SSM - Population Health 11 (2020) 100547

Contents lists available at ScienceDirect

SSM - Population Health

journal homepage: http://www.elsevier.com/locate/ssmph

Article

Beyond recent BMI: BMI exposure metrics and their relationship to health[☆]

Carmen D. Ng[a,*], Michael R. Elliott[b], Fernando Riosmena[c], Solveig A. Cunningham[a]

[a] Emory University, United States
[b] University of Michigan, United States
[c] University of Colorado, United States

ARTICLE INFO

Keywords:
Body mass index
Chronic health
Longitudinal studies
Obesity

ABSTRACT

Body mass index (BMI) is generally used to classify adiposity. Despite the fact that the consequences of adiposity for chronic health accumulate and manifest over time, most population health research exploring the implications of high BMI measures only its recent intensity. Some studies have used retrospective measures involving maximum weight, and even fewer have used BMI at multiple time points to estimate cumulative exposure to adiposity. The goal of this study was to compare BMI exposure metrics that captured different dimensions of body mass – intensity, history, and duration – in models of health indicators linked with adiposity. We used self-reported BMI of young adults (ages 18 – 33 years, n = 8,608) across 11 waves of data from the National Longitudinal Survey of Youth 1997 to evaluate eight BMI exposure metrics: most recent, maximum, mean, and median BMI, proportion of time with overweight/obesity, and excess BMI-years with overweight/obesity. We used these metrics in models of self-reported general health, chronic condition, and diabetes, and ascertained how most recent BMI performed when compared with other metrics that better capture the dynamics of BMI. The Akaike information criteria and Vuong tests were used for model comparison, and the strengths of associations were also compared. Most recent BMI was the best metric for explaining general health. Median BMI was best for explaining diabetes, with most recent BMI under-estimating the association by 13% relative to median BMI. For chronic condition, there was no clear best metric. We concluded that most recent BMI is useful for explaining health outcomes, though other metrics should also be given consideration, particularly for conditions that develop over time. Metrics that accounted for both intensity and history performed quite well, but the duration measures might be less useful.

Introduction

About two-thirds of American adults have body mass index (BMI) classified in the overweight or obesity statuses (Ogden, Carroll, Kit, & Flegal, 2014), and those with high BMI tend to have worse social, economic, and health outcomes (Apovian, 2013; Gortmaker, Must, Perrin, Sobol, & Dietz, 1993; Hammond & Levine, 2010; Lehnert, Sonntag, Konnopka, Riedel-Heller, & Konig, 2013; Must et al., 1999; Ogden, Carroll, & Flegal, 2003; Pi-Sunyer, 2002; Wyatt, Winters, & Dubbert, 2006). The vast majority of studies exploring the implications of high BMI on health use height and weight at survey time. However, BMI is dynamic and fluctuates over time. Effects of elevated BMI, especially at levels classified as obesity, have been found to accumulate over time, resulting in earlier and higher morbidity and mortality (Abdullah et al., 2011; de Lauzon-Guillain et al., 2010; Elliott, Aucott, Hannaford, & Smith, 2005; Myrskyla & Chang, 2009; Yarnell, Patterson, Thomas, & Sweetnam, 2000; Zheng, Tumin, & Qian, 2013). It may thus be informative to consider BMI dynamics and cumulative exposure to elevated BMI in order to develop a better understanding of the relationship between body mass and health. However, to date, there is no established or generally accepted method for taking such aspects into account. While BMI at survey time is an affordable and easy-to-use measure, it is only a proxy of more nuanced and longer-term processes of adiposity accumulation. It thus needs to be compared to other BMI measures in order to

☆ This work was supported by a grant from the National Institute of Health's National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK115937-01). Publication of this article was funded by the University of Colorado Boulder Libraries Open Access Fund.

* Corresponding author. Emory University, Hubert Department of Global Health, 7050-C Claudia Nance Rollins Building, 1518 Clifton Road, Atlanta, GA, 30322, United States.

E-mail address: carmen.ng@emory.edu (C.D. Ng).

https://doi.org/10.1016/j.ssmph.2020.100547
Received 17 September 2019; Received in revised form 26 January 2020; Accepted 26 January 2020
Available online 2 March 2020
2352-8273/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

evaluate its suitability as a predictor of health outcomes.

To go beyond using BMI at the time of survey, which is the most predominantly used, but which accounts for only a person's current status, researchers have taken various approaches to measure previous exposure to high BMI. A number of studies using cross-sectional data have employed the concept of maximum BMI, or BMI as determined by height at survey time and highest recalled weight (Elo, Mehta, & Preston, 2017; Mehta et al., 2014, 2014; Stokes, 2014; Vierboom, 2017). Maximum BMI attempts to measure adiposity at the most intensive moment. Elevated levels of BMI, no matter how temporary, could have a lasting effect on health. In addition to accounting for previous body mass, this metric is less sensitive to illness-induced weight loss, an important consideration at older ages (Stokes, 2014). However, this metric, as defined above, has several weaknesses. First, recalled weight information has been found to be inaccurate, and the inaccuracy is not consistent across population segments (Dahl & Reynolds, 2013; Kyulo, Knutsen, Tonstad, Fraser, & Singh, 2012; Stevens, Keil, Waid, & Gazes, 1990; Tamakoshi et al., 2003). Second, the use of height at survey time in the definition of maximum BMI implicitly assumes that height has not changed since the time when maximum weight was attained, which is especially problematic for children and the elderly. These limitations could be rectified with the use of longitudinal data in which height and weight are measured repeatedly.

Although maximum BMI does account for previous most intensive exposure, it does not take into consideration dynamics as maximum weight represents the weight at only one point in time. Measures of central tendency, such as mean or median BMI, would "average" the fluctuations of BMI over time, which might be more indicative of historical exposure to adiposity than maximum BMI. However, these two measures are only possible with longitudinal data with multiple records over time, and not many studies have gone this route.

While all these metrics are informative, they reflect discrete point measures of BMI, whether it is at the most recent time point, at the time point of maximum BMI, or at a combination of several time points. Longitudinal data would also provide the duration of exposure to elevated BMI, such as the proportion of time spent in obesity or overweight statuses (Mehta et al., 2014, 2014; Muscelli et al., 1998; Santamaria et al., 2011). A more complicated measure is excess BMI-years, a composite measure of the duration when an individual's BMI exceeds a certain threshold and the extent of such excess, and which is analogous to "pack-years" in measuring smoking exposure (Guaraldi et al., 2015; Khan et al., 2006; Kim et al., 2014; Saquib, Stefanick, Natarajan, & Pierce, 2013). A few studies have used excess BMI-years to investigate the risk of diabetes or cardiovascular disease onset (Hu et al., 2017; Lee, Gebremariam, Vijan, & Gurney, 2012; Reis et al., 2015), and while they did observe that excess BMI-years was associated with the risk of such health outcomes, they did not demonstrate how it fared in comparison to other metrics, thus not demonstrating whether it is in fact a better measure. Here, we considered both excess BMI-years above the obesity and overweight thresholds to incorporate intensity, history, and duration into our understanding of adiposity.

We used more than a decade of data from the National Longitudinal Survey of Youth 1997 (NLSY97) to analyze eight BMI exposure metrics: most recent BMI, maximum BMI, mean BMI, median BMI, proportion of time with overweight, proportion of time with obesity, excess BMI-years with overweight, and excess BMI-years with obesity. We discussed how these various operationalizations of BMI differ, and demonstrated their use as covariates in models of health indicators – specifically, general health, chronic condition, and diabetes. These indicators were chosen to capture different dimensions of health, since they likely have varying relationships with adiposity. Comparing different ways of characterizing BMI dynamics over time would help ascertain whether more sophisticated metrics of BMI that account for intensity, history, and/or duration provide different information or yield better-fitting models than metrics that require only cross-sectional data. This assessment would thus help validate the many studies that use recent BMI, or otherwise limit the reach of their interpretations. This information would also be important to help us understand our data needs in population-based studies of cardiometabolic health.

Material and methods

Data

The NLSY97 was conducted by the Bureau of Labor Statistics (BLS) and followed a sample of almost 9,000 American youths (with oversamples of non-Hispanic blacks and Hispanics) who were between the ages of 12 and 16 years at the end of 1996 (Moore, Pedlow, Krishnamurty, & Wolter, 2000). Individuals were interviewed annually from 1997 through 2011, and then in 2013 and 2015 (Moore et al., 2000). Because people were likely still growing at the early waves of the survey, we used the data starting from 2002, when survey participants were between the ages of 18 and 22 years. Additionally, the most recent wave of data with information on the health indicators of interest was in 2013, when survey participants were between 29 and 33 years of age. There were 11 data waves between 2002 and 2013, providing more than a decade of information on early adulthood.

While the NLSY97 contained information on a broad range of subjects, we were most interested in health and contextual variables. At each wave, individuals were asked to report their weights, and at each wave except 2013, individuals were asked to report their heights.

The health indicators used in this analysis were self-reported by respondents in 2013: general health (on a scale of one to five, five being the best), chronic condition diagnosis (to probe, respondents were asked to consider as examples asthma, cardiovascular or heart condition, anemia, diabetes, cancer, epilepsy, HIV/AIDS, sexually transmitted diseases other than HIV/AIDS, or another chronic health condition or life-threatening disease), and diabetes diagnosis. Elevated BMI is a known risk factor for diabetes. However, because diabetes might not have a high prevalence among this age group, we also considered the more general diagnosis of chronic conditions. Finally, general health was chosen to capture the sense of well-being, encompassing aspects of both physical and mental health possibly related to high BMI.

Other variables of interest included sex, race, birth year, and having health care coverage/insurance, all potentially associated with health or reporting about health in ways that could be confounded with higher/lower BMI. Health insurance information was taken from 2013, to match up with the data on health outcomes.

Computing the BMI exposure metrics

BMI was calculated using height and weight information. Not considering 2013, in which height was not asked, we calculated 71,768 BMI values for survey participants in the previous ten waves. This amounted to around 80% of all possible BMI values. Biologically extreme values of height and weight (height <48 in or >84 in, weight <75 lb or >700 lb) were recoded as missing in this study (Noel et al., 2010). If a woman reported being pregnant at a data wave, her weight would not be representative of her true BMI and was therefore recoded as missing. These biologically improbable values and pregnancies resulted in an additional 2,273 missing BMI values, or about 3% of available data.

A missing height or weight value between waves with available height/weight data was linearly interpolated. For a person's height or weight to change from one value to another value, it must pass through all intermediate values. The intermediate value relative to its position on the time scale, i.e., the linearly interpolated value, was thus a reasonable estimate. Linear interpolation assumed a constant rate of gain/loss per unit time. While non-linear interpolation methods could have been used here, they would not have offered superior theoretical support to compensate for their numerical complexity, particularly because the time between measurements was fairly short. Missing height and weight values at the two ends of the study time horizon were replaced by the nearest available height/weight value. As a result, height in 2013 was considered to be the same as height in 2011, or whatever the most recent height value was. Extrapolation was not used so as to avoid making any estimates outside the range with known values. As a robustness check, multiple imputation was used to fill in missing values using the Imputation and Variance Estimation Software (IVEware) version 0.3 (Raghunathan, Solenberger, Berglund, & van Hoewyk, 2016), and there were no substantive differences in conclusions.

BMI in kg/m^2 was then calculated, and improbable values ($BMI < 10\,kg/m^2$ or $> 75\,kg/m^2$) were recoded as missing and filled in using the same replacement procedures. Ultimately, 25,193 BMI values were obtained by the above described methodology, a third of which were due to missing height in 2013. As a result, except for respondents with missing BMI at every wave, all observations in our analyses had valid BMI data spanning the entire length of this study. Most of those who were excluded were no longer in the survey by 2002, the first year of this study.

With valid BMI data spanning 2002 – 2013 for 8,608 respondents, we then proceeded with our eight BMI exposure metrics. *BMI at the most recent wave* (2013) was indicative of current intensity. To incorporate history in addition to intensity, we calculated *maximum BMI*, *mean BMI*, and *median BMI* across all 11 waves.

To take into account the duration of exposure to high BMI, we calculated *proportion of time with obesity*, where obesity was defined as having a BMI $\geq 30\,kg/m^2$. From 2002 to 2010, when the NLSY97 survey was conducted annually, the BMI calculated each year was assumed to last the duration of the entire year. From 2011 to 2013, when the survey was conducted biennially, the BMI calculated in each of these two years was assumed to have a duration of 1.5 years. Proportion of time with obesity was the total duration in the obesity state divided by the total duration of 12 years in this study.

$$\text{Mathematically, this was } \frac{\sum_i duration\ with\ obesity_i}{\sum_i duration_i}$$

where i indexed the wave in the survey.

The strength of such a metric was the incorporation of both history and duration. However, intensity was not accounted for beyond whether someone had a BMI that was above the obesity threshold.

To express history of intensity along with duration, we used *excess BMI-years with obesity*. First, we calculated excess BMI at each wave as BMI – 30, with a minimum of 0. In other words, survey participants with BMI values under 30 were considered not to have excess BMI. In previous studies, excess BMI-years had been computed without a floor function (Hu et al., 2017; Lee et al., 2012; Reis et al., 2015). But without it, an individual could have offsetting positive and negative excesses. Since the purpose of these metrics was to study exposure to elevated levels of BMI, the floor function made sure that the metric focused on this. Excess BMI-years was thus the accumulated excess BMI experienced by a survey participant over the entire time span of the survey. As a result, this metric could be envisioned as the area between the obesity threshold and the BMI curve, when the latter was above the threshold, over the duration of the survey.

$$\text{Mathematically, this was } \sum_i \max(BMI_i - 30,\ 0) * duration_i$$

where i indexed the wave in the survey.

Proportion of time with overweight and *excess BMI-years with overweight* were overweight analogues to the obesity measures. Overweight, indicating less severe yet still elevated adiposity, had a threshold of $25\,kg/m^2$. As a result, overweight was a superset of obesity, and overweight and obesity were not treated as mutually exclusive categories; they simply referred to being above their respective BMI thresholds.

Statistical analysis

We calculated pairwise correlations and descriptive statistics of these eight metrics over the span of early adulthood in the NLSY97.

These BMI exposure metrics were subsequently used as covariates in models of health indicators. General health was treated as a continuous variable in a linear regression model, while reporting any chronic condition and diabetes were treated as dichotomous variables in logistic regression models. General health was treated as both unordered and ordered categorical variables in alternative models, but as substantive results did not differ, linear regression was chosen due to its relative simplicity.

Each of the eight BMI exposure metrics was used individually in models for each of these three health indicators, resulting in 24 models. We investigated the associations between BMI exposure metrics and health indicators. Since these BMI exposure metrics were in different units, they were standardized via centering and scaling before being used as inputs in the models to facilitate comparison of coefficient estimates across models for the same outcome. For each BMI exposure metric, we calculated the mean and standard deviation, and standardized each value of such measure by subtracting the mean and dividing by the standard deviation. This made the interpretation similar across all metrics – the number of standard deviations from the mean. Basic demographic variables likely associated with our health indicators of interest, including sex, race, birth year, and having health care coverage/insurance, were controlled for in these models. Comparison of models was made with the Akaike information criterion (AIC), which is commonly used for model selection; lower values of AIC are indicative of better fit (Burnham & Anderson, 2004).

As a complementary analysis, for each health indicator, Vuong tests for model comparison were then performed for each pair of models to test whether one fit better than the other. Vuong tests allow for comparisons between non-nested models to statistically test whether two competing models are equally close to the true data generating process or whether one model is closer than the other (Vuong, 1989). For each health outcome, there were ${}_8C_2 = 28$ pairwise comparisons. p-values were reported for each comparison and corrections for multiple testing were not used. Bonferroni corrections, and multiple testing more broadly, have been criticized as being too conservative, especially with a

Table 1
Summary statistics of BMI exposure metrics, adjusted by survey weights to be representative of the U.S. young adult population in 2002 – 2013 (n = 8,608).

BMI metric	Mean	Standard error	Range
Most recent BMI	27.837	0.095	14.25 – 74.04
Maximum BMI	29.567	0.110	16.24 – 74.04
Mean BMI	26.602	0.083	15.57 – 65.32
Median BMI	26.504	0.085	15.02 – 62.96
Proportion of time with overweight	0.526	0.006	0 – 1
Proportion of time with obesity	0.226	0.005	0 – 1
Excess BMI-years with overweight	36.573	0.807	0 – 483.85
Excess BMI-years with obesity	15.207	0.561	0 – 423.85

Data: National Longitudinal Survey of Youth 1997.

large number of tests, as in this study (Perneger, 1998; Sedgwick, 2014, 2014; Streiner & Norman, 2011). Consequently, these p-values were not adjusted, and readers could make adjustments appropriate for the purpose of their applications.

The survey design of the NLSY97 was accounted for with sample clustering variables. Custom survey weights for observations in any of the years from 2002 to 2013 were used in the tabulation of descriptive statistics and the correlations to ensure national representativeness. However, survey weights were not used for regression analyses, as advised by the BLS (Bureau of Labor Statistics, 2018).

All analyses were conducted in R version 3.5.1 and the R "survey" package was used for analyses involving survey weights (R Core Team, 2018; Lumley, 2004).

Results

The NLSY97 cohort was 49% female and 51% male. About one-fifth of the cohort was born in each of the five years from 1980 to 1984. The race-ethnicity breakdown was 70.5% non-black/non-Hispanic, 15.4% black/non-Hispanic, 12.9% Hispanic, and 1.2% mixed/non-Hispanic. Three-quarters of the cohort had health insurance in 2013. In 2013, mean self-rated health was approximately 3.7 (on a scale of one to five), about 12% of people reported a chronic condition, and 1.5% of people reported a diabetes diagnosis.

Table 1 shows the summary statistics (mean, standard error, and range) for each of the eight BMI exposure metrics. Of the four metrics in the same unit of measure (kg/m^2), the averages of median and mean BMI (26.5 and 26.6 kg/m^2, respectively) were the lowest, followed by most recent BMI (27.8 kg/m^2), and maximum BMI (29.6 kg/m^2). Thus, people tended to have higher BMI at the most recent wave, consistent with the frequently observed patterns of BMI gain with age. It would be expected for mean BMI to be greater than median BMI because of the effect of outliers (typically higher values). However, the difference between mean and median BMI was minimal. The mean time people spent

Table 3
Coefficient estimates and 95% CIs of the standardized BMI exposure metrics in models of general health, chronic condition, and diabetes diagnosis in 2013 as well as AIC values for the models.

	General health (partial slope)	Chronic condition (log odds)	Diabetes diagnosis (log odds)
Most recent BMI			
Coefficient	**−0.276**	**0.185**	0.569
95% CI	**(−0.297, −0.254)**	**(0.120, 0.249)**	(0.437, 0.698)
AIC	**19236**	**5124.4**	1085.1
Maximum BMI			
Coefficient	−0.273	**0.185**	0.622
95% CI	(−0.294, −0.251)	**(0.119, 0.249)**	(0.490, 0.751)
AIC	19257	**5124.9**	1072.9
Mean BMI			
Coefficient	−0.263	**0.184**	**0.639**
95% CI	(−0.285, −0.241)	**(0.118, 0.248)**	**(0.508, 0.769)**
AIC	19299	**5125.2**	**1067.9**
Median BMI			
Coefficient	−0.255	**0.179**	**0.652**
95% CI	(−0.277, −0.233)	**(0.114, 0.244)**	**(0.520, 0.782)**
AIC	19334	**5126.6**	**1065.1**
Proportion of time with overweight			
Coefficient	−0.213	0.163	0.822
95% CI	(−0.236, −0.190)	(0.088, 0.240)	(0.568, 1.103)
AIC	19520	5136.7	1098.5
Proportion of time with obesity			
Coefficient	−0.258	0.163	0.720
95% CI	(−0.280, −0.236)	(0.095, 0.230)	(0.554, 0.891)
AIC	19327	5133.2	1074.7
Excess BMI-years with overweight			
Coefficient	−0.263	**0.172**	0.563
95% CI	(−0.284, −0.241)	**(0.110, 0.233)**	(0.448, 0.675)
AIC	19287	**5126.7**	1069.8
Excess BMI-years with obesity			
Coefficient	−0.232	0.150	0.442
95% CI	(−0.253, −0.210)	(0.091, 0.207)	(0.345, 0.535)
AIC	19401	5131.1	1083.8

AIC (Akaike information criterion), confidence interval (CI).
Lower AIC values represent stronger evidence for one model over another. Values in bold represent those within three of the minimum since there could be substantial support for such models (Burnham & Anderson, 2004).
Data: National Longitudinal Survey of Youth 1997.

Table 2
Correlations between BMI exposure metrics, adjusted by survey weights to be representative of the U.S. young adult population in 2002 – 2013 (n = 8,608).

	Most recent BMI	Maximum BMI	Mean BMI	Median BMI	Proportion of time with overweight	Proportion of time with obesity	Excess BMI-years with overweight	Excess BMI-years with obesity
Most recent BMI	1.000	0.917	0.928	0.902	0.742	0.814	0.892	0.798
Maximum BMI		1.000	0.952	0.926	0.755	0.838	0.923	0.832
Mean BMI			1.000	0.991	0.800	0.877	0.961	0.861
Median BMI				1.000	0.794	0.870	0.952	0.852
Proportion of time with overweight					1.000	0.664	0.668	0.442
Proportion of time with obesity						1.000	0.900	0.770
Excess BMI-years with overweight							1.000	0.946
Excess BMI-years with obesity								1.000

Data: National Longitudinal Survey of Youth 1997.

Table 4
Results from Vuong tests displaying significance from pairwise comparison of models for each health outcome on each BMI metric and other covariates.

A) General health

	Most recent BMI	Maximum BMI	Mean BMI	Median BMI	Proportion of time with overweight	Proportion of time with obesity	Excess BMI-years with overweight	Excess BMI-years with obesity
Most recent BMI		~ 0.197	>0.002	>0.000	>0.000	>0.002	>0.021	>0.000
Maximum BMI			>0.007	>0.000	>0.000	>0.007	~ 0.059	>0.000
Mean BMI				>0.000	>0.000	~ 0.103	~ 0.175	>0.000
Median BMI					>0.000	~ 0.387	<0.000	>0.002
Proportion of time with overweight						<0.000	<0.000	<0.002
Proportion of time with obesity							<0.036	>0.009
Excess BMI-years with overweight								>0.000
Excess BMI-years with obesity								

B) Chronic condition

	Most recent BMI	Maximum BMI	Mean BMI	Median BMI	Proportion of time with overweight	Proportion of time with obesity	Excess BMI-years with overweight	Excess BMI-years with obesity
Most recent BMI		~ 0.461	~ 0.436	~ 0.339	~ 0.067	~ 0.098	~ 0.324	~ 0.143
Maximum BMI			~ 0.467	~ 0.342	~ 0.061	~ 0.084	~ 0.322	~ 0.131
Mean BMI				~ 0.173	~ 0.051	~ 0.061	~ 0.253	~ 0.102
Median BMI					~ 0.073	~ 0.101	~ 0.481	~ 0.174
Proportion of time with overweight						~ 0.308	~ 0.114	~ 0.283
Proportion of time with obesity							~ 0.100	~ 0.379
Excess BMI-years with overweight								~ 0.069
Excess BMI-years with obesity								

C) Diabetes diagnosis

	Most recent BMI	Maximum BMI	Mean BMI	Median BMI	Proportion of time with overweight	Proportion of time with obesity	Excess BMI-years with overweight	Excess BMI-years with obesity
Most recent BMI		~ 0.054	<0.023	<0.021	~ 0.186	~ 0.217	<0.038	~ 0.443
Maximum BMI			~ 0.190	~ 0.130	>0.032	~ 0.435	~ 0.302	~ 0.059
Mean BMI				~ 0.122	>0.014	~ 0.249	~ 0.229	>0.002
Median BMI					>0.008	~ 0.167	~ 0.106	>0.002
Proportion of time with overweight						<0.022	<0.029	~ 0.185
Proportion of time with obesity							~ 0.321	~ 0.235
Excess BMI-years with overweight								>0.000
Excess BMI-years with obesity								

The vertical measures are in Model 1 and the horizontal measures are in Model 2.
p-values displayed for each pairwise test.
~ means it cannot be determined whether Model 1 fits better than Model 2 or Model 2 fits better than Model 1 at alpha = 0.05.
> means that the null hypothesis of equal fit should be rejected and Model 1 fits better than Model 2 at alpha = 0.05.
< means that the null hypothesis of equal fit should be rejected and Model 2 fits better than Model 1 at alpha = 0.05.
Data: National Longitudinal Survey of Youth 1997.

above the obesity threshold was almost a quarter of the duration of the study, and people accumulated a mean of more than 15 excess BMI-years with obesity in the 12 years from 2002 to 2013. That is, people had an average of $15.207/12 \approx 1.3$ excess BMI units above the obesity threshold per year. The corresponding numbers for overweight were more than half of the duration of the study and more than 36 excess BMI-years.

Table 2 displays the correlations between BMI exposure metrics. While the correlations were all positive and relatively high, there was variation in their values. The highest correlation was between mean BMI and median BMI at 0.991, and the lowest was between proportion of time with overweight and excess BMI-years with obesity at 0.442. Pairwise correlations between most recent, maximum, mean, and median BMI were all above 0.9. Proportion of time with overweight generally had lower correlations with the other metrics, between 0.442 and 0.877, perhaps because it captures different dimensions of adiposity than its counterparts, especially because it does not distinguish between earlier and later onset of overweight status among the young adults in the sample.

As a next step, we examined associations between the BMI metrics and health indicators. Table 3 presents the coefficient estimates (partial slopes for the general health models and log odds for the chronic condition and diabetes diagnosis models) of the standardized BMI exposure metrics from the corresponding models. Those with higher values of any BMI exposure metric rated their general health lower, had a higher likelihood of reporting a chronic condition, and had a higher likelihood of reporting a diabetes diagnosis. (See Appendix for full models.)

Using AIC as a means of goodness-of-fit comparison, we determined the "best" model from the candidate set of models. The model using most recent BMI had the lowest AIC for general health. The coefficient estimate for its standardized BMI metric was larger in magnitude than those of the other BMI metrics, suggesting that the other BMI metrics had weaker relationships with general health. For chronic condition, the models using most recent BMI, maximum BMI, mean BMI, median BMI, and excess BMI-years with overweight had similarly low AIC values. For diabetes diagnosis, the models using median and mean BMI had similarly low AIC values. Standardized most recent BMI had a coefficient of 0.569, an under-estimation of about 13% relative to standardized median BMI with a coefficient of 0.652. While there was not a consensus on which single metric performed the best across health indicators, in all of these models, proportion of time with overweight consistently performed the worst.

Vuong tests were then used to pairwise compare the models (Table 4), to determine whether one model was better than the other.

With general health as a dependent variable, the Vuong tests revealed that models with most recent BMI, maximum BMI, and excess BMI-years with overweight were best, while the model using proportion of time with overweight was consistently worse than models using the other BMI exposure metrics. With each model pair for chronic condition, neither registered as significantly better or worse. With diabetes diagnosis, the models using mean BMI, median BMI, and excess BMI-years with overweight seemed to provide better models, while proportion of time with overweight was the worst.

When we considered the AIC and Vuong test results together, most recent BMI was the best BMI metric in models of general health, though maximum BMI and excess BMI-years with overweight also performed similarly well. For chronic condition, the results were not so clear. In the case of diabetes diagnosis, the models that used median and mean BMI had the lowest AIC values, while the model that used most recent BMI had the second highest value. Furthermore, excess BMI-years with overweight was also found to have a better fit than most recent BMI in the pairwise Vuong test. Most recent BMI, while appropriate for certain health indicators, might not necessarily be the best metric with regard to model fit for others. This made the case for other BMI exposure metrics potentially being more useful in understanding the dynamics between body mass and health outcomes.

Discussion

All BMI-based metrics are proxies for adiposity. Where possible, more precise ways of measuring body fat would allow better measurement of obesity and assessment of its implications (Blundell, Dulloo, Salvador, & Fruhbeck, 2014; Cirulli et al., 2019; Ho et al., 2016; Muller et al., 2012). Nevertheless, the major advantage of BMI-based measures in epidemiologic studies is their easy and non-invasive collection. As a result, it is worthwhile to identify better ways to use BMI to understand adiposity and its relationship with health. To go beyond recent BMI and test whether its ubiquitous use is valid, we used the NLSY97 to study seven other metrics for young adults in the United States between 2002 and 2013: maximum BMI, mean BMI, median BMI, proportion of time with overweight/obesity, and excess BMI-years with overweight/-obesity. While related, these metrics depicted different aspects of exposure to and accumulation of adiposity.

Among the four point-estimate metrics, maximum BMI had the highest average value in the sample, followed by most recent, mean, and median BMI. Since BMI tends to increase with age and most people have difficulty losing weight (Gordon-Larsen et al., 2004, 2010), it is not surprising that the average of most recent BMI was higher than the averages of mean and median BMI. Interestingly, the averages of these four metrics were all above the overweight threshold of 25, indicating that the average person in the population would be classified as overweight, regardless of which of these four metrics was used. Metrics of proportion of time and excess BMI-years accounted for duration of elevated BMI, and the average person spent a quarter of early adulthood in obesity, and accumulated 15 excess BMI-years relative to the obesity threshold. The corresponding figures for overweight were half the duration and 37 excess BMI-years.

These BMI metrics were differently linked with each health indicator. Most recent BMI, the conventionally used metric, resulted in the best model fit for general health. Median BMI resulted in the best model fit for diabetes, while most recent BMI was the second least predictive metric, and the coefficient for standardized most recent BMI was 13% less than the coefficient for standardized median BMI. This points to potentially differential processes in the relationship between adiposity and various aspects of health. While mortality was not studied here, as the survey participants in the NLSY97 were still quite young, previous studies had demonstrated that maximum BMI performs better than BMI at survey time as a predictor of mortality (Stokes, 2014; Vierboom, 2017). Perhaps history is more closely associated with chronic conditions that develop over time, whereas recency is more closely associated with self-rated general health and well-being. In the case of diabetes diagnosis, it is also possible that people who were so diagnosed attempted to lose weight after the diagnosis, rendering most recent BMI not as indicative. Several metrics performed similarly in models for reporting a chronic condition. This could be because of the variety of chronic conditions that exist; for some, such as HIV/AIDS, elevated BMI might not be a risk factor.

Proportion of time with overweight performed the worst for all three health indicators. While this BMI metric does account for duration of time with elevated levels of BMI, it contains less information about the intensity, which many of the other measures do. Though the same could be said about proportion of time with obesity, perhaps the threshold for obesity is already high and detrimental enough and intensity does not need to be further quantified. It has been found in previous studies that obesity is more strongly linked with health than overweight (Flegal, Kit, Orpana, & Graubard, 2013). That being said, the metrics that incorporated a duration component (proportion of time and excess BMI-years) did not provide the best model fit for any particular health indicator. In the case of diabetes, both excess BMI-years metrics performed better than most recent BMI.

A limitation of the NLSY97 data, as well as many other datasets, is the reliance on survey participants' self-reported anthropometrics, which have been shown to be systematically biased, with such biases associated with demographic, social, and economic characteristics (Craig & Adams, 2009; Rowland, 1990). However, using these BMI exposure metrics in models of health indicators might not be a material issue, as previous studies have found that self-reported height and weight perform similarly to direct measurements as covariates in modeling health indicators (Ng, 2019; Preston, Fishman, & Stokes, 2015). It should be noted that chronic condition and diabetes diagnoses in this study were also self-reported, and so there could potentially be incorrect recall of diagnoses.

With regard to the health indicators, when respondents were asked the question on a chronic condition diagnosis, diabetes was suggested as an example of such a condition. Survey participants were also asked whether they had had a diabetes diagnosis. In other words, diabetes was directly asked and also indirectly suggested in another question. The NLSY97 does not allow for separating diabetes from the question on chronic conditions. Furthermore, Type I and Type II diabetes were not distinguished.

Biologically extreme values of height, weight, and BMI were treated as missing. Though the thresholds for physical plausibility could be debated, there were few instances of such extreme values, and sensitivity tests involving slight variations of the cut-off points did not impact substantive results. However, an anthropometric value could still be incorrect even if it does not appear extreme in an absolute sense. For example, even if a height seems reasonable, it would be unlikely that a person has grown several inches between years as an adult. Differences of height, weight, and BMI between each pair of consecutive waves were calculated, and the number of "large" fluctuations was minimal in the grand scheme of the dataset. For example, about 4% of inter-wave height differences were greater than one inch and about 0.3% of inter-wave weight differences were greater than 100 pounds.

Finally, we tested the addition of a square-term for the BMI exposure metrics in our models of health indicators, as it is possible that both high and low levels of BMI could be associated with poor health. Generally, these additional square-terms were not statistically significant. This could be due to low underweight prevalence and the unlikeliness of low BMI due to reverse causation, as the age range in this study was probably too young for the illness-induced weight loss often associated with impending mortality.

The main strengths of this study are two-fold. BMI at survey time is generally used as the main measure of adiposity in epidemiologic studies. Here, we examined the potential of using other BMI-based metrics that account for history, intensity, and duration of exposure to elevated BMI. In addition, with a longitudinal dataset, we were able to calculate such metrics without asking respondents to recall and without having to assume that height remains constant, drawbacks with using cross-sectional data. These BMI exposure metrics could then be used as simple inputs, as BMI at survey time would, in whatever models of interest, bypassing more complicated methodology that would be necessary for longitudinal or trajectory analyses.

A natural next step would be to use these BMI exposure metrics in other applications, with the NLSY97 or other longitudinal datasets. For example, it would be interesting to see how these metrics have changed across cohorts to determine whether there are shifting patterns in adiposity over time. Seeing how these metrics perform at other stages of the life course would likewise be worth studying. It could be that certain BMI exposure metrics are better explanatory variables, say, for older than for younger adults. Such BMI exposure metrics could be calculated with other longitudinal datasets as long as height and weight information was available at multiple time points. Ideally, the years between survey waves would be close to one another, as assumptions must be made on BMI between data collection time points, and the anthropometric measures would also be measured instead of self-reported.

Though all BMI-based metrics produced similar substantive conclusions regarding the significance and direction of association between body mass and health indicators, the choice of a "good" metric based on goodness-of-fit depended on the health indicator of interest, as different facets of body mass might relate differentially to different outcomes of health due to dissimilarities in disease processes. A natural question is whether the extra resources to collect and analyze longitudinal data are worthwhile. Analyses with these BMI exposure metrics could be useful and informative, but starting longitudinal data collection from scratch is expensive and time-consuming. Whether it is worth the effort depends on the research question at hand, as the health indicator of interest might affect this decision. With the exception of diabetes diagnosis, we found that most recent BMI actually does quite well relative to other more complicated measures. This is an important result, as it validates much of the past research using BMI at survey time. However, alternative BMI exposure metrics, especially those that account for both intensity and history, should still be considered as they might be better candidates for conditions that develop over time. Furthermore, beyond statistical criteria, some of these alternative metrics could be used to more directly study theoretical biosocial pathways by which the accumulation of adiposity affects people's health.

Ethics approval

This is an analysis of publicly-available data without individual identifiers.

Declaration of competing interest

There are no conflicts of interest.

Appendix. Results from models of health outcomes on each standardized BMI metric and other covariates

A) Partial slopes from linear regression of general health								
	Most recent BMI	Maximum BMI	Mean BMI	Median BMI	Proportion of time with overweight	Proportion of time with obesity	Excess BMI-years with overweight	Excess BMI-years with obesity
Intercept	−18.182	−17.423	−7.491	−8.074	−9.112	−11.807	−13.854	−23.597
Sex (reference = female)								
Male	0.116 ***	0.101 ***	0.123 ***	0.127 ***	0.170 ***	0.103 ***	0.096 ***	0.082 ***
Birth year	0.011	0.011	0.006	0.006	0.006	0.008	0.009	0.014
Race (reference = non-black/non-Hispanic)								
Black/non-Hispanic	−0.047	−0.040	−0.045	−0.050	−0.072 **	−0.061 *	−0.053 *	−0.073 **
Hispanic	−0.139 ***	−0.133 ***	−0.133 ***	−0.136 ***	−0.144 ***	−0.148 ***	−0.145 ***	−0.166 ***
Mixed/non-Hispanic	−0.067	−0.036	−0.031	−0.034	−0.036	−0.013	−0.041	−0.069
Insurance (reference = no)								
Yes	0.120 ***	0.105 ***	0.112 ***	0.112 ***	0.122 ***	0.110 ***	0.107 ***	0.106 ***
Standardized BMI metric	−0.276 ***	−0.273 ***	−0.263 ***	−0.255 ***	−0.213 ***	−0.258 ***	−0.263 ***	−0.232 ***
B) Log odds from logistic regression of chronic condition								
	Most recent BMI	Maximum BMI	Mean BMI	Median BMI	Proportion of time with overweight	Proportion of time with obesity	Excess BMI-years with overweight	Excess BMI-years with obesity
Intercept	48.314	47.011	39.086	39.380	39.153	43.028	44.041	50.298
Sex (reference = female)								
Male	−0.588 ***	−0.577 ***	−0.592 ***	−0.595 ***	−0.632 ***	−0.583 ***	−0.575 ***	−0.565 ***
Birth year	−0.025	−0.025	−0.021	−0.021	−0.021	−0.023	−0.023	−0.026
Race (reference = non-black/non-Hispanic)								
Black/non-Hispanic	−0.160	−0.166	−0.165	−0.161	−0.143	−0.144	−0.156	−0.141
Hispanic	0.006	0.001	−0.001	0.002	0.006	0.013	0.009	0.022
Mixed/non-Hispanic	0.770 *	0.746 *	0.741 *	0.744 *	0.744 *	0.733 *	0.750 *	0.767 *
Insurance (reference = no)								
Yes	0.286 **	0.296 ***	0.291 **	0.290 **	0.283 **	0.291 **	0.293 **	0.293 **
Standardized BMI metric	0.185 ***	0.185 ***	0.184 ***	0.179 ***	0.163 ***	0.163 ***	0.172 ***	0.150 ***
C) Log odds from logistic regression of diabetes diagnosis								
	Most recent BMI	Maximum BMI	Mean BMI	Median BMI	Proportion of time with overweight	Proportion of time with obesity	Excess BMI-years with overweight	Excess BMI-years with obesity
Intercept	101.544	99.629	54.774	51.121	41.015	45.976	67.659	90.181
Sex (reference = female)								
Male	−0.252	−0.208	−0.237	−0.243	−0.490 *	−0.279	−0.199	−0.186
Birth year	−0.054	−0.053	−0.030	−0.030	−0.023	−0.026	−0.037	−0.048
Race (reference = non-black/non-Hispanic)								
Black/non-Hispanic	0.031	−0.025	−0.051	−0.050	0.065	0.022	−0.031	0.028
Hispanic	0.365	0.328	0.315	0.326	0.338	0.336	0.338	0.385
Mixed/non-Hispanic	−13.274	−13.337	−13.387	−13.404	−13.359	−13.438	−13.372	−13.320
Insurance (reference = no)								
Yes	0.736 **	0.774 **	0.755 **	0.754 **	0.695 **	0.734 **	0.762 **	0.756 **
Standardized BMI metric	0.569 ***	0.622 ***	0.639 ***	0.652 ***	0.822 ***	0.720 ***	0.563 ***	0.442 ***

Significance is denoted by ***p < 0.001, **p < 0.01, and * p < 0.05.
There are potential issues of separation for the mixed/non-Hispanic group in models of diabetes diagnosis, but its inclusion should not affect the validity of the coefficient estimates of the other variables (Allison, 2008).
Data: National Longitudinal Survey of Youth 1997.

References

Abdullah, A., Wolfe, R., Stoelwinder, J. U., et al. (2011). The number of years lived with obesity and the risk of all-cause and cause-specific mortality. *International Journal of Epidemiology, 40*(4), 985–996.

Allison, P. (2008). *Convergence failures in logistic regression*. SAS Global Forum.

Apovian, C. M. (2013). The clinical and economic consequences of obesity. *American Journal of Managed Care, 19*(10 Suppl), s219–228.

Blundell, J. E., Dulloo, A. G., Salvador, J., & Fruhbeck, G. (2014). BMI ESWGo. Beyond BMI–phenotyping the obesities. *Obes Facts, 7*(5), 322–328.

Bureau of Labor Statistics. (2018). National longitudinal survey of youth 1997: Sample weights & design effects. https://www.nlsinfo.org/content/cohorts/nlsy97/using-and-understanding-the-data/sample-weights-design-effects.

Burnham, K. P., & Anderson, C. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(261), 261–304.

Cirulli, E. T., Guo, L., Leon Swisher, C., et al. (2019). Profound perturbation of the metabolome in obesity is associated with health risk. *Cell Metabolism, 29*(2), 488–500 e482.

Craig, B. M., & Adams, A. K. (2009). Accuracy of body mass index categories based on self-reported height and weight among women in the United States. *Maternal and Child Health Journal, 13*(4), 489–496.

Dahl, A. K., & Reynolds, C. A. (2013). Accuracy of recalled body weight-A study with 20-years of follow-up. *Obesity, 21*(6), 1293–1298.

Elliott, A. M., Aucott, L. S., Hannaford, P. C., & Smith, W. C. (2005). Weight change in adult life and health outcomes. *Obesity Research, 13*(10), 1784–1792.

Elo, I. T., Mehta, N., & Preston, S. (2017). The contribution of weight status to black-white differences in mortality. *Biodemography and Social Biology, 63*(3), 206–220.

Flegal, K. M., Kit, B. K., Orpana, H., & Graubard, B. I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories A systematic review and meta-analysis. *The Journal of the American Medical Association, 309*(1), 71–82.

Gordon-Larsen, P., Adair, L. S., Nelson, M. C., & Popkin, B. M. (2004). Five-year obesity incidence in the transition period between adolescence and adulthood: The national longitudinal study of adolescent health. *American Journal of Clinical Nutrition, 80*(3), 569–575.

Gordon-Larsen, P., The, N. S., & Adair, L. S. (2010). Longitudinal trends in obesity in the United States from adolescence to the third decade of life. *Obesity (Silver Spring), 18*(9), 1801–1804.

Gortmaker, S. L., Must, A., Perrin, J. M., Sobol, A. M., & Dietz, W. H. (1993). Social and economic consequences of overweight in adolescence and young adulthood. *New England Journal of Medicine, 329*(14), 1008–1012.

Guaraldi, G., Raggi, P., Gomes, A., et al. (2015). Lung and heart diseases are better predicted by pack-years than by smoking status or duration of smoking cessation in HIV patients. *PloS One, 10*(12), e0143700.

Hammond, R. A., & Levine, R. (2010). The economic impact of obesity in the United States. *Diabetes Metab Syndr Obes, 3*, 285–295.

Ho, J. E., Larson, M. G., Ghorbani, A., et al. (2016). Metabolomic profiles of body mass index in the framingham heart study reveal distinct cardiometabolic phenotypes. *PloS One, 11*(2), e0148361.

Hu, H. H., Nagahama, S., Nanri, A., et al. (2017). Duration and degree of weight change and risk of incident diabetes: Japan epidemiology collaboration on occupational health study. *Preventive Medicine, 96*, 118–123.

Khan, J. C., Thurlby, D. A., Shahid, H., et al. (2006). Smoking and age related macular degeneration: The number of pack years of cigarette smoking is a major determinant of risk for both geographic atrophy and choroidal neovascularisation. *British Journal of Ophthalmology, 90*(1), 75–80.

Kim, S. J., Jee, S. H., Nam, J. M., Cho, W. H., Kim, J. H., & Park, E. C. (2014). Do early onset and pack-years of smoking increase risk of type II diabetes? *BMC Public Health, 14*, 178.

Kyulo, N. L., Knutsen, S. F., Tonstad, S., Fraser, G. E., & Singh, P. N. (2012). Validation of recall of body weight over a 26-year period in cohort members of the Adventist Health Study 2. *Annals of Epidemiology, 22*(10), 744–746.

de Lauzon-Guillain, B., Balkau, B., Charles, M. A., Romieu, I., Boutron-Ruault, M. C., & Clavel-Chapelon, F. (2010). Birth weight, body silhouette over the life course, and incident diabetes in 91,453 middle-aged women from the French Etude Epidemiologique de Femmes de la Mutuelle Generale de l'Education Nationale (E3N) Cohort. *Diabetes Care, 33*(2), 298–303.

Lee, J. M., Gebremariam, A., Vijan, S., & Gurney, J. G. (2012). Excess body mass index-years, a measure of degree and duration of excess weight, and risk for incident diabetes. *Archives of Pediatrics and Adolescent Medicine, 166*(1), 42–48.

Lehnert, T., Sonntag, D., Konnopka, A., Riedel-Heller, S., & Konig, H. H. (2013). Economic costs of overweight and obesity. *Best Practice & Research Clinical Endocrinology & Metabolism, 27*(2), 105–115.

Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software, 9*(1), 1–19.

Mehta, N. K., Stenholm, S., Elo, I. T., Aromaa, A., Heliovaara, M., & Koskinen, S. (2014). Weight histories and mortality among Finnish adults the role of duration and peak body mass index. *Epidemiology, 25*(5), 707–710.

Moore, W., Pedlow, S., Krishnamurty, P., & Wolter, K. (2000). *National longitudinal survey of youth 1997 (NLSY97): Technical sampling report*. Chicago, IL.

Muller, M. J., Lagerpusch, M., Enderle, J., Schautz, B., Heller, M., & Bosy-Westphal, A. (2012). Beyond the body mass index: Tracking body composition in the pathogenesis of obesity and the metabolic syndrome. *Obesity Reviews, 13*(Suppl 2), 6–13.

Muscelli, E., Camastra, S., Gastaldelli, A., et al. (1998). Influence of duration of obesity on the insulin resistance of obese non-diabetic patients. *International Journal of Obesity and Related Metabolic Disorders, 22*(3), 262–267.

Must, A., Spadano, J., Coakley, E. H., Field, A. E., Colditz, G., & Dietz, W. H. (1999). The disease burden associated with overweight and obesity. *The Journal of the American Medical Association, 282*(16), 1523–1529.

Myrskyla, M., & Chang, V. W. (2009). Weight change, initial BMI, and mortality among middle- and older-aged adults. *Epidemiology, 20*(6), 840–848.

Ng, C. D. (2019). Biases in self-reported height and weight measurements and their effects on modeling health outcomes. *SSM - Population Health, 7*, 100405.

Noel, P. H., Copeland, L. A., Perrin, R. A., et al. (2010). VHA corporate data warehouse height and weight data: Opportunities and challenges for health services research. *Journal of Rehabilitation Research and Development, 47*(8), 739–750.

Ogden, C. L., Carroll, M. D., & Flegal, K. M. (2003). Epidemiologic trends in overweight and obesity. *Endocrinology and Metabolism Clinics of North America, 32*(4), 741–760 (vii).

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2014). Prevalence of childhood and adult obesity in the United States, 2011-2012. *Journal of the American Medical Association, 311*(8), 806–814.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ, 316*(7139), 1236–1238.

Pi-Sunyer, F. X. (2002). The obesity epidemic: Pathophysiology and consequences of obesity. *Obesity Research, 10*(Suppl 2), 97S–104S.

Preston, S. H., Fishman, E., & Stokes, A. (2015). Effects of categorization and self-report bias on estimates of the association between obesity and mortality. *Annals of Epidemiology, 25*(12), 907-911 e901-902.

R Core Team. (2018). *for statistical computing*. Vienna, Austria Version 3.5.1. [computer program].

Raghunathan, T., Solenberger, P., Berglund, P., & van Hoewyk, J. (2016). *IVEware: Imputation and variance estimation software*. Ann Arbor, Michigan: University of Michigan Version 0.3. .

Reis, J. P., Allen, N., Gunderson, E. P., et al. (2015). Excess body mass index- and waist circumference-years and incident cardiovascular disease: The CARDIA study. *Obesity, 23*(4), 879–885.

Rowland, M. L. (1990). Self-reported weight and height. *American Journal of Clinical Nutrition, 52*(6), 1125–1133.

Santamaria, F., Montella, S., Greco, L., et al. (2011). Obesity duration is associated to pulmonary function impairment in obese subjects. *Obesity (Silver Spring), 19*(8), 1623–1628.

Saquib, N., Stefanick, M. L., Natarajan, L., & Pierce, J. P. (2013). Mortality risk in former smokers with breast cancer: Pack-years vs. smoking status. *International Journal of Cancer, 133*(10), 2493–2497.

Sedgwick, P. (2014). Multiple hypothesis testing and Bonferroni's correction. *BMJ, 349*, g6284.

Stevens, J., Keil, J. E., Waid, L. R., & Gazes, P. C. (1990). Accuracy of current, 4-year, and 28-year self-reported body-weight in an elderly population. *American Journal of Epidemiology, 132*(6), 1156–1163.

Stokes, A. (2014). Using maximum weight to redefine body mass index categories in studies of the mortality risks of obesity. *Population Health Metrics, 12*.

Streiner, D. L., & Norman, G. R. (2011). Correction for multiple testing: Is there a resolution? *Chest, 140*(1), 16–18.

Tamakoshi, K., Yatsuya, H., Kondo, T., et al. (2003). The accuracy of long-term recall of past body weight in Japanese adult men. *International Journal of Obesity, 27*(2), 247–252.

Vierboom, Y. C. (2017). The contribution of differences in adiposity to educational disparities in mortality in the United States. *Demographic Research, 37*, 1735–1760.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica, 57*(2), 307–333.

Wyatt, S. B., Winters, K. P., & Dubbert, P. M. (2006). Overweight and obesity: Prevalence, consequences, and causes of a growing public health problem. *The American Journal of the Medical Sciences, 331*(4), 166–174.

Yarnell, J. W., Patterson, C. C., Thomas, H. F., & Sweetnam, P. M. (2000). 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. *Journal of Epidemiology & Community Health, 54*(5), 344–348.

Zheng, H., Tumin, D., & Qian, Z. (2013). Obesity and mortality risk: New findings from body mass index trajectories. *American Journal of Epidemiology, 178*(11), 1591–1599.