lbs (70.3 - 84.2 kg)	164 - 196 lbs (74.6 - 89.4 kg)	174 - 208 lbs (79.0 - 94.7 kg)	184 - 220 lbs (83.6 - 99 kg)	195 - 232 lbs (88.3 - 105 kg)	205 - 245 lbs (93.2 - 111 kg)
30 – 34.9	Obese (class I)	164 - 190 lbs (74.4 - 86.7 kg)	175 - 203 lbs (79.3 - 92.4 kg)	186 - 216 lbs (84.3 - 98.3 kg)	197 - 229 lbs (89.5 - 103 kg)	209 - 243 lbs (94.8 - 110 kg)	221 - 257 lbs (100 - 116 kg)	233 - 271 lbs (106 - 123 kg)	246 - 286 lbs (112 - 129 kg)
35 – 39.9	Obese (class II)	191 - 218 lbs (86.8 - 99.1 kg)	204 - 232 lbs (92.5 - 105 kg)	217 - 247 lbs (98.4 - 111 kg)	230 - 262 lbs (104 - 118 kg)	244 - 278 lbs (111 - 125 kg)	258 - 294 lbs (117 - 133 kg)	272 - 310 lbs (124 - 140 kg)	287 - 327 lbs (130 - 148 kg)
≥ 40	Obese (class III)	≥ 219 lbs (≥ 99.2 kg)	≥ 233 lbs (≥ 106 kg)	≥ 248 lbs (≥ 112 kg)	≥ 263 lbs (≥ 119 kg)	≥ 279 lbs (≥ 126 kg)	≥ 295 lbs (≥ 134 kg)	≥ 311 lbs (≥ 141 kg)	≥ 328 lbs (≥ 149 kg)

excluding the underweight range (BMI < 18.5), death risk most likely rises monotonically with body weight (e.g., (11-13)), and that the optimal BMI for longevity is therefore at the low end of the normal-BMI range, with excess weight beyond that causing metabolic dysfunction and cardiovascular strain (7). The new *JAMA* paper by Flegal et al., showing low death rates among the overweight, would seem to challenge that wisdom. Its release in early 2013 has prompted (or renewed) questions over everything from whether BMI is an appropriate measure of fatness (14), to whether mortality is the most interesting outcome to observe (2). But the crux of the scholarly argument is what causes the population-level association observed between overweight status and longevity. That is the topic of this dissertation.

The dissertation tests one hypothesis, in particular, that is frequently given to explain low death rates among the overweight, relative to people with normal weight. This hypothesis is based on, "reverse causation": the extent to which health affects weight, rather than weight affecting health. That is, while we tend to treat BMI as a predictor in studies of body weight and health (with weight causing some health outcome), certain health events, such as disease, medication use, or changes in mobility, may themselves cause weight to change. Reverse causation – or the causal pathway from health to weight – could perhaps explain the longevity advantage among overweight people, if some people with normal weight are slim *only because* they lost weight due to sickness. Consider, for example, a person who develops cancer while overweight, loses weight as a result, and is then observed to be normal weight while at high risk of death. Sickness could make overweight status appear protective to longevity in population-based

studies, when in fact it is poor health (perhaps brought on by being too fat) that drives both low body weight and high risk of death for some segment of the population.

Despite great popularity (15-21), however, this reverse causation hypothesis remains largely untested. To date, there has been limited research on reverse causation, and critics – including, notably, Katherine Flegal at NCHS – argue there is little evidence that sickness-induced weight loss has much impact on observed population-level phenomena (15). This dissertation therefore aims to clarify the role of reverse causation in studies of BMI and mortality. It assesses, in particular, whether sickness-induced weight loss might account for the high observed death rates among people with normal weight, relative to people who are overweight.

Each chapter of the dissertation comprises a separate study, related to reverse causation. Chapter 1 uses data from five high- and middle-income countries to estimate the shape of the BMI–mortality curve, adjusting for recent changes in BMI. This tests whether people losing weight face unusually high death rates, and, if so, whether that might explain why normal-weight people appear to die at higher rates than overweight people. Chapter 2 uses data from Costa Rica and England to test whether sickness can indeed predict weight loss in prospective population-based surveys. Finally, Chapter 3 moves away from data analysis, and uses computer simulation to assess whether it is possible for various hypotheses, including two different models of reverse causation, to explain the contentious association between BMI and mortality.

The three chapters are intended as stand-alone papers. Each gives its own background, and each lays out its conceptual approach and methodology. Because of this, I will not write much more about them here. However, there are a few overarching questions that deserve some attention. These relate to the dissertation as a whole, so I will turn to them now.

***

The first question, and perhaps the most obvious, is, "Why use BMI?"

As noted above, many people question whether BMI is really the best exposure when examining a relationship between fatness and mortality risk (see, e.g., (14)). This is because BMI is based on height and weight only and, consequently, does not distinguish fat mass from lean mass, such as muscle and bone. In recent years, scholars have suggested that in place of BMI, it might therefore be better to analyze body fat percentage (which does distinguish fat from lean mass) or waist-to-hip ratio (measuring how and where the body is storing fat) (e.g, (22, 23)). Both these measures have their benefits. However, BMI is correlated with both, and distinguishes similar categories of body fatness (24). The (comparatively few) studies that assess these measures' links to mortality risk also suggest that they, like BMI's, may be somewhat difficult to interpret (e.g., (25, 26)). Finally, BMI has one chief advantage over other measures, namely ease of data collection. Height and weight measurements are non-invasive and relatively cheap, so BMI data are abundant. Measurements are, for the most part, also

consistent across measurement devices; this means that measured data (although not self-reported values) should be comparable over time and across populations. This is not true of measurements of, say, body fat percentage, which tend to be sensitive to the measurement technique employed (e.g., (27)).

The central problem with BMI then, in my mind, is not that other measures are markedly better, but rather that ambiguity exists about what we aim to measure with it in the first place. Hernán and Taubman argue it is essentially impossible to estimate a true causal effect of obesity on mortality because too many different interventions could allow a given individual to achieve a given BMI value (28). While waist-to-hip ratio and body-fat percentage are somewhat less crude than BMI, neither can avoid this central problem. They too, similarly, reflect countless processes relating to nutrition, physical activity, cardiovascular fitness, accumulation and function of adipose tissue, interaction with microbes, and the hormonal, the metabolic, the genetic, and the epigenetic traits that drive all of the above. Some causal effects within these processes may be inseparable from those of mediating pathways or of correlated factors, related via feedback loops. As a result, BMI and other measures of fatness are perhaps best considered proxies for complex biological systems. That is, they represent real effects, but without a complete understanding of the systems, causal inference is problematic.

In short, then, the answer to, “Why BMI?” is, “largely for convenience.” Dozens, if not hundreds, of studies already look at the association between BMI and mortality (1), and there are many data available with which to assess the impact of reverse causation *on that particular*

association. BMI is therefore a useful and appropriate measure for this dissertation, and the research has implications for perhaps hundreds of previously published studies in medicine, epidemiology, and demography. The answer to whether BMI is a "good" measure in general, however, depends on the exposure of interest. For many of the most exciting questions today in the study of metabolism, that answer remains ambiguous, as we continue to explore and define the mechanisms controlling energy balance.

***

This brings me to the second question: "How is this demography?"

My dissertation is built on work in epidemiology and in public health. Nevertheless, this work is a natural choice for a demographer in two important respects.

First, demographers (more than epidemiologists) tend to be familiar with the concept of heterogeneity outlined by James Vaupel and colleagues (29, 30), demonstrating that population-level associations do not have to represent individual-level associations for any member, or any subgroup, of the population. Typically, therefore, demographers are well prepared to consider non-causal population-level associations, and the possible effects on these of changing population composition.

Second, and building on the previous section's point about causal inference, epidemiologists tend to deal with factors of causal interest (for example, the health effects of arsenic exposure) while social scientists, in contrast, often study exposures of *social* interest (for example, the health effects of race). This means that demographers may be more comfortable, on the whole, considering multifactorial exposures (such as race, sex, immigration status, or educational attainment) that have the same kind of complicated causal interpretation as noted above. This does not mean a demographer finds clarity in multifactorial exposures where an epidemiologist cannot. It means, simply, the demographer may find that ambiguity less noteworthy. Some exposures, and some relationships, are too important to ignore, even if causal pathways cannot be pulled apart completely.

BMI is such an exposure. It is not just the most abundant marker of energy balance in population data, but it, or at least body weight, is also a topic of enormous social significance and popular interest. Today, weight is a perceived marker of beauty, activity preferences, social status, and even moral character (e.g., (6, 31, 32)) – a marker of attractiveness in every sense of the word. The relationship between body weight and mortality is thus not just of medical interest; it is also of broader social interest, as body weight is seen as a meaningful, and highly visible marker of identity.

***

Third, one might wonder, “Why study mortality from all causes?” Reverse causation (or health change leading to weight change) is probably more likely to precede deaths due to cancer than deaths due to traffic accidents. We might then expect that reverse causation has a bigger effect on the relationship between BMI and cancer mortality than it has on, say, the relationship between BMI and mortality due to accidents – or indeed due to some number of other causes. This contention is supported in data. For example, in a 2005 review of 26 studies with information on BMI and cause of death, McGee and the Diverse Populations Collaboration found a longevity advantage among the overweight for all-cause mortality and for cancer mortality, but not for deaths related to heart conditions (33).

Perhaps the most obvious reason to focus on all-cause mortality, then, is that, while BMI could indeed have different relationships with different causes, any given person may only experience one BMI value at a particular time. If, ultimately, the optimal BMI for avoiding cancer death is not the same as the optimal value for avoiding death due to accidents, it would be useful to know how the different cause-of-death risks are balanced. That is, for many public health applications, it is useful to know the overall relationship between BMI and mortality from all causes combined.

In addition, however, many practical issues complicate the study of cause-specific mortality too, and these make all-cause mortality simply a more manageable outcome for the analyses in this dissertation. First is the issue of statistical power. This dissertation tests for reverse causation using data with repeated height and weight observations from the same individuals. The

sample sizes, as a result, are smaller (on the order of a few thousand) than is practical for cause-specific analyses (which requires sub-dividing the data by cause of death). Second, moreover, cause-of-death data are not always high quality. Even in the most developed countries, a large proportion of deaths are routinely attributed in vital statistics data to ill-defined causes or to no cause (34). This makes the data difficult to work with, and open to interpretation. Third, and finally, deaths are often difficult to categorize, even if recorded accurately, since a single death can often be attributed to multiple underlying causes. (Consider, for example, a person who dies with both diabetes and prostate cancer.) For all these reasons, analyses based on all-cause mortality seem more credible, and more feasible in this dissertation, than those based on cause-specific mortality.

***

Finally, it is natural to ask what this dissertation finds, and what I would recommend as a result. People would like to know, "Is it really best to be overweight?"

As will be clear from the chapters that follow, each study in this dissertation finds equivocal support for the hypothesis that high mortality among normal-weight people is due to sickness that caused both weight loss and high risk of death. That is, reverse causation does seem to affect estimates of the relationship between BMI and mortality. But it is unclear how strong and how consistent across populations these effects may be. The optimal BMI for longevity remains elusive as a result, and it is difficult to say whether overweight status is dangerous or desirable.

Since each research chapter is written as a stand-alone paper, as previously noted, each discusses the value and implications of its findings for demographers and epidemiologists engaged in similar research. These chapters do not, in the end, tell us which BMI we should strive for. Nevertheless, the Conclusion explores some implications of this dissertation for individual behavior.

REFERENCES

[1] Flegal KM, Kit BK, Orpana H, et al. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *J Am Med Assoc*. 2013; 309(1):71-82.

[2] Belluck P. Study suggests lower mortality risk for people deemed to be overweight. *New York Times*. January 2, 2013:A10.

[3] NPR. Research: a little extra fat may help you live longer. NPR News, January 2, 2013. Washington, DC. (Transcript at

http://www.npr.org/templates/transcript/transcript.php?storyId=168437030). (Accessed June 26, 2013).

[4] BBC. 'Weight is healthy' study criticised. BBC News, January 2, 2013. London, UK. (http://www.bbc.co.uk/news/health-20889381). (Accessed June 26, 2013).

[5] Blue L. Being overweight is linked to lower risk of mortality. Time.com. January 2, 2013. (http://healthland.time.com/2013/01/02/being-overweight-is-linked-to-lower-risk-of-mortality/). (Accessed July 1, 2013).

[6] Hurd Clarke L. Older women's perceptions of ideal body weights: The tensions between health and appearance motivations for weight loss. *Ageing Soc*. 2002; 22(6):751-773.

[7] Hu F, ed. *Obesity Epidemiology*. Oxford: Oxford University Press; 2008, pp. 149-215.

[8] Harvard School of Public Health. Ask the Expert: Does being overweight really decrease mortality? No! The Nutrition Source, January 2013. Boston, MA. (http://www.hsph.harvard.edu/nutritionsource/questions/ask-the-expert-does-being-overweight-really-decrease-mortality-no/). (Accessed June 26, 2013).

[9] Shetty PS, James WPT. Body mass index: A measure of chronic dietary energy deficiency in adults. Rome, Food and Agriculture Organization of the United Nations, 1994 (FAO Food and Nutrition Paper No. 56). (http://www.fao.org/docrep/t1970e/t1970e00.htm.)

[10] Chandra RK. Nutrition and the immune system: an introduction. *Am J Clin Nutr*. 1997; 66(2):460S-463S.

[11] Manson JE, Stampfer MJ, Hennekens MD, et al. Body weight and longevity: a reassessment. *J Am Med Assoc*. 1987; 257(3):353-358.

[12] Baik I, Ascherio A, Rimm EB, et al. Adiposity and mortality in men. *Am J Epidemiol*. 2000; 152(3):264-271.

[13] Ajani UA, Lotufo PA, Gaziano JM, et al. Body mass index and mortality among US male physicians. *Ann Epidemiol*. 2004; 14(10):731-739.

[14] Heymsfield SB, Cefalu WT. Does body mass index adequately convey a patient's mortality risk? *J Am Med Assoc*. 2013; 309(1):87-88.

[15] Flegal K, Graubard B, Williamson D, et al. Reverse causation and illness-related weight loss in observational studies of body weight and mortality. *Am J Epidemiol*. 2011; 173(1):1-9.

[16] Snih S, Ottenbacher K, Markides K, et al. The effect of obesity on disability vs mortality in older Americans. *Arch Intern Med*. 2007; 167(8):774-780.

[17] Adams K, Schatzkin A, Harris T, et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *N Engl J Med*. 2006; 355(8):763-778.

[18] Ajani U, Lotufo P, Gaziano M, et al. Body mass index and mortality among US male physicians. *Ann Epidemiol*. 2004; 14(10):731-739.

[19] Lindsted KD, Singh PN. Body mass and 26-year risk of mortality among women who never smoked: findings from the Adventist Mortality Study. *Am J Epidemiol*. 1997; 146(1):1-11.

[20] Losonczy K, Harris T, Cornoni-Huntley J, et al. Does weight loss from middle age to old age explain the inverse weight mortality relation in old age? *Am J Epidemiol*. 1995; 141(4):312-321.

[21] Manson J, Willett W, Stampfer M, et al. Body weight and mortality among women. *N Engl J Med*. 1995; 333(11):677-685.

[22] Heitmann BL, Erikson H, Ellsinger BM, et al. Mortality associated with body fat, fat-free mass and body mass index among 60-year-old Swedish men – a 22-year follow-up: The study of men born in 1913. *Int J Obes*. 2000; 24:33-37.

[23] Price GM, Uauy R, Breeze E, et al. 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*. 2006; 84(2):449-460.

[24] Flegal KM, Shepherd JA, Looker AC, et al. Comparisons of percentage body fat, body mass index, waist circumference, and waist-stature ratio in adults. *Am J Clin Nutr*. 2009; 89:500-508.

[25] Bigaard J, Frederiksen K, Tjønneland A, et al. Body fat and fat-free mass and all-cause mortality. *Obes Res*. 2004; 12(7):1042-1049.

[26] Bigaard J, Frederiksen K, Tjønneland A, et al. Waist and hip circumferences and all-cause mortality: Usefulness of the waist-to-hip ratio? *Int J Obes Relat Metab Disord*. 2004; 28(6):741-747.

[27] Nichols J, Going S, Loftin M, et al. Comparison of two bioelectrical impedance analysis instruments for determining body composition in adolescent girls. *J Body Compos Res*. 2006; 4(4):153-160.

[28] Hernán MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *Int J Obes*. 2008; 32(Suppl 3):S8-S14.

[29] Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*. 1979; 16(3):439-454.

[30] Vaupel JW, Yashin AI. Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *Am Stat*. 1985; 39:176-185.

[31] Cossrow NHF, Jeffery RW, McGuire MT. Understanding weight stigmatization: A focus group study. *J Nutr Educ*. 2001; 33(4):208-214.

[32] Falkner NH, French SA, Jeffery RW, et al. Mistreatment due to weight: Prevalence and sources of perceived mistreatment in women and men. *Obes Res*. 2012; 7(6):572-576.

[33] McGee D, Diverse Populations Collaboration. Body mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies. *Ann Epidemiol*. 2005; 15(2):87-97.

[34] Mathers CD, Fat DM, Inoue M, et al. Counting the dead and what they died from: An assessment of the global status of cause of death data. *Bull World Health Organ*. 2005; 83(3):171-177.

Chapter 1

Body weight, weight change, and mortality risk: Evidence from five countries

CHAPTER SUMMARY: *Many studies find that overweight people have lower mortality than people of normal weight, and this result is frequently attributed to "reverse causation": the hypothesis that being thin is not itself a risk factor for ill health, but that ill health may cause people to become thin. This study tests whether recent BMI changes may explain part of the curvilinear relationship between BMI and mortality among adults aged 50-79. Data come from five longitudinal surveys, each with measured height and weight data, and each based in a different high- or middle-income country: Australia, Costa Rica, England, Taiwan, and the USA. The study uses* P-*splines (penalized* B-*splines) to treat BMI and BMI change as continuous variables in penalized logistic regression models, avoiding strong functional-form assumptions. Adjustment for BMI change over the past two to 12 years tends to rotate the BMI–mortality curve, suggesting lower mortality at low BMIs and higher mortality at high BMIs than is predicted in unadjusted analysis. Adjustment may further suggest a slightly lower BMI*

associated with lowest death rates than is apparent in unadjusted analysis. This study provides modest support for the reverse causation hypothesis.

Many studies find a curvilinear relationship between body mass index (BMI) and mortality, with higher-than-average death rates both among people with high BMI and among those with relatively low BMI (1-10). This relationship is counterintuitive, and remains controversial, because in most studies the BMI associated with lowest mortality falls in the "overweight" range or very close to it (at the high end of "normal") under World Health Organization (WHO) classifications. This seems at odds with widely accepted beliefs about healthy body weight, and with evidence linking overweight status to hormonal dysregulation and disease (11).

One hypothesis has become especially popular to explain the curvilinear relationship between BMI and mortality. This hypothesis, often called "reverse causation," posits that being thin is not itself a risk factor for ill health, but that instead ill health may cause people to become thin, with some diseases, such as cancer, inducing weight loss. The hypothesis remains largely untested, however, and has recently faced scrutiny. Katherine Flegal and coauthors (12) write that there is "no good evidence as to how often this [illness-induced weight change] phenomenon occurs," and there is "little theoretical or empirical basis for describing what type

of bias it might cause" in estimates of the BMI–mortality relationship. This paper aims to provide some of the empirical evidence currently lacking.

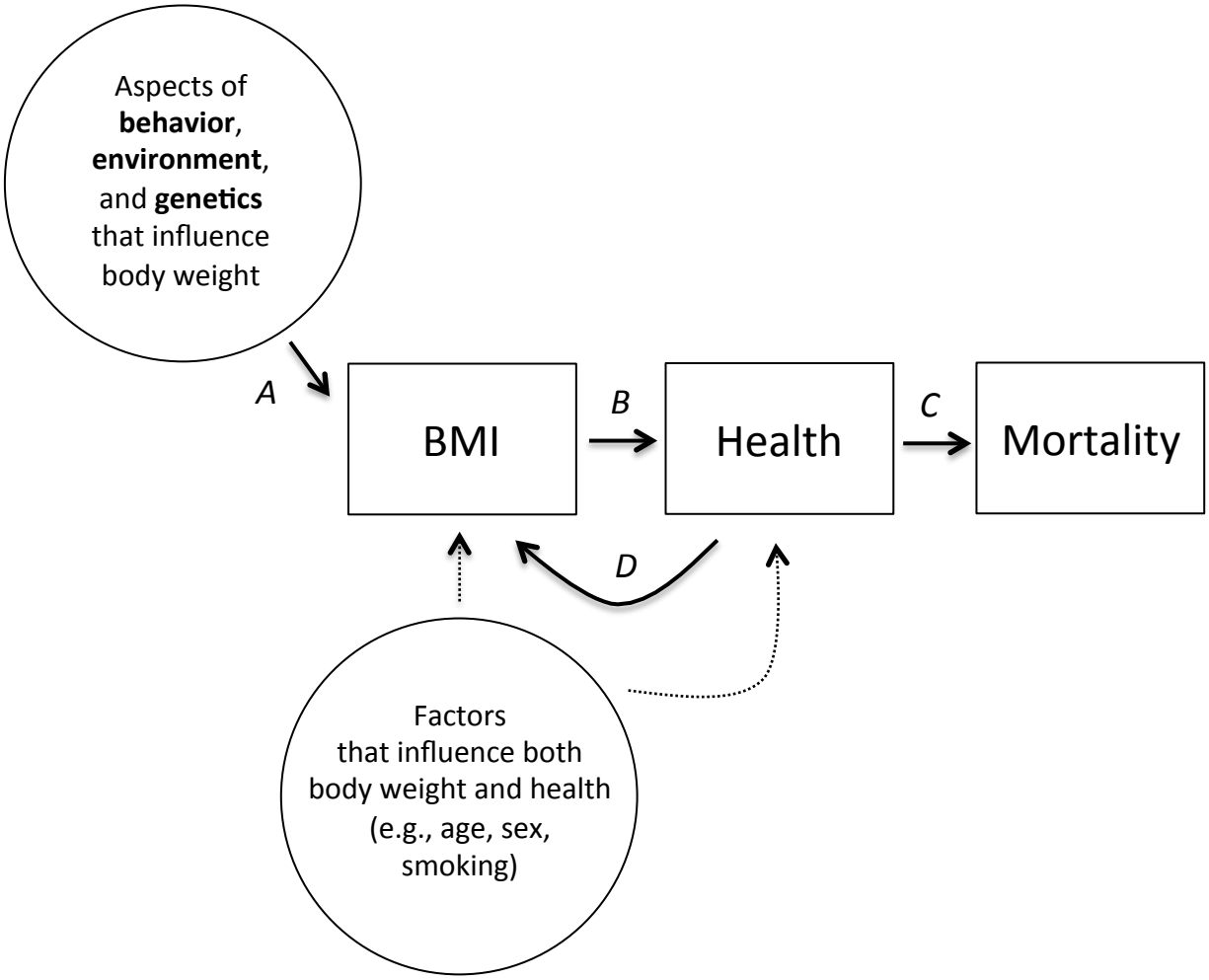
This study uses data from five longitudinal surveys, each based in a different high- or middle-income country, to test whether statistical adjustment for recent BMI change tends to alter the shape of the BMI–mortality curve in a manner consistent with reverse causation. While previous studies of the BMI–mortality relationship attempt to limit potential bias due to illness-induced weight loss (6, 13-17), this study takes the opposite tack. It investigates whether such bias is likely to exist at all, and, if so, whether that bias could be large enough to affect substantive interpretation of the associations between BMI and mortality.

Conceptual approach

This paper assumes a set of causal pathways that link lifestyle, BMI, health, and mortality risk. These are illustrated in Figure 1.1. While scholars would often like to estimate the effect of BMI on mortality, a combination of pathways B and C in the figure, instead they measure a blend of B, C, and D. Pathway D represents the effect of health on BMI: that is, reverse causation. Reverse causation is considered a major source of bias in BMI–mortality estimates (16, 18, 19).

Current techniques to address reverse causation, however, may be inadequate or come at unnecessarily high cost to statistical power or generalizability of results. BMI–mortality studies

Figure 1.1. Presumed Causal Pathways.



commonly restrict samples to never-smokers only, to people without known health problems, or to both (6, 13-17). Many then further limit their samples by ignoring deaths within a specified (but usually arbitrary) follow-up period, with the goal of eliminating from analysis any participant who was sick at baseline without knowing it (thus reducing the potential for pathway D to operate). These restrictions can eliminate up to 90% of deaths from a study (12), with consequences for both power and sample representativeness (20). Nor are the approaches guaranteed to solve the problems they aim to address, however, as large numbers of deletions can themselves distort results (21), and exclusion of deaths early in the follow-up period can, under some circumstances, exacerbate bias from undiagnosed illness (22). Rather than limiting the bias from illness-induced weight change, therefore, this study aims to assess the extent of reverse causation among population-based samples, without restrictions for smoking or poor health. It does so by treating *changes* in BMI as a signal of action along pathways A or D in Figure 1.1 (controlling for age, sex, and smoking), and then estimating the influence of BMI changes on the observed relationship between BMI and mortality.

The paper comprises three sets of analyses. First, it estimates the BMI–mortality relationship (unadjusted for recent BMI changes) to reproduce the central finding of previous studies: a curvilinear relationship between BMI and death risk. One reason that this relationship remains contentious is that most studies to date have analyzed BMI in categories: e.g., quartiles or quintiles of a population distribution (5), or bins corresponding to the WHO "underweight," "normal," "overweight," and "obese" classifications (respectively, BMI < 18.5, 18.5 – <25, 25 – <30, or ≥30) (10). This makes it almost impossible to discern the minimum of the BMI–mortality

curve or to distinguish regions that are relatively flat from those where the slope changes quickly. This study uses splines to produce a more detailed picture of the relationship, using measured BMI instead of self-reported values.

Second, this study estimates the relationship between BMI changes and mortality (unadjusted for BMI level) to determine whether body-weight fluctuations, and especially weight losses, are predictive of mortality as hypothesized. Here it is assumed that BMI change may have a different association with mortality than one would expect simply from observing the slope of the BMI–mortality curve estimated in the first set of analyses. That is, while changes in BMI due to behavior or environment (pathway A in Figure 1.1) will affect mortality risk solely through their impact on BMI level, other changes (through pathway D) instead signal unobserved health processes, such as disease or medication use, that can exert an independent effect on mortality risk through pathway C.

Third, and finally, this study estimates the relationship between BMI and mortality net of recent BMI change. Crucially, while this study assumes that BMI change may *signal* health changes (as pathway D does), it does not assume that BMI change will itself *cause* health changes independent of BMI level. That is, the causal-pathways schematic in Figure 1.1 does not include "BMI change" explicitly. It assumes, instead, that changes in BMI (which occur due to changes in lifestyle or health, e.g., via pathways A or D) affect health exclusively through their impact on achieved BMI level (represented by the "BMI" box) and, by extension, on pathway B. This means that, if one were to control perfectly for factors that influence both body weight and

health (e.g., age, sex, and smoking), there should be no association between mortality and BMI changes due to lifestyle, net of achieved BMI level. Based on this principle, the third set of analyses aims to highlight the impact, if any, of reverse causation on the BMI–mortality relationship. Any association detected here between BMI change and mortality should reflect the mortality impact of health conditions or processes that induced the weight change (through pathway D), or, perhaps, of BMI history (if, e.g., past obesity has a lingering effect once a person becomes lean, or past leanness has an effect once a person becomes heavy (23)). As a result, adjusting the BMI–mortality curve for BMI change should reduce the impact of both existing health conditions and BMI history on the estimated association, and make the adjusted curve reflect better the causal BMI–mortality relationship.

This study does not attempt to determine which weight changes in particular are caused by sickness (that is, which occur via pathway D rather than pathway A) because of probable errors identifying sickness-induced weight changes from survey data. Recorded sickness status may be inaccurate, as respondents may have undiagnosed or preclinical disease, may misremember the name for a diagnosed condition, or may choose not to report a known disease. Furthermore, data on weight-change intention have limited use because they cannot distinguish sickness-induced weight gain from other, lifestyle-related unintentional weight gain, and because some people who plan to lose weight via lifestyle modification may succeed only because of illness (19).

MATERIALS AND METHODS

Data

Data come from five longitudinal surveys, each based in a different high- or middle-income country:

1. The Australian Longitudinal Study of Ageing (ALSA);
2. The Costa Rican Longevity and Healthy Aging Study (CRELES);
3. The English Longitudinal Study of Ageing (ELSA) (24);
4. The Social Environment and Biomarkers of Aging Study (SEBAS), based in Taiwan; and
5. The U.S. National Health and Nutrition Examination I Epidemiologic Follow-Up Study (NHEFS) (25).

Data are described in Appendix 1A. In brief, all five data sources are longitudinal, population-based health and social surveys. Four of the five (all but NHEFS) focus on older adults so, to ensure rough comparability, this study limits analyses to respondents aged 50-79. Each survey conducts periodic physical examinations. Each measures participants' body weight at two or more waves, and four of the five also measure height at multiple waves; NHEFS measures height at baseline only. BMI is calculated as weight (in kilograms) divided by the square of height (in meters). Change in BMI is defined as BMI at one survey wave minus BMI at a previous

wave, so that positive values reflect weight gain and negative values reflect weight loss. Each survey tracks mortality through vital registration and proxy interviews.

Figure 1.2 summarizes the timing of height and weight measurements and the length of mortality tracking in each survey. Table 1.1 shows sample characteristics at baseline.

Overview of methods

This study tests whether adjustment for recent BMI change tends to alter the shape of five populations' BMI–mortality curves in a manner consistent with the reverse causation hypothesis. The study comprises three sets of analyses. Each is described below, but they share many methodological features.

Logistic regression with P-*splines*.

All three sets of analyses use logistic regression. In all cases, the outcome is a binary variable for death within a specified follow-up period, with that period's duration – hereafter referred to as *n* years – set as long as available data will allow to a maximum of $n = 8$. This maximum reflects a focus on the years immediately following BMI measurement, when it is generally presumed that bias from illness-induced weight loss should be greatest (22). Hazard models, or time-to-event models, are used for sensitivity analysis in surveys with appropriate detail on timing of death (see Appendix 1B).

Figure 1.2. Timing of Data Collection in Five Population Surveys. Waves without BMI measurement are not shown. "X" denotes end of mortality follow-up. This study does not estimate a BMI–mortality curve from measurements at ALSA Wave 6, CRELES Wave 3, or ELSA Wave 4, due to short follow-up and few deaths. ELSA draws its sample from participants in a previous national study, the Health Survey for England, whose linked records are called ELSA Wave 0.

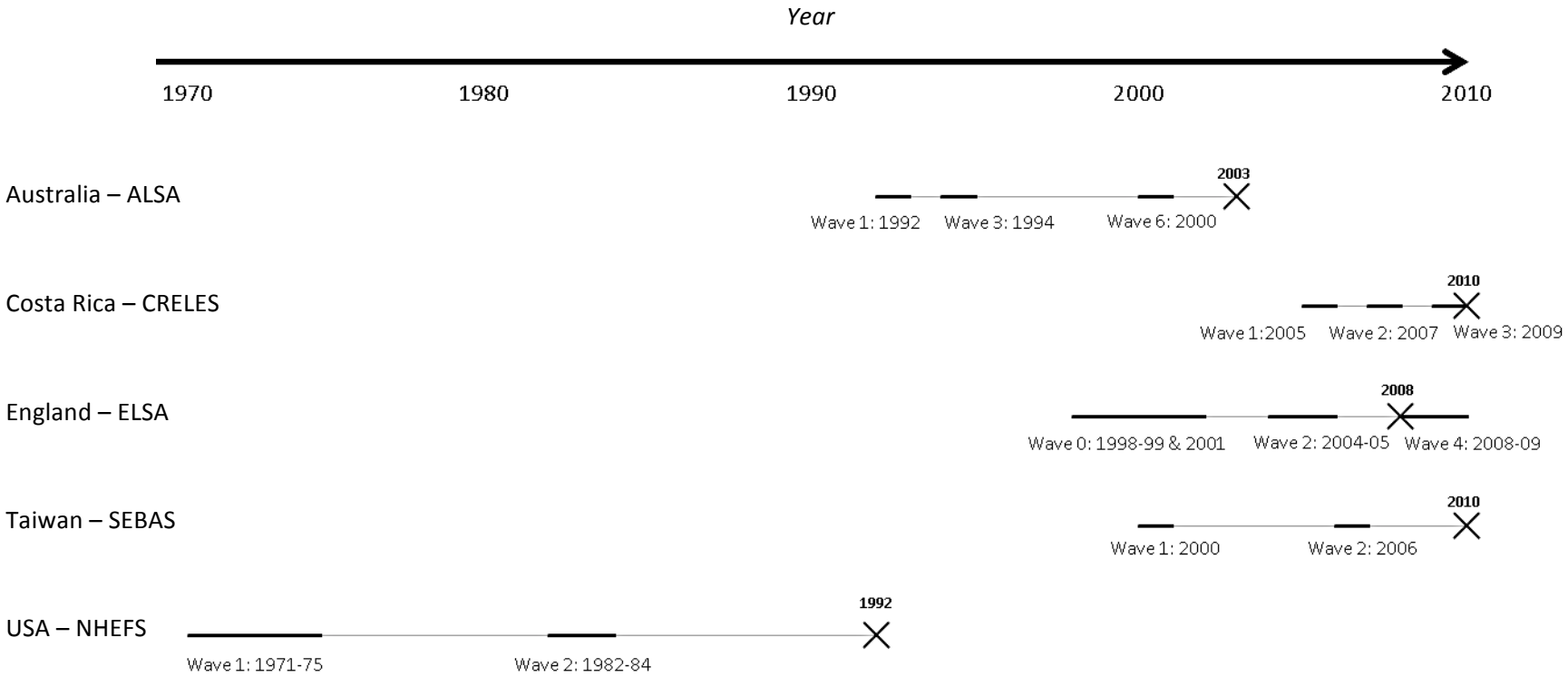


Table 1.1. Participant Characteristics at Study Baseline: Five Surveys.

Study population	*Year of baseline BMI measurement*	*Median BMI [interquartile range]*	*Median age [interquartile range]*	*Proportion female*	*Proportion current smoker*	*Sample size (N)*
Australia – ALSA	1992	26.13 [23.57, 28.80]	73 [71, 76]	52.1%	9.4%	995
Costa Rica – CRELES	2005	26.68 [23.96, 29.88]	70 [65, 75]	52.6%	10.2%	1,651
England – ELSA	1998-99	27.02 [24.47, 29.98]	62 [55, 69]	54.0%	21.4%	7,896
Taiwan – SEBAS	2000	24.19 [22.10, 26.67]	67 [70, 73]	42.2%	24.5%	916
USA – NHEFS	1971-75	25.81 [22.96, 28.96]	61 [55, 67]	51.9%	29.0%	3,192

All three sets of analyses treat BMI or BMI change (or both) as predictor variables. This study uses *P*-splines, or penalized *B*-splines, as described by Eilers and Marx (26), to treat these predictors as continuous, without imposing strong functional-form assumptions. This allows for a detailed picture of the relationships, including the minima. As with other regression splines, it is possible to include (parametric) control variables. However, *P*-splines may be better suited to the analyses here than other splines would be, as *P*-splines appear less affected by noise at the tails of the predictor distributions (results not shown). Spline construction and smoothing parameter selection (based on the Akaike Information Criterion) are described in Appendix 1C. All analyses are performed in Stata 12 (StataCorp LP, College Station, TX).

Adjustment for possible confounders.

All three sets of analyses control for age, age squared, sex, and smoking status (current, former, or never), measured at the start of the mortality follow-up period. Collectively these variables are referred to throughout the rest of this paper as "the controls" or "the covariates."

Controls are included in (penalized) logistic regression models in the conventional way: i.e., with log odds of death during some period as a linear function of the variables. Smoking status is included because smoking is known both to affect death risk and to be correlated with body weight (18, 27). The study is not restricted to never-smokers only as its goal is to test for evidence of reverse causation among the full population. Age is modeled with linear and quadratic terms because all-cause death rates are known to increase non-linearly with age

within the range considered here (28). Potential effect modification of the BMI–mortality relationship by age is mitigated by restricting the study sample to adults aged 50-79.

Presentation of results.

For each set of analyses, results are shown separately from each country in each time period, rather than pooled together. Pooling is possible in the analyses of BMI and mortality (by settling *n* equal across samples, e.g.). However, pooling is not straightforward in analyses of BMI change because, as shown in Figure 1.2, the interval between weight measurements – hereafter referred to as *w* years – differs in duration across survey waves. (Surveys with more than two measurements per person suggest that people rarely experience a constant rate of weight change; the data do not suggest an obvious equivalence between BMI-change measurements taken over different values of *w*.) The variation across samples in values of *n* and *w* also means that results are not easily compared across time and place.

For all analyses, results are shown as predicted probabilities of death within *n* years for a 72-year-old never-smoking male. Prediction lines are shown for the middle 95% of the BMI or BMI change distributions (although results are generated from the full samples), since the tails of these distributions are typically very long and the data are sparse, so that confidence intervals are wide at extreme values.

Three sets of analyses

1) BMI and mortality.

Across five data sources, there are eleven survey waves with body-weight measurements and mortality follow-up. This study therefore produces eleven different time- and place-specific estimates of the BMI–mortality relationship. A separate (penalized) logistic regression model is built for each. Odds of death within *n* years are modeled with *P*-splines for BMI, adjusting for covariates, as described above. 95% confidence intervals are calculated at each level of BMI conditional on a smoothing parameter (see Appendix 1C).

2) BMI change and mortality.

The data allow for twelve separate country- and time-specific estimates of the relationship between BMI change and mortality. Odds of death within *n* years are modeled by penalized logistic regression with *P*-splines for BMI change, adjusting for covariates. 95% confidence intervals are calculated at each level of BMI change, conditional on a smoothing parameter (see Appendix 1C).

3) BMI, BMI change, and mortality.

Statistical power is limited when considering BMI level and change simultaneously, so the study includes just five time- and place-specific estimates of the association: one per country. Data are selected from each survey source by choosing the period with the largest sample size.

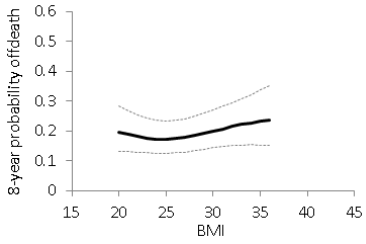
For each estimate, BMI level is treated as in the BMI–mortality analysis above, using *P*-splines to represent the relationship between BMI and mortality, and adjusting for covariates. However, to conserve power, the impact of BMI change is modeled with just two parameters: one for BMI change and one for BMI change squared (an approach informed by results, shown below, of the BMI-change–mortality analyses, for which a two-parameter model tends to fit quite well). As before, results are shown for a 72-year-old never-smoking male; however, the reference person is further assumed to have stable weight (BMI change = 0) over the *w* years since the previous BMI measurement. 95% confidence intervals are calculated conditional on a smoothing parameter (see Appendix 1C). Appendix 1D tests for effect modification between BMI and BMI change in their joint association with mortality risk.

RESULTS

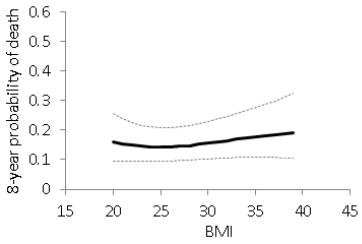
Figure 1.3 shows estimated BMI–mortality relationships for eleven different time- and place-specific samples. It is not straightforward to compare risk magnitude across different values of *n* (i.e., different durations of mortality follow-up), and this study is not designed to address why some curves may be flatter than others. However, in all countries and all survey waves considered, death risk appears elevated at low BMIs and, in most cases, at high BMIs as well. Except in Costa Rica, the minimum-mortality BMI is centered somewhere on the “overweight” range, a finding consistent with previous studies (10). Sensitivity analyses with hazard models show similar results (see Appendix 1B).

Figure 1.3. BMI and Mortality in Five Populations (with Controls for Age, Sex, and Smoking). Solid lines show predicted probabilities of death (within a specified interval) for a 72-year-old never-smoking male. Dotted lines show 95% confidence intervals, estimated conditional on a smoothing parameter. *N* is sample size; *n* is duration of mortality follow-up.

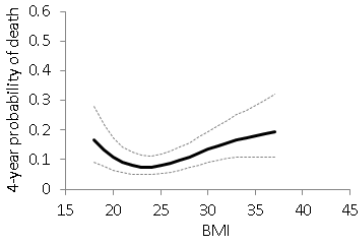
Australia, 1992
N = 995; *n* = 8 years



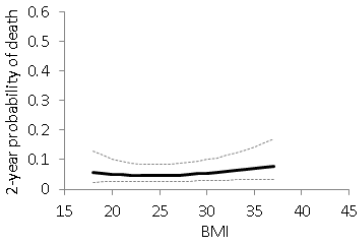
Australia, 1994
N = 727; *n* = 8 years



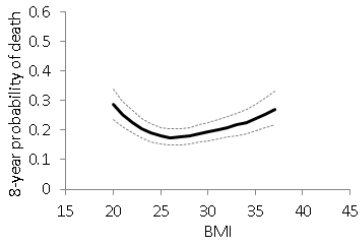
Costa Rica, 2005
N = 1,651; *n* = 4 years



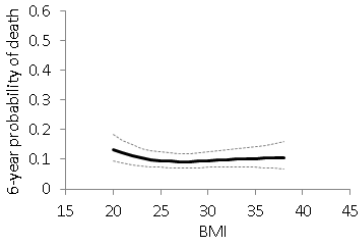
Costa Rica, 2007
N = 1,322; *n* = 2 years



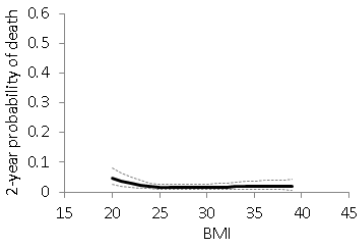
England, 1998-99
N = 7,896; *n* = 8 years



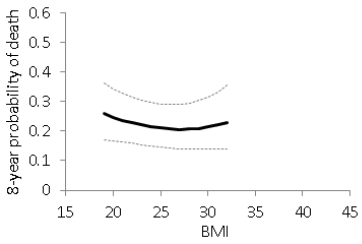
England, 2001
N = 5,337; *n* = 6 years



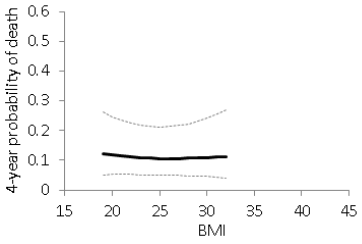
England, 2004-05
N = 6,385; *n* = 2 years



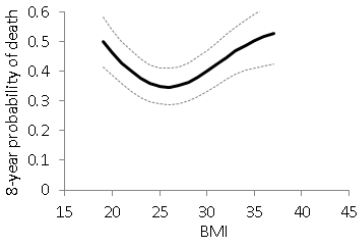
Taiwan, 2000
N = 916; *n* = 8 years



Taiwan, 2006
N = 501; *n* = 4 years



USA, 1971-75
N = 3,192; *n* = 8 years



USA, 1982-84
N = 5,217; *n* = 8 years

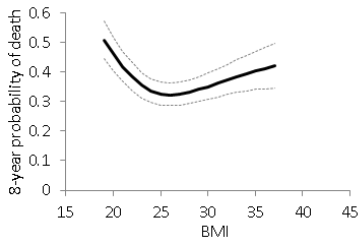


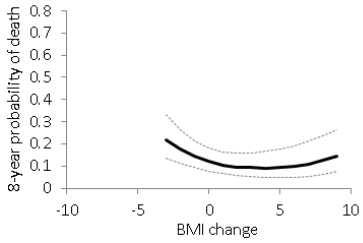
Figure 1.4 shows estimated relationships between BMI change and mortality in twelve time- and country-specific samples. As before, it is not simple to compare associations across different values of *n*; nor is it straightforward to compare across different values of *w* (i.e., different-length intervals over which BMI change was measured). Nevertheless, Figure 1.4 shows associations between BMI loss and mortality in all samples. Large BMI gain is often associated with mortality too, but the relationship between mild BMI gain and death risk appears flat or even negative. Interpretation is delicate without adjustment for BMI level, as observed BMI-change–mortality associations will reflect both the impact of whatever caused BMI to change as well as the difference in mortality risk at the old and new BMI levels.

Finally, Figure 1.5 and Table 1.2 show results from the analyses of BMI and mortality net of BMI change. In all five populations, normal-weight people are more likely to have lost weight than heavier people (see Table 1.2). As a result, with BMI-change adjustment, the BMI–mortality curve tends to rotate counter-clockwise from the unadjusted curve. This suggests that, in the absence of weight change, people would have lower mortality at low BMIs and higher mortality at high BMIs than is implied by the unadjusted curve. Furthermore, the minimum-mortality BMI appears slightly lower following adjustment for BMI change. There is no evidence of effect modification: i.e., differing relationships between BMI change and mortality at different levels of BMI (see Appendix 1D).

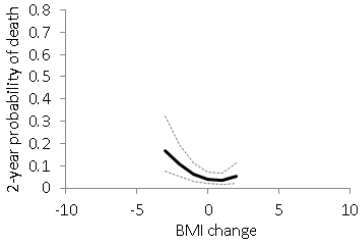
These findings are consistent with the reverse causation hypothesis. In each country, a X^2 (Wald) test is used to assess the two BMI-change adjustment parameters (change and change

Figure 1.4. BMI Change and Mortality in Five Populations (with Controls for Age, Sex, and Smoking). Solid lines show predicted probabilities of death (within a specified interval) for a 72-year-old never-smoking male. Dotted lines show 95% confidence intervals, estimated conditional on a smoothing parameter. *N* is sample size; *n* is duration of mortality follow-up; *w* is the interval (in the past) over which BMI change has occurred.

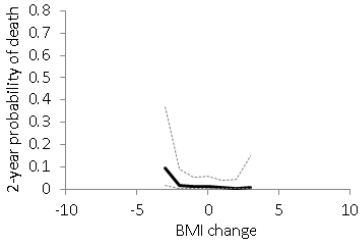
Australia, 1994
N = 656; *n* = 8 years; *w* = 2 years



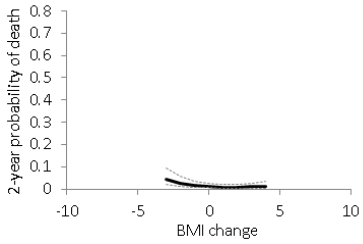
Costa Rica, 2007
N = 1,211; *n* = 2 years; *w* = 2 years



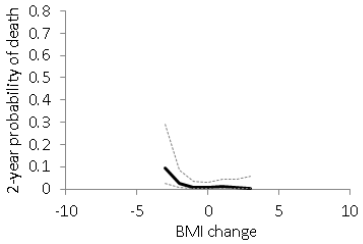
England, 2004
N = 919; *n* = 2 years; *w* = 3 years



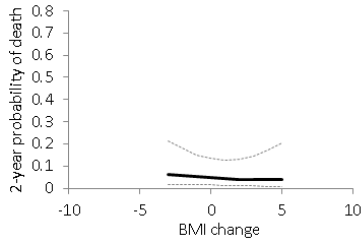
England, 2004-05
N = 2,761; *n* = 2 years; *w* = 6 years



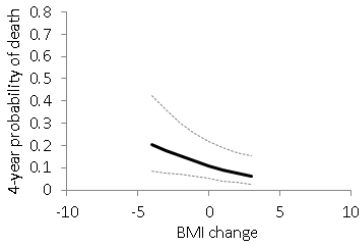
England, 2005
N = 1,480; *n* = 2 years; *w* = 4 years



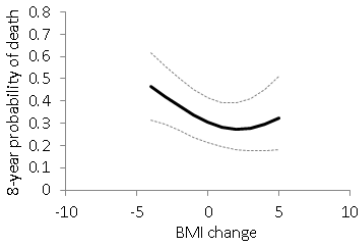
England, 2005
N = 850; *n* = 2 years; *w* = 7 years



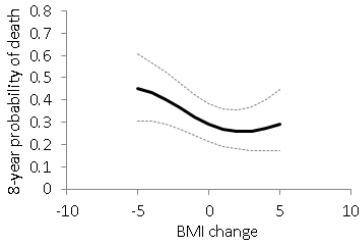
Taiwan, 2006
N = 501; *n* = 4 years; *w* = 6 years



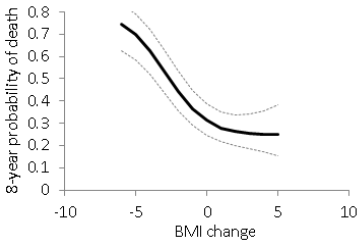
USA, 1982-83
N = 787; *n* = 8 years; *w* = 8 years



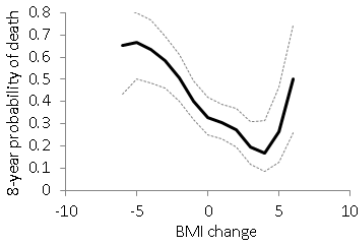
USA, 1982-84
N = 829; *n* = 8 years; *w* = 9 years



USA, 1982-84
N = 1,416; *n* = 8 years; *w* = 10 years



USA, 1982-84
N = 1,498; *n* = 8 years; *w* = 11 years



USA, 1983-84
N = 519; *n* = 8 years; *w* = 12 years

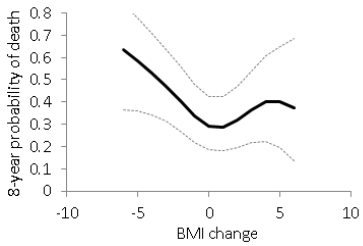
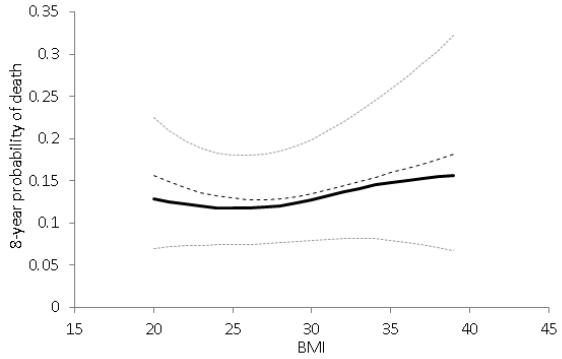
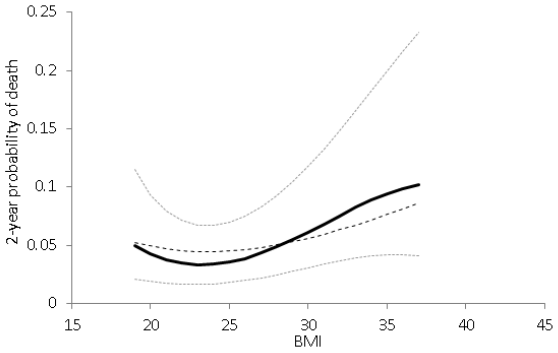


Figure 1.5. BMI and Mortality in Five populations, with and Without Adjustment for BMI Change (with Controls for Age, Sex, and Smoking). Dashed black lines show the BMI–mortality curve for a 72-year-old never-smoking male. Solid black lines show the prediction (with 95% confidence intervals, in grey) adjusted for BMI change, assuming BMI change = 0. *N* is sample size; *n* is duration of mortality follow-up; *w* is the interval (in the past) over which BMI change has occurred. Note different y-axis scales in different panels.

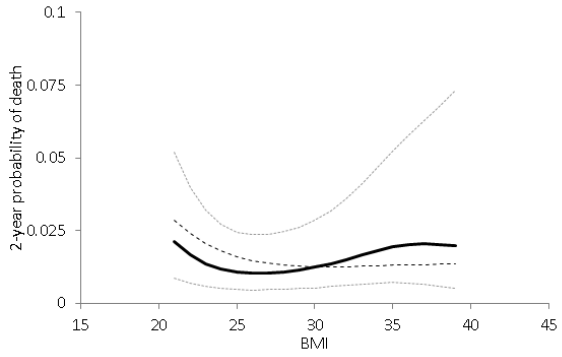
Australia, 1994
N = 665; *n* = 8 years; *w* = 2 years



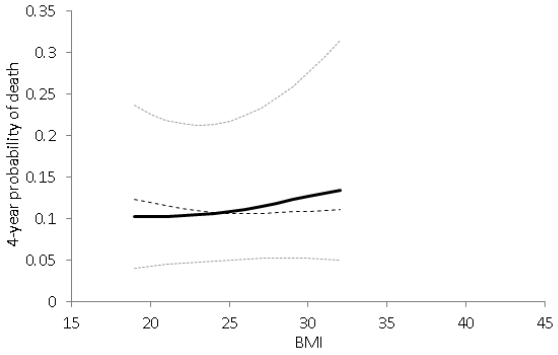
Costa Rica, 2007
N = 1,211; *n* = 2 years; *w* = 2 years



England, 2004-05
N = 2,761; *n* = 2 years; *w* = 6 years



Taiwan, 2006
N = 501; *n* = 4 years; *w* = 6 years



USA, 1982-84
N = 1,498; *n* = 8 years; *w* = 11 years

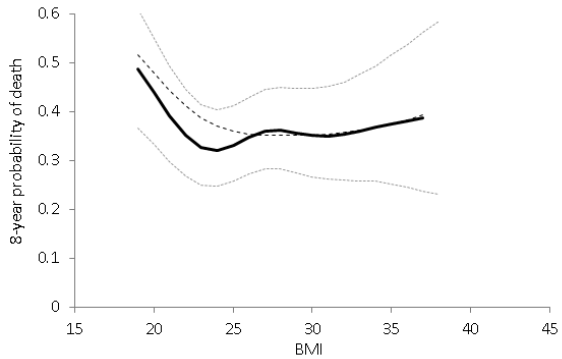


Table 1.2. The Impact of BMI-Change Adjustment on the Relationship Observed Between BMI and Mortality.

Study population	*Sample size (N)*	*Proportion of normal-weight respondents (BMI ≥18.5 to < 25) with BMI loss ≥1.0 point over previous* w *years*	*Proportion of overweight and obese respondents (BMI ≥25) with BMI loss ≥1.0 point over previous* w *years*	*Lowest-mortality BMI before adjustment for BMI change*	*Lowest-mortality BMI after adjustment for BMI change*	P-*value for difference in the shape of the curve, before and after adjustment (from* X^2 *test)*
Australia – ALSA 1994	656	30.89% (*w* = 2)	26.42%	27	25	0.003
Costa Rica – CRELES 2007	1,211	21.47% (*w* = 2)	16.48%	24	23	0.011
England – ELSA 2004-05	2,761	31.84% (*w* = 6)	15.93%	31	26	<0.001
Taiwan – SEBAS 2006	501	25.87% (*w* = 6)	14.54%	26	20	0.083
USA – NHEFS 1982-84	1,498	32.87% (*w* = 11)	21.53%	29	24	<0.001

squared), and the tests suggest that, overall, differences between curves with and without adjustment are unlikely to be due to chance (see Table 1.2). However, with a large, relatively flat portion of the BMI–mortality curve near its minimum, and with wide confidence intervals due to limited sample sizes, it is not possible to distinguish differences in relative-risk estimates before and after BMI-change adjustment. Nor are decreases in minimum-mortality BMI statistically significant at a conventional level. In short, it seems likely that reverse causation affects estimates of the BMI–mortality relationship, but this study cannot determine whether effects might be large enough to have a meaningful impact on interpretation of the curve.

DISCUSSION

This paper assesses plausibility of the reverse causation hypothesis in body-weight studies, testing the impact of recent BMI changes on observed relationships between BMI and death risk. This study is among the first to evaluate empirical evidence for reverse causation, although many scholars claim it produces important bias (12).

This study finds modest evidence to support reverse causation. As hypothesized, adjusting the BMI–mortality curve for recent BMI changes tends to rotate the curve, suggesting higher mortality at high BMIs and lower mortality at low (but clinically normal) BMIs than is observed without adjustment. Nevertheless, even in adjusted analyses, death risk appears to rise steeply with decreasing BMI within the normal BMI range.

The modest impact of BMI-change adjustment may signal that illness-induced weight loss, while important for some people, has little impact on the population-level association between BMI and mortality. Alternatively, it may be attributable to data limitations: in particular, crude measurement of BMI change. Changes in weight are observed only over long periods, two to 12 years, so some illness-induced weight change may go unrecorded (e.g., if a person's weight rises for several years but then falls sharply before the final measurement). Moreover, this study considers BMI change irrespective of cause. For these reasons, one may consider BMI change, as measured, to be a noisy indicator of all sickness-related change, suggesting that estimates in this paper should be conservative. In sum, the modest impact of BMI-change adjustment may not suggest that reverse causation, itself, has a modest effect on BMI–mortality estimates. The results of this study would be consistent either with a population-level impact that is very large or one that is quite small.

This study has assumed that BMI change may signal health changes, such as disease onset, but that BMI change affects mortality risk only through achieved BMI level (see Figure 1.1). This assumption is important for treating the linear and quadratic terms on BMI change as indicators of possible reverse causation in models of the BMI–mortality relationship (that is, pathway D in Figure 1.1). It is possible, however, that the assumption is incorrect and that change itself affects death rates, as would be the case, for example, if weight fluctuation affects metabolism in a way that affects longevity, or if weight change affects mental health in a way that affects longevity. The implication for results would depend on exact circumstances. However, it is

doubtful that such instances are systematic enough and common enough to change the substantive conclusions of this study.

This study has four main advantages over previous research. First, as noted above, this study offers a distinctive conceptual approach, and is among the first to produce empirical evidence with which to evaluate the reverse causation hypothesis. Second, this paper uses relatively strong data. All BMI values are based on objective height and weight measurements, whereas most past studies incorporating weight change have used self-reported values (16, 29-34), which are often inaccurate (35, 36). Contrary to popular belief, measurement error from faulty self-reports is not guaranteed to attenuate statistical estimates, since the error may not be random or with a mean of 0 (36, 37). Third, this study produces unusually detailed representations of the BMI–mortality curve, even before adjustment for BMI change, using *P*-splines to avoid arbitrary grouping of study participants or strong functional-form assumptions. Fourth, and finally, this paper provides limited information about body weight and mortality risk in middle-income countries in addition to high-income countries. The WHO estimates that more than half of the world's overweight- and obesity-related deaths now occur in middle-income countries (38), so it is increasingly important to understand patterns beyond the most developed nations.

The study has two major limitations. First, sample sizes are only moderate, and there is limited statistical power to compare results across the five countries or to test for effect modification, either between BMI level and BMI change, or by covariates. Second, as previously noted,

although the BMI data are comparatively good, they are still imperfect for testing reverse causation. Data limitations make it difficult to pool across samples (with differently spaced BMI measurements), to capture all weight changes that occur (as weight-cyclers, e.g., can be indistinguishable from people with stable weight) or to assess the cause of any particular weight change.

Ultimately, this study provides some of the first empirical evidence on reverse causation. The results, for adults 50 to 79, are broadly consistent with the hypothesis that illness-induced weight loss may cause elevated death rates at clinically normal BMIs. However, this study cannot assess whether illness-induced weight change is common, or has a large enough impact at the population level to meaningfully affect epidemiological inference. It will be difficult to assess this without population data that include much more frequent weight measurement than is common today.

REFERENCES

[1] Durazo-Arvizu R, McGee D, Cooper R, et al. Mortality and optimal body mass index in a sample of the US population. *Am J Epidemiol*. 1998; 147(8):739-749.

[2] Calle E, Thun M, Petrelli J, et al. Body-mass index and mortality in a prospective cohort of US adults. *N Engl J Med*. 1999; 341(15):1097-1105.

[3] McGee D, Diverse Populations Collaboration. Body mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies. *Ann Epidemiol*. 2005; 15(2):87-97.

[4] Gu D, He J, Duan X, et al. Body weight and mortality among men and women in China*. J Am Med Assoc*. 2006; 295(7):776-783.

[5] Dolan C, Kraemer H, Browner W, et al. Associations between body composition, anthropometry, and mortality aged 65 years and older. *Am J Public Health*. 2007; 97(5):913-918.

[6] Snih S, Ottenbacher K, Markides K, et al. The effect of obesity on disability vs mortality in older Americans. *Arch Intern Med*. 2007; 167(8):774-780.

[7] Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900,000 adults: collaborative analyses of 57 prospective studies. *Lancet*. 2009; 373(9669):1083-1095.

[8] Lin W, Tsai S, Albu J, et al. Body mass index and all-cause mortality in a large Chinese cohort. *Can Med Assoc J*. 2011; 183(6):E329-E336.

[9] Zheng W, McLerran D, Rolland B, et al. Association between body-mass index and risk of death in more than 1 million Asians. *N Engl J Med*. 2011; 364(8):719-729.

[10] Flegal KM, Kit BK, Orpana H, et al. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *J Am Med Assoc*. 2013; 309(1):71-82.

[11] Hu F, ed. *Obesity Epidemiology*. Oxford: Oxford University Press; 2008, pp. 149-215.

[12] Flegal K, Graubard B, Williamson D, et al. Reverse causation and illness-related weight loss in observational studies of body weight and mortality. *Am J Epidemiol*. 2011; 173(1):1-9.

[13] Adams K, Schatzkin A, Harris T, et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *N Engl J Med*. 2006; 355(8):763-778.

[14] Ajani U, Lotufo P, Gaziano M, et al. Body mass index and mortality among US male physicians. *Ann Epidemiol*. 2004; 14(10):731-739.

[15] Lindsted KD, Singh PN. Body mass and 26-year risk of mortality among women who never smoked: findings from the Adventist Mortality Study. *Am J Epidemiol*. 1997; 146(1):1-11.

[16] Losonczy K, Harris T, Cornoni-Huntley J, et al. Does weight loss from middle age to old age explain the inverse weight mortality relation in old age? *Am J Epidemiol*. 1995; 141(4):312-321.

[17] Manson J, Willett W, Stampfer M, et al. Body weight and mortality among women. *N Engl J Med*. 1995; 333(11):677-685.

[18] Manson JE, Stampfer MJ, Hennekens MD, et al. Body weight and longevity: a reassessment. *J Am Med Assoc*. 1987; 257(3):353-358.

[19] Willett W. Weight loss in the elderly: cause or effect of poor health? *Am J Clin Nutr* 1997; 66(4)**:**737-738.

[20] Mehta N, Chang V. Obesity and mortality. In Cawley J, ed. *The Oxford Handbook of the Social Science of Obesity*. New York: Oxford University Press; 2011:502-516.

[21] BMI in Diverse Populations Collaborative Group. Effect of smoking on the body mass index-mortality relation: empirical evidence from 15 studies. *Am J Epidemiol*. 1999; 150(12):1297-1308.

[22] Allison D, Heo M, Flanders D, et al. Examination of 'early mortality exclusion' as an approach to control for confounding by occult disease in epidemiologic studies of mortality risk factors. *Am J Epidemiol*. 1997; 146(8):672-680.

[23] Preston SH, Mehta NK, Stokes A. Modeling obesity histories in cohort analyses of health and mortality. *Epidemiology*. 2013; 24(1):158-166.

[24] Marmot M, et al. English Longitudinal Study of Ageing: Wave 0 (1998, 1999 and 2001) and Waves 1-4 (2002-2009) [computer file]. 15th Edition. Colchester, Essex: UK Data Archive [distributor]; April 2011. SN: 5050. (http://dx.doi.org/10.5255/UKDA-SN-5050-1).

[25] National Center for Health Statistics. NHANES I Epidemiologic Follow-up Study (NHEFS) [machine-readable data file and documentation]. Hyattsville, MD: National Center for Health Statistics [distributor].

[26] Eilers PHC, Marx BD. Flexible smoothing with *B*-splines and penalties. *Stat Sci*. 1996; 11(2):89-121.

[27] Albanes D, Jones DY, Micozzi MS, et al. Associations between smoking and body weight in the US population: analysis of NHANES II. *Am J Public Health*. 1987; 77(4):439-444.

[28] Finch CE. *Longevity, Senescence, and the Genome.* Chicago: The University of Chicago Press; 1990.

[29] Lee I, Paffenbarger R Jr. Change in body weight and longevity. *J Am Med Assoc*. 1992; 268(15):2045-2049.

[30] Andres R, Muller D, Sorkin J. Long-term effects of change in body weight on all-cause mortality: a review. *Ann Intern Med*. 1993; 119(7 Pt 2):737-743.

[31] Mikkelsen K, Heitmann B, Keiding N, et al. Independent effects of stable and changing body weight on total mortality. *Epidemiology*. 1999; 10(6):671-678.

[32] Yarnell J, Patterson C, Thomas H, et al. Comparison of weight in middle age, weight at 18 years, and weight change between, in predicting subsequent 14 year mortality and coronary events: Caerphilly Prospective Study*. J Epidemiol Community Health*. 2000; 54(5):344-348.

[33] Sauvaget C, Ramadas K, Thomas G, et al. Body mass index, weight change and mortality risk in a prospective study in India. *Int J Epidemiol*. 2008; 37(5):990-1004.

[34] Myrskyla M, Chang V. Initial BMI, weight change, and mortality among middle-and older-aged adults. *Epidemiology*. 2009; 20(6):840-848.

[35] Connor Gorber S, Tremblay M, Moher D, et al. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev*. 2007; 8(4):307-326.

[36] Shields M, Connor Gorber S, Tremblay MS. Estimates of obesity based on self-report versus direct measures. *Health Rep*. 2008; 19(2**)**:1-16.

[37] Chiolero A, Peytremann-Bridevaux I, Paccaud F. Associations between obesity and health conditions may be overestimated if self-reported body mass index is used. *Obes Rev*. 2007; 8(4):373-374.

[38] World Health Organization. *Global health risks: mortality and burden of disease attributable to selected major risks.* 2009; http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf. Accessed May 2, 2011.

Chapter 2

Disease and weight loss: A prospective study of middle-aged and older adults in Costa Rica and in England

Joint work with Noreen Goldman and Luis Rosero-Bixby

CHAPTER SUMMARY: *Doctors commonly recommend weight loss to improve patients' health. However, in population-based studies, weight loss is associated with higher mortality than weight gain or weight maintenance. We test whether disease predicts weight loss in two longitudinal surveys, as disease may confound, or account for, the unexpected relationship between weight loss and mortality. Data come from the Costa Rican Longevity and Healthy Aging Study (CRELES) and the English Longitudinal Study of Ageing (ELSA). We define two overlapping categories of weight loss: (a) a measured loss of more than 1.0 BMI point (formally, a change < -1.0 point) or (b) a measured loss of more than 2.0 BMI points (formally, a change < -2.0 points). In separate logistic regression analyses by country and weight-loss outcome, we estimate associations between weight loss and self-reported diseases, adjusting for age (range 52-79), sex, smoking, and initial BMI. In ELSA, diabetes onset is associated with BMI change*

*< -1.0 point (*OR *= 1.94 [95% CI: 1.29 – 2.93]), and onset of diabetes, cancer, or lung disease is associated with change < -2.0 points (respectively,* OR *= 2.25 [95% CI: 1.34 – 3.80];* OR *= 2.70 [95% CI: 1.49 – 4.89];* OR *= 1.82 [95% CI: 1.02 – 3.26]). In CRELES, disease-onset reports are not associated with weight loss at 5% significance. These findings suggest disease may predict weight loss in population surveys. However, associations vary considerably by survey and weight-loss definition.*

Weight loss is a subject of intense popular and scholarly interest. In developed countries, most adults claim they have tried to lose weight, either to improve their health or their looks (1-3). Researchers test interventions for weight loss promotion (4, 5), and aim to quantify the benefits of weight loss for blood pressure management (6), diabetes risk (7), and other cardiovascular risk factors (8). Despite widespread belief that moderate weight loss would benefit most adults living in rich countries, however, population-based studies find that people losing weight have higher death rates, not lower ones, than weight-gainers or weight-maintainers (e.g. (9-12)). This paper aims to reconcile the apparent conflict: a presumed benefit of weight loss for most individuals, but a penalty associated with weight loss at the population level.

In particular, this study tests whether clinical disease may predict weight loss in population-based studies of middle-aged and older adults. Clinicians consider unintentional weight loss a

symptom of many life-threatening conditions, including cancer, chronic lung disease, and congestive heart failure, (13). But while there is evidence that disease leads to weight loss in select clinical populations (e.g., (14, 15)), it is unclear whether disease also commonly drives weight loss in non-clinical samples, in which extreme sickness may be rare. Two retrospective studies in the USA link unintentional weight loss to disease or poor health (16, 17), but these results have not been replicated in prospective data. Moreover, among participants in the Americans' Changing Lives Survey, Kahng, Dunkle, and Jackson find no effect of health on change in body mass index (BMI), and find, instead, that falling BMI tends to *precede* health decline (18). Prospective studies of weight loss and physical function or mobility yield ambiguous results, varying by participant characteristics (19, 20).

This study tests for relationships between self-reported clinical diagnoses and subsequent or concurrent changes in BMI, using prospective data from two population-based surveys: one in Costa Rica and one in England. Crucially, these data include multiple waves with measured height and weight: an advantage over previous population studies of weight change, which have typically relied on self-reported values for at least one weight measurement. Self-reports can be biased, as people tend to overstate height and understate weight, but also noisy, with a good deal of random error (21). Survey respondents may be especially unlikely to know and report their true weight if weight is changing rapidly, such that surveys could collect the least accurate information about the events that would otherwise be most informative. Avoiding this pitfall, our study allows a more robust population-level investigation of disease and weight loss than has previously been possible.

MATERIALS AND METHODS

Data

Data come from two longitudinal surveys: the Costa Rican Longevity and Healthy Aging Study (CRELES) and the English Longitudinal Study of Ageing (ELSA) (22). These data are described in Appendix 2A. In brief, each survey is nationally representative, within certain age limits, of the population from which it is drawn. Each measures participants' height and weight over three waves, and each collects detailed participant reports about health history and recent diagnoses. Because relationships between BMI and health may vary by age, especially among the elderly (23), we limit the sample to participants below age 80 at the end of the weight-change measurement interval (see below). By studying two culturally distinct populations, we hope to assess the constancy of relationships across societies.

Survey waves.

Table 2.1 shows measurement timing in CRELES and ELSA. For simplicity, we refer to waves with initial height and weight measurements as "time 1," to waves with the second height and weight measurements as "time 2," and to waves with the third as "time 3," even though these "times" do not occur in the same years in England as in Costa Rica. While intervals from time 1 to time 2 and from time 2 to time 3 are essentially the same length *within* either population, they are shorter in Costa Rica (1.5 to 2.5 years) than they are in England (3 to 4 years). This prevents direct comparison of the magnitude of results between countries.

Table 2.1. Survey Measurement in Costa Rica (CRELES) and in England (ELSA).

	CRELES	*ELSA*
Country of study	Costa Rica	England
Year of first BMI measurement ("time 1")	2004-06	2001
Year of second BMI measurement ("time 2")	2006-08	2004-05
Year of third BMI measurement ("time 3")	2009-10	2008-09
Length of interval between BMI measurements	1.5 – 2.5 years	3 – 4 years
Age range (at the end of each BMI-change interval)	61-79	52-79
Number of participants with valid BMI at two adjacent waves	1,285	4,423

Eligible sample.

Because this study deals with weight change, we include only participants with valid height and weight measurements at two adjacent waves: i.e., at both times 1 and 2, or at both times 2 and 3. We assume there is no underlying time trend in the propensity to lose weight, so that within each population the relationship between disease and weight loss from times 1 to 2 will be the same as from times 2 to 3. (Analysis of each period separately supports this view; results not shown.) We include all observations from participants observed during only one weight-change interval (i.e., from time 1 to 2 or from time 2 to 3). For participants observed during both intervals, we include only the first observation. This is because, in both CRELES and ELSA, individuals' weight-change outcomes are negatively correlated. (I.e., weight cycling across the three waves is more common than continued gain or loss.) Common techniques for handling repeated measures data are thus not practical, given our observations. Random-intercept models and fixed-effects models are designed for positive, rather than negative, correlation among observations from the same individual. More flexible random-effects models (such as growth-curve models) are theoretically justified but difficult to implement using our data, which include few observations per person: a maximum of three BMI measurements per respondent (two weight-change intervals), but often fewer.

Survey attrition.

As noted above, the sample is a subset of initial CRELES and ELSA respondents: those who participate at two adjacent waves. This sample may not be representative of the study populations at baseline. In each country, 3% to 5% of participants die between any two waves.

Attrition among the living ranges from just under 10% (between times 1 and 2 in CRELES) to roughly one third (between ELSA's first wave, collected as part of another national study not designed to be longitudinal, and the first ELSA follow-up).

Because attrition is likely correlated with disease, we do not describe either typical magnitude or prevalence of illness-induced weight loss within a population. Instead, we examine how health status predicts weight loss over 1.5 to 4 years among a group with continued participation in voluntary, longitudinal surveys. These results may not be generalizable to the full population. However, understanding links between weight loss and disease in population-based surveys is important, as previous studies link weight loss over similar intervals to increased mortality (9-12). Few studies examine the cause (24).

Definition of "weight loss."

We treat weight loss as a categorical variable, examining associations with disease at the tails of the weight-change distribution, rather than at the mean. Using a categorical marker instead of a continuous one will lead to some loss of information (see sensitivity analysis, below). However, we choose to analyze weight loss as dichotomous because many health conditions appear more common among both weight losers *and* weight gainers (see Appendix 2A). It is possible, therefore, that a disease could be associated with weight loss beyond a given threshold without being associated with mean weight change. This may occur either because a disease causes weight variability (gain or loss), or because disease causes weight loss but

observations of net weight change between waves capture periods both of sickness (weight loss) and of recovery (weight gain).

We calculate BMI at each wave as weight in kilograms divided by the square of height in meters. We then define two (overlapping) categories of weight loss: (a) a measured BMI change between two waves of < -1.0 BMI point (i.e., a loss of more than 1.0 BMI point), and (b) a measured BMI change of < -2.0 BMI points (i.e., a loss of more than 2.0 BMI points). BMI change is driven overwhelmingly by weight change, as participants' heights fluctuate only very slightly (see Appendix 2A). We select two different thresholds for BMI loss because some traits could be associated with extreme loss, but not with milder loss. Changes of 1.0 and 2.0 BMI points are both large enough they are unlikely to be achieved through daily variation in water mass. For a person who is 5'10" (1.78 m), for example, a change of 1.0 BMI point corresponds to a weight change of 7.0 lbs (3.2 kg) and a change of 2.0 BMI points corresponds to a change of 13.9 lbs (6.3 kg).

Disease diagnoses.

CRELES and ELSA participants report lifetime and recent disease diagnoses at time 1, and all new diagnoses since the previous interview at times 2 and 3. We focus on six major conditions: cancer, diabetes, heart attack, stroke, arthritis, and lung disease. We consider a condition present "at baseline" if a respondent reports it at the survey wave starting the weight-change interval. We consider "disease onset" to occur if the respondent does not report the disease at the initial wave, but reports it at the next wave.

Disease definitions are almost identical in CRELES and ELSA, with three exceptions. Diabetes status in ELSA is determined solely by self-report, whereas in CRELES it is determined in part by fasting glucose and glycosylated hemoglobin tests administered during the survey. In ELSA, study participants are asked if a doctor has ever diagnosed "chronic lung disease such as chronic bronchitis or emphysema"; in CRELES, the lung-disease question includes asthma and tuberculosis as well. Finally, in CRELES, questions about heart attack are specific to that event, whereas ELSA combines three serious manifestations of cardiovascular disease – "heart attack, angina, and congestive heart failure" – into a single marker. This study therefore tests for associations of weight loss with heart attack in Costa Rica, but with heart disease more broadly in England.

Regression analyses

In four separate logistic regression analyses – one for each weight-loss outcome in each country – we estimate the effects of clinical diagnoses on weight loss, adjusting for age, sex, smoking status (current, former, never, measured at the start of the weight-change interval), and initial BMI. All analyses are performed in Stata 12 (StataCorp LP, College Station, TX).

Odds ratios associated with continuous variables (i.e., age, BMI) are calculated for a two-standard-deviation change (approximately), as this makes the magnitude of estimates from continuous variables comparable to those from binary variables, at least when the mean of the

binary variable is close to 0.5 (25). Confidence intervals may be used to test whether any given odds ratio is different from one at a specified level of statistical significance. However, due to the large number of parameters tested (68, excluding sensitivity analysis, below), one would expect to find roughly three associations significant at 5% even if there were no true relationships. Each association, taken alone, should therefore be viewed as suggestive.

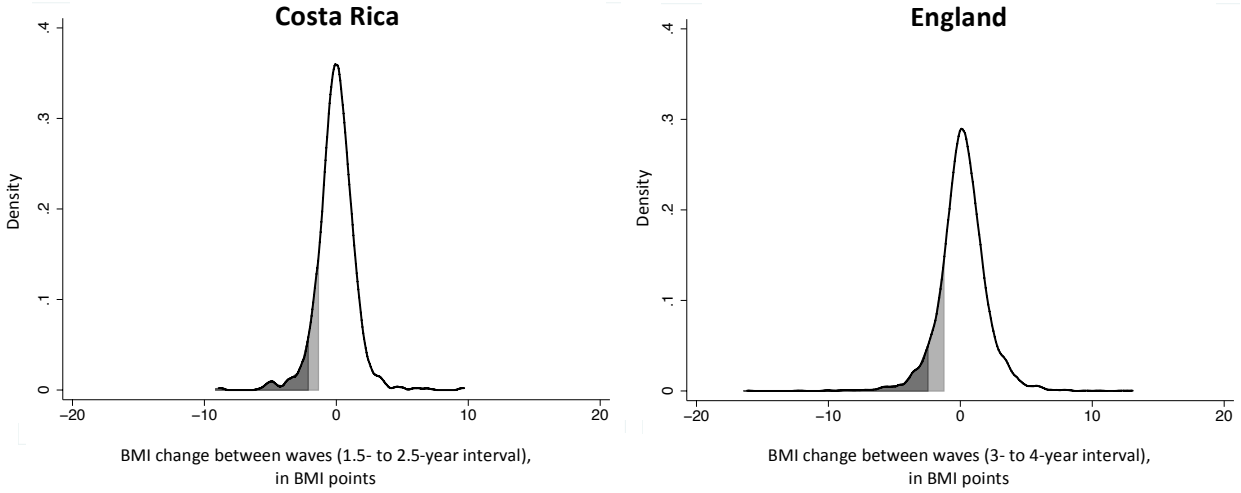
Sensitivity analysis.

To gauge robustness of results, we test alternative model specifications. First, we examine how results change with different definitions of disease or ill health as predictors (including composite measures of disease and participants' self-assessments of their overall health, rated as excellent, very good, good, fair, or poor). Next, we examine how results change with a different functional form: either linear regression or multinomial logistic regression (with outcomes for both weight loss and weight gain). These robustness checks assess, among other things, the degree to which important information is lost by modeling BMI change as a dichotomous outcome rather than as a continuous one. Robustness checks are described in Appendix 2B.

RESULTS

Figure 2.1 shows density plots of BMI change. These plots show a wider range of observed changes in England (where BMI change is measured over 3 to 4 years) than in Costa Rica (where

Figure 2.1. BMI Changes Observed in Costa Rica and in England. Values from Costa Rica represent net changes in BMI between the CRELES waves in 2005 and 2007, or between the waves in 2007 and 2009. Values from England represent net changes between ELSA survey waves in 2001 and 2004-05, or between the waves in 2004-05 and 2008-09. Dark grey shaded areas show values < -2.0 BMI points; light grey areas show BMI-change values between -1.0 points and -2.0 points.



the interval is 1.5 to 2.5 years). Table 2.2 summarizes these BMI-change distributions and shows total self-reported disease prevalence. Prevalence of each condition is listed in Appendix 2A.

Figure 2.2 shows odds ratios from logistic regression of BMI loss more than 1.0 point on self-reported clinical conditions, adjusted for age, sex, smoking status, and starting BMI. Figure 2.3 shows results from the analyses of BMI loss more than 2.0 points. In Figure 2.2, we observe associations between diabetes in ELSA and BMI loss more than 1.0 point (*OR* = 1.63 for diabetes at baseline [95% CI: 1.23 – 2.17], and *OR* = 1.94 for onset during the weight-change interval [95% CI: 1.29 – 2.93]). No other conditions are associated at a 5% level of significance in either country. However, as shown in Figure 2.3, when we limit the outcome to more extreme weight loss greater than 2.0 BMI points, associations with diabetes in England appear stronger (*OR* = 2.08 for disease at baseline [95% CI: 1.45 – 2.99], and *OR* = 2.25 for onset [95% CI: 1.34 – 3.80]). We further observe associations at 5% in England between cancer onset and weight loss (*OR* = 2.70 [95% CI: 1.49 – 4.89]) and between lung disease onset and weight loss (*OR* = 1.82 [95% CI: 1.02 – 3.26]). In Costa Rica, no reports of new disease onset are significantly associated with weight loss, but Figure 2.3 shows a relationship between baseline arthritis and weight loss more than 2.0 BMI points (*OR* = 2.00 [95% CI: 1.08 – 3.72]).

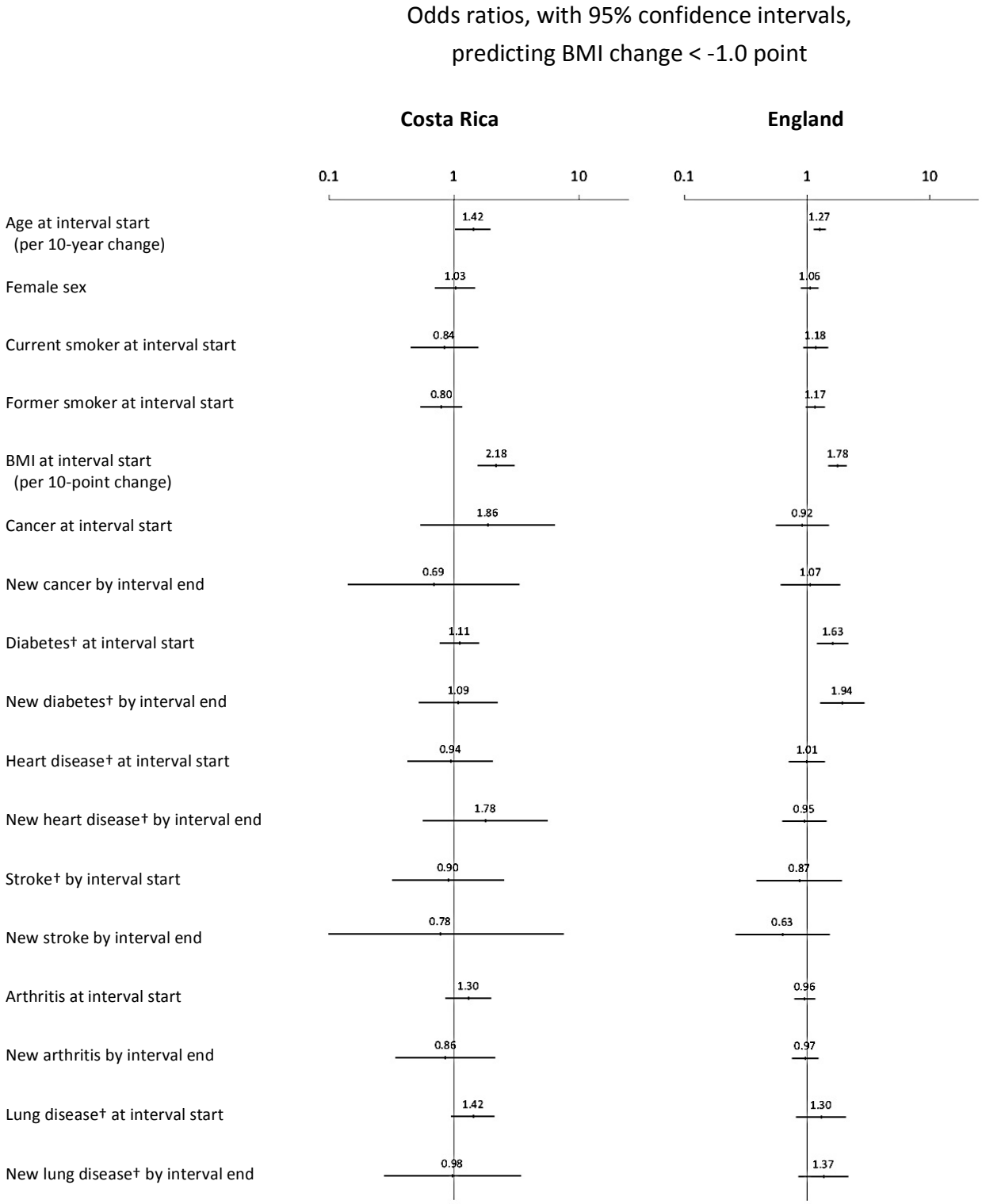
Sensitivity analyses (Appendix 2B) yield consistent results. When health status is measured either by self-rated general health or by composite measures of disease, we find no additional

Table 2.2. BMI Changes Between Adjacent Survey Waves. Values from CRELES reflect changes between the waves in 2005 and 2007, or between the waves in 2007 and 2009. Values from ELSA represent changes between waves in 2001 and 2004-05, or between the waves in 2004-05 and 2008-09.

	CRELES	*ELSA*
Mean BMI change (in BMI points)	-0.02	+0.20
Proportion with BMI change < - 1.0 BMI point	18.5%	19.8%
Proportion with BMI change < - 2.0 BMI points	5.6%	8.4%
Proportion reporting any disease*	55.9%	49.3%

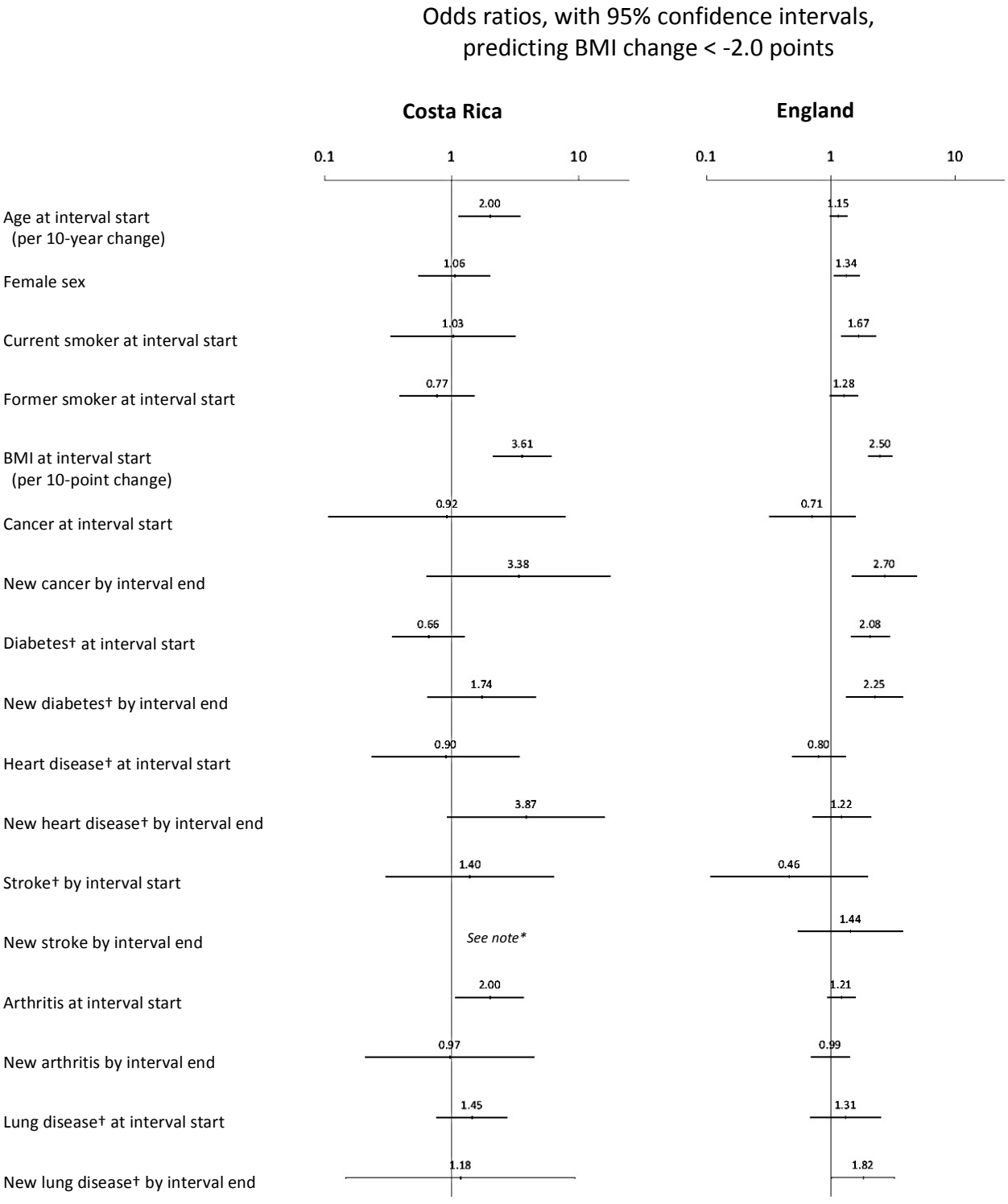
* Respondent reports any of the six conditions analyzed (diabetes, cancer, heart attack, stroke, arthritis, or chronic lung disease), either at baseline or newly occurring during the weight-change interval

Figure 2.2. Self-Reported Disease Diagnoses and BMI Change < -1.0 BMI Point.



† See Chapter 2 text for differences in disease definitions between CRELES and ELSA.

Figure 2.3. Self-Reported Disease Diagnoses and BMI Change < -2.0 BMI Points.



† See Chapter 2 text for differences in disease definitions between CRELES and ELSA.

* No CRELES participant experiences new stroke onset and BMI change < - 2.0 BMI points.

relationships in CRELES; in ELSA, we find associations between BMI change < -2.0 points and composite measures ("any chronic illness" or "any chronic illness that limits activity") and also between self-rated health (as a continuous marker) and BMI change of this magnitude. These findings, like those from the primary analysis, suggest a link in ELSA between extreme weight loss and poor health, but provide little evidence of relationships between disease and milder weight loss in ELSA, or between disease and any weight loss in CRELES.

Results from linear regression, shown in full in Appendix 2B, are generally weaker than those from the primary analysis. Diabetes onset in England is the sole condition associated at 5% significance with BMI change as a continuous variable (associated with a change of -0.44 BMI points [95% CI: -0.79 – -0.09]). However, changes associated with baseline diabetes and cancer onset in England are marginally significant (respectively, -0.23 [95% CI: -0.47 – 0.00]) and -0.40 [95% CI: -0.81 – 0.01]). Results from multinomial logistic models are very similar to those from the primary analysis.

DISCUSSION

This study is among the first to examine clinical characteristics of people losing weight in population-based, prospective data. We analyze two national surveys, one in Costa Rica and one in England, each collecting measured height and weight and detailed self-reports of clinical diagnoses.

We find a relatively complicated relationship between health status and weight change. Associations between disease and weight loss may vary both by survey and by the scale of weight loss considered, with more apparent relationships between disease and weight loss in ELSA than in CRELES, and more apparent relationships of disease with major BMI loss greater than 2.0 points than with more broadly defined BMI loss greater than 1.0 point. The larger number of significant associations between disease and major weight loss may suggest that mild weight loss can have many causes (including, e.g., improved diet and physical activity), while people who suffer major weight loss are unusually likely to be sick. Observed variation by country could reflect differences in the underlying populations of Costa Rica and of England (e.g., in the burden of disease), or could reflect simply the longer interval between BMI measurements in ELSA than in CRELES or differences in survey sampling (e.g., the slightly different age ranges covered, or difference in sample sizes). Most likely, it is some combination of all of these.

A major goal of this study is to isolate specific clinical conditions that could explain the population-level association between weight loss and mortality (9-12). Existing empirical evidence on the subject is scarce, but few alternative explanations, besides disease, have been offered for the weight-loss–mortality relationship. In ELSA, we find links between major weight loss and new cancer diagnosis, new lung disease diagnosis, and diabetes diagnosis, whether reported at baseline or newly emerging during the weight-change interval. Cancer and lung disease are recognized causes of unintentional weight loss (26), so these associations are

consistent with the theory that weight loss predicts mortality because weight loss can signal serious illness. An association between diabetes and BMI loss may appear less consistent, since weight loss among diabetics has been shown to improve prognosis (27). However, diabetes status may confound the relationship between weight loss and mortality, if, for example, diabetics are unusually likely to try to lose weight and succeed. In CRELES, the relationship between baseline arthritis and BMI loss more than 2.0 points could be causal if mobility restrictions limit participants' ability to buy, prepare, or consume food. However, the number of disease associations observed in Costa Rica as significant at 5% is no greater than one would expect by chance alone: one of 24 parameters tested.

Together, these results provide equivocal support for the hypothesis that clinical disease drives weight loss in most population-based surveys. It may thus be worth considering, at least, other possible explanations for the link between weight loss and mortality, especially since these have been underdeveloped to date. It is possible that attempts to lose weight are themselves harmful on occasion (28). However, it seems unlikely they would frequently be fatal. A more plausible alternative, perhaps, is that weight loss can signal "frailty": in gerontology, a broadly defined syndrome, characterized by loss of muscle mass and bone density in addition to body fat (29, 30). Distinct from clinical disease and functional limitations, frailty has nonetheless been linked to lethargy, weakened immune response, and increased risk of falls and hospitalization (29, 30). This syndrome could explain, for example, the results of Kahng, Dunkle, and Jackson, who find that weight loss often precedes health decline, rather than health decline preceding weight loss (18).

Our findings about disease and weight loss are in some ways exploratory, as we test for a large number of possible associations in two different countries. This strategy is motivated by a lack of theoretical or empirical work to inform methodological decisions about models of disease and weight loss. In our primary analysis, we use categorical outcomes for weight loss rather than continuous ones. However, robustness checks with linear regression and continuous outcomes do not yield substantively different or more informative results. Because we find stronger associations between disease and extreme weight loss than between disease and milder weight loss or between disease and mean BMI change, future scholars may wish to test, in particular, for relationships between health and weight change at the left tails of the weight-change or BMI-change distributions.

This study has four limitations. First, despite relatively high-quality data, measurements of weight change and disease onset are both crude because we observe people only 1.5 to 4 years apart. Among CRELES and ELSA respondents who participate at two adjacent waves, we observe only net BMI change between the waves. Some people who lose weight due to sickness may therefore still be observed as net gainers or maintainers. In addition, we do not observe the exact sequence of disease and weight-change events between waves. That is, when disease is diagnosed between survey waves, we cannot determine whether weight change during the interval occurs simultaneously with disease progression, whether disease precedes weight change, or whether weight change precedes disease (either because preclinical disease is causing weight loss, or because weight loss makes people susceptible to disease). Our study therefore highlights the importance of measurement timing. To fully assess relationships

between disease and weight change, even among a sample of survivors without attrition, one would need almost continuously collected data on weight, disease, and preclinical disease indicators.

Second, disease diagnoses are self-reported. Self-reports may be imperfect measures of true disease status, either because participants do not know they are sick, do not know the name of the disease that affects them, or choose not to report their illness in a survey. Each disease-diagnosis category may also represent a broad range of disease severity. It is possible we would observe different relationships between clinical conditions and weight loss if, for example, we could distinguish stage I cancers from stage IV cancers.

Third, we may lack statistical power to detect some relationships. Low power is a particular problem in CRELES (*N* = 1,285), where power to reject the null hypothesis is less than 50% for three quarters of the statistical tests, assuming that the true odds ratio associated with a predictor is 2. In contrast, in England, power is greater than 70% for nearly two thirds of the statistical tests, and greater than 90% for one third of them. Appendix 2C lists results from power tests.

Fourth, results may not be widely generalizable. We estimate associations between disease and weight loss among survey respondents who participate at two adjacent waves, spaced 1.5 to 4 years apart. These associations are valuable for explaining previous findings that link weight loss over similar periods to high death rates (9-12). However, results cannot be interpreted as

the causal effects of disease on weight loss among the general population. In the absence of survey attrition, associations between disease and weight loss would most likely be larger than observed here, since survey exit and death are very likely correlated with sickness.

Despite these limitations, we believe our findings represent the clinical determinants of weight loss among Costa Rican and English adults, aged 52 to 79, who participate in longitudinal, population-based surveys. Epidemiologists rely on such voluntary surveys for inference about the effects of BMI on health. Yet our findings suggest both that one-time BMI measurements are noisy measures of long-term BMI exposure (because weight change is relatively common), and that the impact of health on BMI (reverse causation) may be complex and difficult to measure. Ultimately, disease may predict weight loss in population-based studies, but estimated associations appear sensitive both to the survey sample and to the magnitude of weight loss considered.

REFERENCES

[1] Saad L, Gallup. To lose weight, Americans rely more on dieting than exercise. Princeton, NJ: Gallup; 2011. (http://www.gallup.com/poll/150986/lose-weight-americans-rely-dieting-exercise.aspx). (Accessed July 1, 2013).

[2] The Heart and Stroke Foundation of Canada. Canadians trapped on weight-loss rollercoaster. Ottawa, ON: Heart and Stroke Foundation of Canada; 2010. (http://www.heartandstroke.com/site/apps/nlnet/content2.aspx?c=ikIQLcMWJtE&b=3485819&ct=8136053). (Accessed July 1, 2013).

[3] Yoong SL, Carey ML, Sanson-Fisher RW, et al. A cross-sectional study assessing the self-reported weight loss strategies used by adult Australian general practice patients. *BMC Family Practice*. 2012; 13:48.

[4] Jakicic JM, Clark K, Coleman E, et al. American College of Sports Medicine position stand. Appropriate intervention strategies for weight loss and prevention of weight regain for adults. *Med Sci Sport Exer*. 2001; 33(12):2145-2156.

[5] Miller WC, Koceja DM, Hamilton EJ. A meta-analysis of the past 25 years of weight loss research using diet, exercise or diet plus exercise intervention. *Int J Obesity*. 1997; 21:941-947.

[6] Neter JE, Stam BE, Kok FJ, et al. Influence of weight reduction on blood pressure: a meta-analysis of randomized controlled trials. *Hypertension*. 2003; 42:878-884.

[7] Sjostrom CD, Peltonen M, Wedel H, et al. Differentiated long-term effects of intentional weight loss on diabetes and hypertension. *Hypertension*. 2000; 36:20-25.

[8] Esposito K, Pontillo A, Di Palo C, et al. Effect of weight loss and lifestyle changes on vascular inflammatory markers in obese women: a randomized trial. *J Am Med Assoc*. 2003; 289(14):1799-1804.

[9] Lee I, Paffenbarger RS. Change in body weight and mortality. *J Am Med Assoc*. 1992; 268(15):2045-2049.

[10] Losonczy K, Harris T, Cornoni-Huntley J, et al. Does weight loss from middle age to old age explain the inverse weight mortality relation in old age? *Am J Epidemiol*. 1995; 141:312-321.

[11] Wannamethee SG, Shaper AG, Lennon L. Reasons for intentional weight loss, unintentional weight loss, and mortality in older men. *Arch Intern Med*. 2005; 165(9):1035-1040.

[12] Myrskyla M, Chang VW. Weight change, initial BMI, and mortality among middle- and older-aged adults. *Epidemiol*. 2009; 20(6):840-848.

[13] Willett W. Weight loss in the elderly: cause or effect of poor health? *Am J Clin Nutr*. 1997; 66:737-738.

[14] Tisdale MJ. Mechanisms of cancer cachexia. *Physiol Rev*. 2009; 89:381-410.

[15] Von Haehling S, Lainscak M, Springer J, et al. Cardiac cachexia: a systematic overview. *Pharmacol Therapeut*. 2009; 121(3):227-252.

[16] Meltzer AA, Everhart JE. Unintentional weight loss in the United States. *Am J Epidemiol*. 1995; 142(10):1039-1046.

[17] French SA, Jeffery RW, Folsom AR, et al. Relation of weight variability and intentionality of weight loss to disease history and health-related variables in a population-based sample of women aged 55-69 years. *Am J Epidemiol*. 1995; 142(12):1306-1314.

[18] Kahng SK, Dunkle RE, Jackson JS. The relationship between the trajectory of body mass index and health trajectory among older adults: multilevel modeling analyses. *Res Aging*. 2004; 26(1):31-61.

[19] Lee JS, Kritchevsky SB, Tylavsky F, et al. Weight change, weight change intention, and the incidence of mobility limitation in well-functioning community-dwelling older adults. *J Gerontol A Biol Sci Med Sci*. 2005; 60(8):1007-1012.

[20] St-Arnaud-McKenzie D, Payette H, Gray-Donald K. Low physical function predicts either 2-year weight loss or weight gain in healthy community-dwelling older adults: The NuAge Longitudinal Study. *J Gerontol A Biol Sci Med Sci*. 2010; 65A(12):1362-1368.

[21] Connor Gorber S, Tremblay M, Moher D, et al. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev*. 2007; 8:307-326.

[22] Marmot M, et al. English Longitudinal Study of Ageing: Wave 0 (1998, 1999 and 2001) and Waves 1-4 (2002-2009) [computer file]. 15th Edition. Colchester, Essex: UK Data Archive [distributor]; April 2011. SN: 5050. (http://dx.doi.org/10.5255/UKDA-SN-5050-1).

[23] Stevens J, Cai J, Pamuk ER, et al. The effect of age on the association between body-mass index and mortality. *New Eng J Med*. 1998; 338:1-7.

[24] Wannamethee SG, Shaper AG, Walker M. Weight change, body weight and mortality: the impact of smoking and ill health. *Int J Epidemiol*. 2001; 30(4):777-86.

[25] Gelman A, Hill J. *Data analysis using regression and multilevel/hierarchical models*. Cambridge, UK: Cambridge University Press; 2007:57.

[26] Goldman L, Ausiello D, eds. *Cecil Medicine.* 23rd ed. Philadelphia, PA: Saunders Elsevier; 2008:350, 619, 1353.

[27] Lean ME, Powrie JK, Anderson AS, et al. Obesity, weight loss and prognosis in type 2 diabetes. *Diabet Med*. 1990; 7(3):228-233.

[28] Kassirer JP, Angell M. Losing weight – an ill-fated New Year's resolution. *New Engl J Med*. 1998; 338(1):52-54.

[29] Xue Q. The frailty syndrome: definition and natural history. *Clin Geriatr Med*. 2011; 27(1):1-15.

[30] Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001; 56(3):M146-M157.

Chapter 3

Nutritional reserve or reverse causation? Comparing rival hypotheses of the curvilinear relationship between body mass index and mortality

CHAPTER SUMMARY: *Many scholars believe that sickness-induced weight loss can explain high death rates observed among people with clinically normal body weight, relative to people who are overweight. However, concepts of "reverse causation" – that is, sickness leading to weight loss, rather than normal weight leading to sickness – are rarely formalized and rarely described in detail. This paper presents three competing models to explain the curvilinear relationship between body mass index (BMI) and mortality. The paper then simulates hypothetical populations living under each of the three, demonstrating that any one of the models could explain real-world phenomena. In the future, empirical studies that rely heavily on reverse causation for interpretation of the BMI–mortality curve may wish to specify how, exactly, reverse causation is expected to operate. The complexity of such models has been underappreciated. By clarifying presumed causal mechanisms, however, scholars may also elucidate the implied consequences of their models. This, in turn, may generate new empirical predictions, creating new opportunities for hypothesis testing.*

Body mass index (BMI) has been linked to cardiovascular disease, diabetes, and some types of cancer, with higher BMI associated with higher disease risk (1). Because of this, many scholars suspect that all-cause mortality risk, too, rises monotonically with body weight (e.g., (2-4)) – at least for the large majority of developed-world adults (5, 6) who are not underweight (where "underweight," a sign of chronic malnutrition (7), is defined as BMI < 18.5). Claims of a monotonic relationship between BMI and mortality, however, are at odds with most population-based studies worldwide, which instead find a curvilinear, or non-monotonic, relationship between BMI and mortality among people who are not underweight (8-16). Moreover, in such population-based studies, the lowest death rates are typically observed among people who are *over*weight (BMI ≥ 25 and < 30), rather than among those with clinically normal body weight (BMI ≥ 18.5 and < 25) (5). The association between BMI and mortality is therefore hotly contested, as scholars argue about possible data artifacts that could make overweight status appear protective to longevity, even if it were in fact harmful (17, 18).

This paper addresses some of these possible data artifacts. To date, empirical research on the causes of the BMI–mortality curve has been inconclusive (see Chapters 1 and 2). This study therefore moves away from data analysis and, instead, uses simple mathematical models and simulation to explore *potential* explanations for the curvilinear BMI–mortality relationship. That is, instead of asking, "Do we find empirical support for an explanation of low death rates observed among overweight people?," the paper poses a new question: "Are different proposed hypotheses even capable of producing the relationships we observe?"

The question is not trivial because, often, proposed hypotheses go un-formalized, with key details omitted. As others note elsewhere (19), one example of an indistinct hypothesis like this is “reverse causation”: the hypothesis that, all else being equal, death risk increases with BMI (excluding the underweight range), but that sickness may lead to weight loss and to low body weight, thus creating a spurious association between clinically normal BMI and mortality. Reverse causation of this type may seem intuitive. However, the all-important details of sickness-induced weight loss – including which people suffer, when, how much, and with what mortality penalty – are rarely stated explicitly (19). Without these details, it is difficult to judge whether reverse causation is sufficient to explain the longevity advantage observed among the overweight.

This paper assesses three competing hypotheses for the curvilinear relationship between BMI and mortality among people who are not underweight. Of these three rival hypotheses, one assumes it is optimal for longevity to be overweight, all else being equal; two assume it is in fact optimal to be normal weight. The three hypotheses are outlined below with supporting evidence and, for each, a simple model is constructed to describe its causal mechanisms. The paper then simulates three hypothetical populations, each one living under the model implied by one of the three hypotheses.

The paper thus makes a number of contributions. First, it clarifies the assumptions of popular hypotheses about body weight and mortality risk, lending badly needed conceptual clarity to a

field with many contentious findings (17, 18). Second, it demonstrates by simulation whether proposed ideas are sufficient to explain the phenomena they are supposed to explain; this provides theoretical backing, if not empirical evidence, to support or reject various hypotheses. Finally, the paper briefly examines implications of each model considered, as these may be used to assess further the plausibility of different hypotheses. The results may eventually be used to inform new empirical tests, or decisions about realism v. parsimony in models of body weight and health.

MATERIALS AND METHODS

Three competing hypotheses

Consider three alternative hypotheses for the curvilinear relationship between BMI and mortality. These are not exhaustive, but each has some biological support. Each, furthermore, may be presented as a simple model, with just a few parameters.

For simplicity, the three models constructed to describe these hypotheses may share some basic features. Assume in each case a single-sex population of 30-year-olds. Assume that each member of the population, at that age, falls into one of three weight categories: I, II, or III. These may be likened to the World Health Organization's "normal," "overweight," and "obese" classifications: respectively, BMI ≥ 18.5 and < 25; BMI ≥ 25 and <30; and BMI ≥ 30.

("Underweight" is ignored here because less than 2% of U.S. adults are underweight (6), and because, as noted above, controversy over the BMI–mortality relationship centers on high death rates at clinically normal weight, relative to overweight.) Over time, the population of 30-year-olds ages. People may move from one weight category to another; people may die. Death rates can vary with age and with weight category. Beyond these shared features, model details are dictated by the assumptions of the three hypotheses.

Hypothesis #1

The "nutritional reserve" explanation.

This is perhaps the simplest explanation for the curvilinear relationship between BMI and mortality: namely, that observed patterns represent a true, causal relationship. All else being equal, a person has lower mortality if he is overweight than if he either is obese or has clinically normal body weight.

This hypothesis is not popular, given strong evidence to link overweight status to hormonal dysregulation and to chronic disease (1). However, humans have evolved, clearly, to gain weight in an environment of abundance, and this suggests that extra fat must be adaptive in at least some circumstances. Fat stores could serve as an emergency energy source – a nutritional reserve – during times of health crisis such as infection or food shortage. These stores could also provide padding against falls or other injury; for example, overweight and obese women

appear to have lower rates of hip fracture than women who are thinner (20). Finally, it has been suggested that the same adipose tissue that increases disease risk among the healthy may actually work to reduce systemic inflammation and its damage in the presence of life-threatening diseases such as heart failure (21). This may help to explain the well-known "obesity paradox," in which heavier people appear to have higher risk of disease onset, but lower risk of death once sick (22-24). In short, it is possible that a great deal of body fat is deleterious to health, but that some is beneficial, and that overweight status (but not obesity) is therefore optimal for longevity.

This explanation is straightforward to model. For the model population described above – a single-sex cohort, aged 30, with members distributed among three weight categories – assume a conventional Gompertz model of mortality (25). That is, as the cohort ages, the logarithm of the death rate rises linearly with age (i.e., death rates increase exponentially). Assume also, for convenience, that death rates are proportional by weight category (i.e., proportional hazards by weight category), and that, on average, people gain weight as they age, but that the rate of transition from weight category I to weight category II, or from weight category II to weight category III, will decrease with age.

Given this, we may define a death rate, d, for an individual or group, i, as follows:

$$d_i = \begin{cases} \mu \exp(z x_i) & \text{if } i \text{ is weight category I} \\ \mu \exp(z x_i)\, \beta_2 & \text{if } i \text{ is weight category II} \\ \mu \exp(z x_i)\, \beta_3 & \text{if } i \text{ is weight category III} \end{cases}$$

where

μ is the baseline mortality hazard, the death rate at age 30

z describes the increase in mortality risk with age, assuming that the multiplicative effect on baseline hazard, μ, at age $x + 30$ is $\exp(zx)$

β_2 describes the multiplicative effect of weight category II on mortality risk, relative to category I, and

β_3 describes the multiplicative effect of weight category III on mortality risk, relative to category I.

A rate of transitioning up a weight category, g, for an individual or group, i, may be defined as follows:

$$g_i = \begin{cases} \gamma \exp(\gamma_a x_i) & \text{if } i \text{ is weight category < III} \\ 0 & \text{otherwise} \end{cases}$$

where

γ is the baseline rate – i.e., the rate at age 30 – of moving up one weight category (to a maximum of weight category III), and

γ_a describes the change with age in the rate of moving up one weight category (to a maximum of weight category III), assuming that the multiplicative effect on rate, γ, at age $x + 30$ is $\exp(\gamma_a x)$.

Since this model assumes that nobody loses weight over time, a rate of transitioning down a weight category, l, is defined for all individuals as follows:

$$l_i = 0$$

In sum, the nutritional reserve model may be described with just 6 essential parameters: μ, z, β_2, β_3, γ, and γ_a. A nice feature of this model, furthermore, is that most of these parameter values are known, since the model assumes explicitly that observed relationships are true relationships. Parameter values for μ, z, β_2, and β_3 may all thus be set from empirical observation.

Hypothesis #2

Conventional reverse causation: Serious illness leads to weight loss.

This is probably the most common interpretation of "reverse causation" with respect to body weight. Under this hypothesis, all else being equal, a person has lower mortality if he has clinically normal weight than if he is overweight. However, some people become sick, and sickness may cause weight loss, creating a spurious association between normal weight and mortality. For example, we may imagine a person who is overweight, gets cancer, loses 40 lbs (18 kg), and is then observed to be both normal weight and at high risk of death – even though thinness itself is beneficial for longevity.

The hypothesis is intuitive. Unintentional weight loss is a well-recognized symptom of several clinical diseases, including cancer, chronic lung disease, and congestive heart failure (26). This explanation also provides a simple solution to the aforementioned obesity paradox: Among those with disease, the sickest may lose the most weight, such that heavy patients are disproportionately likely to have a mild case, and, by extension, likely to have a good prognosis. Many scholars consider pre-existing disease to be a crucial confounder in studies of body weight and health outcomes (2, 27, 28).

Nevertheless, this reverse causation hypothesis requires a surprisingly complex model. Assume, as in the nutritional reserve model above, that there is a causal, linearly increasing relationship between age and logged death rates, and that death rates are proportional by weight category. Now, however, assume that in addition to the risk of death, people also face some risk of developing serious illness. For simplicity, assume that this process is similar in some respect to the mortality process, such that logged sickness rates, too, rise linearly in age, and that sickness rates are proportional by weight category, with heavier people more likely to get disease. Rates of weight change now vary by health status. Healthy people (without serious illness) tend to gain weight as they age, but the rate of transitioning up a weight category declines with age. Sick people tend to lose weight. Sick people also face a substantially increased risk of death. For model simplicity, assume that everyone is healthy at age 30. Assume that nobody who becomes sick ever recovers.

Under these conditions, a death rate, d, for an individual or group, i, can be defined as follows:

$$d_i = \begin{cases} \mu \exp(zx_i) & \text{if } i \text{ is weight category I and not sick} \\ \mu \exp(zx_i)\, \beta_2 & \text{if } i \text{ is weight category II and not sick} \\ \mu \exp(zx_i)\, \beta_3 & \text{if } i \text{ is weight category III and not sick} \\ \mu \exp(zx_i)\, \theta & \text{if } i \text{ is weight category I and sick} \\ \mu \exp(zx_i)\, \beta_2\, \theta & \text{if } i \text{ is weight category II and sick} \\ \mu \exp(zx_i)\, \beta_3\, \theta & \text{if } i \text{ is weight category III and sick} \end{cases}$$

where

μ is the baseline mortality hazard, the death rate at age 30

z describes the increase in mortality risk with age, assuming that the multiplicative effect on baseline hazard, μ, at age $x + 30$ is $\exp(zx)$

β_2 describes the multiplicative effect of weight category II on mortality risk, relative to category I

β_3 describes the multiplicative effect on mortality risk associated with weight category III, relative to category I, and

θ describes the multiplicative effect of sickness on mortality risk.

A rate of sickness onset, s, for individual or group, i, can also be defined:

$$s_i = \begin{cases} \rho \exp(z_s x_i) & \text{if } i \text{ is weight category I and not sick} \\ \rho \exp(z_s x_i)\, \beta_{s2} & \text{if } i \text{ is weight category II and not sick} \\ \rho \exp(z_s x_i)\, \beta_{s3} & \text{if } i \text{ is weight category III and not sick} \\ 0 & \text{otherwise} \end{cases}$$

where

ρ is the baseline sickness hazard, the rate of sickness onset at age 30

z_s describes the increase in sickness risk with age, assuming that the multiplicative effect on sickness onset rate, ρ, at age $x + 30$ is $\exp(z_s x)$

β_{s2} describes the multiplicative effect of weight category II on sickness risk, relative to weight category I, and

β_{s3} describes the multiplicative effect of weight category III on sickness risk, relative to weight category I.

A rate of transitioning up a weight category, g, is defined as follows:

$$g_i = \begin{cases} \gamma \exp(\gamma_a x_i) & \text{if } i \text{ is weight category < III and not sick} \\ 0 & \text{otherwise} \end{cases}$$

where

γ is the baseline rate – i.e., the rate at age 30 – of moving up one weight category among the healthy (to a maximum of weight category III), and

γ_a describes the change with age in the rate of moving up one weight category among the healthy (to a maximum of weight category III), assuming that the multiplicative effect on rate, γ, at age $x + 30$ is $\exp(\gamma_a x)$.

Finally, a rate of transitioning down a weight category, l, is defined for individual or group, i:

$$l_i = \begin{cases} \gamma_s & \text{if } i \text{ is weight category > I and sick} \\ 0 & \text{otherwise} \end{cases}$$

where

γ_s is the rate of moving down one weight category among the sick (to a minimum of weight category I).

This model is therefore described with 12 essential parameters: the 6 that were included in the nutritional reserve model, and an additional 6 to describe sickness onset, sickness-induced weight loss, and the effect of sickness on mortality risk (i.e., θ, ρ, z_s, β_{s2}, β_{s3}, and γ_s). Two major studies (3, 4) find that, in the absence of pre-existing disease, the association between BMI and mortality among non-smokers is linear (excluding the underweight range). Therefore, to simplify the model, assume that $\beta_3 = 2(\beta_2 - 1) + 1$. (That is, the hazard ratio increases linearly with weight category. For example, if the hazard ratio associated with β_2 is 2.0, then the hazard ratio associated with β_3 is 3.0; if the ratio for β_2 is 1.25, the ratio for β_3 is 1.5; etc.) For further simplicity, assume $\beta_{s3} = 2(\beta_{s2} - 1) + 1$ as well.

The resulting model then has 10 parameters – only four more than the model for Hypothesis #1, nutritional reserve. Almost none of the parameter values are observed directly, however,

since observed relationships generally reflect multiple underlying processes. For example, the observed relationship between age and mortality reflects both (a) the direct effect of age on mortality, $\exp(zx)$, and (b) the effect of age on sickness, $\exp(z_s x)$, in combination with the effect of sickness on mortality, θ.

Hypothesis #3

A modified theory of reverse causation: All declines in health, whether due to illness or to aging, may cause weight loss.

This is perhaps a less obvious interpretation of reverse causation: the idea that anyone in poor health is at risk of weight loss, even if "poor health" is simply due to old age. The model implied by the hypothesis, however, is in fact a simplification of the model used to describe Hypothesis #2, conventional reverse causation.

To appreciate this explanation, consider a population at risk of two major causes of death: one, which does not depend on health and thus does not have to vary by age – for example, accidents or violence – and a second, which depends heavily on health status, for example, almost all diseases, whether chronic or infectious. In this population, we may then treat death risk as depending *only* on health and on a baseline hazard (to represent, respectively, deaths due to illness and deaths due to accidents or violence; the model has some similarities to the well-known Gompertz-Makeham model, which treats death risk as a function of an age-related

component and an age-invariant "background" mortality component (25)). Importantly, in this model, *deaths are not directly related to age or to body weight*. Onset of poor health, however, or "sickness" (broadly defined) *is* assumed to depend on age and weight category. To reiterate, age and body weight influence death risk exclusively through their influence on health.

The implication of defining "sickness" so broadly is that "sickness"-induced weight loss must be common among the elderly. There is, in fact, support for this claim. Gerontologists often refer to a wide-ranging syndrome known as "frailty" (29) or geriatric "failure to thrive" (30), characterized by unintentional weight loss. This syndrome, which is distinct from functional disabilities or clinical disease (although often accompanied by both), has been linked to weakened immunity, increased risk of falls, and increased rates of hospitalization and death (29-31). Furthermore, Fried et al. find that frailty and its precursor symptoms are indeed common among Americans, even as young as ages 65-74 (29).

To model the hypothesis, assume (as in the model for Hypothesis #2) that every person is healthy at age 30, but that, as people age, they face some risk of sickness. Assume that the logged rates of sickness onset increase linearly with age. Sick people face a higher death rate than the healthy. Assume that healthy people tend to gain weight, while the sick tend to lose weight. Assume, further, that nobody who becomes sick later recovers.

The death rate, d, can then be defined for individual or group, i, as follows:

$$d_i = \begin{cases} \mu & \text{if } i \text{ is not sick} \\ \mu\theta & \text{if } i \text{ is sick} \end{cases}$$

where

μ is the baseline mortality hazard, the death rate among the healthy at every age, and

θ describes the multiplicative effect of sickness on mortality risk.

A rate of sickness onset, s, for individual or group, i, can be defined as follows:

$$s_i = \begin{cases} \rho \exp(z_s x_i) & \text{if } i \text{ is weight category I and not sick} \\ \rho \exp(z_s x_i)\, \beta_{s2} & \text{if } i \text{ is weight category II and not sick} \\ \rho \exp(z_s x_i)\, \beta_{s3} & \text{if } i \text{ is weight category III and not sick} \\ 0 & \text{otherwise} \end{cases}$$

where

ρ is the baseline sickness hazard, the rate of sickness onset at age 30

z_s describes the increase in sickness risk with age, assuming that the multiplicative effect on sickness onset rate, ρ, at age $x + 30$ is $\exp(z_s x)$

β_{s2} describes the multiplicative effect of weight category II on sickness risk, relative to weight category I, and

β_{s3} describes the multiplicative effect of weight category III on sickness risk, relative to weight category I.

The rate of transitioning up a weight category, g, is as follows:

$$g_i = \begin{cases} \gamma & \text{if } i \text{ is weight category < III and not sick} \\ 0 & \text{otherwise} \end{cases}$$

where

γ is the rate of moving up one weight category among the healthy (to a maximum of weight category III).

Finally, the rate of transitioning down a weight category, l, is defined:

$$l_i = \begin{cases} \gamma_s & \text{if } i \text{ is weight category > I and sick} \\ 0 & \text{otherwise} \end{cases}$$

where

γ_s is the rate of moving down one weight category among the sick (to a minimum of weight category I).

In other words, this model representing "modified" reverse causation is described with 8 essential parameters: μ, θ, ρ, z_s, β_{s2}, β_{s3}, γ, and γ_s. We no longer need a parameter for γ_a (present, above, in the models for Hypotheses #1 and #2), since a decline in the propensity to gain weight is built into this model via high sickness rates among the elderly. Furthermore, as in the model for Hypothesis #2, we may eliminate a parameter by assuming that $\beta_{s3} = 2(\beta_{s2} - 1) + 1$. This leaves a model with 7 parameters, only one more than in the nutritional reserve model.

Simulation

This paper simulates three populations, each one living under the conditions implied by one of the three hypotheses above. Simulation allows a deeper exploration of model consequences than may be apparent from descriptions of model construction alone. In particular, simulation can demonstrate whether any of the three models – or perhaps all of them – is able to produce observations similar to those from real-world data.

Three simulated populations.

Each simulated population comprises 10,000 people, tracked from age 30 to age 90 (or until death, whichever comes first). The simulations are macro-simulations; they are not stochastic. This means that calculated probabilities of an event (i.e., death, sickness onset, or weight change) can be interpreted as the fixed proportion of the population experiencing the event during some period. This can result in fractional people changing states.

Although the models are defined above in rates (for continuous processes), for ease of computation, all calculations are conducted using discrete transition probabilities. Each simulation population is updated one year at a time. Figure 3.1 shows the transition matrix used to translate (approximately) the rates to probabilities for the models that correspond to Hypotheses #2 and #3 (reverse causation). Specifically, the matrix shows how rates d, s, g, and l, defined above for each model, transform a simulation population at year x (the vector on the right of the figure) into the population at year $x + 1$ (the vector on the left). Figure 3.1 can also be considered to represent the simulation corresponding to Hypothesis #1 (nutritional reserve), if we simply set all values of s, the rate of sickness onset, to 0 (such that no member of the population can fall into groups the labeled n_{s1}, n_{s2}, or n_{s3}, which are not at risk of transitioning up a weight category).

The conversion from rates to probabilities shown in Figure 3.1 is an approximation only, as the once-yearly transitions shown do not account for any partial-year exposures (for instance, a person who becomes sick halfway through the year, spending half the year at risk of transitioning up a weight category and half the year at risk of transitioning down). Importantly, then, Figure 3.1 illustrates the order of events assumed to occur during each transition. First, some proportion of the population dies during the year ($1 - exp(-d)$), with the death rate, d, for each group based on the group's sickness status and weight category at the beginning of the year. Next, some proportion of the population that survives may experience new sickness onset within the year ($1 - exp(-s)$), with the sickness onset rate, s, based on weight category at the

Figure 3.1. Transition Matrix: Using Rates of Population Processes to Calculate Annual Changes in Population Composition. Transition probabilities are used to transform the population at time x (the vector on the right) into the population at time $x + 1$ (the vector on the left). See text for equations to define rates of death, d; of sickness onset, s; of weight gain, g; and of weight loss, l. Rate values are calculated separately for each sub-group of the population (i.e., for each element in the right-hand vector). However, this figure omits subscripts to improve readability.

$$\begin{bmatrix} n_{dead} \\ n_{h1} \\ n_{h2} \\ n_{h3} \\ n_{s1} \\ n_{s2} \\ n_{s3} \end{bmatrix}_{x+1} = \begin{bmatrix} 1 & 1-e^{-d} & 1-e^{-d} & 1-e^{-d} & 1-e^{-d} & 1-e^{-d} & 1-e^{-d} \\ 0 & e^{-d}\,e^{-s}\,e^{-g} & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{-d}\,e^{-s}\,(1-e^{-g}) & e^{-d}\,e^{-s}\,e^{-g} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{-d}\,e^{-s}\,(1-e^{-g}) & e^{-d}\,e^{-s} & 0 & 0 & 0 \\ 0 & e^{-d}\,(1-e^{-s}) & e^{-d}\,(1-e^{-s})\,(1-e^{-l}) & 0 & e^{-d} & e^{-d}\,(1-e^{-l}) & 0 \\ 0 & 0 & e^{-d}\,(1-e^{-s})\,e^{-l} & e^{-d}\,(1-e^{-s})\,(1-e^{-l}) & 0 & e^{-d}\,e^{-l} & e^{-d}\,(1-e^{-l}) \\ 0 & 0 & 0 & e^{-d}\,(1-e^{-s})\,e^{-l} & 0 & 0 & e^{-d}\,e^{-l} \end{bmatrix} \begin{bmatrix} n_{dead} \\ n_{h1} \\ n_{h2} \\ n_{h3} \\ n_{s1} \\ n_{s2} \\ n_{s3} \end{bmatrix}_{x}$$

x is age in years minus 30.
n_{h1} is the number of healthy people in weight category I.
n_{h2} is the number of healthy people in weight category II.
n_{h3} is the number of healthy people in weight category III.
n_{s1} is the number of sick people in weight category I.
n_{s2} is the number of sick people in weight category II.
n_{s3} is the number of sick people in weight category III.

beginning of the year. Finally, people face a risk of weight change (usually $1 - exp(-g)$ or $1 - exp(-l)$, but sometimes 0), as if they had spent the entire year in the sickness-status group *in which they end the year*. Only people assigned to a weight category at the start of the year are at risk of transitioning out of it. (That is, a person may change weight categories no more than once per year.) All simulations are performed in R.2.15.2 (32).

Selecting parameter values.

Simulations in this paper illustrate *possible* relationships between body weight and mortality. Parameter values for each model are thus chosen for convenience, in that they produce illustrations to resemble observations from real-world data. These values are not estimated from data.

Three criteria guide the selection of parameter values. First, ideally, the simulated populations' death rates by age would approximate a real population's observed death rates by age. Second, the simulated populations' weight category distributions at each age would resemble a real population's distribution across BMI categories (normal, overweight, obese) by age. Third, naturally, the simulated populations would exhibit a relationship between body-weight category and mortality that resembles an observed curvilinear BMI–mortality relationship. In this study, the "target" or designated patterns for each simulation population to reproduce are based loosely on data from U.S. males born around 1950, but they are ultimately arbitrary. A combination of parameter values is selected for each model from a pre-specified range of values (also arbitrary) as the combination of these that best reproduces the designated targets.

The choice of these targets and selection of parameter values are both described in Appendix 3A. However, neither is central to this paper, as it is not necessary to explore, for a given model, every possible parameter-value combination (of which there are infinitely many) to demonstrate that at least one combination can satisfy some criteria.

Initial conditions.

For all simulations, the proportions in weight categories II and III at age 30, when the simulation begins, are set equal to the target proportions in these weight categories. The baseline hazard, μ, is set equal in all simulations to the target death rate at age 30: 1.69 per 1,000 person-years.

RESULTS

Figure 3.2 presents logged death rates by age in each of the three simulated populations. Figure 3.3 shows proportions of these simulated populations in each weight category by age, and Figure 3.4 shows hazard ratios associated with weight categories II and III by age, relative to category I. In each figure, the stipulated (target) relationship to reproduce is also shown for comparison.

Overall, the simulated populations exhibit relationships between age, body weight, and mortality that roughly resemble one other, as well as the designated targets. All three populations reproduce the stipulated age pattern of mortality relatively well (Figure 3.2). In all

Figure 3.2. Death Rates, by Age, in Three Simulated Populations.

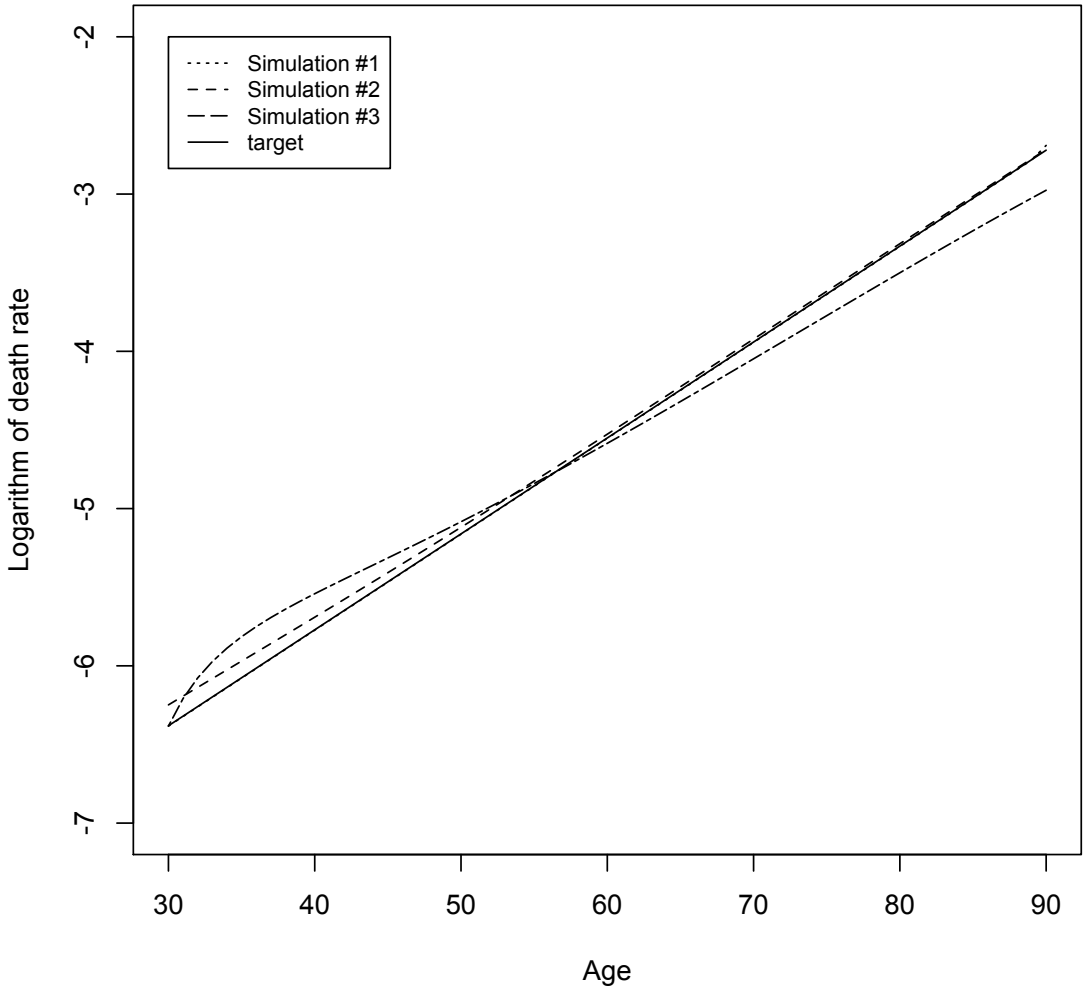


Figure 3.3. Weight Category Distribution, by Age, in Three Simulated Populations.

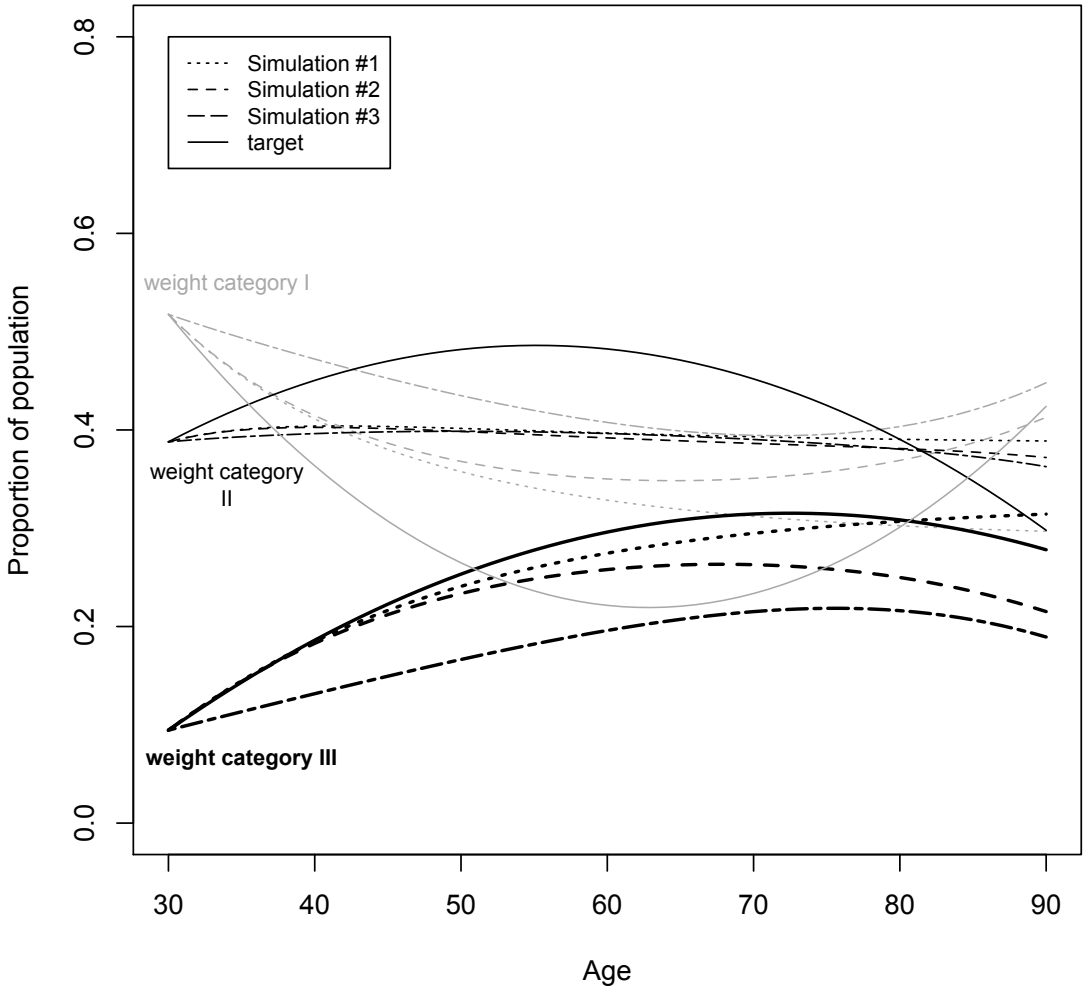
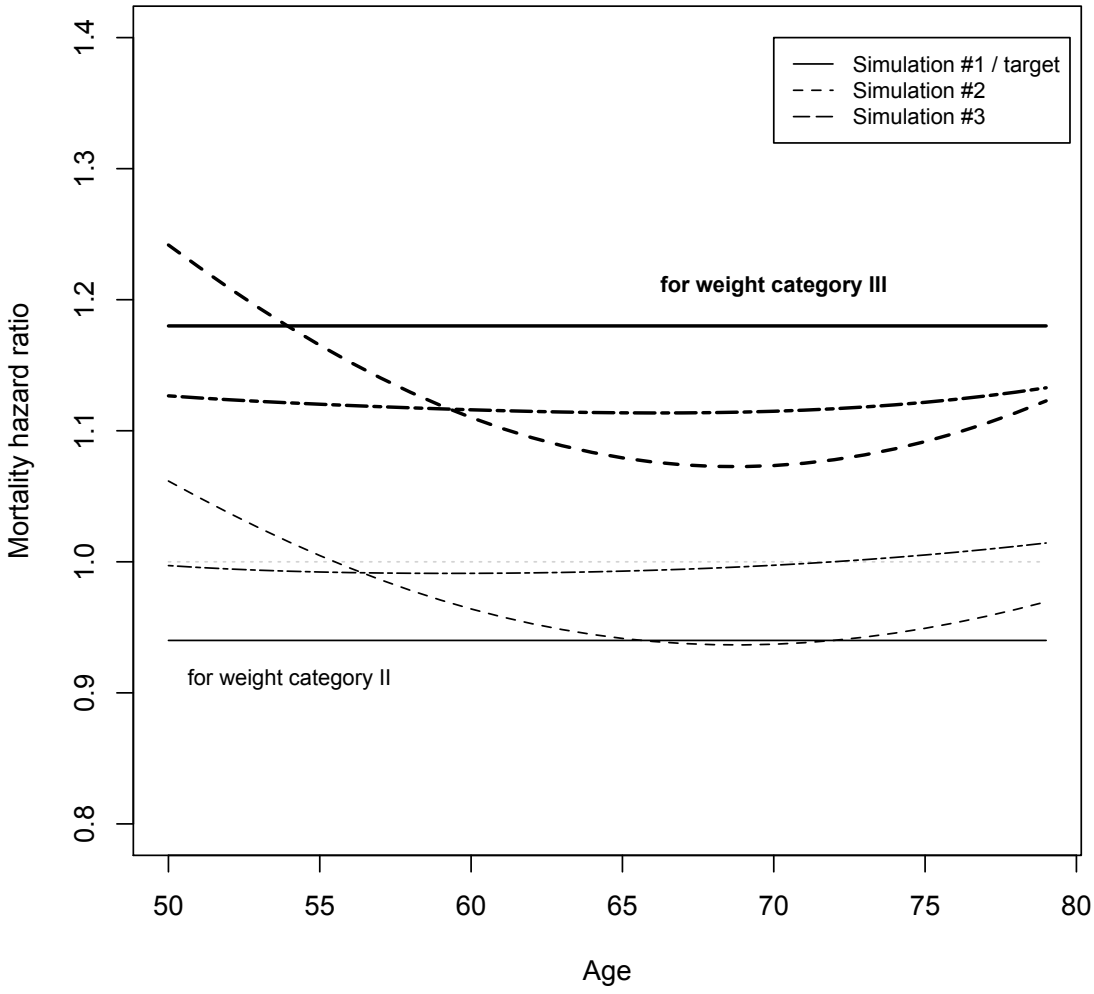


Figure 3.4. Mortality Hazard Ratios Associated with Weight Categories II and III, by Age, in Three Simulated Populations.



three, the fit to the target weight category distribution is somewhat weaker (Figure 3.3). However, this may reflect some unrealistic assumptions that are common to all three simulations, such as, most obviously, the inability for most people to lose weight. (Only sick people, under Hypotheses #2 and #3, may do so.)

Figure 3.4 shows the paper's most notable finding: All three hypotheses can explain a curvilinear relationship between BMI and mortality (at least within the age range where one is typically observed). That is, all three simulated populations have a mortality hazard ratio < 1 associated with weight category II, relative to weight category I, averaged over ages 50-79. Table 3.1 presents the mean mortality hazard ratios simulated for each hypothetical population. Note that Simulation #1 is guaranteed to reproduce exactly the target hazard ratios; that model assumes explicitly that observed (target) relationships reflect causal processes, so the simulation hazard ratios are fixed at the target levels. Simulations of reverse causation (Hypotheses #2 and #3), however, are not guaranteed to reproduce this relationship, and there has been some debate over whether it is possible to do so (19). Figure 3.4 and Table 3.1 show that, indeed, it is. Moreover, while the hazard ratios associated with Hypotheses #2 and #3 may not appear to match the target values very closely, it is possible to achieve a better match by compromising slightly on the match either to the age pattern of mortality or to the weight-category distribution by age (results not shown).

Table 3.1. Mean Hazard Ratios Associated with Weight Categories II and III, for Three Simulated Populations. Means calculated over age range 50-79.

Hazard ratios	*Simulation #1: Nutritional reserve hypothesis*	*Simulation #2: Conventional reverse causation hypothesis*	*Simulation #3: Modified reverse causation hypothesis*
Ratio associated with weight category II (relative to category I)	0.940	0.969	0.997
Ratio associated with weight category III (relative to category I)	1.180	1.117	1.119

Parameter values for each simulation are listed in Table 3.2. Most of these parameters have an intuitive interpretation, so the values may be used to compare details of presumed causal processes across the simulated populations: for example, which hypothetical population has the greatest propensity to gain weight at a given age, and which experiences the greater longevity penalty associated with sickness (in the simulations corresponding to Hypotheses #2 and #3). Furthermore, one may analyze the simulated populations to calculate a number of additional quantities of interest: for example (again in Simulations #2 and #3), sickness prevalence at a given age, the frequency of sickness-induced weight loss by age (i.e., transitioning down a weight category), and the association between weight loss and mortality.

Table 3.2. Selected Parameter Values: Three Models.

Simulation #1: Nutritional reserve hypothesis		*Simulation #2: Conventional reverse causation hypothesis*		*Simulation #3: Modified reverse causation hypothesis*	
μ	0.00169	μ	0.00169	μ	0.00169
z	0.061	z	0.04		
β_2	0.94	β_2	1.25		
β_3	1.18	β_3	1.5		
γ	0.03	γ	0.03	γ	0.01
γ_a	-0.05	γ_a	-0.05		
		θ	16	θ	128
		ρ	0.0005	ρ	0.001
		z_s	0.07	z_s	0.06
		β_{s2}	1.5	β_{s2}	2.0
		β_{s3}	2.0	β_{s3}	3.0
		γ_s	0.1	γ_s	0.1

DISCUSSION

This paper has used simulation to demonstrate that observed phenomena are broadly consistent with more than one hypothesis for the curvilinear relationship between BMI and mortality. The competing hypotheses presented here are not exhaustive, and others may also be capable of explaining empirical findings (e.g., 33). Nevertheless, this paper shows a range of possible causal mechanisms, all with biological support, and demonstrates that any one of them could produce typical observations of the relationships between body weight, mortality, and age.

Proponents of the conventional reverse causation hypothesis may not find this surprising. However, by defining a model to represent a given hypothesis, one not only clarifies the presumed causal process – a useful conceptual task in its own right – but also elucidates the implications of that hypothesis. Reverse causation has recently been criticized as an explanation for the BMI–mortality curve because, at present, there are no good estimates of the frequency or magnitude of sickness-induced weight loss that could produce observed phenomena (19). However, the simulations in this paper, or more sophisticated versions of them, can provide exactly those estimates.

This paper, in other words, demonstrates two important points about reverse causation. First, it shows that reverse causation is sufficient to explain a population-level longevity advantage among the overweight. The second and more unexpected result, however, is that the major

implications of this explanation – in terms of sickness-induced weight loss – may in fact be quite plausible. For example, in the simulation corresponding to Hypothesis #2 (conventional reverse causation), sickness prevalence at age 60 is 4.2%. The proportion of sick people that transitions down a weight category in that population that year is 4.9%, which corresponds to a population-level prevalence of sickness-induced weight loss (transitioning down one weight category) of 0.2%. By age 80, these figures are 10.5%, 5.3%, and 0.6%, respectively. It would appear, then, that sickness-induced weight loss does not have to be extensive for it to fully explain the observed BMI–mortality relationship. It is conceivable that 4% of a population could have major clinical disease by age 60 (although this is difficult to test empirically, of course, as appropriate definitions of "major clinical disease" are themselves open to interpretation). It is further plausible that some fraction of 1% of the population may experience sickness-induced weight loss at age 60.

Simulations in this paper demonstrate mere possibilities for the relationships between body weight, age, and mortality, however. As such, caution is needed when interpreting the apparent implications of a model (such as those described above), or when comparing parameter values across models (listed in Table 3.2). That is, for a given model, it is always possible, and perhaps even likely, that another combination of parameter values exists that would similarly match the target relationships between age, weight, and death rates, but with different consequences for other population traits (for example, for sickness prevalence). Comparing parameter values in Table 3.2, therefore, allows one to compare population

processes across the *simulated populations*, but does not, in general, compare features intrinsic to the models themselves.

It is nevertheless useful to explore these populations' major similarities and differences, as these highlight some range of the different causal mechanisms that may produce observed phenomena. The clearest difference between Simulation #1 (the nutritional reserve hypothesis) and Simulations #2 and #3 (reverse causation) is, by design, the "true" association between weight category and health, captured by the parameters β_2, β_3, β_{s2}, and β_{s3} (see Table 3.2). In Simulation #1, all else being equal, it is optimal for longevity to be in weight category II (β_2 = 0.94). In contrast, in Simulations #2 and #3, it is optimal to be in weight category I, and the penalties associated with higher body weight are substantial. For example, in Simulation #2, irrespective of sickness status (which itself is more likely to strike people who are heavy), people in weight category II have death rates 25% higher than those in weight category I (ignoring the
discrepancy, explained above, that is introduced by the approximate conversion of rates to annual transition probabilities). People in weight category III have death rates 50% higher than the lowest-risk group.

As noted above, the model for Hypothesis #3 (modified reverse causation) is a simplification of the model for Hypothesis #2 (conventional reverse causation). Because of this, Simulation #2 is guaranteed to be able to reproduce observed phenomena as long as Simulation #3 can. (That is, a model with added parameters will always fit at least as well as a simpler model.) The main

value of analyzing both of these hypotheses, therefore, is to demonstrate whether reverse causation, as conventionally understood, can be simplified mathematically with some slight reinterpretation of the assumed underlying biology. Table 3.2 gives some sense of the differences in this assumed biology between the two simulated populations. The most apparent difference, perhaps, is the penalty associated with sickness, θ. This penalty is large in Simulation #2 for conventional reverse causation (θ = 16), but enormous in Simulation #3 for modified reverse causation (θ = 128). This difference may be easily explained, however, by the fact that all healthy people in Simulation #3 experience the same death rate as a healthy 30-year-old, whereas people in Simulation #2 experience increasing death rates with age, whether or not they are sick. That is, in Simulation #3, θ must be large enough to account for a good deal of the increase in death rates with age.

In the end, it is not straightforward to say which of the three hypotheses presented may be closest to the truth. Ideally, each model would be adjudicated on some combination of its ability to reproduce observed phenomena, its biological plausibility, and, most likely, its parsimony. Depending on how one weights these various criteria, however, one could reasonably argue that any of the three hypotheses presented here is "best." The nutritional reserve hypothesis can explain observed relationships very well. However, it is to some degree constructed post-hoc to fit them, and some scholars doubt its biological plausibility (e.g., 1, 2). The conventional reverse causation hypothesis, in contrast, is generally viewed as more biologically plausible (1, 2), but the complexity required to model it has been greatly underappreciated. Even a simplistic model like the one in this paper, with no recovery from

sickness and no weight loss among the healthy, demands more than half a dozen parameters that essentially cannot be estimated from data. Modified reverse causation presents something of a compromise: more parsimonious than the conventional hypothesis, but retaining its central explanatory features.

In the future, empirical studies that rely heavily on one hypothesis for interpretation of the BMI–mortality relationship may wish to specify how, exactly, causal processes are expected to operate. Lack of specificity is a near-universal problem in descriptions of reverse causation, but is not unique to that explanation. Specificity would be helpful for assessing hypothesis plausibility along every dimension: empirical grounding, biological support, and simplicity.

Furthermore, it is only by describing models in greater detail that scholars will generate new testable hypotheses. This study shows that, at present, it is not straightforward to rule out any of the competing theories of BMI and mortality risk. However, if a simulation were to predict more specific phenomena, these could be verified or rejected with additional empirical study: for example, by pinpointing a particular association between weight change and mortality or between sickness measures and weight change, or, alternatively, by demonstrating changes in the BMI–mortality relationship itself, either with time since BMI measurement (34), or with age. The true challenge will be finding predictions that are particular to a given model, however. As this paper demonstrates, very different hypotheses can sometimes explain the same observations.

REFERENCES

[1] Hu F, ed. *Obesity Epidemiology*. Oxford: Oxford University Press; 2008:149-215.

[2] Manson JE, Stampfer MJ, Hennekens MD, et al. Body weight and longevity: a reassessment. *J Am Med Assoc*. 1987; 257(3):353-358.

[3] Baik I, Ascherio A, Rimm EB, et al. Adiposity and mortality in men. *Am J Epidemiol*. 2000; 152(3):264-271.

[4] Ajani UA, Lotufo PA, Gaziano JM, et al. Body mass index and mortality among US male physicians. *Ann Epidemiol*. 2004; 14(10):731-739.

[5] Flegal KM, Kit BK, Orpana H, et al. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *J Am Med Assoc*. 2013; 309(1):71-82.

[6] Fryar CD, Ogden CL. Prevalence of underweight among adults aged 20 years and over: United States, 2007-2008. Hyattsville, MD: National Center for Health Statistics; October 2010. (Health E-stat: http://www.cdc.gov/nchs/data/hestat/underweight_adult_07_08/underweight_adult_07_08.pdf). (Accessed June 5, 2013).

[7] Shetty PS, James WPT. Body mass index: A measure of chronic dietary energy deficiency in adults. Rome, Food and Agriculture Organization of the United Nations, 1994 (FAO Food and Nutrition Paper No. 56). (http://www.fao.org/docrep/t1970e/t1970e00.htm.)

[8] Durazo-Arvizu R, McGee D, Cooper R, et al. Mortality and optimal body mass index in a sample of the US population. *Am J Epidemiol*. 1998; 147(8):739-749.

[9] Calle E, Thun M, Petrelli J, et al. Body-mass index and mortality in a prospective cohort of US adults. *N Engl J Med*. 1999; 341(15):1097-1105.

[10] McGee D, Diverse Populations Collaboration. Body mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies. *Ann Epidemiol*. 2005; 15(2):87-97.

[11] Gu D, He J, Duan X, et al. Body weight and mortality among men and women in China. *J Am Med Assoc*. 2006; 295(7):776-783.

[12] Dolan C, Kraemer H, Browner W, et al. Associations between body composition, anthropometry, and mortality aged 65 years and older. *Am J Public Health*. 2007; 97(5):913-918.

[13] Snih S, Ottenbacher K, Markides K, et al. The effect of obesity on disability vs mortality in older Americans. *Arch Intern Med*. 2007; 167(8):774-780.

[14] Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900,000 adults: collaborative analyses of 57 prospective studies. *Lancet*. 2009; 373(9669):1083-1095.

[15] Lin W, Tsai S, Albu J, et al. Body mass index and all-cause mortality in a large Chinese cohort. *Can Med Assoc J*. 2011; 183(6):E329-E336.

[16] Zheng W, McLerran D, Rolland B, et al. Association between body-mass index and risk of death in more than 1 million Asians. *N Engl J Med*. 2011; 364(8):719-729.

[17] Hughes V. The big fat truth [news feature]. *Nature*. 2013; 497:428-430.

[18] Willett WC, Hu FB, Thun M. Overweight, obesity, and all-cause mortality [letter to the editor]. *J Am Med Assoc*. 2013; 309(16):1681-1682.

[19] Flegal K, Graubard B, Williamson D, et al. Reverse causation and illness-related weight loss in observational studies of body weight and mortality. *Am J Epidemiol*. 2011; 173(1):1-9.

[20] Gnudi S, Sitta E, Lisi L. Relationship of body mass index with main limb fragility fractures in postmenopausal women. *J Bone Miner Metab*. 2009; 27(4): 479-484.

[21] Arena R, Lavie CJ. The obesity paradox and outcomes in heart failure: Is excess body weight truly protective? *Future Cardiol*. 2010; 6(1):1-6.

[22] Curtis JP, Selter JG, Wang Y, et al. The obesity paradox: Body mass index and outcomes in patients with heart failure. *Arch Intern Med*. 2005; 165(1):55-61.

[23] Andersen KK, Olsen TS. The obesity paradox in stroke: Lower mortality and lower risk of readmission for recurrent stroke in obese stroke patients. *Int J Stroke*. 2013; doi: 10.1111/ijs.12016. [Epub ahead of print].

[24] Pikkers P, de Keizer N, Dusseljee J, et al. Body mass index is associated with hospital mortality in critically ill patients: An observational cohort study. *Crit Care Med*. 2013; doi: 10.1097/CCM.0b013e31828a2aa1. [Epub ahead of print].

[25] Preston SH, Heuveline P, Guillot M. Demography: Measuring and modeling population processes. Oxford, UK: Blackwell Publishers Ltd.; 2001:192.

[26] Goldman L, Ausiello D, eds. *Cecil Medicine. 23rd edn*. Philadelphia: Saunders Elsevier; 2008:350, 619, 1353.

[27] Willett W. Weight loss in the elderly: cause or effect of poor health? *Am J Clin Nutr*. 1997; 66:737-738.

[28] Losonczy K, Harris T, Cormoni-Huntley J, et al. Does weight loss from middle age to old age explain the inverse weight mortality relation in old age? *Am J Epidemiol*. 1995; 141(4):312-321.

[29] Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001; 56(3):M146-M157.

[30] Robertson RG, Montagnini M. Geriatric failure to thrive. *Am Fam Physician*. 2004; 70(2):343-350.

[31] Xue Q. The frailty syndrome: definition and natural history. *Clin Geriatr Med*. 2011; 27(1):1-15.

[32] R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2012. ISBN 3-900051-07-0. (http://www.R-project.org/).

[33] Childers DK, Allison DB. The 'obesity paradox': A parsimonious explanation for relations among obesity, mortality rate and aging? *Int J Obes*. 2010; 34(8):1231-1238.

[34] Allison D, Heo M, Flanders D, et al. Examination of ‘early mortality exclusion’ as an approach to control for confounding by occult disease in epidemiologic studies of mortality risk factors. *Am J Epidemiol*. 1997; 146(8):672-680.

Conclusion

This dissertation has assessed the role of "reverse causation" in studies of body weight and mortality, testing whether sickness-induced weight loss can explain low death rates among overweight people, relative to people with clinically normal body weight. All three chapters of the dissertation show modest support for this explanation.

Chapter 1 shows that people losing weight are at higher risk of death than people with stable or rising weight, and that this accounts for some – but maybe not all – of the apparent longevity advantage associated with being overweight. Chapter 2 shows that clinical disease can predict subsequent or concurrent weight loss in prospective, population-based samples, but that measured associations are sensitive both to the survey and to the magnitude of weight loss considered. This hints that reverse causation could play a different role from one sample to the next. Finally, Chapter 3 demonstrates by simulation that it is possible at least for reverse causation to produce an apparent longevity advantage among the overweight, even if every

individual in the population would be better off with normal weight. However, these simulated relationships are merely possible ones, not necessarily true ones.

In sum, while it seems that sickness-induced weight loss does affect estimates of the optimal BMI for longevity, the optimal value itself remains unclear. It could be in the overweight range; it could be in the normal-weight range. Each chapter has discussed the research implications of its findings in detail. However, given strong popular interest in body-weight topics (1), I would like to close with a brief discussion about what these findings might suggest for individual risk, and for individuals' choices about desired body weight. I have studied reverse causation with respect to mortality only, and not to disease, disability, discrimination, or any other outcome. This limits the range of reasonable recommendations. Nevertheless, I would offer four suggestions related to body weight, drawing together results from this dissertation and from previous related work.

1. Avoid becoming obese.

While this dissertation does not identify an optimal BMI, it confirms that obesity most likely harms longevity. Ample existing evidence links obesity to a range of other negative health and socioeconomic outcomes as well (2, 3). In short, there are few obvious reasons to become obese, and plenty of reasons to avoid it.

In the samples analyzed in this dissertation, however, body weight tends to increase with age until well into one's 60s or 70s. (This is consistent with previous research on weight gain and age (4).) Therefore, younger adults – say, people under age 50 – may benefit from keeping their weight well below the obese range, with a buffer to avoid obesity in the case that some weight gain with age cannot be prevented. Successfully avoiding obesity, therefore, may take a lifetime of effort.

2. Avoid being underweight.

My research papers do not report much on the dangers of underweight (BMI < 18.5), since few people in the studies' age range, 50-79, fall into this very low weight category. However, in the data I use, as in previous studies (e.g., (5, 6)), mortality risk is usually highest of all among the underweight.

Some excess deaths among this group are very likely due to reverse causation. Still, even for people without clinical disease, being underweight can be dangerous. Underweight status requires sufficiently low weight for height (see Table I.1, in the Introduction, for values) that, almost by definition, it implies low muscle mass in addition to low fat mass (7). Extreme malnutrition can of course lead to death by starvation. But, even well before that point, malnutrition will compromise immune function, increasing the risks of illness and of death (8). Shetty and James of the United Nations Food and Agriculture Organization argue that (despite

the arbitrariness of all cut-points) the definition of underweight, BMI < 18.5, is a useful marker among adults for chronic energy deficiency (7): that is, inadequate food intake. No one should strive for such a low BMI.

3. Put body weight in perspective.

This may seem tongue in cheek, but the point is a serious one. Americans worry about their weight a great deal (1, 9), and they spend billions each year on weight-loss attempts (10). However, some of this concern and money might be better focused elsewhere.

For most American adults (11) – say, anyone with a BMI between 21 (a "normal" value) and 34 (an "obese" one) – the association between BMI and mortality, while not trivial, is only moderate (12). For example, differences within this BMI range are typically small compared to the difference in death rates between smokers and never-smokers. Smokers have been observed to die at two to three times the rate of never-smokers (13, 14). At least when it comes to longevity, therefore, body weight is often not a person's most important modifiable risk factor.

4. Be wary of weight loss, no matter how much you want it.

Lastly, one of the strongest findings in this dissertation is an association between weight loss and subsequent mortality among middle-aged and older adults. In every population studied here, people losing weight suffer higher death rates than people who either gain weight or maintain their weight. This association may seem surprising, given abundant medical and commercial advice on how to lose weight. But the link between weight loss and mortality is very well documented (15-18). It is also nothing new. Some 400 years ago even, William Shakespeare wrote about the weight loss that typically precedes death. (As part of the seven stages of life, he wrote, a middle-aged man, "In fair round belly with good capon lined," eventually becomes frail and old, "His youthful hose, well saved, a world too wide / For his shrunk shank" (19).) Among the elderly, weight loss is a common corollary of aging, and has been linked to increased risks of falls, infection, and hospitalization (20).

With the rise of obesity over the past 50 years, millions of Americans are now reaching old age having struggled with their weight throughout adulthood. For these people who are also overweight or obese, easy weight loss later in life may seem very welcome. However, weight loss that comes without improvement in diet quality or in physical fitness may confer few health benefits – and may, more importantly, signal grave health problems. This dissertation has focused on a scientific debate surrounding weight loss and its effect on observed relationships between body weight and mortality. But the less novel result that weight loss predicts mortality, ultimately, may be the one that is most relevant to public health, or to individuals concerned about their own health. No matter how desirable weight loss may seem, it is very often a very bad omen.

REFERENCES

[1] Cossrow NHF, Jeffery RW, McGuire MT. Understanding weight stigmatization: A focus group study. *J Nutr Educ*. 2001; 33(4):208-214.

[2] Hu F, ed. *Obesity Epidemiology*. Oxford: Oxford University Press; 2008, pp. 149-215.

[3] Crandall CS, Schiffhauer KL. Anti-fat prejudice: Beliefs, values, and American culture. *Obes Res*. 1998; 6(6):458-460.

[4] Williamson DF. Descriptive epidemiology of body weight and weight change in U.S. adults. *Ann Intern Med*. 1993; 119(7):646-649.

[5] Flegal KM, Graubard BI, Williamson DF, et al. Excess deaths associated with underweight, overweight, and obesity. *J Am Med Assoc*. 2005; 293(15):1861-1867.

[6] Weitoft GR. Underweight, overweight and obesity as risk factors for mortality and hospitalization. *Scand J Public Health*. 2008; 36(2):169-176.

[7] Shetty PS, James WPT. Body mass index: A measure of chronic dietary energy deficiency in adults. Rome, Food and Agriculture Organization of the United Nations, 1994 (FAO Food and Nutrition Paper No. 56). (http://www.fao.org/docrep/t1970e/t1970e00.htm.)

[8] Chandra RK. Nutrition and the immune system: an introduction. *Am J Clin Nutr*. 1997; 66(2):460S-463S.

[9] Hurd Clarke L. Older women's perceptions of ideal body weights: The tensions between health and appearance motivations for weight loss. *Ageing Soc*. 2002; 22(6):751-773.

[10] ABC News. 100 million dieters, $20 billion: The weight-loss industry by the numbers. 20/20, May 8, 2012. (http://abcnews.go.com/Health/100-million-dieters-20-billion-weight-loss-industry/story?id=16297197#.UdL4q-u9wzN). (Accessed July 2, 2013).

[11] Flegal KM, Carroll MD, Kit BK, et al. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *J Am Med Assoc*. 2012; 307(5):491-497.

[12] Flegal KM, Kit BK, Orpana H, et al. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *J Am Med Assoc*. 2013; 309(1):71-82.

[13] Doll R, Peto R, Wheatley K, et al. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ*. 1994; 309:901-911.

[14] Doll R, Peto R, Boreham J, et al. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*. 2004; 328:1519-1533.

[15] Lee I, Paffenbarger RS. Change in body weight and mortality. *J Am Med Assoc*. 1992; 268(15):2045-2049.

[16] Losonczy K, Harris T, Cornoni-Huntley J, et al. Does weight loss from middle age to old age explain the inverse weight mortality relation in old age? *Am J Epidemiol*. 1995; 141:312-321.

[17] Wannamethee SG, Shaper AG, Lennon L. Reasons for intentional weight loss, unintentional weight loss, and mortality in older men. *Arch Intern Med*. 2005; 165(9):1035-1040.

[18] Myrskyla M, Chang VW. Weight change, initial BMI, and mortality among middle- and older-aged adults. *Epidemiol*. 2009; 20(6):840-848.

[19] Shakespeare W. *As you like it*. Act 2, scene 7, lines 1051-1059.

[20] Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001; 56(3):M146-M157.

Appendices

APPENDIX 1A: DESCRIPTION OF DATA

This Appendix describes five longitudinal surveys used in Chapter 1:

1. The Australian Longitudinal Study of Ageing (ALSA);
2. The Costa Rican Study of Longevity and Healthy Aging (CRELES);
3. The English Longitudinal Study of Ageing (ELSA);
4. The Social Environment and Biomarkers of Aging Study (SEBAS), based in Taiwan; and
5. The U.S. National Health and Nutrition Examination I Epidemiologic Follow-Up Study (NHEFS).

Sampling

With the exception of ALSA, each survey is nationally representative, within certain age limits, of the country from which it is drawn. The ALSA sample is drawn from the Adelaide area. Two of the surveys oversample from select subpopulations. ALSA oversamples men (since males are relatively scarce at the older ages). NHEFS oversamples people living in poverty. Because there is no consensus on whether to include survey weights in regression analyses (1), this study ignores the selective oversampling. Sex is a covariate in all regressions, which means the extra men in ALSA should not affect results unless there are interactions between sex and BMI in the relationship between BMI and mortality, or between sex and BMI change in the

relationship between BMI change and mortality. No such interactions are observed in these data (results not shown). The large proportion of NHEFS participants living in poverty may lead to unusually high mortality estimates for the USA. However, as with sex in the ALSA sample, oversampling will not affect the shapes of the BMI–mortality curve or the BMI-change–mortality curve, unless the relationships differ by income.

In each figure, sample sizes (*N*) represent the number of study participants with non-missing values for all variables: i.e., BMI and/or BMI change, survival, age (within the range 50-79 at BMI measurement), sex (male or female), and smoking status (current, former, or never). It is possible for a study to have more observations in later waves than in earlier ones, as in NHEFS, if a greater number of study participants age into the 50-79 age range between survey waves than age out of it.

Measurement of BMI and BMI change

This study uses measured height and weight rather than self-reported values. To avoid undue influence of outliers or recording errors in the data, participants are excluded from analysis if they have a recorded weight below 32 kg (70 lbs) or above 273 kg (600 lbs), or if they have a recorded height below 1.22 m (48 in) or above 2.03 m (80 in). These thresholds are somewhat arbitrary, but were chosen in part because they exclude only a very small proportion of study participants from each survey: roughly 1 in 1,000. The upper weight limit set is not, in practice,

restrictive because each survey has a maximum weight up to which their scales are accurate (generally around 135 kg, or roughly 300 lbs), and in all cases it falls below the cut-off set by this study. ALSA, CRELES, ELSA and SEBAS all measure participants' height every time they measure weight. Because of a systematic measurement error in CRELES wave 2, height at wave 2 in Costa Rica is calculated as the average of height at waves 1 and 3, or, if wave 3 height is missing, as height at wave 1. NHEFS measures height only at baseline. Height among NHEFS participants is therefore treated as constant across waves (i.e., assuming that adult height is fixed, at least from middle age until age 80). This may, on average, overestimate adult height in the U.S.

BMI is calculated as weight (in kilograms) divided by the square of height (in meters). Change in body mass index is defined as BMI at one survey wave minus BMI at a previous wave, so that positive values reflect weight gain and negative values reflect weight loss. For all surveys besides ELSA, weight-change intervals (the periods between BMI measurements) are calculated as the number of days between one measurement and another, and rounded to the nearest year. ELSA does not provide exact dates of interview or measurement, so that weight change intervals for ELSA are calculated to the nearest year as, simply, the difference between the year of second measurement and the year of the first.

Mortality

All five surveys track mortality through vital registration systems and proxy interviews. ALSA, SEBAS, and NHEFS provide exact dates of death and of BMI measurement, so it is straightforward to determine survival to exactly *n* years after BMI measurement. CRELES provides exact date of BMI measurements but only month and year of death. Therefore, in that survey, deaths within *n* years are estimated assuming that every death occurred on the fifteenth day of the month in which it occurred.

Survival estimates are more complicated for ELSA. To protect participant privacy, ELSA provides only year of death (not month or day). It further provides year (but not month or day) of first BMI measurement, and year and month (but not day) of subsequent BMI measurements. This means it is not possible to determine survival status at precisely *n* years for all participants. E.g., a person interviewed in 1998 who dies in 2006 may have died within 8 years of BMI measurement, or may have died slightly after the 8-year mark. A small fraction of ELSA participants, roughly 0.02% in each analysis, has unknown mortality status at exactly *n* years.

To deal with survival uncertainty, this study assumes that death dates and BMI-measurement dates are all evenly distributed throughout the year[1]. It can then be shown that each

[1] Although exact date of BMI measurement is not given in ELSA, the month of measurement is given for waves in 2004-05 and in 2008-09 (but not at the baseline waves, in 1998-99 or in 2001). Available data on the month of measurement support assumptions that interview/measurement dates are evenly spaced throughout the year. In countries with data on exact date of death (e.g., the USA), there is evidence that deaths are more likely to occur in winter than in summer. However, the mean and the median dates of death are almost exactly at the middle of the year, so that deviation from a flat distribution of deaths should not strongly affect applicability of the proof below.

participant with unknown mortality status at *n* years has exactly 50% probability of survival to that point and 50% probability of death (see proof that follows). Survival to *n* years is therefore assigned within the unknown-survival group, at random with 50% probability, using the random number generator in Stata 12.

Proof: Assigning survival probabilities where exact date of death is unknown

Consider deaths that occur in the calendar year *n* years after the year of interview. Deaths that occur before this are known to take place within *n* years; deaths that occur later are known not to fall within *n* years.

For these cases where year of death is equal to year of interview + *n,* survival status at exactly *n* years is known only for some fraction of the decedents: i.e., those for which age at death is (*n*-1) or (*n*+1) years higher than age at interview (since a person cannot survive *n* years without having at least *n* birthdays before death, and cannot die within *n* years if he or she has had more than *n* birthdays). For the other cases, survival status at exactly *n* years is unknown.

Assume that death days and interview dates are both evenly distributed throughout the year. Each possible ordering of these event dates (by month and day, with no regard for year) will then occur with equal probability. There are only two possible orderings, listed below in Table 1A.1. Either the death day falls earlier in the year than the anniversary of the interview date,

and the decedent is dead within *n* years, or the death day falls later in the year than the anniversary of the interview date, and the decedent survives slightly more than *n* years. There is therefore a 50% probability of death within *n* years.

Table 1A.1. Orderings of Events, with Mortality Outcomes.

Possible orderings of dates (month and day only, ignoring year)			*Dead within* n *years*
a	death day	day of interview	yes
b	day of interview	death day	no

APPENDIX 1A REFERENCES

[1] Winship C, Radbill L. Sampling weights and regression analysis. *Sociol Methods Res*. 1994; 23(2):230-257.

APPENDIX 1B: TIME-TO-EVENT ANALYSES (HAZARD MODELS)

Three of the five surveys give detailed information on date of death. However, to protect participant privacy, the other two (CRELES and ELSA) give only month or year or death, rather than exact day. This makes it difficult to conduct time-to-event analyses in these data, especially in the ELSA survey. In that survey, timing information is limited to the year of death, and mortality follow-up is relatively short (only two years in the analyses that include BMI change).

For the three surveys that do include detailed information on date of death, time-to-event analysis (survival analysis) is used to assess robustness of results from penalized logistic regression. This Appendix shows results from three example analyses:

1) the BMI–mortality relationship from Australia in 1992 (with mortality follow-up to 2000),
2) the BMI–mortality relationship from Taiwan in 2000 (with mortality follow-up to 2008), and
3) the BMI–mortality relationship from the USA in 1971-75 (with mortality follow-up to 1979-83).

Results are shown both with and without substantial smoothing, and compared to results from logistic regression analyses like those used in Chapter 1. This allows some assessment of the similarities and differences between results generated using different model types.

Hazard models

In these data, the log of the all-cause mortality hazard appears to increase linearly with age within the age range 50 to 79, just as it does in virtually all human populations (1). Therefore, this Appendix defines the logged mortality hazard in each country as a linear function of age (time-varying), with dummy-variable controls for female sex, current-smoker status at baseline (time-invariant) and former-smoker status at baseline (time-invariant): i.e., assuming proportional hazards by sex and by smoking status. BMI at survey measurement (time-invariant) is included in the model also as a system of linear equations; the implementation, described below, depends on smoothing.

$$\ln(\text{hazard}) = \beta_0 + \beta_1[\text{age}] + \beta_2[\text{female}] + \beta_3[\text{baseline current smoker}] + \beta_4[\text{baseline former smoker}] + f(\text{measured BMI}) \qquad \textbf{\textit{Eqn. 1B.1}}$$

As in Chapter 1, analyses are performed in Stata 12 (StataCorp LP, College Station, TX).

B-*splines without penalty.*

Hazards are first generated with minimal smoothing and compared to results from logistic regression with minimal smoothing. These results without substantial smoothing allow for clear comparison of the hazard models' findings with findings from logistic regression models.

As in the primary analyses in Chapter 1, in hazard models in each country, *B*-splines are used to represent the relationship between BMI and mortality. However, no penalty is introduced to smooth results. (See Appendix 1C for description of *B*-splines and penalty). *B*-splines are constructed as in the primary analyses, with knots at every two BMI points throughout the BMI range (or as close as estimation will allow and still converge, given that there is no penalty). These hazards are then used to generate predicted probabilities of death within eight years, which are compared to predicted probabilities of death within eight years from logistic regression models using *B*-splines (also with no penalty, i.e., equivalent to the primary analyses, but with $\lambda = 0$; see Appendix 1C).

Results with smoothing.

There is currently no Stata program to implement penalized *B*-splines, or *P*-splines (see Appendix 1C), in hazard models. This Appendix therefore uses regression splines to emulate hazard smoothing, generating results comparable to those from Chapter 1's primary analyses using *P*-splines. Regression splines appear more sensitive to the range of the predictor

distribution than *P*-splines are (in logistic regression models), and (both in logistic regression and in hazard regression) the results based on regression splines are sensitive to knot placement (results not shown). Nevertheless, regression splines provide an adequate robustness check.

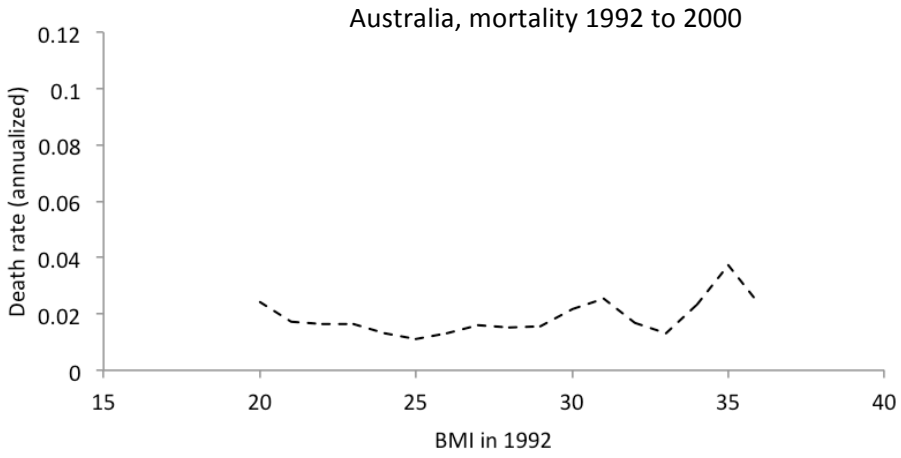
For these analyses, hazards are modeled as in Equation 1B.1, with natural cubic splines (2) for BMI. Internal knots are placed at BMI values of 18.5, 25, 30, and 35: values that correspond to thresholds of the World Health Organization body-weight categories.

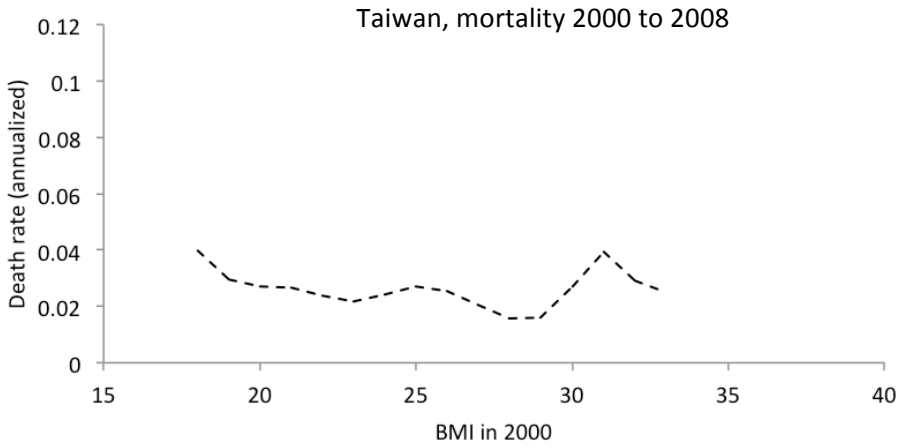
Results

As in the Chapter 1, results are shown for a 72-year-old never-smoking male, and displayed for the middle 95% of the BMI distribution. (I.e., the full sample is used for estimation, but results are not displayed at the tails of the BMI distribution, which are typically very long, and for which results are imprecise due to sparse data).

Figures 1B.1 and 1B.2 show results without substantial smoothing. They are noisy representations of the BMI–mortality relationship, but are useful for comparing results across model type. 1B.1 shows the hazards estimated using *B*-splines. 1B.2 shows predicted probabilities of death within eight years generated using these hazards, alongside the results from logistic regression models with no penalty (i.e., equivalent to results in Chapter 1, except

Figure 1B.1. Mortality Hazards in Three Countries. Results shown for a 72-year-old never-smoking male.





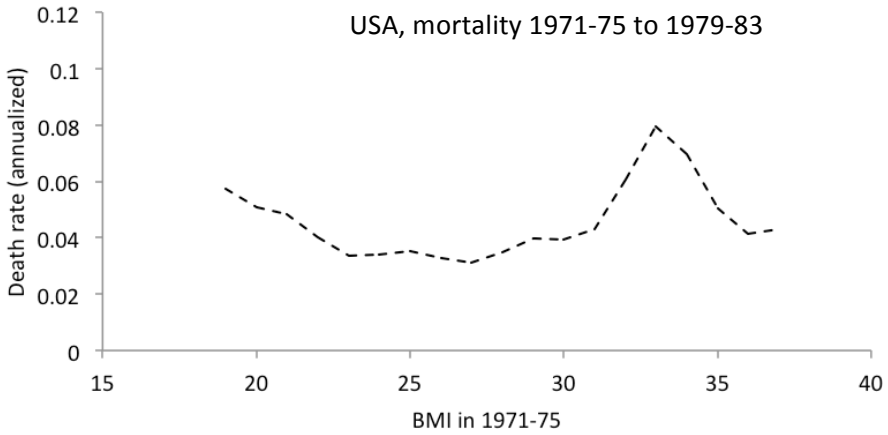
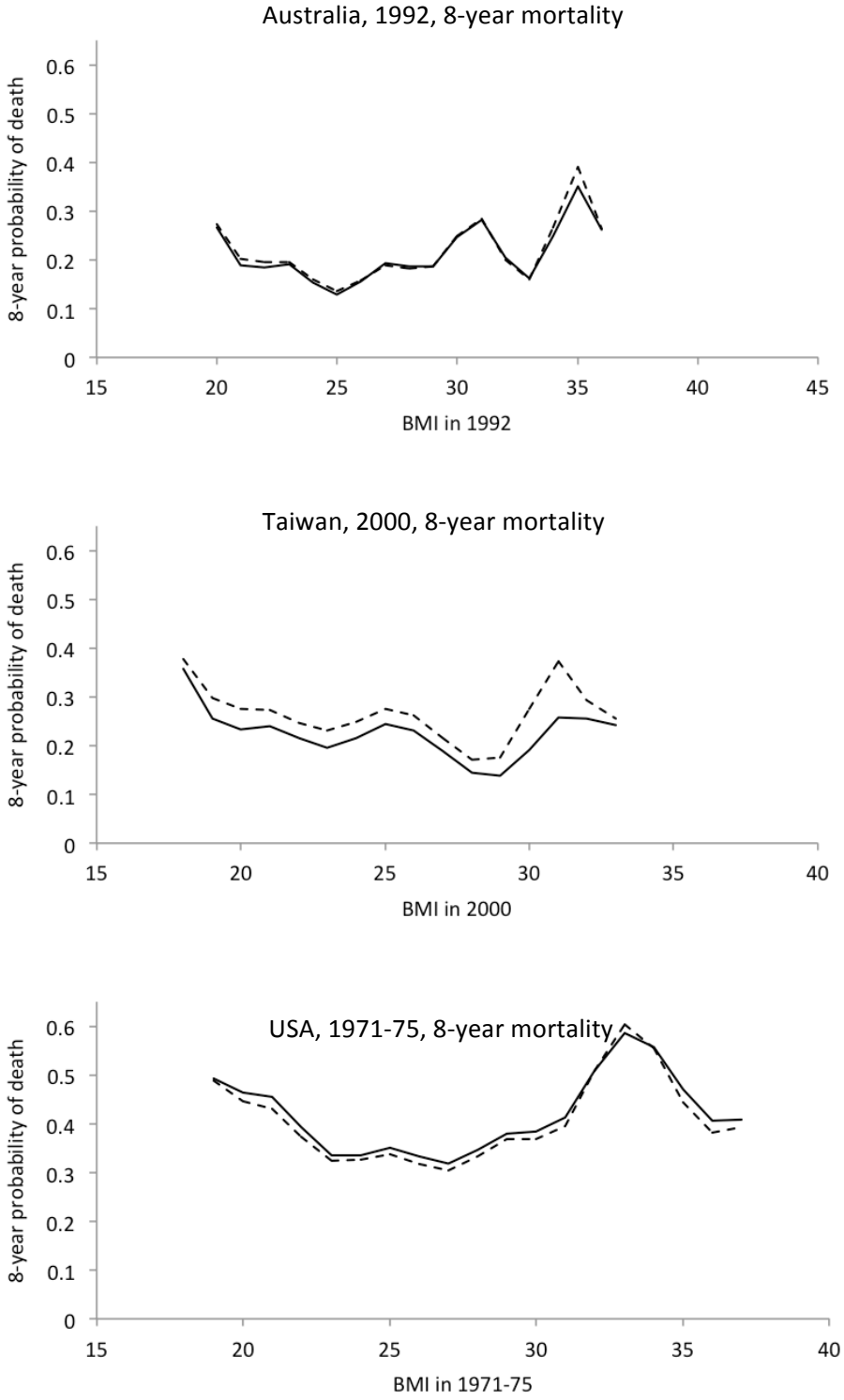


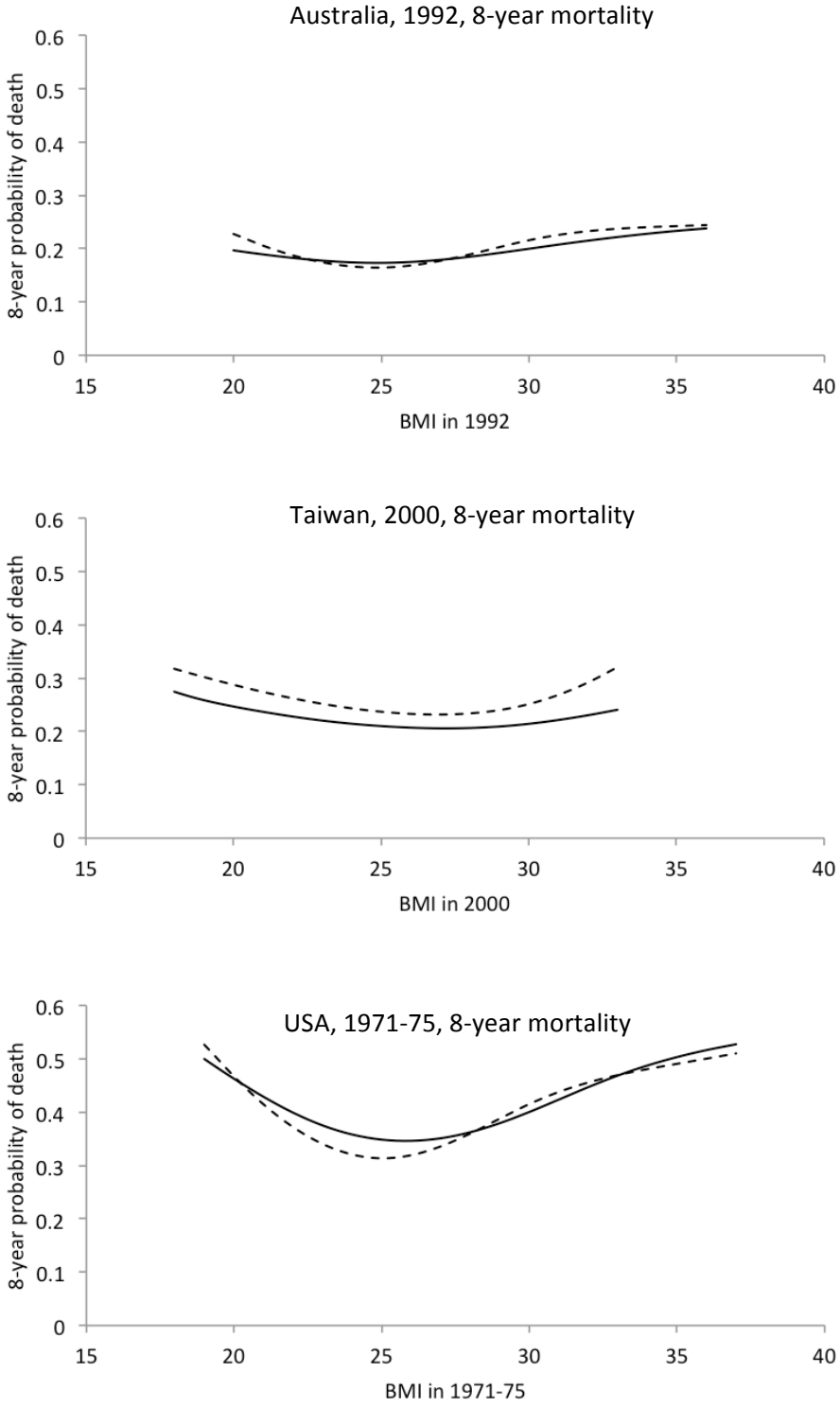
Figure 1B.2. Predicted Probabilities of Death Within Eight Years, Without Smoothing, in Three Countries. Generated from hazard models (dotted lines) and logistic models (solid lines) with *B*-splines, no penalty. Results shown for a 72-year-old never-smoking male.



without smoothing: $\lambda = 0$; see Appendix 1C). In each country, the shape of the relationship between BMI and mortality is similar across the two types of analysis.

Since results look similar across both types of analyses before substantial smoothing, it would be surprising if the smoother results did not. Figure 1B.3 presents these results. Predicted probabilities of death from penalized logistic regression models (solid lines) are the same as those shown in Chapter 1. Predicted probabilities from hazard models (dotted lines) are generated from models with regression splines. The curves from time-to-event analyses may appear to have sharper changes in slope in some regions, but these are quite sensitive to knot placement (results not shown). As in the previous figure, the overall shape of the curves is similar.

Figure 1B.3. Predicted Probabilities of Death Within Eight Years, with Smoothing, in Three Countries. Generated from hazard models (dotted lines) and logistic models (solid lines) with smoothing. Results shown for a 72-year-old never-smoking male.



APPENDIX 1B REFERENCES

[1] Finch CE. *Longevity, Senescence, and the Genome.* Chicago: The University of Chicago Press; 1990.

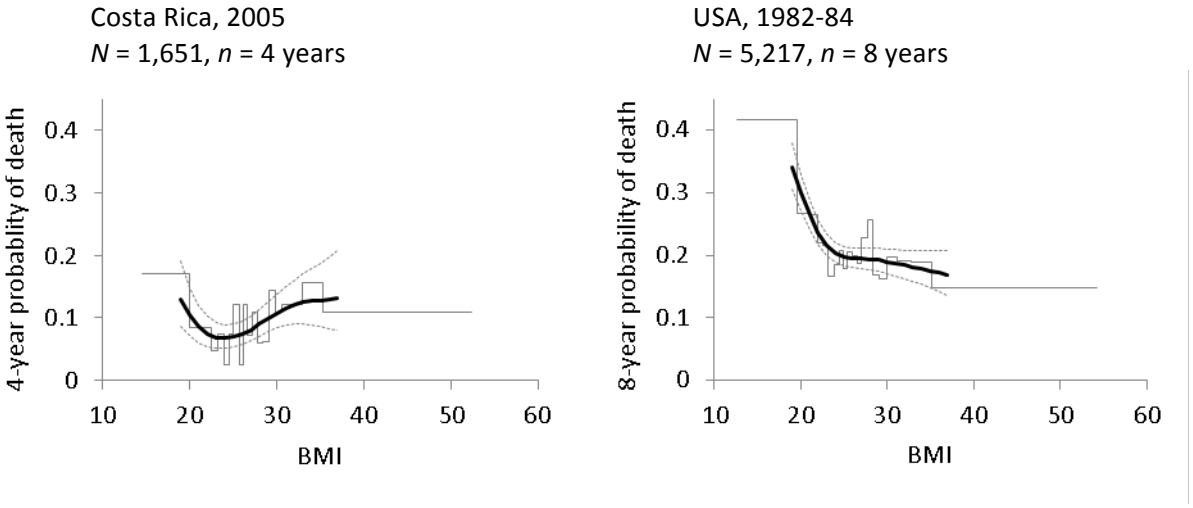
[2] Hastie TJ, Tibshirani RJ. *Generalized Additive Models.* London: Chapman & Hall; 19

APPENDIX 1C: DESCRIPTION OF *P*-SPLINES

Chapter 1 uses *P*-splines, or penalized *B*-splines, as described by Eilers and Marx (1) to produce data-driven estimates of the relationships between BMI and mortality, or between BMI change and mortality. *P*-splines may be considered a compromise of sorts between regression splines and smoothing splines, as they are constructed from a series of polynomial pieces (like regression splines) with difference penalties on regression coefficients of adjacent splines (for smoothing). Using *P*-splines allows a researcher to sidestep issues in selection of knot number and location that are so crucial to regression splines, but the splines are computationally cheaper than smoothing splines.

This Appendix briefly describes the *P*-spline method and selection of the smoothing parameters used in Chapter 1, based on the Akaike Information Criterion. Figure 1C.1 shows two examples of smoothed BMI–mortality curves generated using the technique. These examples, comparing model results to observed proportion dead by BMI, demonstrate that predictions from a *P*-spline model can fit data quite closely, even when the data are noisy and when the predictor–outcome relationship is not easily described by a low-order polynomial.

Figure 1C.1. Smoothing the Body Mass Index – Mortality Relationship with *P*-Splines: Two Examples.



Solid grey lines show the observed proportion dead by body mass index (BMI), for each 5% of the BMI distribution. Black lines show predicted probabilities of death (within specified period) from penalized logistic regressions using *P*-splines, without adjustment for covariates. Dotted lines represent 95% confidence intervals, conditional on a smoothing parameter (see below). *N* is sample size; *n* is duration of mortality follow-up.

Description of method: An overview

The *P*-spline method makes use of a large number of *B*-splines: collections of polynomial pieces, joined together at "knots" placed at designated values of a predictor variable. Construction of *B*-splines is described below. In any given analysis, enough *B*-splines are constructed to cover the entire range of the predictor of interest. By then regressing a given outcome on the *B*-splines, one can produce a relatively detailed representation of the association between predictor and outcome, without prior knowledge of where inflection points should occur along the fitted curve. Using many *B*-splines and many knots allows a more or less exact fit to the

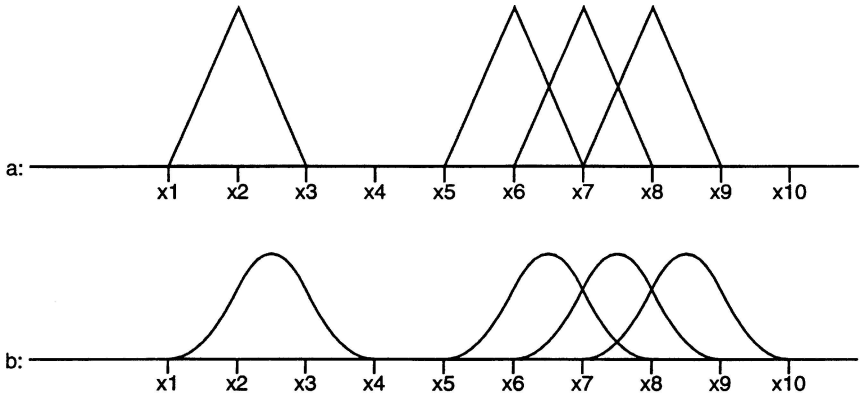
data. Since the *P*-spline method uses relatively many *B*-splines, it then allows for smoothing to avoid excessive noise in the fitted relationship. A single tuning parameter, called λ, is incorporated into the (penalized) regression likelihood function, penalizing differences on coefficients of adjacent B-splines.

Construction of B-splines

Readers may wish to consult a reference on *B*-splines, such as one by de Boor (e.g., (2)), or the Eilers and Marx paper on *P*-splines (1). However, these splines are illustrated below in Figure 1C.2, and their quintessential features (paraphrased from Eilers and Marx (1)) are as follows:

- Each *B*-spline of degree q is positive on a domain spanned by $q + 2$ knots, and is 0 elsewhere.
- The *B*-spline is constructed from $q + 1$ polynomial pieces, each of degree q.
- The polynomial pieces are joined together at q inner knots.
- Wherever the pieces join, their derivatives to order $q - 1$ are continuous.

Figure 1C.2. Examples of *B*-Splines. Reproduced from Eilers and Marx ((1), Figure 1).



"Illustrations of one isolated *B*-spline and several overlapping ones: a) degree one; b) degree two."

All analyses in Chapter 1 use cubic *B*-splines. Each *B*-spline therefore covers a range spanning five knots, and consists of four cubic pieces, joined together at three internal knots. At each of these internal knots, the first and second derivatives of adjoining pieces are continuous.

For this study, *B*-splines are constructed with knots at every 2 points of the predictor variable of interest (e.g., with knots at BMI = 15, BMI = 17, …, for analyses of BMI and mortality, and, e.g., with knots at BMI change = -8, BMI change = -6, …, for the analyses of BMI change and mortality), and with one additional knot at the maximum value of the predictor variable and one knot at the minimum. The method does not require knots to be evenly spaced.

From *B*-splines to *P*-splines

P-splines are fit by penalized log-likelihood estimation. In particular, to smooth the estimated relationship between predictor and outcome, the *P*-spline method penalizes higher-order differences between the regression coefficients of adjacent *B*-splines. This is done with a single tuning parameter, called λ.

One forms the penalized regression likelihood function simply by subtracting a penalty term (including λ) from the log-likelihood, $l(y; a)$:

$$L = l(y; a) - (\lambda / 2) \sum_{j=k+1}^{n} (\Delta^{k} a_j)^2$$

where *y* is the outcome variable,

a is vector of coefficients corresponding to the constructed *B*-splines,

a_j is the coefficient for the *j*th *B*-spline,

and $\Delta^{k} a_j$ is the *k*th-order difference between a_j and a_{j-1}.

In this study, second-order differences are penalized, so $\Delta^{k} a_j = \Delta\Delta a_j = a_j - 2a_{j-1} + a_{j-2}$. Eilers and Marx demonstrate that such second-order difference penalties produce qualitatively similar results to somewhat better-known smoothing-spline techniques that penalize the second derivatives of each spline (3-5). However, penalizing differences on adjacent coefficients allows for simpler calculation.

Note, crucially, that smaller values of λ represent a smaller penalty in the likelihood function; larger values represent a larger penalty, and thus produce smoother fitted lines. E.g., if λ = 0 there is no smoothing. As λ goes to infinity, the fitted line approaches a straight line.

Selecting a value for λ, the smoothing parameter

The smoothing parameter, λ, is chosen based on Akaike Information Criterion (AIC) values. For each analysis in Chapter 1, a model is fit and AIC calculated using a number of possible λ values: 0.001, 0.01, 0.1, 1, 10, 100, and 1,000. Results in the chapter are shown for each analysis using the λ value (of those listed) that yields the optimal (lowest) AIC, with two exceptions. First, the maximum λ value selected is 100, since larger values yield lowest AIC only when the sample size is unusually small (effective degrees of freedom are low) and, in those cases, a value of λ = 1,000 appears to lead to oversmoothing: i.e., monotonically increasing or decreasing functions, when a smaller λ value would suggest more complicated relationships. Second, in the analysis of BMI and mortality net of BMI change in the United States, the λ value selected based on AIC appears to produce unusually noisy results, relative both to results from other countries and to results from the United States in other years during roughly the same period (e.g., adjusting the BMI–mortality curve in 1982-84 for 10-year or 9-year BMI change rather than for 11-year BMI change); results for this analysis are therefore shown in the Chapter 1 for λ = 10 rather than λ = 1. Figure 1C.3 displays these USA results using λ = 10 alongside those using λ = 1 for

comparison. Table 1C.1 presents all of the AIC values calculated for the analyses reported in Figures 1.3, 1.4, and 1.5 in Chapter 1.

Figure 1C.3. Analysis of BMI and Mortality Net of BMI Change (over 11 Years), USA, 1982-84.

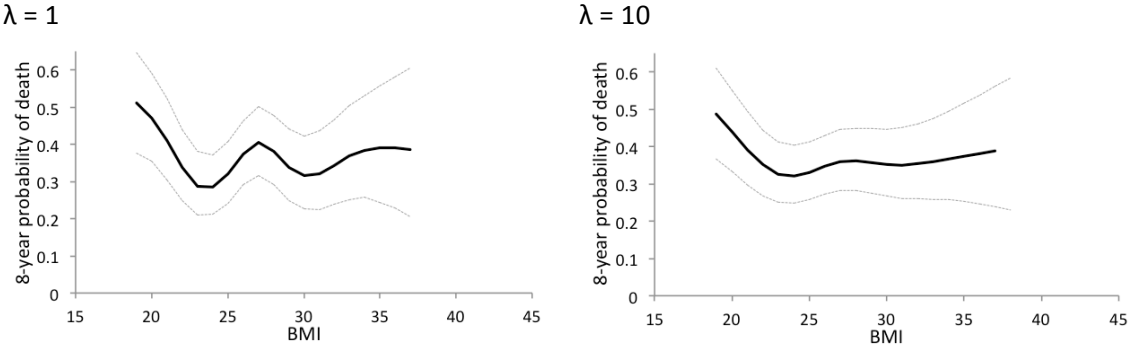


Table 1C.1. AIC Values for the Analyses Shown in Figures 1.3, 1.4, and 1.5, by Select Values of λ. For each analysis, the best-fit AIC (among these choices of λ) is shown in bold type. The value selected to produce the figure is shown with grey shading.

	Values of λ						
	0.001	0.01	0.1	1	10	100	1000
Analyses of Figure 1.3: BMI & mortality							
Australia, 1992	1089.0	1086.6	1084.4	1080.7	1076.9	**1075.7**	1076.4
Australia, 1994	795.2	792.9	789.8	786.9	782.8	780.2	**779.3**
Costa Rica, 2005	1015.2	1015.0	1014.3	1011.6	**1008.9**	1010.8	1017.2
Costa Rica, 2007	521.1	519.9	517.1	512.7	508.8	**508.1**	508.9
England, 1998-99	5219.6	5217.8	5217.2	5213.8	5210.2	**5208.7**	5219.8
England, 2001	2468.2	2466.6	2462.6	2458.1	2454.1	**2451.3**	2452.9
England, 2004-05	1175.1	1175.8	1172.9	1171.6	**1171.5**	1171.8	1177.2
Taiwan, 2000	852.8	852.3	849.9	845.6	843.3	**841.7**	842.6
Taiwan, 2006	309.7	308.6	307.0	304.7	303.0	302.1	**302.0**
USA, 1971-75	2555.2	2555.3	2552.5	2547.9	2544.1	**2543.0**	2555.6
USA, 1982-84	4479.8	4477.2	4473.3	4469.5	4466.0	**4464.1**	4476.5
Analyses of Figure 1.4: BMI change & mortality							
Australia, 1992 to 1994	692.6	690.5	686.9	683.8	**680.8**	683.5	691.3
Costa Rica, 2005 to 2007	475.8	474.4	**472.9**	474.2	477.5	479.3	479.6
England, 2001 to 2004	**183.7**	184.4	186.2	187.5	188.0	188.2	188.3
England, 1998-99 to 2004-05	533.5	532.2	530.1	**528.6**	529.2	532.1	532.8
England, 2001 to 2005	204.4	**202.6**	203.4	205.3	206.7	206.8	206.8
England, 1998 to 2005	173.8	172.9	172.3	171.5	**170.9**	171.1	171.1
Taiwan, 2000 to 2006	301.2	299.9	298.7	298.5	297.7	297.2	**297.1**
USA, 1974-75 to 1982-83	632.4	630.3	627.2	624.2	**622.9**	626.1	628.4
USA, 1973-74 to 1982-83	752.1	749.8	746.3	742.7	**741.7**	744.1	745.7
USA, 1972-74 to 1982-84	1195.4	1194.4	1193.3	1191.4	**1189.7**	1193.7	1199.2
USA, 1971-73 to 1982-84	**1252.7**	1259.4	1261.9	1264.0	1265.3	1269.0	1271.8
USA, 1971-72 to 1983-84	456.7	454.9	452.4	**451.6**	452.2	453.2	454.1
Analyses of Figure 1.5: BMI, BMI change, & mortality							
Australia, 1992 to 1994	688.9	687.6	688.4	687.2	684.9	683.1	**682.3**
Costa Rica, 2005 to 2007	490.3	488.8	485.3	481.1	**477.2**	477.7	479.0
England, 1998-99 to 2004-05	533.3	533.7	533.1	529.8	**527.0**	528.5	533.6
Taiwan, 2000-2006	306.0	305.6	305.1	303.4	301.9	301.1	**300.1**
USA, 1971-73 to 1982-84	1268.2	1267.6	1266.2	**1264.2**	1266.2	1267.5	1270.5

APPENDIX 1C REFERENCES

[1] Eilers PHC, Marx BD. Flexible smoothing with *B*-splines and penalties. *Stat Sci*. 1996; 11:89-121.

[2] De Boor C. *A Practical Guide to Splines*. Berlin: Springer; 1978.

[3] Hastie TJ, Tibshirani RJ. *Generalized Additive Models.* London: Chapman & Hall; 1991.

[4] O'Sullivan F. A statistical perspective on ill-posed inverse problems (with discussion). *Stat Sci*. 1986; 1(4):505-527.

[5] O'Sullivan F. Fast computation of fully automated log-density and log-hazard estimators. *SIAM J Sci Stat Comput*. 1988; 9(2):363-379.

APPENDIX 1D: TESTS FOR EFFECT MODIFCATION BETWEEN BMI LEVEL AND BMI CHANGE

BMI change could affect mortality risk differently at different levels of BMI. For example, a person whose weight loss leaves him underweight may have higher mortality than someone whose weight loss leaves him overweight, even if both experience the same total BMI loss. Similarly, people who gain weight and become obese may have higher mortality than people who experience comparable weight gain (from a lower starting BMI) to become overweight.

This Appendix tests formally for effect modification between BMI level and BMI change in the analyses (described in Chapter 1) of BMI and mortality net of BMI change. It considers the following five scenarios:

a) BMI change < 0 resulting in BMI < 18.5
    - *weight loss leading to underweight*

b) BMI change < 0 resulting in BMI < 22
    - *weight loss leading to underweight or low normal weight*

c) BMI change < 0 resulting in BMI < 25
    - *weight loss leading to underweight or normal weight*

d) BMI change > 0 resulting in BMI > 30

- *weight gain leading to obesity*

e) BMI change > 0 resulting in BMI > 35

- *weight gain leading to "class II" or "class III" obesity, by WHO standards*

For each scenario, a dummy variable is generated. Values are assigned to each study participant as dummy = 1 if scenario occurs; 0 otherwise. Each of the dummy variables is then added, one by one, without the others, to the analyses described in section, "3) BMI, BMI change, and mortality" in Materials and Methods.

In total, this yields 25 tests for effect modification: one for each of five dummy variables (five scenarios of possible effect modification) in five different analyses (one per country). Table 1D.1 shows results of the 25 tests: odds ratios predicting death within *n* years associated with each scenario in each country (controlling, as described in Chapter 1, for age, age squared, sex, smoking status, and the main effects of BMI and BMI change). The table also shows 95% confidence intervals.

There is no evidence of effect modification, as no test yields an odds ratio statistically significantly different from 1.0 at 5%. While statistical power for these tests is low, the point estimates are about as likely to be in the hypothesized direction as the opposite. There is no discernible pattern among the odds ratios.

Table 1D.1. Odds Ratios, Predicting Death Within *n* Years, for Five Possible Effect-Modification Scenarios.

Country	Scenario (a)		Scenario (b)		Scenario (c)		Scenario (d)		Scenario (e)	
	OR	95% CI	*OR*	95% CI	*OR*	95% CI	*OR*	95% CI	*OR*	95% CI
Australia	0.20	<0.01, 8.66	1.07	0.45, 2.51	1.05	0.59, 1.86	1.31	0.54, 3.16	0.64	0.17, 2.45
Costa Rica	0.77	0.08, 7.53	0.64	0.17, 2.39	0.54	0.21, 1.39	1.30	0.49, 3.41	2.85	0.60, 13.58
England	0.34	0.04, 3.22	1.16	0.35, 3.88	1.67	0.65, 4.32	0.98	0.35, 2.79	3.22	0.62, 16.88
Taiwan	1.44	0.23, 9.16	1.20	0.40, 3.60	0.72	0.29, 1.81	2.67	0.60, 11.95	<0.01	near ∞
USA	0.73	0.23, 2.31	1.22	0.67, 2.22	0.70	0.45, 1.08	1.05	0.53, 2.09	1.14	0.31, 4.24

OR is odds ratio; CI is confidence interval; see previous page to define scenarios (a) through (e)

APPENDIX 2A: DETAILS OF THE CRELES AND ELSA DATA

Data come from two longitudinal population surveys of older adults: the Costa Rican Longevity and Healthy Aging Study (CRELES) and the English Longitudinal Study of Ageing (ELSA). These data are well-suited to studying determinants of weight change, as each survey measures participants' height and weight over three different survey waves, rather than relying on self-reported values. Both surveys also collect detailed participant reports about health history and current health conditions. Both link to national death registration systems to track mortality.

Participants surveyed

Both ELSA and CRELES are nationally representative, within certain age limits, of the populations from which they are drawn. Both focus on older adults. To ensure rough comparability across the two population samples – especially among the elderly, for whom the relationship between BMI and health may vary by age (1) – we limit the study samples to participants who are at least 50 years old and no more than 79 years old at the time when weight change is measured. (See below for description of weight-change measurement.) CRELES has no study participants in their 50s; the age range for the Costa Rican sample is 61-79. The age range for the English sample is 52-79.

Timing of survey waves

Both ELSA and CRELES collect height and weight measurements at three separate survey waves. In CRELES, these measurements occur roughly two years apart, in 2004-06 (concentrated in 2005), in 2006-08 (concentrated in 2007), and in 2009-10 (concentrated in 2009). To ensure comparable measurements of BMI change over time, we include in analysis only those observations for which BMI measurements were spaced at least 18 months apart and no more than 30 months apart: i.e., all weight-change intervals in Costa Rica are between 1.5 and 2.5 years long. In ELSA, the survey waves with height and weight measurements occur first in 1998-99 and 2001 (since ELSA's first wave was split), then in 2004-05, and finally in 2008-09. We restrict the ELSA sample to include only observations in which BMI measurements are spaced three to four years apart (i.e., BMI change is measured between 2001 and 2004 or 2005; between 2004 and 2008; or between 2005 and 2008 or 2009). Following these sample restrictions, weight-change measurements should be comparable *within* each of the two surveys. However, it is important to note that the intervals between measurements are substantially shorter in Costa Rica (1.5 to 2.5 years) than they are in England (3 to 4 years).

ELSA has two additional survey waves, collected in 2002-03 (between times 1 and 2) and in 2006-07 (between times 2 and 3), but neither of these collect height and weight information. Data from these waves are used to determine some changes in disease status between times 1 and 2 or between times 2 and 3. The waves are otherwise not used in the present study.

Health information collected

At each survey wave, CRELES and ELSA ask survey participants to rate their health overall, and to provide details about recent and past medical diagnoses. Participants' reports of medical diagnoses may not provide a full picture of population health, as some survey participants may not know that they are sick, may not know the clinical name for the conditions affecting them, or may choose to withhold information from survey-takers. However, both England and Costa Rica have relatively well-functioning healthcare systems that provide universal coverage to citizens and permanent residents. As a result, it is likely that survey respondents with major, symptomatic conditions will have interacted with the formal healthcare sector, and have had opportunities for diagnosis.

All of the health analyses in Chapter 2 focus on major clinical conditions – i.e., cancer, diabetes, heart disease (excluding hypertension only), stroke, chronic lung disease, and arthritis – or (in the sensitivity analysis described in Appendix 2B) on self-rated health. Except in the case of diabetes, explained below, all clinical conditions from the CRELES survey are determined by respondents' answers to a series of questions of the form, "Has a doctor ever told you that you had […]?" We assume that CRELES participants only have active cancer if they report it being diagnosed in the past two years. (Survey respondents with earlier diagnoses are assigned missing values for the presence of cancer, since it is unknown whether the cancer is still active.) For all other disease diagnoses, we assume that the condition is still present if the respondent claims it has ever been diagnosed by a doctor. The survey questions are similar in ELSA. In that

survey, study participants are asked at times 2 and 3 if a doctor has ever diagnosed various conditions. At time 1, however, participants are not asked about each condition specifically; instead, they are asked to list each "long-standing illness, disability or infirmity" that affects them. At time 1, we assume that a condition is present if an ELSA respondent reports that it currently affects them. At later waves, we assume, as we do in CRELES, that most conditions are present as long as they have ever been diagnosed by a doctor, but we treat cancer as present at time 2 or time 3 only if it has been diagnosed during the previous measurement interval. For context, we include the unadjusted proportion of survey respondents reporting each condition, along with sample demographic characteristics, in Table 2A.1 (Costa Rica) and Table 2A.2 (England).

There are a few important differences in disease definitions between CRELES and ELSA. In ELSA, study participants are asked about "chronic lung disease such as chronic bronchitis or emphysema"; in CRELES the question about lung disease includes asthma and tuberculosis as well as chronic bronchitis and emphysema. Diabetes diagnosis in ELSA is determined solely by participant self-report, whereas in CRELES it is determined in part by fasting glucose and glycosylated hemoglobin tests that were administered during the study; this means that diabetes prevalence appears higher in Costa Rica than it does in England. As a result, diabetes status is likely to be more accurate in CRELES than in ELSA, but ELSA participants who report diabetes may well be, on average, sicker than those in CRELES. Finally, in CRELES, the question about heart attack is specific to that type of event, whereas in ELSA the survey combines a few different serious manifestations of cardiovascular disease – "heart attack, angina, and

Table 2A.1. Descriptive Statistics, Costa Rica. Demographic and health traits of Costa Rican study participants, by BMI change, measured over an interval of 1.5 to 2.5 years (CRELES: 2005-07 or 2007-09).

CRELES AGE, SEX, SMOKING STATUS, & BMI

	N	Mean age at interval start	% female	% current smoker at interval start	% former smoker at interval start	Mean BMI at interval start	Mean BMI at interval end
Total sample	1,285	69.07	52.7%	10.3%	32.1%	27.16	27.14
Weight losers							
BMI change < -1	238	69.54	59.7%	7.2%	28.8%	28.94	26.93
BMI change < -2	72	70.10	63.9%	5.6%	26.8%	30.17	26.87
Weight maintainers							
BMI change > -1, < 1	804	69.13	49.8%	11.4%	33.6%	26.69	26.68
BMI change > -2, < 2	1,149	69.04	51.7%	10.7%	32.2%	26.97	26.98
Weight gainers							
BMI change > 1	243	68.41	55.6%	9.9%	30.6%	26.99	28.87
BMI change > 2	64	68.48	57.8%	9.4%	37.5%	27.21	30.44

CRELES HEALTH

	Cancer at interval start	New cancer by interval end	Diabetes† at interval start	New diabetes† by interval end	Heart disease† at interval start	New heart disease by interval end	Ever had stroke by interval start	New stroke by interval end	Arthritis at interval start	New arthritis by interval end	Lung disease† at interval start	New lung disease† by interval end	Self-rated health* at interval start
Total sample	1.1%	1.1%	23.9%	4.3%	4.1%	1.2%	2.7%	0.5%	14.6%	3.1%	16.1%	2.0%	3.30
Weight losers													
BMI change < -1	1.8%	0.8%	29.7%	5.0%	5.0%	2.1%	3.4%	0.4%	19.1%	3.0%	23.2%	1.3%	3.44
BMI change < -2	1.5%	2.8%	23.6%	8.3%	5.6%	4.2%	4.2%	0.0%	25.4%	2.8%	26.4%	1.4%	3.42
Weight maintainers													
BMI change > -1, < 1	1.3%	1.2%	23.3%	4.2%	3.6%	1.1%	2.2%	0.6%	12.5%	2.9%	14.8%	2.4%	3.25
BMI change > -2, < 2	1.2%	1.0%	24.4%	4.1%	3.8%	1.0%	2.7%	0.6%	13.7%	3.1%	15.3%	2.2%	3.28
Weight gainers													
BMI change > 1	0.0%	0.8%	20.2%	3.7%	4.9%	0.4%	3.7%	0.4%	17.4%	3.7%	13.6%	1.6%	3.32
BMI change > 2	0.0%	0.0%	15.6%	3.1%	7.8%	0.0%	1.6%	0.0%	18.8%	3.1%	18.8%	0.0%	3.47

* Self-rated health is given on five-point scale, with 1 = excellent and 5 = poor.
† See Chapter 2 or Appendix 2A for differences in disease definitions between CRELES and ELSA.

Table 2A.2. Descriptive Statistics, England. Demographic and health traits of English study participants, by BMI change, measured over an interval of 3 to 4 years (ELSA: 2001-04/05 or 2004/05-2008/09).

ELSA AGE, SEX, SMOKING STATUS, & BMI

	N	Mean age at interval start	% female	% current smoker at interval start	% former smoker at interval start	Mean BMI at interval start	Mean BMI at interval end
Total sample	4,423	61.61	54.4%	17.7%	43.7%	27.98	28.17
Weight losers							
BMI change < -1	865	62.65	55.3%	17.7%	46.8%	29.14	26.89
BMI change < -2	373	62.53	61.4%	21.9%	46.2%	30.37	27.02
Weight maintainers							
BMI change > -1, < 1	2,290	61.66	50.4%	16.4%	43.6%	27.36	27.39
BMI change > -2, < 2	3,519	61.75	52.0%	16.1%	43.9%	27.61	27.72
Weight gainers							
BMI change > 1	1,258	60.80	60.9%	20.1%	41.7%	28.29	30.49
BMI change > 2	531	60.06	65.1%	25.3%	40.6%	28.68	31.91

ELSA HEALTH

	Cancer at interval start	New cancer by interval end	Diabetes† at interval start	New diabetes† by interval end	Heart disease† at interval start	New heart disease by interval end	Ever had stroke by interval start	New stroke by interval end	Arthritis at interval start	New arthritis by interval end	Lung disease† at interval start	New lung disease† by interval end	Self-rated health* at interval start
Total sample	2.5%	1.8%	6.0%	2.5%	5.6%	3.8%	0.9%	0.9%	24.0%	11.7%	2.4%	2.6%	2.35
Weight losers													
BMI change < -1	2.6%	2.1%	9.5%	4.3%	6.5%	4.2%	0.9%	0.7%	26.2%	11.5%	3.1%	3.2%	2.43
BMI change < -2	2.4%	4.3%	13.0%	5.4%	6.2%	5.6%	0.5%	1.3%	32.2%	11.6%	3.2%	4.2%	2.60
Weight maintainers													
BMI change > -1, < 1	2.4%	2.1%	4.7%	2.1%	5.1%	3.8%	0.8%	0.8%	21.5%	11.4%	2.0%	2.3%	2.29
BMI change > -2, < 2	2.6%	1.7%	5.0%	2.2%	5.4%	3.7%	0.9%	0.7%	22.4%	11.4%	2.1%	2.4%	2.31
Weight gainers													
BMI change > 1	2.4%	1.2%	6.0%	1.8%	5.8%	3.3%	0.9%	1.4%	27.0%	12.4%	2.6%	2.7%	2.40
BMI change > 2	1.7%	1.3%	7.8%	2.3%	6.4%	3.0%	0.8%	1.9%	28.6%	14.3%	2.7%	2.7%	2.48

* Self-rated health is given on five-point scale, with 1 = excellent and 5 = poor.

† See Chapter 2 or Appendix 2A for differences in disease definitions between CRELES and ELSA.

congestive heart failure" – into a single marker. This means that our study analyzes the association of heart attack with weight loss in Costa Rica, but the association of heart disease more broadly with weight loss in England.

Measurement of height, weight, and body mass index

All height and weight values come from direct measurement rather than self-reports. Due to a systematic error in height measurement at CRELES wave 2, we estimate each CRELES participant's height at that time as the mean of his height at waves 1 and 3, or, for participants with no measurement at wave 3, as equal to height at wave 1.

To limit the impact of extreme outliers and presumed recording errors, we restrict the sample of observed BMI-change values in both CRELES and ELSA to those from survey respondents with recorded heights of at least 1.22 m (48 in) and no more than 2.03 m (80 in) at both the beginning and the end of the weight-change interval. We further restrict the sample to observations from people with recorded weights of no less than 32 kg (70 lbs). These cut-offs are somewhat arbitrary, but were chosen in part because they eliminate only a very small proportion of survey participants (roughly 1 per 1,000 in ELSA, e.g.). Both CRELES and ELSA also have a maximum possible weight beyond which their scales were not accurate, around 130 kg. Heavier participants are not included in the sample as they have no measured weight values.

In these data, BMI change is due almost exclusively to weight change and not to changes in height. All the study participants are middle-aged or older (aged 50-79 at the end of the weight-change interval). As a result, observed heights change little over time: on average well under 1 cm between one survey wave and the next, and with very slight "growth" about as common as very slight "shrinking" – a finding that suggests that observed height change over time is driven largely by random measurement error. We eliminate 13 BMI-change observations in ELSA and 13 in CRELES as presumed recording errors; they show either implausibly large changes in height between two survey waves (growing or shrinking by more than 10 cm [4 in]), or extreme changes in weight (more than 25kg [55 lbs]) that are inconsistent with waist-circumference measurements or interviewer assessments of whether the respondent is underweight or overweight.

Total sample size, after restrictions

The analyses in this study use one valid BMI-change observation per survey respondent. Following the sample restrictions described above – by age, by duration of intervals between survey waves, and by height and weight measurements – there are 1,285 in CRELES and 4,423 in ELSA.

APPENDIX 2A REFERENCES

[1] Stevens J, Cai J, Pamuk ER, et al. The effect of age on the association between body-mass index and mortality. *New Eng J Med.* 1998; 338:1-7.

APPENDIX 2B: SENSITIVITY ANALYSIS

To gauge robustness of results, we test several model specifications beyond those described in the main text of Chapter 2.

First, we examine how results change when we use different definitions of disease or ill health. Although we do not expect diseases in the primary analysis to be highly collinear, we test for the association between BMI loss and composite measures of disease: in England, self-reports of "any chronic illness" or "any chronic illness that limits activities" at the start of the weight-change interval or emerging during the interval; and in both Costa Rica and in England a composite of the 12 disease diagnoses variables that are considered in the primary analysis (i.e., six conditions at baseline, and six disease-onset variables). As in the primary analysis, we test each composite measure for associations with BMI change < -1.0 point and with BMI change < -2.0 points, using logistic regression with controls (measured at the start of the interval) for age, sex, smoking status (current, former, never), and initial BMI. In separate analyses in each country, we then also test for associations between BMI loss (of the two different magnitudes) and starting self-rated health, operationalized either as a continuous marker (on a five-point scale) or as a dichotomous one (first, [poor or fair] v. [good, very good, or excellent], and, second, in a new model, [poor] v. [fair, good, very good, or excellent]). We use logistic regression analyses to test for associations of each self-rated-health marker with

each BMI-loss outcome, adjusting, as above, for age, sex, smoking status, and BMI as they were measured at the beginning of the weight-change interval.

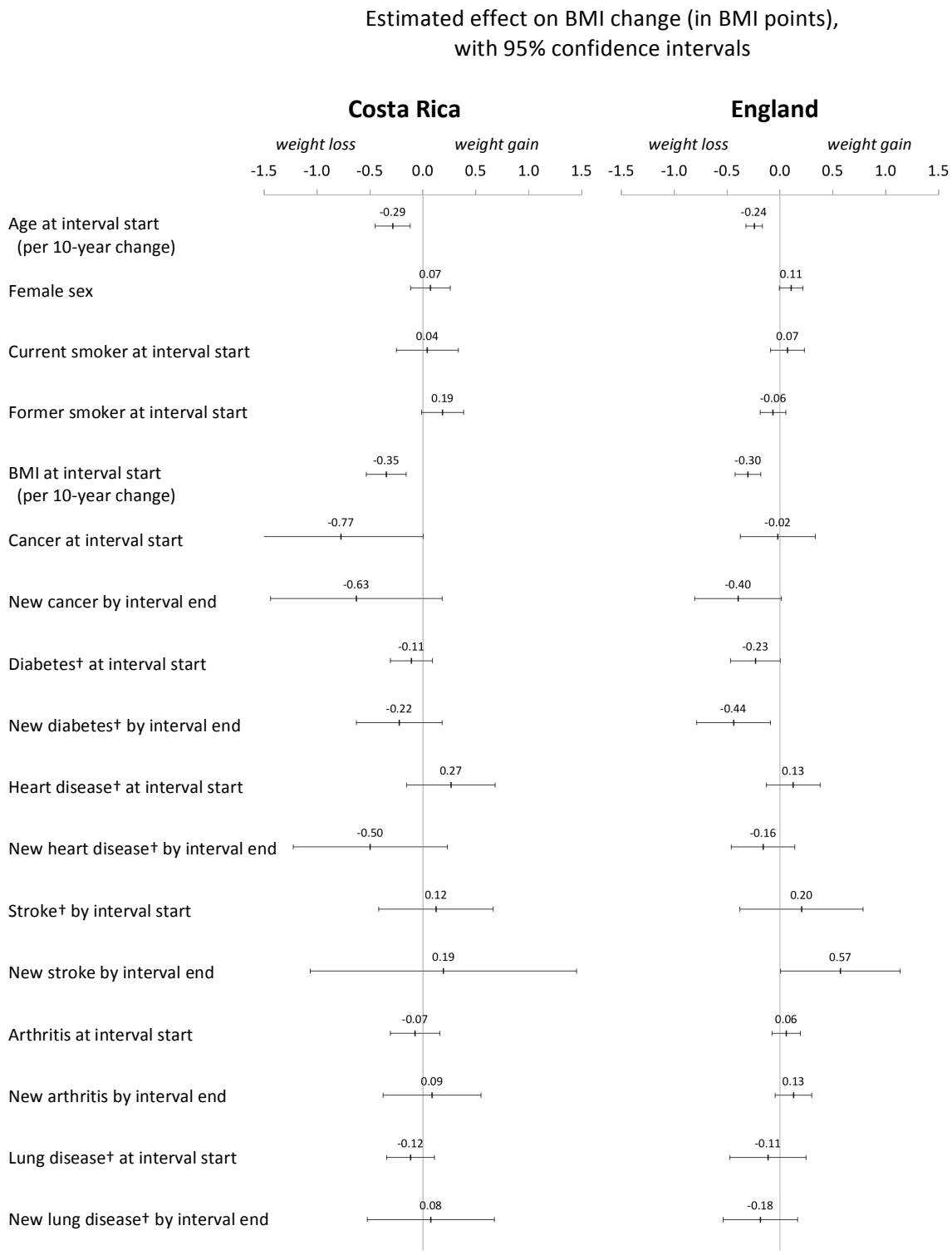
In the second set of robustness checks, we examine alternatives to logistic regression, and test whether a different functional form might better detect relationships between BMI loss and disease. Linear regression has the disadvantage that it does not distinguish a variable that has no effect on weight change from one that affects equally the probability of all extreme changes, whether positive or negative. However, if we treat BMI change as a continuous variable (in linear regression) rather than as a dichotomous one (in logistic regression), we may obtain richer outcome information from the same observations. In our robustness checks, therefore, we test for linear associations of clinical conditions with BMI change (continuous) and with weight change (continuous, in kilograms), as well as for linear associations of BMI change or weight change with self-rated health measures (either continuous and dichotomous, as described above). Lastly, we use multinomial logistic models to test whether individual disease conditions, the composite measures of disease, or self-rated health have an association with the probabilities of both weight gain and weight loss, relative to the probability of maintaining roughly stable weight. We might expect this to occur, for example, if some diseases lead to weight gain while others lead to weight loss, or if people who have been diagnosed with disease are likely to lose weight when they are first sick but then to regain it once they recover. All robustness-check models, as above, include controls (measured at the start of the interval) for age, sex, smoking status (current, former, never), and starting BMI.

Results

Results are largely consistent with the primary findings shown in Chapter 2. In the first series of checks, examining health definitions, we find no evidence of additional relationships in Costa Rica between health markers and BMI loss more than 1.0 point or more than 2.0 points, although many statistically insignificant associations (with composite measures of disease or with self-rated general health) are in the hypothesized direction. In England, we find one additional association with BMI loss more than 1.0 point at a significance level of 5%; baseline self-rated general health, defined as a dichotomous variable for poor v. any other report, is associated with a 50% increase in the odds of BMI change < -1.0 point (*OR* = 1.50 [95% CI: 1.02 – 2.21]). We see further associations in England between BMI loss of more than 2.0 points and composite measures of disease ("any chronic illness" or "any chronic illness that limits activity") and also between self-rated general health at baseline (continuous) and BMI loss more than 2.0 points. These findings are broadly consistent with the primary analysis, which found relationships between several individual disease conditions and BMI change more than 2.0 points.

The second set of robustness checks examines functional form. Overall, the results from linear regression are weaker than those from the primary logistic models. Figure 2B.1 shows results from a linear regression of BMI change (a continuous variable) on self-reported diseases. Diabetes onset in England is the sole condition associated at 5% with BMI change as a continuous variable (associated with a change of -0.44 BMI points [95% CI: -0.79 – -0.09]).

Figure 2B.1. Self-Reported Disease Diagnoses and BMI Change (Continuous).



† See Chapter 2 or Appendix 2A for differences in disease definitions between CRELES and ELSA.

However, in England, the changes associated with baseline diabetes and with cancer onset are close to significant at 5% (respectively, -0.23 [95% CI: -0.47 – 0.00]) and -0.40 [95% CI: -0.81 – 0.01]). Linear regression of weight change in kilograms on self-reported diseases also reveals a borderline significant association in ELSA between arthritis onset and weight change, suggesting that disease may lead to mild weight *gain* (a change of +0.47 kg [95% CI: 0.01 – 0.92]), although there is no such apparent association with BMI change (weight change adjusted for height squared).

Results from multinomial logistic models look similar to those from the primary analysis. The multinomial results do reveal some additional associations in CRELES between self-rated general health and BMI changes (gains or losses) of more than 1.0 point, relative to roughly stable weight (defined as a change > -1.0 point and < 1.0 point). Each two-point worsening in self-rated general health (continuous) is associated with a 37% increase in the probability of BMI loss of at least 1.0 point relative to the probability of roughly stable weight (*OR* = 1.37 [95% CI: 1.00 – 1.87]). When self-rated general health is treated as a dichotomous variable, fair or poor health is similarly associated with a 37% increase in the probability of BMI loss of at least 1.0 point relative to roughly stable weight (*OR* = 1.37 [95% CI: 1.02 – 1.85]). However, there is no apparent relationship in the multinomial models between self-rated general health and BMI loss more than 2.0 points.

APPENDIX 2C: POWER TO DETECT ASSOCIATIONS BETWEEN DISEASES AND WEIGHT LOSS

We calculate power using standard-error estimates from models described in Chapter 2: i.e., with all diseases included simultaneously, and controls for age, sex, smoking, and initial BMI.

Costa Rica — Power to detect odds ratio = 2

	BMI change < -1.0 point	*BMI change < -2.0 points*
Cancer at baseline	19.8%	9.2%
Cancer onset	13.6%	12.7%
Diabetes at baseline	96.6%	54.8%
Diabetes onset	46.8%	28.4%
Heart disease at baseline	41.2%	17.4%
Heart disease onset	22.1%	15.9%
Stroke at baseline	26.4%	14.2%
Stroke onset	8.7%	–
Arthritis at baseline	90.7%	59.1%
Arthritis onset	31.6%	14.2%
Lung disease at baseline	93.1%	56.0%
Lung disease onset	19.2%	9.7%

England — Power to detect odds ratio = 2

	BMI change < -1.0 point	*BMI change < -2.0 points*
Cancer at baseline	78.2%	40.1%
Cancer onset	69.2%	62.9%
Diabetes at baseline	99.7%	96.4%
Diabetes onset	91.2%	74.5%
Heart disease at baseline	98.2%	78.5%
Heart disease onset	91.3%	71.2%
Stroke at baseline	39.8%	15.2%
Stroke onset	33.6%	28.8%
Arthritis at baseline	> 99.9%	> 99.9%
Arthritis onset	> 99.9%	96.6%
Lung disease at baseline	82.9%	56.0%
Lung disease onset	84.4%	64.4%

APPENDIX 3A: SELECTION OF PARAMETER VALUES

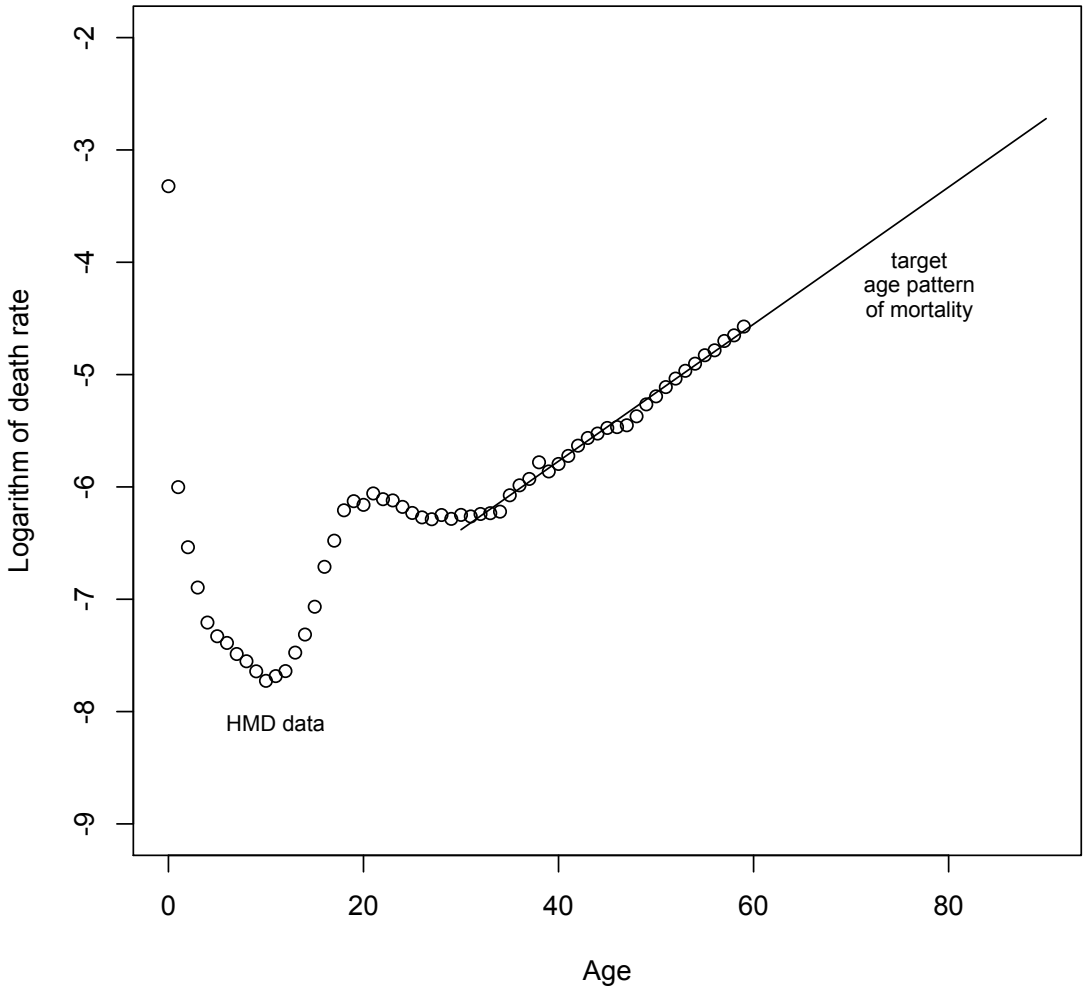
Simulation populations are designed to resemble, very roughly, the population of U.S. males born in 1950 (i.e., aged 30 in 1980, aged 40 in 1990, aged 50 in 2000, etc.). For each simulation, parameter values are selected so that the simulated population will match designated (or "target") relationships between age, weight, and mortality risk.

Designated relationships for the simulated populations to reproduce

A "target," or stipulated, age pattern of mortality for the simulated populations is based on death rates reported in the Human Mortality Database (HMD) (1). The HMD provides cohort death rates for U.S. males born in 1950 by single year of age, up to age 59. The target relationship for the simulated populations to reproduce is then modeled from these data, assuming a linear increase in logged death rates with age, from age 30 to 90: i.e., a Gompertz pattern of adult mortality (2). Figure 3A.1 shows HMD observations (from age 0 to 59) alongside target mortality for the simulations to reproduce (from age 30 to 90).

A target population distribution across weight categories at each age is based loosely on BMI observations in the U.S. National Health Interview Survey (3), calculated from self-reported height and weight, and on measured BMI values in the English Longitudinal Study of Ageing (4).

Figure 3A.1. Observed Death Rates by Age, with Target Age Pattern of Mortality for the Simulation Populations to Reproduce



Target slope: 0.061
Target value at age 30: 0.00169

HMD: Human Mortality Database

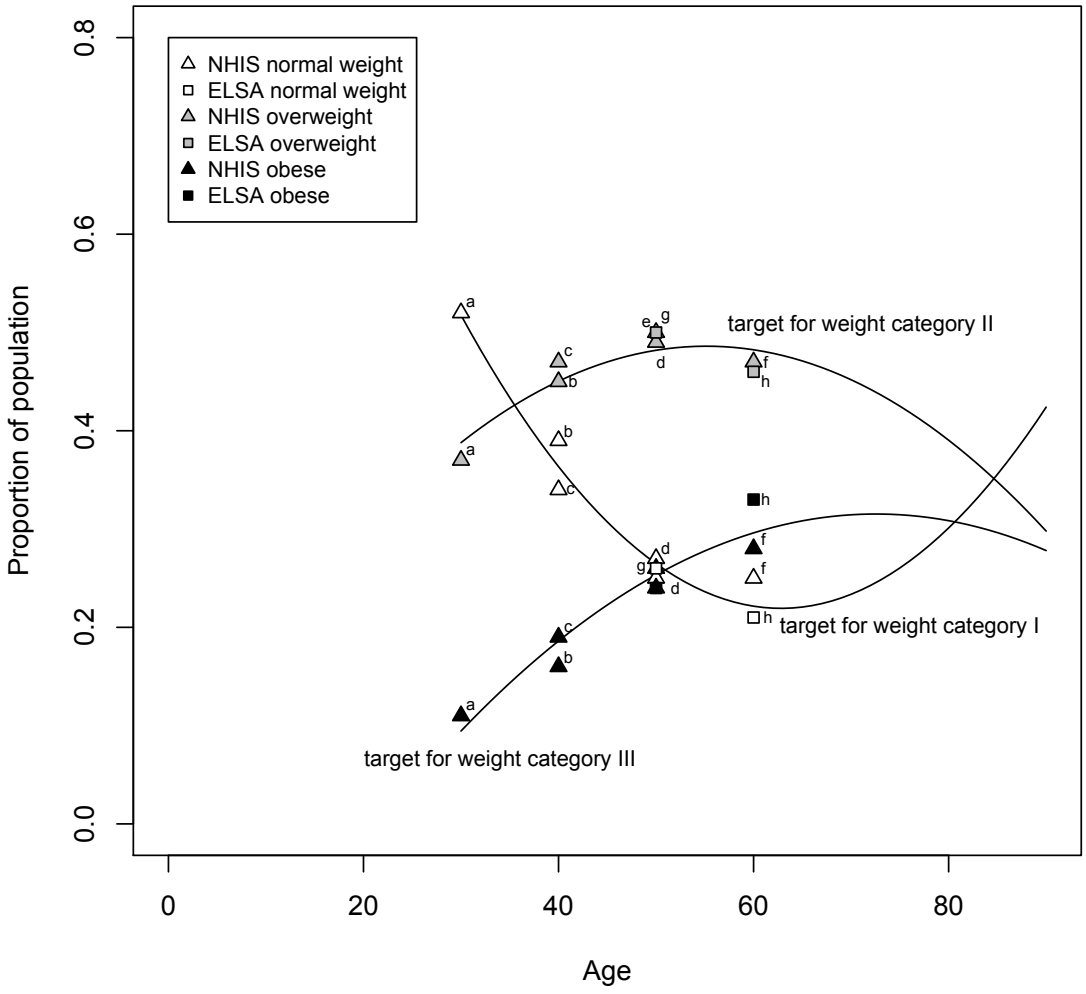
Figure 3A.2 shows observed proportions of the non-underweight male population that are normal weight, overweight, and obese in these two data sources; the target for the simulation populations is superimposed. This designated relationship for the simulations is essentially arbitrary above age 60, as data are not yet available beyond that age for the cohort born in 1950, and there is no validated method to project future prevalence of obesity or overweight.

Target relationships between weight category and mortality risk are based on a recent meta-analysis of nearly 100 studies worldwide of BMI and mortality (5). That paper finds (for both sexes combined) a mortality hazard ratio of 0.94 associated with overweight status relative to normal weight, and a hazard ratio of 1.18 associated with obesity relative to normal weight. Because there is limited information about the shape of the BMI–mortality curve before middle age or at the oldest ages, this study stipulates a target relationship between weight category and mortality only for the age range 50-79. For simplicity, target hazard ratios are assumed to be constant by age within this range. They are shown in Figure 3A.3: a hazard ratio of 0.94 associated with weight category II and of 1.18 with weight category III, as in the meta-analysis.

Parameter search

The parameter-value selection process has two steps. First, a wide range of possible (arbitrary) values is assigned to each parameter in each model. These possible values are listed in Table 3A.1. Then, for each model, every possible combination of different parameter values is

Figure 3A.2. Observed BMI Category by Age, with Target Weight Category Distributions for the Simulation Populations

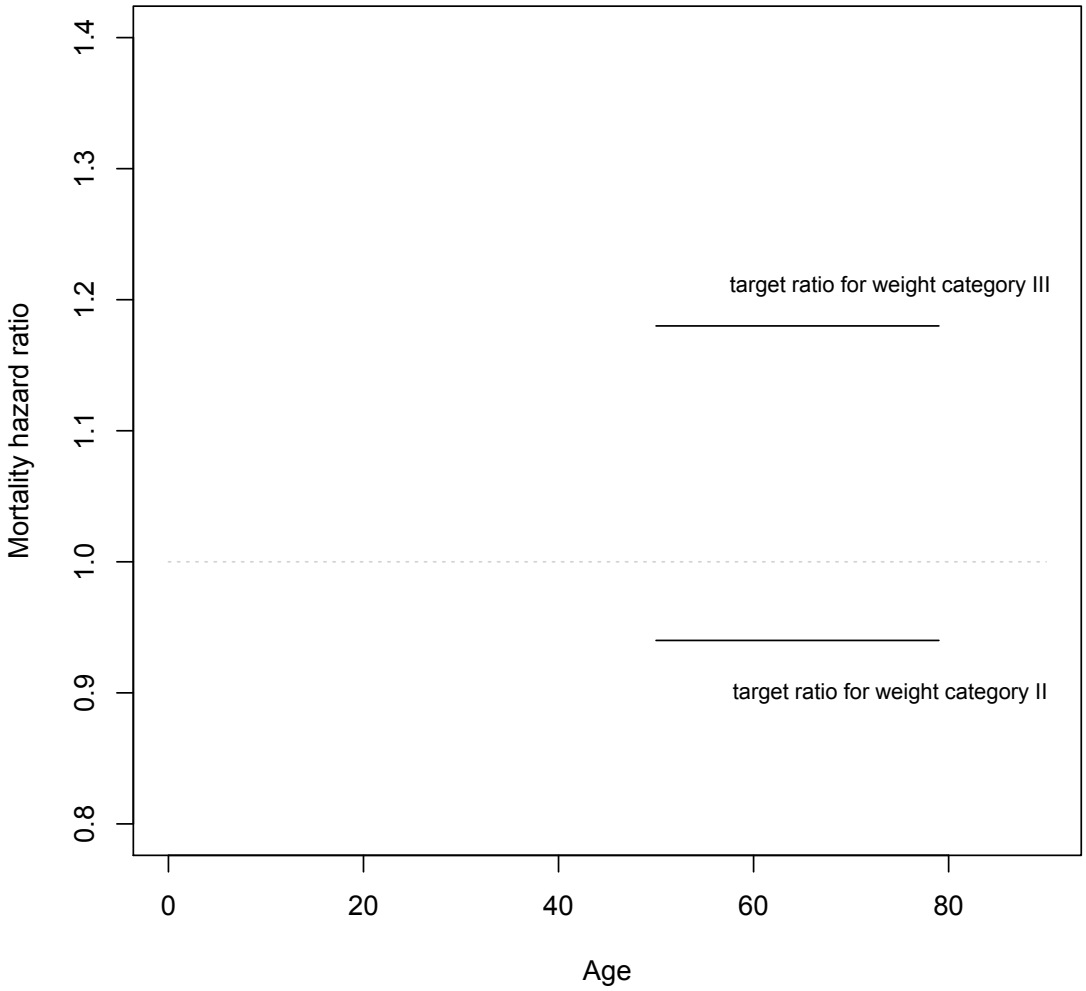


Target proportion in category I: $1.31 - 3.48*10^{-2}*age + 2.77*10^{-4}*age^2$
Target proportion in category II: $1.43*10^{-2} + 1.71*10^{-2}*age - 1.55*10^{-4}*age^2$
Target proportion in category III: $-3.27*10^{-1} + 1.77*10^{-2}*age + 1.22*10^{-4}*age^2$

NHIS: National Health Interview Survey
ELSA: English Longitudinal Study of Ageing

a: Men aged 27-33 in 1986.
b: Men aged 37-43 in 1990.
c: Men aged 37-43 in 1996.
d: Men aged 47-53 in 2000.
e: Men aged 47-53 in 2004.
f: Men aged 57-63 in 2004.
g: Men aged 47-53 in 1998, 1999, or 2001.
h: Men aged 57-63 in 2008 or 2009.

Figure 3A.3. Target Relationship Between Weight Category and Mortality, by Age, for the Simulation Populations



Target ratio for category II: 0.94 (constant across age range 50-79)
Target ratio for category III: 1.18 (constant across age range 50-79)

Table 3A.1. Values Considered in the Parameter-Value Search.

Model 1: Nutritional reserve

μ	z	β_2	β_3	γ	γ_a
0.00169	0.061	0.94	1.18	0.005	0
				0.01	-0.0005
				0.02	-0.001
				0.03	-0.005
				0.04	-0.01
				0.05	-0.05
				0.06	-0.1
				0.07	
				0.08	
				0.09	

Model 2: Conventional reverse causation

μ	z	β_2	γ	γ_a	θ	ρ	z_s	β_{s2}	γ_s
0.00169	0.01	1.06	0.005	0	2	0.0005	0.03	1.06	0.05
	0.02	1.12	0.01	-0.0005	4	0.001	0.04	1.12	0.1
	0.03	1.25	0.02	-0.001	8	0.002	0.05	1.25	0.2
	0.04	1.5	0.03	-0.005	16	0.004	0.06	1.5	0.4
	0.05	2	0.04	-0.01	32	0.008	0.07	2	0.6
	0.06	2.5	0.05	-0.05	64	0.016	0.08	2.5	1
		3	0.06	-0.1	128	0.032	0.09	3	2
			0.07		256	0.064	0.10		4
			0.08		512		0.11		
			0.09						

Model 3: Modified reverse causation

μ	γ	θ	ρ	z_s	β_{s2}	γ_s
0.00169	0.005	2	0.0005	0.03	1.06	0.05
	0.01	4	0.001	0.04	1.12	0.1
	0.02	8	0.002	0.05	1.25	0.2
	0.03	16	0.004	0.06	1.5	0.4
	0.04	32	0.008	0.07	2	0.6
	0.05	64	0.016	0.08	2.5	1
	0.06	128	0.032	0.09	3	2
	0.07	256	0.064	0.10		4
	0.08	512		0.11		
	0.09					

assessed for its "goodness of fit" to the target relationships (see "Testing 'goodness of fit,'" below), and the best-fit combination is selected. In the model for Hypothesis #1, for example (the nutritional reserve explanation), there are 70 different possible combinations of parameter values to be assessed: one for each of 10 possible values of γ by 7 possible values of γ_a. This model has relatively few parameter-value combinations to choose among because the model assumes that observed relationships are true relationships – and thus values for μ, β_2, β_3, and z may be selected directly from the target relationships. In contrast, for the conventional reverse causation model (Hypothesis #2), there are 106,686,720 different combinations to be assessed. This is because that model has 10 parameters, rather than 6, and because almost none of the parameter values may be observed directly in data (since observed relationships can reflect multiple simultaneous causal processes). The model for Hypothesis #3 (modified reverse causation) lies in between, with 362,880 parameter-value combinations to assess.

Testing "goodness of fit"

For each model, each combination of parameter values is first assessed for its ability to explain a non-monotonic relationship between BMI and mortality. Specifically, a combination of values is rejected if, in the simulated population it produces, the mortality hazard ratio for weight category II relative to weight category I, averaged over ages 50-79, is > 1, or if the hazard ratio for weight category III relative to category I is < 1.1.

Next, for each model, all combinations satisfying the above constraint are ranked on how closely they match the three target relationships shown in Figures 3A.1, 3A.2, and 3A.3. One combination of values is considered to fit better than another if it yields a lower value of the distance measure, D, defined as follows (constructed in a form analogous to Pearson's χ^2 test statistic):

$$
\begin{aligned}
D \quad = \quad & \sum_{i=31}^{90} \frac{(OM_x - EM_x)^2}{EM_x} \\
& + 0.5 \left(\sum_{i=31}^{90} \frac{\left(OP_{II,x} - EP_{II,x}\right)^2}{EP_{II,x}} + \sum_{i=31}^{90} \frac{\left(OP_{III,x} - EP_{III,x}\right)^2}{EP_{III,x}} \right) \\
& + \left(\sum_{i=50}^{79} \frac{\left(OM_{II,x} - EM_{II,x}\right)^2}{EM_{II,x}} + \sum_{i=50}^{79} \frac{\left(OM_{III,x} - EM_{III,x}\right)^2}{EM_{III,x}} \right)
\end{aligned}
$$

where "OM_x" is the observed number of deaths at year x in a population simulated using the combination of parameter values, and "EM_x" is the expected number of deaths, or the number of deaths that would occur in that population in that year if the simulated population still alive were subject to the target death rates; "$OP_{II,\,x}$" (or "$OP_{III,\,x}$") is the observed percentage (between 0 and 100) of the simulated population in weight category II (or III) at year x, and "$EP_{II,\,x}$" (or "$EP_{III,\,x}$") is 100 times the target proportion in that weight category at year x; "$OM_{II,x}$" (or "$OM_{III,\,x}$") is the observed number of deaths at year x in the simulated population in weight category II (or III), and "$EM_{II,\,x}$" (or "$EM_{III,\,x}$") is the expected number, or the number of

deaths that would occur at that age in that weight category if death rates were equal to the target hazard ratio, multiplied by the simulation's death rate in weight category I at year x.

In other words, the distance measure, D, may be used to assess how well a simulated population fits the target patterns of mortality by age (shown in Figure 3A.1), of weight-category distribution by age (Figure 3A.2), and of mortality hazard ratios by weight category (Figure 3A.3). These three criteria are all incorporated into the one distance measure. The component of this statistic for weight-category distribution by age (i.e., the second line, comparing observed to expected percentages of the population in each weight category) is multiplied by 0.5 so that, like the other two criteria, this one contributes an equivalent of 60 pieces of information (despite 60 years of observation from two different weight categories). The distance measure assesses death rates and weight-category distributions starting only at age 31, since, as mentioned in Chapter 3, simulated conditions at age 30 (i.e., the initial conditions for each simulation) are already matched to the target conditions at that age.

APPENDIX 3A REFERENCES

[1] Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). (http://www.mortality.org or http://www.humanmortality.de). (Data downloaded on 15 May 2013.)

[2] Finch CE. *Longevity, Senescence, and the Genome.* Chicago: The University of Chicago Press; 1990.

[3] Minnesota Population Center and State Health Access Data Assistance Center. Integrated Health Interview Series: Version 5.0. Minneapolis: University of Minnesota; 2012. (http://www.ihis.us).

[4] Marmot M, et al. English Longitudinal Study of Ageing: Wave 0 (1998, 1999 and 2001) and Waves 1-4 (2002-2009) [computer file]. 15th Edition. Colchester, Essex: UK Data Archive [distributor]; April 2011. SN: 5050. (http://dx.doi.org/10.5255/UKDA-SN-5050-1).

[5] Flegal KM, Kit BK, Orpana H, et al. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *J Am Med Assoc*. 2013; 309(1):71-82.