PREDICTING ADULT HEALTH AND MORTALITY FROM ADOLESCENT FACIAL CHARACTERISTICS IN YEARBOOK PHOTOGRAPHS*

ERIC N. REITHER, ROBERT M. HAUSER, AND KAREN C. SWALLEN

Several important longitudinal studies in the social sciences have omitted biomarkers that are routinely recorded today, including height and weight. To account for this shortcoming in the Wisconsin Longitudinal Study (WLS), an 11-point scale was developed to code high school senior class yearbook photographs of WLS participants for relative body mass (RBM). Our analyses show that although imperfect, the RBM scale is reliable ($\alpha = .91$) and meets several criteria of validity as a measure of body mass. Measured at ages 17–18, the standardized relative body mass index (SRBMI) was moderately correlated ($r = .31$) with body mass index (BMI) at ages 53–54 and with maximum BMI reported between ages 16 and 30 ($r = .48$). Overweight adolescents ($\geq$ 90th percentile of SRBMI) were about three times more likely than healthy-weight adolescents (10th–80th percentile of SRBMI) to be obese in adulthood and, as a likely consequence, significantly more likely to report health problems such as chest pain and diabetes. Overweight adolescents also suffered a twofold risk of premature death from all nonaccidental causes as well as a fourfold risk of heart disease mortality. The RBM scale has removed a serious obstacle to obesity research and lifelong analyses of health in the WLS. We suggest that other longitudinal studies may also be able to obtain photos of participants at younger ages and thus gain a prospectively useful substitute for direct measures of body mass.

Longitudinal research in the social sciences has enriched our understanding of the antecedents and consequences of obesity (Crossman, Sullivan, and Benin 2006; Goodman and Whitaker 2002; Himes 2000; Mannino et al. 2006; Novak, Ahlgren, and Hammarstrom 2006; Sundquist and Johansson 1998). Unfortunately, for much of the twentieth century, social scientists did not routinely record the height and weight of study participants (but see, e.g., Ayres 1909). As a consequence, a number of important longitudinal studies that might otherwise provide excellent resources for obesity research—and, more generally, for research on the later consequences of early life characteristics and conditions—are limited by the lack of body mass indicators early in life.

The Wisconsin Longitudinal Study (WLS) fits the previous description. The WLS is a cohort study of over 10,000 individuals that has been used to study a wide range of health issues, including cognitive function (Krahn et al. 2003; Seltzer et al. 2005), mental health (Carr 1997; Johnson and Breslau 2006), reproductive health (Marks and Shinberg 1997; Shinberg 1998), hormone therapy (Marks and Shinberg 1998), immune function (Rosenkranz et al. 2003), and healthcare decisions (Flynn, Smith, and Vanness 2006). Armed with data on height and weight from the 1993 and 2004 waves, researchers have begun to conduct obesity research with WLS data (e.g., Reither 2005). However, the WLS did not

collect data on height or weight either at baseline (1957) or in a subsequent wave in 1975, thus limiting its potential contributions to research on adult health and other consequences of early life characteristics and conditions.

The WLS is not alone in this respect. For instance, the National Longitudinal Study of the Class of 1972 (NLS-72) surveyed 22,652 high school seniors about their educational attainment and career aspirations. Subsequent waves in 1973, 1974, 1976, 1979, and 1986 created an exceptionally rich archive for scholars to investigate influences such as family background, high school curriculum, and racial/ethnic identification on educational and occupational outcomes (U.S. Department of Education 1992). Unfortunately, the NLS-72 did not measure height or weight, making it impossible to determine the impact of obesity on the educational or career trajectories of participants in the study. The capacity of longitudinal studies like the NLS-72 to contribute to obesity research would expand tremendously if there were some way to add new measures of body mass, particularly at baseline. Similar observations apply to other longitudinal studies, including the National Education Longitudinal Study of 1988, the National Longitudinal Survey of Youth 1979 (which initially collected data on height and weight in 1981), and the National Longitudinal Survey of Young Women 1968 (which did not collect data on height or weight until 1991). Is there any recourse for such investigations other than potentially unreliable retrospective questioning that excludes members of the sample who, for one reason or another, no longer participate?

This study attempts to fill this void by developing a new, substitute measure of body mass that codifies the facial features of WLS participants when they were high school seniors in 1957. As we discuss in some detail, extant research has shown that characteristics of the face and neck are related to body mass and central adiposity (fattiness). Findings from this body of literature suggest that the methodical examination of photographs may provide a novel and valid way to assess the body mass of human subjects.

FACIAL CHARACTERISTICS, ADIPOSITY, AND BODY MASS

A number of clinical studies have investigated the relationship between facial characteristics and body habitus in samples of human subjects. These studies have demonstrated that deposits of adipose tissue in the cheeks and neck, neck circumference, and craniofacial morphology are all related to body mass and central adiposity. For instance, Levine, Ray, and Jensen (1998) used computer tomography to measure cheek, visceral abdominal, and abdominal subcutaneous fat in 25 patients who were being treated at the Mayo Clinic for various conditions. Despite the small sample of subjects, Levine et al. (1998) demonstrated that the quantity of cheek fat was strongly related to deposits of visceral abdominal fat ($r = .54$, $p < .01$).

In another study, Laakso, Matilainen, and Keinänen-Kiukaanniemi (2002) evaluated associations between neck circumference and obesity. Results from this study of 541 Finns clearly demonstrated that neck circumference was associated with various measures of obesity. Among men, neck circumference was significantly correlated with waist-to-hip ratio ($r = .41$, $p < .01$), waist circumference ($r = .65$, $p < .01$), and BMI ($r = .69$, $p < .01$). These correlations were nearly identical among women.

Like neck circumference, deposits of subcutaneous adipose tissue (SAT) in the neck are related to central- and upper-body adiposity. To study the relationships between SAT deposits in 15 different body locations, Möller et al. (2000) utilized the LIPOMETER to assess SAT-topography in 590 healthy adult subjects. Factor analyses of these data indicated that SAT deposits in the neck, upper back, front chest, lateral chest, upper abdomen, lower abdomen, and hip were highly interrelated.

This evidence suggests that neck circumference and deposits of adipose tissue in the face and neck may provide useful indicators of general adiposity. In addition, studies of craniofacial morphology have demonstrated that bone structures in the face are associated

with body mass (Yu et al. 2003). For instance, in an investigation of facial measurements in obese and nonobese adolescents, Sadeghianrizi et al. (2005) found that the facial skeletal structures of obese adolescents tend to be relatively large. Notably, obese adolescents exhibited jaws (mandibles) that were, on average, 8–10 millimeters longer than those of control subjects.

HEALTH OUTCOMES ASSOCIATED WITH ADIPOSITY OF THE FACE AND NECK

The distribution of adipose tissue affects the probability of disease incidence. For instance, central adiposity is an important determinant of hypertension, insulin resistance, and heart disease (Rexrode et al. 1998). Although not widely recognized, research has also shown that characteristics of the face and neck are associated with health complications such as hypertension (Laakso et al. 2002), Type 2 diabetes (Tafeit et al. 2000), and sleep apnea (Mortimore et al. 1998). To illustrate, Laakso et al. (2002) found that neck circumference was an independent predictor of hypertension among northern Finns. Regardless of gender, Finns in the highest quintile of neck circumference were approximately three times more likely to have hypertension than Finns in the lowest quintile, even after controlling for BMI. This finding was corroborated by Tafeit et al. (2000), who found that neck adiposity was better at discriminating Type 2 diabetics from nondiabetics than adiposity in 14 other body locations (e.g., upper abdomen), BMI, or body fat percentage. Significantly, this suggests that facial mass may carry information about future health conditions above and beyond that in a baseline measure of BMI.

A PROMISING NEW APPROACH

The literature just reviewed suggests that the methodical examination of photographs of the face and neck may offer a novel and valid way to measure the relative body mass of human subjects. Furthermore, because facial adiposity is associated with chronic conditions such as hypertension and sleep apnea, the development of new methodologies that isolate facial characteristics may lead to other scientific advancements, such as an improved ability to predict the development of certain health conditions. However, while this review of the literature has provided scientific justification for expecting a solid relationship between facial characteristics and body mass, it has not addressed the ability of humans to distinguish lean from heavy persons solely on the basis of facial photographs—and with good reason. A review of the literature produced only one study (Rudin 1996) that required participants to estimate weight from photographs. Unfortunately, because this study used crude weight categories and provided no formal assessment of reliability or validity, it did not provide substantial evidence on the capacity of humans to estimate weight from photographs.

Furnham and Radley (1989), however, found that a group of young participants (ages 16–21) were able to rank order drawings depicting naked persons of varying adiposity along a continuum from very thin to obese. Also, Kato and Higashiyama (1998) found that undergraduate participants were generally able to provide accurate estimates of height based on full-length photographs. Obviously, there are many important differences between the tasks required in these studies and determining the relative body mass of subjects via photographs that depict the faces and necks of clothed individuals. Nonetheless, these results are encouraging because they imply that humans have a refined capacity for differentiation of physical traits based on visual stimuli.

On balance, extant research suggests real promise in developing a new methodology to estimate relative body mass from yearbook photographs. This study represents an initial attempt to develop and use such a methodology. A random sample of 3,027 photographs was selected from the yearbooks of WLS participants and assigned values for relative body mass by a team of six coders. Results of this study demonstrate that the new measure of

relative body mass is reliable and meets several criteria of predictive validity. This is good news for the WLS and for other longitudinal investigations that currently lack baseline measures of BMI but may be able to access photographs of research participants.

METHODS

Study Population and Photographs

The Wisconsin Longitudinal Study (WLS) is a random sample of 10,317 persons who graduated from a public, private, or parochial high school in Wisconsin in 1957 (Sewell et al. 2004). From 1957 to the latest wave of data collection in 2004, the WLS has collected information on educational and occupational histories, socioeconomic status, military service, marital status, family characteristics, social participation, psychological well-being, health behaviors, health outcomes, and many contextual factors (Hauser 2005; Sewell et al. 2004). Data on the height and weight of WLS participants were gathered in the 1993 and 2004 waves, making it possible to assess BMI change during part of the U.S. obesity epidemic, which began around 1980 (Flegal et al. 1998). Although the WLS is not nationally representative, its participants resemble over two-thirds of Americans, now entering retirement age, in terms of educational attainment and ethnic background (Hauser 2005).

Yearbooks from 1957 were not obtained early in the study. Rather, they were collected only beginning in 2000—more than 40 years after high school graduation—from various sources in the towns where selected schools were located (Meland 2002). The initial sources included high schools, public libraries, and personal contacts made through the schools. We began with the larger schools in urban areas and gradually added smaller schools in rural areas and small towns. These yearbooks yielded 7,500 photographs covering three-quarters of the WLS graduates. However, beginning early in 2007, we joined distribution of a respondent report with tailored requests to borrow, scan, and return the yearbooks from the remaining schools, and this effort has been very successful. We think this demonstrates the feasibility of obtaining early-life photographs in a large population-based sample long after they were taken.

Through a cluster sampling design that selected schools based on probabilities proportional to size (PPS), a subsample of 93 schools representing 3,130 WLS participants was randomly chosen for inclusion in this study. Photographs of WLS participants, all high school seniors in 1957, were subsequently extracted through computerized scanning technologies. Because of a small number (103) of missing photographs, 3,027 photographs were available for this study.

Scale Development

Separate scales for boys and girls were developed to measure the relative body mass (RBM) of WLS participants in 1957. The term "relative body mass" was chosen because the scales were hypothesized to result in proxies for body mass index (BMI = weight(kg) / height(m)^2). This hypothesis arises from the assumption that yearbook photographs do not provide clear visual evidence about the height of participants. That is, as coders rate photographs for relative mass, they should do so independently of height. In this way, the RBM scale is analogous to BMI, which is also a measure of weight that controls for the height of subjects.

We designed the RBM scale to display 11 points, two verbal anchors, and five photographs. Figures 1 and 2 provide examples of RBM scales; the individuals shown are not in the WLS sample, nor were these specific pictures used in the coding operation. Eleven scale points permit us to account for the wide array of body types among adolescents in U.S. society. Also, previous research suggests that using 11 scale points affords an optimal level of reliability (Alwin 1992, 1997).

Figure 1.   **Model of Relative Body Mass (RBM) Scale Used to Code Yearbook Photographs of Male Participants in the Wisconsin Longitudinal Study (WLS)**
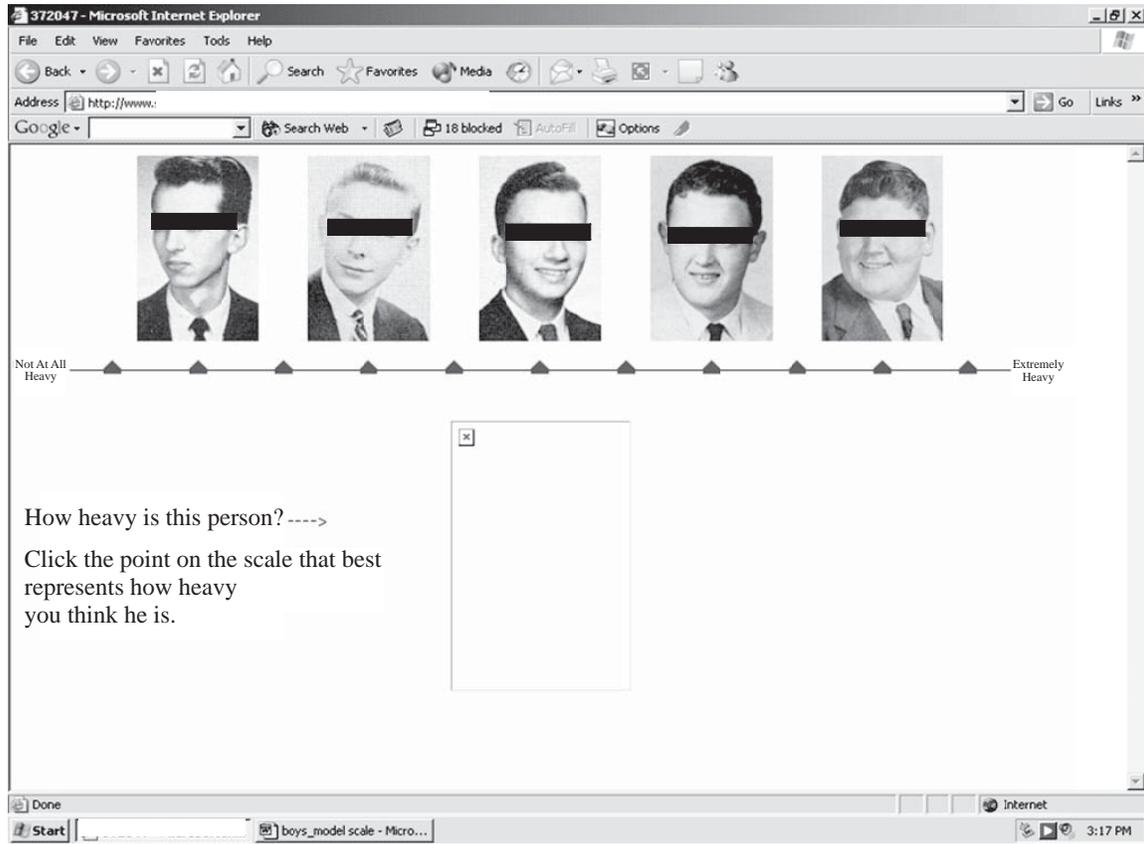

Measures

For each of the 3,027 yearbook photographs, each of the six coders recorded a RBM scale score ranging from 1 to 11. RBM scores were combined to form the standardized relative body mass index (SRBMI). SRBMI was calculated separately for male and female photos by (1) generating coder-specific z scores, (2) summing z scores across coders, and (3) dividing the sum of z scores by the number of coders in the study. That is, for an individual participant,

$$\text{SRBMI}_{jk} = \frac{\sum_{i=1}^{n}\left[(x_{ijk} - \overline{x}_{ik})/s_{ik}\right]}{n},$$

where i is an individual coder, n is the number of coders in the study, j is one of the 3,027 WLS participants, k is the participant's gender, and x_{ijk} is the series of RBM scale scores for coder i and participant j of gender k, with mean $\overline{x}_{ik}$ and standard deviation s_{ik}. We standardized the RBM scores in this way to eliminate differences among coders in the location (mean) and variance of the scores that they assigned.

In some applications, SRBMI was treated as both a continuous and a categorical variable. This permitted the evaluation of possible nonlinear associations between SRBMI and

Figure 2. Model of Relative Body Mass (RBM) Scale Used to Code Yearbook Photographs of Female Participants in the Wisconsin Longitudinal Study (WLS)
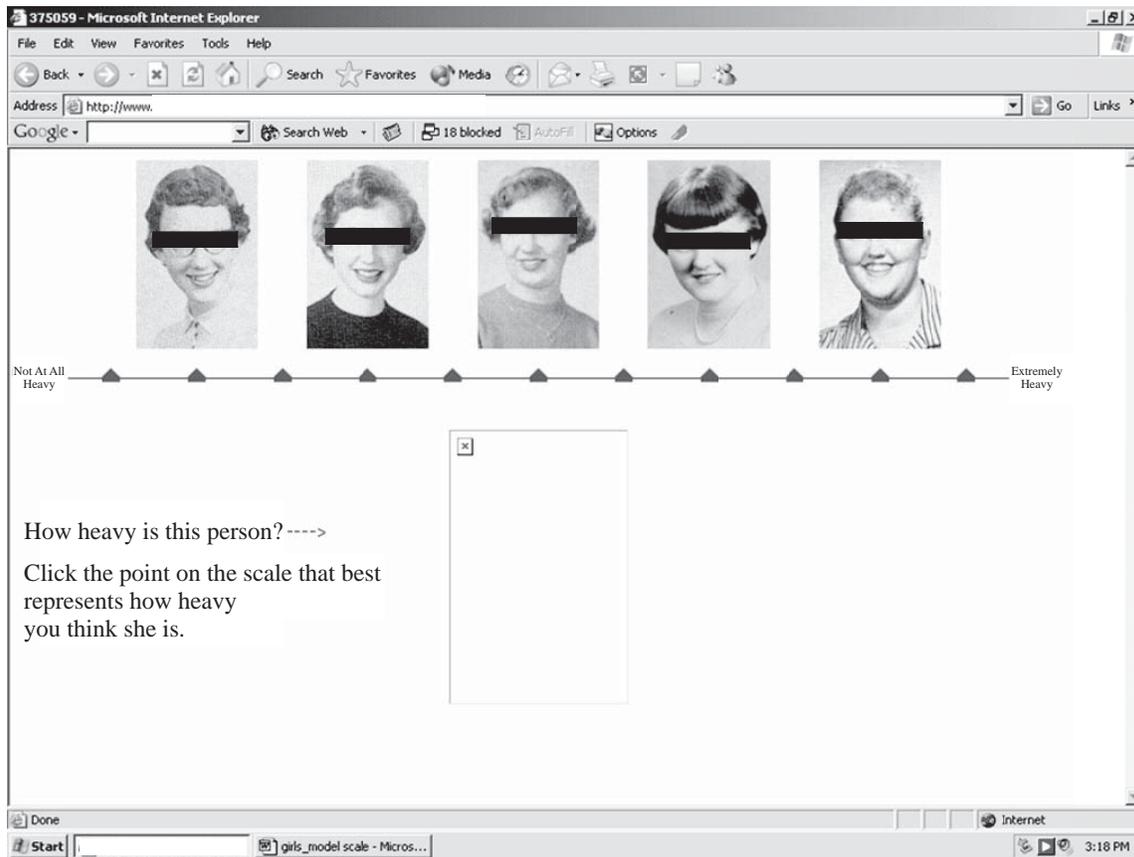
health outcomes. It also permitted SRBMI to be divided into an approximation of standard BMI classifications for adolescents: underweight, healthy weight, at risk for overweight, and overweight. Previous research has used BMI percentiles from Centers for Disease Control and Prevention (CDC) growth charts to define underweight at or below the 5th percentile, healthy weight between the 5th and 85th percentiles, at risk for overweight between the 85th and 95th percentiles, and overweight at or above the 95th percentile (Ogden et al. 2002). To provide sufficient statistical power for each subgroup, we altered these percentile ranges slightly for categories derived from SRBMI; underweight was defined at or below the 10th percentile of SRBMI; healthy weight, between the 10th and 80th percentiles; at risk for overweight, between the 80th and 90th percentiles; and overweight, at or above the 90th percentile.

Measures from the 1993 WLS used to assess the predictive validity of the RBM scale included self-reported height and weight, health symptoms, and chronic conditions. The WLS measured height and weight in inches and pounds, which were then converted into a measure of BMI (U.S. Department of Health and Human Services 2004). Indicators of obesity (BMI $\geq$ 30) and class II obesity (BMI $\geq$ 35) were created to estimate the association between RBM in 1957 and obesity in 1993.

Additionally, we included measures from the 2004 WLS that asked respondents about their maximum weight, and the age at which they reached it. By combining maximum

weight with the 1993 estimate of height (which we assumed not to change appreciably), we created a variable that estimated maximum BMI. We limited our analyses of this new estimate of BMI to respondents who reported a maximum weight between the ages of 16 and 30, thus providing an estimate of BMI that was more proximate to high school graduation than the 1993 measure of BMI.

Participants in the 1993 WLS responded to a series of questions probing whether they had experienced certain health-related symptoms in the past six months. Four symptoms (muscle aches, back pain or strain, chest pain, and shortness of breath) were included in this investigation because of their known associations with obesity (Lean, Han, and Seidell 1999; Stunkard 1996). WLS participants also responded to questions about chronic conditions that had been diagnosed by a medical professional. Four conditions (arthritis, high blood pressure, diabetes, and heart trouble) were retained because of known associations with obesity (Manson, Skerrett, and Willett 2002; Pi-Sunyer 2002; Stunkard 1996). Health symptoms and chronic conditions were coded as indicator variables (i.e., 1 = symptom/ condition present; 0 = symptom/condition not present).

Measures of all-cause and cause-specific mortality were used to provide another assessment of the criterion-related validity of the RBM scale. Determination of death and cause of death were made by searching the National Death Index (NDI) for WLS participants (National Center for Health Statistics 1999). WLS staff searched the NDI database for causes of death in October of 2001. Of the 3,027 WLS participants in this study, 159 died between 1979 (the first year covered by the NDI) and 1998. Data from the NDI were matched with cause-of-death codes from the International Classification of Diseases (ICD-9) to construct measures of mortality resulting from (1) all causes, (2) all nonaccidental causes, (3) all major diseases of the heart, and (4) all malignant cancers (World Health Organization 1977).

Coding the Photos

Three males and three females were recruited to code photographs of WLS participants. Coders were non-Hispanic, white, predoctoral students in sociology and varied in age from 26 to 33 years. Prior to coding photographs, coders signed confidentiality agreements and participated in a brief seminar, where they received orientation training and a basic set of written coding instructions that were retained for reference. The written instructions provided coders with technical information (e.g., how to log in to the computer program) and prompted them to examine photographs carefully for clues about RBM, such as neck diameter, fatty deposits in the cheeks and neck, and noticeable bone structures (such as protruding cheek bones). Coding took place in the WLS research office at the University of Wisconsin–Madison over a three-week period in January 2005.

Coders were instructed in a standard protocol: to compare the photograph of the WLS participant to the scale photographs located along the top of the computer monitor (see Figure 2). Note that the box at the bottom of the monitor is where the photograph of the WLS participant appeared. Coders were given 10 seconds to make this initial comparison and, using a computer mouse, click the point that corresponded to their impression of where the photographed person fit along the weight-scale continuum. The chosen scale point changed color from black to red, and coders were presented with the question, "Does the red symbol indicate your choice?" Before answering, coders were instructed to study certain characteristics of the WLS participant (e.g., fatty deposits in the cheeks) in more detail. If coders subsequently answered "Yes," they were presented with the next photograph. Coders answering "No" were provided with an opportunity to change their selection before moving on to the next photograph.

To assess intrarater reliability, each coder reevaluated 100 male and 100 female photographs that were randomly chosen from the sample of 3,027. Reevaluation of these photographs was not conducted until a minimum of one week after project completion to reduce the likelihood of recall bias.

Statistical Analyses

SAS 9.1 and Stata 9.0 were used to manage and analyze data (SAS Institute Inc. 2003; StataCorp LP 2005). Coefficient α was used to evaluate interrater reliability (Cronbach 1951). Under the assumption of parallel tests, intrarater reliability was estimated as the correlation coefficient between RBM scores assigned at two different times.

Discriminant and criterion-related validity of the RBM scale was assessed by examining correlations between SRBMI, height, and BMI. Additionally, the validity of the RBM scale was evaluated by conducting logistic regression analyses to estimate the risk of obesity, health symptoms, and chronic conditions in 1993 as a function of RBM classification in 1957. In these analyses, robust standard errors were generated by accounting for clustering within schools from which WLS participants were sampled. Logistic regression was also used to evaluate whether RBM in 1957 predicted all-cause and cause-specific mortality in subsequent years. Interaction effects between gender and RBM were explored in logistic regression analyses but are not reported here because in no instance did interaction terms significantly improve model fit.

RESULTS

Analyses confirmed that the RBM scale is a reliable measure. Intrarater reliability was acceptable, as evidenced by correlations of .66, .71, .80, .82, .85, and .88 for the sample of 200 photos that were double-coded (Nunnally 1978). Also, the number of coders and the agreement between coders were sufficiently high. For the entire sample of 3,027 photos, Cronbach's α was .91. Reliability was not affected by the gender of the WLS participant, as shown by equal alpha values for female photos (α = .91) and male photos (α = .91).

Analyses also confirmed that the RBM scale is correlated with later measures of body mass and is not correlated with height. The correlation between SRBMI and height was .00 ($p > .05$) in the full sample. This finding varied little by gender of the WLS participant, demonstrating that height did not confound the RBM scale. The correlation between adolescent SRBMI and BMI at ages 53–54 was positive and moderate in strength ($r = .31$, $p < .05$), despite the passage of 36 years between measurements. The correlation between SRBMI and BMI at ages 53–54 was slightly stronger among males ($r = .34$, $p < .05$) than among females ($r = .30$, $p < .05$). In the 2004 WLS surveys, participants were asked to report their highest weight and the age at which they weighed the most. Among 224 participants included in a subsample who reported their highest weight at ages 16–30, the correlation between BMI (using height at ages 53–54) and SRBMI was .48 ($p < .05$). Among 125 males, it was .48 ($p < .05$), and among 99 females, it was .55 ($p < .05$).

Adolescent RBM also predicted the risk of obesity (BMI ≥ 30) in adulthood. Study participants classified as underweight at ages 17–18 were 40% less likely than participants classified as healthy in weight to report obesity at ages 53–54 (odds ratio [OR] = 0.60, 95% confidence interval [CI] = 0.38, 0.96). Relative to healthy-weight adolescents, the risk of adult obesity was 2.74 (95% CI = 1.92, 3.90) times greater among adolescents at risk for overweight and 3.36 (95% CI = 2.44, 4.64) times greater among overweight adolescents, showing the graded effect of adolescent RBM on the odds of obesity in middle age. This gradient increased when class II obesity (BMI ≥ 35) was considered. Compared with healthy-weight adolescents, adolescents at risk for overweight had over a threefold risk (OR = 3.33; 95% CI = 2.05, 5.41) and overweight adolescents had nearly a *fivefold* risk (OR = 4.76; 95% CI = 2.90, 7.84) of exhibiting class II obesity at ages 53–54. These associations did not vary substantially by gender.

Participants with elevated RBM in 1957 tended to report more health symptoms and chronic conditions in 1993 than other WLS participants. Results from a series of logistic regression analyses of health symptoms on SRBMI (both in continuous and categorical form) are shown in Table 1. Although the continuous form of SRBMI was a significant

Table 1.   Odds Ratios for Associations Between Relative Body Mass in 1957 and Reported Health Symptoms in 1993

	Health Symptoms Experienced in Prior Six Months			
	Muscle Aches	Back Pain or Strain	Chest Pain	Shortness of Breath
Continuous Predictor				
Standardized relative body mass index (SRBMI)	1.13 (1.02, 1.26)	1.07 (0.95, 1.21)	1.21 (0.96, 1.53)	1.10 (0.93, 1.31)
Categorical Predictors				
Underweight	0.96 (0.68, 1.36)	0.88 (0.64, 1.21)	0.99 (0.54, 1.84)	1.02 (0.59, 1.75)
Healthy weight (reference)	—	—	—	—
At risk for overweight	1.07 (0.82, 1.40)	1.06 (0.77, 1.44)	0.85 (0.45, 1.61)	0.56 (0.33, 0.94)
Overweight	1.38 (1.05, 1.81)	1.19 (0.92, 1.55)	2.16 (1.35, 3.47)	1.81 (1.24, 2.65)

Notes: Numbers in parentheses are 95% confidence intervals. $n = 2,049$.

predictor of muscle aches only, the categorical form of SRBMI significantly predicted three of the four health symptoms. Specifically, persons classified as overweight in 1957 were 1.38 (95% CI = 1.05, 1.81) times more likely to report muscle aches, 1.81 (95% CI = 1.24, 2.65) times more likely to report shortness of breath, and 2.16 (95% CI = 1.35, 3.47) times more likely to report chest pain than persons classified as healthy in weight. Unexpectedly, participants at risk for overweight at ages 17–18 were significantly less likely than healthy-weight participants to report shortness of breath at ages 53–54.

Analyses of chronic conditions diagnosed by a health professional showed that the continuous form of SRBMI was a significant predictor of arthritis, high blood pressure, and diabetes (see Table 2). Also, adolescents classified as underweight in 1957 were 36%

Table 2.   Odds Ratios for Associations Between Relative Body Mass in 1957 and Reported Health Conditions in 1993

	Health Conditions Diagnosed by a Medical Professional			
	Arthritis	High Blood Pressure	Diabetes	Heart Trouble
Continuous Predictor				
Standardized relative body mass index (SRBMI)	1.19 (1.08, 1.31)	1.25 (1.11, 1.41)	1.44 (1.09, 1.90)	1.20 (0.96, 1.51)
Categorical Predictors				
Underweight	0.82 (0.56, 1.20)	0.64 (0.42, 0.99)	1.06 (0.45, 2.52)	1.04 (0.56, 1.95)
Healthy weight (reference)	—	—	—	—
At risk for overweight	0.95 (0.68, 1.34)	1.03 (0.72, 1.47)	1.77 (0.94, 3.34)	0.64 (0.32, 1.31)
Overweight	1.17 (0.87, 1.58)	1.23 (0.89, 1.69)	2.38 (1.33, 4.27)	1.79 (1.12, 2.86)

Notes: Numbers in parentheses are 95% confidence intervals. $n = 2,049$.

less likely than adolescents classified as healthy in weight to report high blood pressure in 1993 (OR = 0.64; 95% CI = 0.42, 0.99). Overweight adolescents were significantly more likely than healthy-weight adolescents to report diabetes (OR = 2.38; 95% CI = 1.33, 4.27) and heart trouble (OR = 1.79; 95% CI = 1.12, 2.86) in adulthood. When the continuous and categorical forms of SRBMI are viewed together, the only health symptom or chronic condition *not* significantly associated with adolescent RBM was back pain or strain.

To determine whether associations between adolescent SRBMI and adult health outcomes were mediated by adult BMI, we reexamined regression models while controlling for the 1993 measure of BMI (results not shown). In most instances, introducing current BMI substantially attenuated associations between adolescent SRBMI and health conditions in later life. However, current BMI did not explain the presence of certain health outcomes among overweight adolescents. For instance, overweight adolescents were 1.81 (95% CI = 1.09, 2.98) times more likely than healthy-weight adolescents to report chest pain in 1993, even after the 1993 measure of BMI was taken into account.

Data from the NDI indicate that SRBMI predicted all-cause mortality (OR = 1.25; 95% CI = 1.05, 1.49) among WLS participants between 1979 and 1998 (see Table 3). The removal of accidental deaths strengthened this association. The continuous form of SRBMI was a particularly strong predictor of heart disease mortality (OR = 1.69; 95% CI = 1.10, 2.58). Relative to healthy-weight adolescents, overweight adolescents had roughly a two-fold risk of premature mortality from nonaccidental causes (OR = 1.91; 95% CI = 1.24, 2.94) and a *fourfold* risk of premature mortality from heart disease (OR = 3.99; 95% CI = 1.85, 8.61). Neither continuous nor categorical SRBMI variables were significant predictors of cancer mortality, although it is worth noting that the observed coefficients were all in the expected direction. For instance, relative to healthy-weight adolescents, underweight adolescents were about 12% less likely (OR = 0.88; 95% CI = 0.39, 1.96) and overweight adolescents were about 50% more likely (OR = 1.49; 95% CI = 0.76, 2.94) to die from cancer between 1979 and 1998.

Table 3. **Odds Ratios for Associations Between Relative Body Mass in 1957 and Mortality Indicators From the National Death Index (NDI)**

	Mortality Indicator			
	NDI, All Cause	NDI, Nonaccidental	NDI, Heart Disease	NDI, Cancer
Continuous Predictor				
Standardized relative body mass index (SRBMI)	1.25 (1.05, 1.49)	1.32 (1.10, 1.59)	1.69 (1.10, 2.58)	1.16 (0.91, 1.49)
Categorical Predictors				
Underweight	0.90 (0.53, 1.52)	0.86 (0.48, 1.53)	0.98 (0.29, 3.32)	0.88 (0.39, 1.96)
Heathy weight (reference)	—	—	—	—
At risk for overweight	1.22 (0.77, 1.94)	1.38 (0.86, 2.21)	1.28 (0.49, 3.37)	1.01 (0.49, 2.07)
Overweight	1.69 (1.10, 2.58)	1.91 (1.24, 2.94)	3.99 (1.85, 8.61)	1.49 (0.76, 2.94)

Notes: Numbers in parentheses are 95% confidence intervals. n = 3,010.

DISCUSSION

Despite many laudable qualities, the WLS has been limited in etiological investigations of health by the lack of a baseline measure of body mass. Through the development of the RBM scale, this study has removed that limitation. The RBM scale is reliable, and it meets several criteria for validity. An index summarizing the RBM scores of individual coders (i.e., SRBMI) was uncorrelated with height, as it should be. The RBM scale also has criterion (predictive) validity. It is moderately correlated with BMI after the passage of 36 years and more highly correlated with BMI among graduates who reached their maximum weight between the ages of 16 and 30. Continuous and categorical measures of adolescent SRBMI predicted the midlife risk of obesity, health symptoms, chronic conditions, and mortality. Current BMI mediated some of the associations between adolescent SRBMI and adult health outcomes, emphasizing the importance of weight control after adolescence but also the unfortunate persistence of some deleterious health consequences of overweight in adolescence. To be sure, it would be desirable to validate RBM against a direct, contemporaneous measure of BMI, but that has not been possible in the WLS.

Available evidence indicates that the RBM scale is comparable with adolescent BMI in terms of its ability to predict mid-life BMI, morbidity, and mortality. For instance, a British cohort study reported a correlation of .39 between measures of BMI taken at ages 13 and 50 (Wright et al. 2001). The British cohort was followed for approximately the same length of time (38 years) as the WLS cohort (36 years), so it is encouraging to note that the correlation observed by Wright et al. was similar to the correlation observed between adolescent RBM and adult BMI in the WLS.

Analyses of 508 participants in the Harvard Growth Study found that overweight adolescents (defined as BMI ≥ 75th percentile) were significantly more likely to report diabetes, coronary heart disease (CHD), arthrosclerosis, and hip fracture than lean adolescents (25th–50th percentile of BMI) after 55 years of follow-up (Must et al. 1992). To illustrate, participants who were overweight as adolescents were about two times more likely to report CHD and diabetes than participants who were lean. Childhood BMI is also a significant predictor of CHD risk factors, such as cholesterol, triglycerides, insulin imbalance, and blood pressure (Freedman et al. 2001). These results are similar to our own findings. Adolescent measures of RBM were significant predictors of several health symptoms (muscle aches, chest pain, and shortness of breath) and chronic conditions (arthritis, high blood pressure, diabetes, and heart trouble) that are indicative of cardiovascular, pulmonary, and musculoskeletal distress in adulthood.

Studies that have investigated the relationship between adolescent BMI and premature mortality are generally consistent with our results. For example, Must et al. (1992) found elevated risks of all-cause and cause-specific (e.g., heart disease) mortality among men—but not women—who were obese as adolescents. However, a longitudinal study of 128,121 Norwegians found that overweight adolescents were at increased risk of premature mortality, regardless of gender (Engeland et al. 2004). Relative to adolescents with medium BMI (25th–74th percentiles), adolescents with very high BMI (≥ 85th percentile) were 40% more likely to die within 29 years of follow-up. By way of comparison, overweight adolescents in the WLS were 69% more likely than their healthy-weight peers to die within 41 years of follow-up.

On several occasions, coders raised questions about the proper coding of large individuals—particularly males—with athletic characteristics (e.g., wide necks) and relatively lean builds. Because the RBM scale was developed as a proxy for BMI, coders were instructed to code such persons as heavy on the RBM scale, regardless of whether this was caused by lean mass or fat mass. This raises the possibility of gender differences in RBM but, as noted, interaction effects between gender and RBM did not significantly improve the fit of models explored here.

The inability of the RBM scale to differentiate lean mass from fat mass may be viewed as a limitation of the measure, but it is one shared with the most common measure of body mass: BMI. Other limitations of this study include the following. First, there were relatively few coders of RBM, making it difficult to assess systematic differences in coding tendencies by gender, age, ethnicity, or socioeconomic background. Second, WLS participants are almost all non-Hispanic whites, and we do not know whether the reliability of our codes can be generalized to other populations. Third, although extracted photographs of WLS participants were generally of similar type and quality, there were some inconsistencies that likely contributed to random error. For instance, coders remarked that certain poses (e.g., profile shots) and clothing options (e.g., turtlenecks) obfuscated some features, and particularly the neck. Also, all WLS photographs were black-and-white (not color), which could be disadvantageous relative to color photographs in the measurement of RBM. On a handful of occasions, coders expressed difficulty discerning whether apparent features of the face and neck (e.g., folds of adipose tissue in the neck) were real or the by-product of shadows. If available, color photographs could help resolve those difficulties. Fourth, adult BMI, health symptoms, and chronic conditions were based on self-reports, which may have led to some attenuation in the associations between these outcomes and RBM. Last, although the sample size in this study was relatively large, it nevertheless covered less than one in three WLS participants. This limited the statistical power of this investigation and will present some limitations to future investigations until a measure of RBM is available for many more WLS participants.

In future studies, it could be useful to develop a measure of facial adiposity to complement the RBM scale. A measure designed to isolate deposits of fat in the cheeks and in the neck from lean facial tissue has the potential to provide a superior indicator of obesity, per se. Understanding the relative impact of components of lean, fat, and total facial mass on the incidence of health symptoms, chronic disease, and mortality could also help improve the understanding of disease etiology and lead to the development of new applications in research and clinical practice. Recall, for instance, that some studies (e.g., Tafeit et al. 2000) have found that characteristics of the face and neck are strong predictors of Type 2 diabetes, even when compared with traditional measures, such as BMI or body fat percentage.

Future studies may also benefit from the inclusion of a larger and more diverse group of coders. The inclusion of additional coders would help maximize the reliability of the RBM scale, although it is worth reiterating that the degree of internal consistency achieved in this study was excellent. Additional coders would also provide more flexibility to retain only individuals who perform well in terms of interrater reliability, intrarater reliability, and criterion-related validity. Clearly, humans must possess varying levels of ability with regard to the visual assessment of RBM. Future studies could take advantage of this variation by identifying persons with the highest levels of ability. Furthermore, the inclusion of a more diverse set of coders with respect to age, ethnicity, and socioeconomic background could help determine whether important differences exist across these groups. Limited existing research suggests that gender differences in body image do not distort intersubjective agreement regarding the relative weight of persons in photographs (Rudin 1996), but more research is needed to answer questions regarding the effects of gender and other personal characteristics on coding tendencies.

Despite these limitations and intriguing prospects for future study, the development of the RBM scale represents an important methodological development for the WLS. Other longitudinal studies without baseline measures of body mass may also benefit from the development of similar RBM indicators. In the WLS, we have been able to obtain this measure for individuals who died in early adulthood, as well as for those who are still participating in the study. In American populations, it should be feasible to obtain photographic evidence of the kind used in the WLS, even many years after secondary schooling has been

completed. In such applications, the RBM scale would likely represent a valid and reliable form of retrospective measurement, as it has in the WLS. Furthermore, emerging prospective studies of physical and psychosocial health from childhood through adulthood and older ages could contribute to knowledge in this area by incorporating RBM measures. Importantly, this would provide opportunities to evaluate RBM measures contemporaneously with traditional measures of body mass, providing more definitive evidence regarding their similarities and differences. Finally, it is encouraging to consider that while the RBM scale performed admirably in this study, it is a new construct. With additional research, refinements to RBM measures promise to unlock their full potential as legitimate alternatives or complements to traditional measures of body mass in human populations.

REFERENCES

Alwin, D.F. 1992. "Information Transmission in the Survey Interview: Number of Response Categories and the Reliability of Attitude Measurement." *Sociological Methodology* 22:83–118.

———. 1997. "Feeling Thermometers Versus 7-Point Scales: Which Are Better?" *Sociological Methods & Research* 25:318–40.

Ayres, L.P. 1909. *Laggards in Our Schools: A Study of Retardation and Elimination in City School Systems*. New York: Charities Publication Committee.

Carr, D. 1997. "The Fulfillment of Career Dreams at Midlife: Does It Matter for Women's Mental Health?" *Journal of Health and Social Behavior* 38:331–44.

Cronbach, L.J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16: 297–334.

Crossman, A., D.A. Sullivan, and M. Benin. 2006. "The Family Environment and American Adolescents' Risk of Obesity as Young Adults." *Social Science & Medicine* 63:2255–67.

Engeland, A., T. Bjørge, A. Tverdal, and A.J. Søgaard. 2004. "Obesity in Adolescence and Adulthood and the Risk of Adult Mortality." *Epidemiology* 15:79–85.

Flegal, K.M., M.D. Carroll, R.J. Kuczmarski, and C.L. Johnson. 1998. "Overweight and Obesity in the United States: Prevalence and Trends, 1960–1994." *International Journal of Obesity and Related Metabolic Disorders* 22:39–47.

Flynn, K.E., M.A. Smith, and D. Vanness. 2006. "A Typology of Preferences for Participation in Healthcare Decision Making." *Social Science & Medicine* 63:1158–69.

Freedman, D.S., L.K. Khan, W.H. Dietz, S.R. Srinivasan, and G.S. Berenson. 2001. "Relationship of Childhood Obesity to Coronary Heart Disease Risk Factors in Adulthood: The Bogalusa Heart Study." *Pediatrics* 108:712–18.

Furnham, A. and S. Radley. 1989. "Sex Differences in the Perception of Male and Female Body Shapes." *Personality and Individual Differences* 10:653–62.

Goodman, E. and R.C. Whitaker. 2002. "A Prospective Study of the Role of Depression in the Development and Persistence of Adolescent Obesity." *Pediatrics* 110:497–504.

Hauser, R.M. 2005. "Survey Response in the Long Run: The Wisconsin Longitudinal Study." *Field Methods* 17:3–29.

Himes, C.L. 2000. "Obesity, Disease, and Functional Limitation in Later Life." *Demography* 37:73–82.

Johnson, E.O. and N. Breslau. 2006. "Is the Association of Smoking and Depression a Recent Phenomenon?" *Nicotine & Tobacco Research* 8:257–62.

Kato, K. and A. Higashiyama. 1998. "Estimation of Height for Persons in Pictures." *Perception and Psychophysics* 60:1318–28.

Krahn, D., J. Freese, R. Hauser, K. Barry, and B. Goodman. 2003. "Alcohol Use and Cognition at Mid-Life: The Importance of Adjusting for Baseline Cognitive Ability and Educational Attainment." *Alcoholism: Clinical and Experimental Research* 27:1162–66.

Laakso, M., V. Matilainen, and S. Keinänen-Kiukaanniemi. 2002. "Association of Neck Circumference With Insulin Resistance-Related Factors." *International Journal of Obesity and Related Metabolic Disorders* 26:873–75.

Lean, M.E., T.S. Han, and J.C. Seidell. 1999. "Impairment of Health and Quality of Life Using New US Federal Guidelines for the Identification of Obesity." *Archives of Internal Medicine* 159:837–43.

Levine, J.A., A. Ray, and M.D. Jensen. 1998. "Relation Between Chubby Cheeks and Visceral Fat." *New England Journal of Medicine* 339:1946–47.

Mannino, D.M., J. Mott, J.M. Ferdinands, C.A. Camargo, M. Friedman, H.M. Greves, and S.C. Redd. 2006. "Boys With High Body Masses Have an Increased Risk of Developing Asthma: Findings From the National Longitudinal Survey of Youth (NLSY)." *International Journal of Obesity* 30:6–13.

Manson, J.E., P.J. Skerrett, and W.C. Willett. 2002. "Epidemiology of Health Risks Associated With Obesity." Pp. 422–32 in *Eating Disorders and Obesity: A Comprehensive Handbook,* 2nd ed., edited by C.G. Fairburn and K.D. Brownell. New York: The Guilford Press.

Marks, N.F. and D.S. Shinberg. 1997. "Socioeconomic Differences in Hysterectomy: The Wisconsin Longitudinal Study." *American Journal of Public Health* 87:1507–14.

———. 1998. "Socioeconomic Status Differences in Hormone Therapy." *American Journal of Epidemiology* 148:581–93.

Meland, S.A. 2002. *Objectivity in Perceived Attractiveness: Development of a New Methodology for Rating Facial Attractiveness*. Master's thesis. Department of Sociology, University of Wisconsin–Madison.

Möller, R., E. Tafeit, T.R. Pieber, K. Sudi, and G. Reibnegger. 2000. "Measurement of Subcutaneous Adipose Tissue Topography (SAT-Top) by Means of a New Optical Device, LIPOMETER, and the Evaluation of Standard Factor Coefficients in Healthy Subjects." *American Journal of Human Biology* 12:231–39.

Mortimore, I.L., I. Marshall, P.K. Wraith, R.J. Sellar, and N.J. Douglas. 1998. "Neck and Total Body Fat Deposition in Nonobese and Obese Patients With Sleep Apnea Compared With That in Control Subjects." *American Journal of Respiratory and Critical Care Medicine* 157:280–83.

Must, A., P.F. Jacquez, G.E. Dallal, C.J. Bajema, and W.H. Dietz. 1992. "Long-Term Morbidity and Mortality of Overweight Adolescents. A Follow-Up of the Harvard Growth Study of 1922 to 1935." *New England Journal of Medicine* 327:1350–55.

National Center for Health Statistics. 1999. "National Death Index." Available online at http://www.cdc.gov/nchs/ndi.htm.

Novak, M., C. Ahlgren, and A. Hammarstrom. 2006. "A Life-Course Approach in Explaining Social Inequity in Obesity Among Young Adult Men and Women." *International Journal of Obesity* 30:191–200.

Nunnally, J.C. 1978. *Psychometric Theory*. New York: McGraw-Hill.

Ogden, C.L., K.M. Flegal, M.D. Carroll, and C.L. Johnson. 2002. "Prevalence and Trends in Overweight Among US Children and Adolescents, 1999–2000." *Journal of the American Medical Association* 288:1728–32.

Pi-Sunyer, X.F. 2002. "Medical Complications of Obesity in Adults." Pp. 467–72 in *Eating Disorders and Obesity: A Comprehensive Handbook,* 2nd ed., edited by C.G. Fairburn and K.D. Brownell. New York: The Guilford Press.

Reither, E.N. 2005. "Why Are Our Waistlines Expanding? Age-Period-Cohort Analyses of the Obesity Epidemic and a Critical Examination of Mass Preparation Theory." Ph.D. dissertation. Department of Sociology, University of Wisconsin–Madison.

Rexrode, K.M., V.J. Carey, C.H. Hennekens, E.E. Walters, G.A. Colditz, M.J. Stampfer, W.C. Willett, and J.E. Manson. 1998. "Abdominal Adiposity and Coronary Heart Disease in Women." *Journal of the American Medical Association* 280:1843–48.

Rosenkranz, M.A., D.C. Jackson, K.M. Dalton, I. Dolski, C.D. Ryff, B.H. Singer, D. Muller, N.H. Kalin, and R.J. Davidson. 2003. "Affective Style and *in Vivo* Immune Response: Neurobehavioral Mechanisms." *Proceedings of National Academy of Sciences of the United States of America* 100:11148–52.

Rudin, J.P. 1996. "Perceived Weight of Other Persons." *Journal of Social Psychology* 136: 719–26.

Sadeghianrizi, A., C.-M. Forsberg, C. Marcus, and G. Dahllöf. 2005. "Craniofacial Development in Obese Adolescents." *European Journal of Orthodontics* 27:550–55.

SAS Institute Inc. 2003. *SAS for Windows, Version 9.1*. Cary, NC: SAS Institute Inc.

Seltzer, M.M., F. Floyd, J. Greenberg, J. Lounds, M. Lindstromm, and J. Hong. 2005. "Life Course Impacts of Mild Intellectual Deficits." *American Journal on Mental Retardation* 110:451–68.

Sewell, W.H., R.M. Hauser, K.W. Springer, and T.S. Hauser. 2004. "As We Age: A Review of the Wisconsin Longitudinal Study, 1957–2001." Pp. 3–114 in *Research in Social Stratification and Mobility,* Vol. 20, edited by K.T. Leicht. New York: Elsevier.

Shinberg, D.S. 1998. "An Event History Analysis of Age at Last Menstrual Period: Correlates of Natural and Surgical Menopause Among Midlife Wisconsin Women." *Social Science & Medicine* 46:1381–96.

StataCorp LP. 2005. *Stata Statistical Software: Release 9*. College Station, TX: StataCorp LP.

Stunkard, A.J. 1996. "Current Views on Obesity." *American Journal of Medicine* 100:230–36.

Sundquist, J. and S.-E. Johansson. 1998. "The Influence of Socioeconomic Status, Ethnicity and Lifestyle on Body Mass Index in a Longitudinal Study." *International Journal of Epidemiology* 27:57–63.

Tafeit, E., R. Möller, K. Sudi, and G. Reibnegger. 2000. "ROC and CART Analysis of Subcutaneous Adipose Tissue Topography (SAT-Top) in Type-2 Diabetic Women and Healthy Females." *American Journal of Human Biology* 12:388–94.

U.S. Department of Education, Center for Education Statistics. 1992. *National Longitudinal Study of the Class of 1972*. [Computer file]. Bibliographic Citation: ICPSR version. Chicago, IL: National Opinion Research Center [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1999.

U.S. Department of Health and Human Services. Centers for Disease Control and Prevention. 2004. "BMI — Body Mass Index: About BMI for Adults." Available online at http://www.cdc.gov/nccdphp/dnpa/bmi/adult_BMI/about_adult_BMI.htm.

World Health Organization. 1977. *Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death, The Ninth Revision (ICD-9)*. Geneva: WHO.

Wright, C.M., L. Parker, D. Lamont, and A.W. Craft. 2001. "Implications of Childhood Obesity for Adult Health: Findings From Thousand Families Cohort Study." *BMJ* 323:1280–84.

Yu, X., K. Fujimoto, K. Urushibata, Y. Matsuzawa, and K. Kubo. 2003. "Cephalometric Analysis in Obese and Nonobese Patients With Obstructive Sleep Apnea Syndrome." *Chest* 124:212–18.