The timing of the rise in U.S. obesity varies with measure of fatness

Richard V. Burkhauser a,b, John Cawley a,b,*, Maximilian D. Schmeiser c

a Department of Policy Analysis and Management, Cornell University, MVR Hall, Ithaca, NY 14853, United States
b National Bureau of Economic Research, Cambridge, MA 02138, United States
c Department of Consumer Science, University of Wisconsin-Madison, Madison, WI 53706, United States

1. Introduction

According to the U.S. Surgeon General, obesity in the United States has risen to “nationwide epidemic proportions” (U.S. Department of Health and Human Services, 2001, p. v). This is troubling because obesity is associated with elevated risks of morbidity (U.S. Department of Health and Human Services, 2001) and mortality (Flegal et al., 2007), and therefore higher health care costs (Finkelstein et al., 2003; Monheit et al., 2009) and lower health-related quality of life (Sach et al., 2007). In addition, obesity is associated with adverse labor market outcomes such as lower wages (Cawley, 2004), job absenteeism (Cawley et al., 2007; Finkelstein et al., 2005) and not being employed (Burkhauser and Cawley, 2008). Obese individuals also may suffer social stigma (Puhl and Heuer, 2009).

Recently, economists have sought to determine the causes of the rise in obesity, studying in particular the impact of falling food prices (Lakdawalla and Philipson, 2002; Christian and Rashad, 2009; Rashad, 2006), increasing maternal employment (Anderson et al., 2003; Courtemanche, 2009), technological change (Lakdawalla et al., 2005; Cutler et al., 2003; Lakdawalla and Philipson, 2002; Philipson and Posner, 1999), rising incomes (Cawley et al., Forthcoming; Schmeiser, Forthcoming), changes in the rate of time discount (Komlos et al., 2004; Borghans and Golsteyn, 2006; Zhang and Rashad, 2008), density of, or proximity to, food stores or restaurants (Currie et al., 2009; Powell et al., 2007; Chou et al., 2004), and reduced smoking (Chou et al., 2004; Gruber and Frakes, 2006; Rashad, 2006; Baum, 2009; Courtemanche, Forthcoming). For example, Lakdawalla and Philipson (2002) estimate that 41–43% of the rise in body mass index (BMI) between 1981 and 1994 was due to the declining price of food, and Courtemanche (2009) estimates that changes in adult labor force participation are responsible for 1.4% of the rise in adult obesity between 1961 and 2004 and 10.4% of the rise in overweight children between 1968 and 2001.

To date these efforts to determine the causes of the recent rise in obesity have defined obesity using body mass index (BMI). This paper compares long-run trends in the prevalence of obesity when obesity is defined using skinfold thickness instead of BMI, using data from the full series of U.S. National Health Examination Surveys. The results indicate that when one uses skinfold thickness rather than BMI to define obesity, the rise in the prevalence of obesity is detectable 10–20 years earlier. This underscores the importance of examining multiple measures of fatness when monitoring or otherwise studying obesity.

* Corresponding author. Tel.: +1 607 255 0952; fax: +1 607 255 4071.
E-mail address: JHC38@cornell.edu (J. Cawley).
index, which is calculated as weight in kilograms divided by height in meters squared. This is due to the practical consideration that the data needed to compute BMI (weight and height) are widely available. When obesity is defined using BMI, its prevalence was relatively constant for cohorts measured from roughly 1960 to roughly 1980, after which it rose considerably (Flegal et al., 1998, 2002).¹

Aside from BMI, there is one additional measure of fatness that has been consistently collected in nationally representative health surveys: skinfold thickness. The purpose of this paper is to date the beginning of the rise in obesity if obesity is defined using skinfold thickness, and to determine whether the timing of that rise differs from when obesity is defined using BMI.

BMI and skinfold thickness are both widely accepted indices for measuring fatness and defining obesity (World Health Organization, 1995; Mei et al., 2002). For example, both are frequently used endpoints to evaluate anti-obesity interventions (Flegal et al., 2006; Doak et al., 2006). A comparison of the two measures concluded that BMI and skinfold thickness are “interchangeable for many epidemiologic research applications” (Must et al., 1991). A study assessing BMI as a predictor of body fat used skinfold thickness as the gold standard for subcutaneous fat (Mei et al., 2002). Studies have used both BMI and skinfold thickness to track short-term trends in obesity, although studies using skinfold thickness are far less common (Freedman et al., 1997; Kromeyer-Hauschild and Jaeger, 1998; Moreno et al., 2001; Dollman and Pilgrim, 2005), and no previous study has tracked the prevalence of obesity defined using skinfolds for the full series of surveys that we examine.

BMI and skinfold thickness each have their strengths and weaknesses as measures of fatness and obesity. BMI is widely used because it is easy to calculate, but it is recognized as a noisy measure of fatness because it does not distinguish fat from muscle, bone, and other lean body mass (Burkhauser and Cawley, 2008; Kragelund and Omland, 2005; Gallagher et al., 1996; Smalley et al., 1990; Garn et al., 1986). As a result, BMI overestimates fatness among those who are muscular (U.S. Department of Health and Human Services, 2001). Given these limitations of BMI, economic research is beginning to use more accurate measures of fatness (e.g. Johansson et al., 2009; Burkhauser and Cawley, 2008).

A strength of skinfold thickness is that it distinguishes fat from muscle. Because between 70 and 90% of all adipose tissue is subcutaneous, skinfold thicknesses accurately measure total body fat (Heysmfield et al., 2004). It is preferable that the skinfold thickness be measured at the triceps and subscapular sites because of accessibility, ease of measurement, and high correlation with total body fat (Heysmfield et al., 2004). A limitation is that triceps and subscapular skinfolds do not measure central adiposity, which has been shown to be associated with a greater risk of morbidity and mortality (National Institutes of Health, 1998). Skinfolds are less accurate in the very lean and very obese (Deurenberg and Deurenberg-Yap, 2004), but that is less of a problem for this analysis, which focuses on obesity status rather than variation in fatness within clinical weight classifications. Differences across surveys in the way skinfolds are measured may limit the accuracy of estimated trends (Flegal et al., 1990), although others have argued that there is no evidence of meaningful differences in measurement procedures across the surveys we examine (Gortmaker and Dietz, 1990). Another limitation is that, as we explain in a subsequent section, using skinfolds to define obesity for adults involves the use of two conversion equations (from skinfold thickness to body density, and body density to percent body fat); consequently, the body fat percentage calculated from skinfolds reflects the cumulative error from the application of two models. A limitation of both BMI and skinfolds is that they may ignore differences across ethnic groups in body fat or its distribution (Burkhauser and Cawley, 2008; Deurenberg and Deurenberg-Yap, 2004).

This paper does not take the position that skinfold thickness is superior to BMI as a measure of fatness. Rather, we argue that given the enormous attention devoted to trends in obesity based on BMI, the trends in obesity based on the only alternative measure of fatness that has been regularly collected – skinfold thickness – should be examined in order to ascertain whether new perspectives can be gained regarding the onset of the obesity epidemic.

2. Data and methods

2.1. Surveys

This study utilizes the series of nationally representative, cross-sectional health surveys sponsored by the National Center for Health Statistics, Centers for Disease Control and Prevention (CDC, 2009). Prior to 1999 four discrete surveys were conducted. The National Health Examination Survey, Cycle 1 (NHES I) was conducted on a sample of persons aged 18–79 years during 1959–1962. The NHES Cycle 3 (NHES III) was conducted on a sample of youths aged 12–17 years during 1966–1970. The National Health and Nutrition Examination Surveys (NHANES) program began with NHANES I, which was conducted over 1971–1975. It was followed by NHANES II conducted over 1976–1980 and then by NHANES III conducted over 1988–1994. It was not until 1999 that the modern series of NHANES Continuous cross-sectional surveys was launched with a new cross-sectional survey initiated every other year with data collection lasting two years. At the time of our report, four NHANES Continuous surveys were publicly available: NHANES 1999–2000, NHANES 2001–2002, NHANES 2003–2004, and NHANES 2005–2006 (National Center for Health Statistics, 1965, 1973, 1994, 2000; McDowell et al., 1981). In this paper we refer to each survey by its title, and when the title does not include the dates of the survey (e.g. NHANES I does not have the dates of the survey in its title but NHANES 2005–2006 does), we list the years of the survey in parentheses; e.g. NHANES I (1971–1975). In each of these surveys, a complex, stratified, multistage probability cluster sampling design

¹ For example, Flegal et al. (2002) write: “The prevalence of obesity was relatively constant from 1960 to 1980, then increased as reported by NHANES III in 1988–1994.” (page 1724).
is used to select individuals representative of the civilian non-institutionalized U.S. population.

For adults, there is variation in the upper end of the age range sampled in the various NHES and NHANES surveys. In NHES I and each NHANES survey, data are always collected for those up to age 74, so for the sake of consistency we limit each adult sample to those aged 20–74 years. Categories of race and ethnicity were not consistent across the surveys, so in our primary estimates we pool all race and ethnicity groups. However, we also present results for White and Black adults separately by gender. Final sample sizes for the 8 cross-sectional surveys, consistent across the surveys, so in our primary estimates we limit each adult sample to those aged 20–74 years, so for the sake of consistency we use the term obesity to refer to the maximum weight is clinically obese.

Adult obesity was defined in accordance with the National Institutes of Health and the World Health Organization as a BMI of 30 or higher (NHBLI Expert Panel 1998; World Health Organization Consultation on Obesity 2000). Youth obesity was defined using smoothed age- and gender-specific 85th percentile thresholds for BMI from the NHANES I distribution (Must et al., 1991).²

Skinfolds were assessed using calipers at the triceps and subscapular region (below the shoulder blade). In NHES and NHANES I and II, skinfold thicknesses were measured by Lange calipers manufactured by Cambridge Scientific Instruments that have a maximum value of 65 mm. For the examinations associated with NHANES III, a Holtain T/W caliper with a 50 mm maximum was used. For NHANES

2 We follow Gortmaker et al. (1987) and Must et al. (1991) in classifying youths as obese if their tricep skinfold exceeds the historic 85th percentile. For a time, the highest weight classification for youths was “overweight” (e.g. Kuczmarski et al., 2002) but in 2007 the American Academy of Pediatrics recommended that the term “obesity” be used even for youths (Barlow and the Expert Committee, 2007). For the sake of consistency across age groups we use the term obesity to refer to the highest weight classification for both youths and adults.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All males</td>
<td>12–17</td>
<td>Skinfold</td>
<td>(n = 6710)</td>
<td>(n = 2111)</td>
<td>(n = 1970)</td>
<td>(n = 2377)</td>
<td>(n = 1738)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>(10.1)</td>
<td>(14.5)</td>
<td>(15.2)</td>
<td>(20.8)</td>
<td>(27.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(14.6)</td>
<td>(18.4)</td>
<td>(16.6)</td>
<td>(26.9)</td>
<td>(30.0)</td>
</tr>
<tr>
<td>All males</td>
<td>12–17</td>
<td>Skinfold</td>
<td>(n = 3514)</td>
<td>(n = 1064)</td>
<td>(n = 1043)</td>
<td>(n = 1129)</td>
<td>(n = 889)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>(10.5)</td>
<td>(13.5)</td>
<td>(13.8)</td>
<td>(20.8)</td>
<td>(28.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(14.1)</td>
<td>(15.8)</td>
<td>(16.7)</td>
<td>(26.9)</td>
<td>(29.7)</td>
</tr>
<tr>
<td>White males</td>
<td>12–17</td>
<td>Skinfold</td>
<td>(n = 3024)</td>
<td>(n = 806)</td>
<td>(n = 856)</td>
<td>(n = 276)</td>
<td>(n = 184)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>(11.4)</td>
<td>(14.5)</td>
<td>(14.5)</td>
<td>(22.5)</td>
<td>(26.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(14.6)</td>
<td>(16.4)</td>
<td>(17.1)</td>
<td>(28.0)</td>
<td>(24.5)</td>
</tr>
<tr>
<td>Black males</td>
<td>12–17</td>
<td>Skinfold</td>
<td>(n = 471)</td>
<td>(n = 250)</td>
<td>(n = 158)</td>
<td>(n = 400)</td>
<td>(n = 259)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>(5.4)</td>
<td>(8.1)</td>
<td>(8.6)</td>
<td>(16.5)</td>
<td>(28.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(10.5)</td>
<td>(12.1)</td>
<td>(15.0)</td>
<td>(22.9)</td>
<td>(37.3)</td>
</tr>
<tr>
<td>All females</td>
<td>12–17</td>
<td>Skinfold</td>
<td>(n = 3196)</td>
<td>(n = 1047)</td>
<td>(n = 927)</td>
<td>(n = 1248)</td>
<td>(n = 849)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>(9.7)</td>
<td>(15.5)</td>
<td>(16.7)</td>
<td>(20.7)</td>
<td>(28.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(15.1)</td>
<td>(21.1)</td>
<td>(16.6)</td>
<td>(26.9)</td>
<td>(30.3)</td>
</tr>
<tr>
<td>White females</td>
<td>12–17</td>
<td>Skinfold</td>
<td>(n = 2668)</td>
<td>(n = 789)</td>
<td>(n = 749)</td>
<td>(n = 337)</td>
<td>(n = 173)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>(10.0)</td>
<td>(15.6)</td>
<td>(17.1)</td>
<td>(20.1)</td>
<td>(22.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(14.7)</td>
<td>(20.8)</td>
<td>(16.2)</td>
<td>(23.8)</td>
<td>(24.1)</td>
</tr>
<tr>
<td>Black females</td>
<td>12–17</td>
<td>Skinfold</td>
<td>(n = 513)</td>
<td>(n = 251)</td>
<td>(n = 162)</td>
<td>(n = 435)</td>
<td>(n = 235)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>(8.7)</td>
<td>(14.3)</td>
<td>(15.4)</td>
<td>(25.2)</td>
<td>(29.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(18.1)</td>
<td>(22.5)</td>
<td>(19.3)</td>
<td>(35.0)</td>
<td>(43.9)</td>
</tr>
</tbody>
</table>

² NHES indicates National Health Examination Survey; and NHANES, National Health and Nutrition Examination Survey.

b Estimated prevalences for ages 12–17 years combined races were age-standardized by the direct method to the 2000 Census population using each individual age from 12 to 17 years.
Continuous, the medical examinations used a Holtain-brand caliper with a 45 mm maximum. In general variability in compression can affect the reliability of skinfold measurements, but the Lange and Holtain calipers used in NHES and NHANES exert a constant pressure (Heymsfield et al., 2004). The NHES and NHANES medical technicians were trained in taking skinfold thicknesses to ensure accuracy and reliability (National Center for Health Statistics, 2000).

The NHANES III and NHANES Continuous noted when a skinfold exceeded the capacity of the calipers. The percentage of the examined adult sample with tricep skinfolds larger than the maximum caliper size was as follows: 1.9% in NHANES III and from 3.6% to 5.8% in each of the three surveys in NHANES Continuous. The percentage of the examined adult sample with subcapular skinfolds larger than the maximum caliper size was as follows: 2.4% in NHANES III, 4.5% in NHANES 1999–2000, 1.9% in NHANES 2001–2002, 1.6% in NHANES 2003–2004, and 2.1% in NHANES 2005–2006. We do not make use of the subcapular skinfolds for our analysis of youth obesity. For youths, the percentage of the examined sample with tricep skinfolds larger than the maximum caliper size was as follows: 1.4% in NHANES III, 1.5% in NHANES 1999–2000, 1.2% in NHANES 2001–2002, 1.4% in NHANES 2003–2004, and 0.9% in NHANES 2005–2006. For all respondents whose skinfold exceeds the maximum caliper size we recode their skinfold thickness to be equal to the maximum caliper size. This top-coding of skinfold thickness does not affect our estimates of the prevalence of obesity because such individuals are obese whether their skinfold is set equal to the maximum caliper size or an even larger number.

In addition to recording whether the skinfold exceeded the maximum caliper size, the NHANES Continuous indicated if the examiner could not obtain a measurement (presumably for reasons other than the skinfold exceeding the maximum caliper size). For adults, the percentage of the examined sample for whom triceps skinfold thickness could not be obtained was 2.8% in NHANES 1999–2000, 5.6% in NHANES 2001–2002, 7.1% in NHANES 2003–2004, and 10.2% in NHANES 2005–2006. The percentage of the examined adult sample for which subscapular skinfold thickness could not be obtained was 15.3% in NHANES 1999–2000, 21.4% in NHANES 2001–2002, 20.8% in NHANES 2003–2004, and 27.0% in NHANES 2005–2006. For youths, the percentage of the examined sample for whom triceps skinfold thickness could not be obtained was 3.2% in NHANES III, 1.2% in NHANES 1999–2000, 1.9% in NHANES 2001–2002, 3.2% in NHANES 2003–2004, and 3.9% in NHANES 2005–2006. When the skinfold could not be obtained, we impute it using the other skinfold thickness, measured BMI, age, age squared, and race and ethnicity; this prediction equation is based on respondents with complete information. These regression models explain 91.2–96.4% of the variance in skinfolds, implying that the imputation procedure accurately predicts the missing skinfolds.

The definition of obesity based on skinfold thickness differs for youth and adults. Following the literature, we define youth obesity using tricep (but not subscapular) skinfold thickness; specifically, a youth is classified as obese if the tricep skinfold thickness exceeds the smoothed age- and gender-specific 85th percentile threshold from the NHANES I distribution (Gortmaker and Dietz, 1990; Must et al., 1991).

Adults were classified as obese or non-obese using the following steps. First, body density was predicted using tricep and subscapular skinfold thicknesses (Durnin and Womersley, 1974). Second, percent body fat was computed using body density (Durnin and Womersley, 1974; Siri, 1956). (The Durnin and Womersley (1974) conversion equations, which account for nonlinearities, are among the most widely used for this purpose (Heymsfield et al., 2004). Use of an alternate equation for converting body density to percent body fat (Brozek et al., 1963) yielded similar trends.) Finally, men were classified as obese if their percent body fat exceeded 25%, and women were classified as obese if their percent body fat exceeded 30% (National Institute of Diabetes and Digestive and Kidney Diseases, Accessed 2007).

2.3. Methods

Statistical analyses were carried out using SAS for Windows software (SAS for Windows, Version 9.1.3 Service Pack 3). All analyses excluded pregnant females. The prevalence of obesity was calculated for each age group and in each survey. To avoid having changes in the age distribution of the population affect trends in obesity, the obesity rates in each survey are age- and sex-adjusted based on population counts in the 2000 Census. Specifically, estimates were age-standardized to Census 2000 counts of age groups 12–17 years, 20–39 years, 40–59 years, and 60–74 years. (We use the age group 12–17 years because that was the age range of NHES III, and we use the adult age groups that are the convention in recent studies of the prevalence of obesity (e.g. Flegal et al., 2002). Adhering to previously used age ranges facilitates com-

Notes:
2) Obesity rates are age- and sex-adjusted based on population counts in the 2000 Census.

Fig. 1. Trend in youth obesity measured using skinfold thickness and body mass index. NHANES III (1966–1970) to NHANES 2005–2006.
Table 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All males and females</td>
<td>12–17</td>
<td>Skinfold</td>
<td>5.1 (3.3–6.8)</td>
<td>P &lt; .001</td>
<td></td>
<td>11.0 (8.3–13.7)</td>
<td>P &lt; .001</td>
<td></td>
<td>16.1 (13.8–18.4)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>2.1 (0.2–3.9)</td>
<td>0.01</td>
<td></td>
<td>15.6 (12.8–18.3)</td>
<td>P &lt; .001</td>
<td></td>
<td>17.7 (15.2–20.2)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td>All males</td>
<td>12–17</td>
<td>Skinfold</td>
<td>3.3 (1.0–5.6)</td>
<td>P &lt; .001</td>
<td></td>
<td>15.8 (12.0–19.6)</td>
<td>P &lt; .001</td>
<td></td>
<td>19.1 (15.8–22.5)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>2.6 (0.1–5.1)</td>
<td>0.02</td>
<td></td>
<td>15.1 (11.1–19.0)</td>
<td>P &lt; .001</td>
<td></td>
<td>17.7 (14.2–21.1)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td>White males</td>
<td>12–17</td>
<td>Skinfold</td>
<td>3.2 (0.5–5.8)</td>
<td>0.01</td>
<td></td>
<td>14.6 (7.9–21.2)</td>
<td>P &lt; .001</td>
<td></td>
<td>17.7 (11.4–24.0)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>2.4 (–0.4–5.2)</td>
<td>0.05</td>
<td></td>
<td>11.5 (4.8–18.2)</td>
<td>P &lt; .001</td>
<td></td>
<td>13.9 (7.6–20.2)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td>Black males</td>
<td>12–17</td>
<td>Skinfold</td>
<td>3.3 (–1.6–8.1)</td>
<td>0.09</td>
<td></td>
<td>12.3 (5.7–18.9)</td>
<td>P &lt; .001</td>
<td></td>
<td>15.6 (10.2–20.9)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>4.6 (–1.7–10.8)</td>
<td>0.08</td>
<td></td>
<td>14.2 (6.3–22.1)</td>
<td>P &lt; .001</td>
<td></td>
<td>18.8 (12.6–25.0)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td>All females</td>
<td>12–17</td>
<td>Skinfold</td>
<td>7.0 (4.4–9.6)</td>
<td>P &lt; .001</td>
<td></td>
<td>5.9 (2.1–9.8)</td>
<td>P &lt; .001</td>
<td></td>
<td>12.9 (9.7–16.1)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>1.5 (–1.2–4.2)</td>
<td>0.14</td>
<td></td>
<td>16.2 (12.1–20.4)</td>
<td>P &lt; .001</td>
<td></td>
<td>17.7 (14.2–21.3)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td>White females</td>
<td>12–17</td>
<td>Skinfold</td>
<td>7.2 (4.3–10.1)</td>
<td>P &lt; .001</td>
<td></td>
<td>3.5 (–2.8–9.8)</td>
<td>0.14</td>
<td></td>
<td>10.7 (4.9–16.5)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>1.5 (–1.5–4.4)</td>
<td>0.16</td>
<td></td>
<td>15.3 (8.2–22.3)</td>
<td>P &lt; .001</td>
<td></td>
<td>16.7 (10.1–23.4)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td>Black females</td>
<td>12–17</td>
<td>Skinfold</td>
<td>6.7 (0.6–12.8)</td>
<td>0.02</td>
<td></td>
<td>13.8 (5.7–21.8)</td>
<td>P &lt; .001</td>
<td></td>
<td>20.5 (14.2–26.7)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>1.1 (–5.8–8.1)</td>
<td>0.37</td>
<td></td>
<td>20.9 (12.2–29.6)</td>
<td>P &lt; .001</td>
<td></td>
<td>22.1 (15.0–29.1)</td>
<td>P &lt; .001</td>
<td></td>
</tr>
</tbody>
</table>

a NHES indicates National Health Examination Survey; NHANES, National Health and Nutrition Examination Survey; and CI, confidence interval.
b Estimated prevalences for ages 12–17 years combined races were age-standardized by the direct method to the 2000 Census population using each individual age from 12 to 17 years.

Comparisons with the previous literature but has the unfortunate consequence of omitting youths aged 18–19. For all surveys, estimates are based on the sampling weights associated with the sample that underwent medical examinations. Estimates also take into account the complex survey design, including unequal probabilities of selection, non-response, and deliberate over-sampling of certain groups. Tests of the hypothesis of equality of obesity prevalence over time were conducted, using a significance level of .05.

3. Results

3.1. Youths aged 12–17 years

Fig. 1 presents trends in the prevalence of obesity between NHES III (1966–1970) and NHANES 2005–2006 defined using BMI and skinfold thickness for youths aged 12–17. Because each survey involved data collection over several years, the prevalence of obesity in each survey is plotted at the mid-point year of each survey. The specific estimates by survey, sex, and definition of obesity are provided in Table 1.


Earlier studies that examined changes in the prevalence of youth obesity defined using BMI concluded that obesity was relatively constant from NHES III (1966–1970) to NHANES II (1976–1980) (Harlan et al., 1988). Limiting our attention to this range, Fig. 1 confirms that obesity defined using BMI experienced little net increase between NHES III (1966–1970) and NHANES II (1976–1980) because it first rose between NHES III (1966–1970) and NHANES I (1971–1975) and then fell between NHANES I (1971–1975) and NHANES II (1976–1980). In contrast, obesity defined using skinfolds rose monotonically from the first to the second to the third survey. Table 1 provides the exact statistics; for both genders pooled BMI-defined obesity first rose from 14.6% to 18.4%, then fell to 16.6%, while skinfold-defined obesity rose from 10.1 to 14.5% to 15.2%.

Table 2 reports the changes in prevalence from survey to survey in percentage points and also indicates whether the change was statistically significant. Table 2 confirms that the 5.1 percentage point change between NHES III (1966–1970) and NHANES II (1976–1980) in skinfold-defined obesity was statistically significant (P < .001), and that the overall change of 2.1 percentage points over the same period in BMI-defined obesity was also statistically significant at the 1% level (P = 0.01). When the changes are examined separately by sex, the point estimates of the change from NHES III (1966–1970) to NHANES II (1976–1980) remain positive for both measures of obesity, but they are only statistically significant for obesity defined using skinfolds.

3 The standard errors do not take into account error introduced by estimating percent body fat using the two-step process of Durnin and Womersley (1974), which can predict body fat with errors of 3.5–5.0% and a 95% confidence interval between plus or minus 7–10% (Heymsfield et al., 2004).
Table 1 also presents the prevalence of obesity for youths disaggregated by race and gender, with the trends depicted in Fig. 2. Among male youth (both whites and blacks) obesity measured either by skinfold or BMI rose between NHES III (1966–1970) and NHANES II (1976–1980) with no decline in obesity between NHANES I (1971–1975) and NHANES II (1976–1980) using either measure. Only among female youths (both whites and blacks) is there a rise in skinfold-defined obesity but a drop in BMI-defined obesity from NHANES I (1971–1975) to NHANES II (1976–1980).

Turning to the latest five surveys, Fig. 1 confirms that the trend in both measures of obesity is generally (though not monotonically) upward. Table 2 reports the change from NHANES II (1976–1980) to NHANES 2005–2006. An 11 percentage point rise in skinfold-defined obesity and a 15.6 percentage point rise in BMI-defined obesity are both statistically significant at the 1% level.

The last columns of Table 2 indicate that the long-term trend from NHES III (1966–1970) to NHANES 2005–2006 is a rise of 16.1 percentage points in skinfold-defined obesity and a rise of 17.7 percentage points in BMI-defined obesity. Both values are statistically significant (P < .001).

3.2. Adults aged 20–74 years

The prevalence of adult obesity defined using BMI and skinfold thickness are presented in Tables 3A and 3B. Table 4 documents the percentage-point changes in obesity between NHES I (1959–1962) and NHANES II (1976–1980), NHANES II (1976–1980) and NHANES 2005–2006, and the overall change between NHES I (1959–1962) and NHANES 2005–2006; it also lists P values associated with a test of the hypothesis that each trend is equal to zero.

Earlier studies that examined changes in the prevalence of adult obesity defined using BMI have concluded that obesity was relatively constant from NHES I (1959–1962) to NHANES II (1976–1980) (Flegal et al., 1998, 2002). The first rows of Tables 3A and 4 confirm that finding; obesity defined using BMI rose from 14.4% to 14.7% between NHES
I (1959–1962) and NHANES II (1976–1980) \((P = .34)\) before rising to 31.8% in NHANES 2005–2006 \((P < .001)\). However, when skinfold thicknesses are used to define obesity, the rise in obesity is apparent beginning with the earliest surveys, which for NHANES II (1976–1980) \((P < .001)\) and continues to 43.4% in NHANES 2005–2006 \((P < .001)\). Fig. 4 illustrates that the rise in obesity defined using skinfold thickness began before the rise in obesity defined using BMI.

Table 3A also shows the prevalence of obesity by subgroups: by gender and by gender and race. For each gender, with all ages 20–74 pooled, the rise in skinfold-defined obesity is apparent between NHES I (1959–1962) and NHANES II (1976–1980), while the rise in BMI-defined obesity is only apparent between NHANES II (1976–1980) and NHANES III (1988–1994).

Fig. 4 also indicates that, for females, the prevalence of obesity defined by skinfold thickness is much higher than the prevalence of obesity defined by BMI. This is because a greater percentage of fat is carried subcutaneously (as opposed to viscerally) in women than in men \((Enzi et al., 1986)\).

Considering age groups by gender in Table 3B, the same pattern is present for men 40–59 and 60–74 and women 20–39 and 40–59: the rise in skinfold-defined obesity is apparent between NHES I (1959–1962) and NHANES II (1976–1980), while the rise in BMI-defined obesity is only apparent between NHANES II (1976–1980) and NHANES III (1988–1994). Among men 20–39 and women 60–74 there is little increase in skinfold-defined obesity and a decline in BMI-defined obesity between NHES I (1959–1962) and NHANES II (1976–1980), although the long-term trend shows an increase in each measure of obesity between NHES I (1959–1962) and NHANES 2005–2006.

4. Discussion

Previous research that used BMI to define obesity dated the beginning of the rise in U.S. obesity, among both youths and adults, from NHANES II (1976–1980) to NHANES III (1988–1994). This article finds that, when skinfold thickness is used to define obesity, the rise in obesity is apparent starting from the earliest surveys, which for

### Table 3A

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All males</td>
<td>20–74</td>
<td>Skinfold</td>
<td>26.1</td>
<td>(n = 6126)</td>
<td>(n = 15904)</td>
<td>(n = 11777)</td>
<td>(n = 14041)</td>
<td>(n = 3528)</td>
<td>(n = 3823)</td>
<td>(n = 3557)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>14.4</td>
<td>14.4</td>
<td>14.7</td>
<td>22.7</td>
<td>29.8</td>
<td>29.6</td>
<td>31.0</td>
<td>31.8</td>
</tr>
<tr>
<td>All males</td>
<td>20–74</td>
<td>Skinfold</td>
<td>13.5</td>
<td>(n = 2895)</td>
<td>(n = 6309)</td>
<td>(n = 5613)</td>
<td>(n = 6614)</td>
<td>(n = 1747)</td>
<td>(n = 1938)</td>
<td>(n = 1822)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>11.7</td>
<td>12.5</td>
<td>12.4</td>
<td>19.9</td>
<td>25.0</td>
<td>25.1</td>
<td>27.8</td>
<td>27.8</td>
</tr>
<tr>
<td>White males</td>
<td>20–74</td>
<td>Skinfold</td>
<td>13.9</td>
<td>(n = 2502)</td>
<td>(n = 5325)</td>
<td>(n = 4891)</td>
<td>(n = 2430)</td>
<td>(n = 759)</td>
<td>(n = 963)</td>
<td>(n = 922)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>11.5</td>
<td>12.0</td>
<td>11.9</td>
<td>19.5</td>
<td>26.7</td>
<td>26.6</td>
<td>29.3</td>
<td>31.3</td>
</tr>
<tr>
<td>Black males</td>
<td>20–74</td>
<td>Skinfold</td>
<td>12.1</td>
<td>(n = 332)</td>
<td>(n = 903)</td>
<td>(n = 608)</td>
<td>(n = 1872)</td>
<td>(n = 337)</td>
<td>(n = 399)</td>
<td>(n = 381)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>14.9</td>
<td>17.6</td>
<td>14.6</td>
<td>20.2</td>
<td>26.0</td>
<td>23.6</td>
<td>31.2</td>
<td>33.7</td>
</tr>
<tr>
<td>All females</td>
<td>20–74</td>
<td>Skinfold</td>
<td>39.1</td>
<td>(n = 3231)</td>
<td>(n = 9595)</td>
<td>(n = 6164)</td>
<td>(n = 7151)</td>
<td>(n = 1781)</td>
<td>(n = 1885)</td>
<td>(n = 1735)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>17.0</td>
<td>16.2</td>
<td>16.9</td>
<td>25.4</td>
<td>33.0</td>
<td>32.4</td>
<td>32.6</td>
<td>32.4</td>
</tr>
<tr>
<td>White females</td>
<td>20–74</td>
<td>Skinfold</td>
<td>39.1</td>
<td>(n = 2766)</td>
<td>(n = 7895)</td>
<td>(n = 5331)</td>
<td>(n = 2725)</td>
<td>(n = 707)</td>
<td>(n = 926)</td>
<td>(n = 881)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>15.9</td>
<td>15.5</td>
<td>15.2</td>
<td>22.9</td>
<td>30.2</td>
<td>31.0</td>
<td>30.6</td>
<td>31.7</td>
</tr>
<tr>
<td>Black females</td>
<td>20–74</td>
<td>Skinfold</td>
<td>45.5</td>
<td>(n = 415)</td>
<td>(n = 1594)</td>
<td>(n = 722)</td>
<td>(n = 2161)</td>
<td>(n = 377)</td>
<td>(n = 390)</td>
<td>(n = 381)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMI</td>
<td>25.2</td>
<td>28.7</td>
<td>29.5</td>
<td>36.4</td>
<td>48.7</td>
<td>45.4</td>
<td>52.0</td>
<td>48.2</td>
</tr>
</tbody>
</table>

\* NHES indicates National Health Examination Survey; and NHANES, National Health and Nutrition Examination Survey.

\* Estimated prevalences for ages 20–74 years combined races were age-standardized by the direct method to the 2000 Census population using age groups 20–39, 40–59, and 60–74 years.
adults is NHES I (1959–1962) and for youths is NHES III (1966–1970). This implies that the rise in obesity may be due to more gradual and longer-run influences than was previously appreciated. Avenues for future research include determining whether previously studied factors such as falling food prices and changes in labor force participation explain a different percentage of the rise in obesity when it is defined using skinfold thickness instead of BMI. Future research should also explore whether the earlier rise of obesity defined using skinfold thickness suggests additional early or long-run factors that may be important in explaining the increase in obesity. The relative importance of previously studied factors may change, and previously unappreciated factors may prove to be predictive, when one uses skinfold thickness rather than BMI to define obesity.

Our findings also relate to a 20-year-old controversy over the use of skinfold thicknesses and BMI to define adolescent obesity. A study published in 1987 (Gortmaker et al., 1987) found a rise between NHES III (1966–1970) and NHANES II (1976–1980) in the prevalence of adolescent obesity defined using skinfold thickness. A study published the next year (Harlan et al., 1988) opposed the use of skinfolds to measure trends, arguing: “The main point at issue is how to interpret the observation that, for adolescents, body mass index does not show a trend across surveys but skinfold measurements do show a trend. One possibility is that from 1967 to 1980 the adolescent population got fatter at a constant body mass index . . . Another possibility is that the small differences in skinfold measurements across surveys are an artifact of measurement error. Neither of these possibilities can be ruled out with the available data.” (Flegal et al., 1990).

This paper, which uses 25 years of data from five surveys that were not available during the original exchange, sheds light on the main point of that debate (as described by Flegal et al., 1990). First, adolescent obesity defined using BMI was rising during at least part of the period (see Table 1). For adolescent females, obesity defined using BMI rose from 15.1% to 21.1% between NHES III (1966–1970) and NHANES I (1971–1975) before falling from 21.1% to 16.6% between NHANES I (1971–1975) and NHANES II (1976–1980). For adolescent males, obesity

---

4 Gortmaker et al. (1987) and Harlan et al. (1988) define youth obesity using the 85th percentile values in the NHES I (1959–1962). For the sake of consistency between skinfolds and BMI in the percentile threshold for obesity and for comparisons with the previous literature such as Harlan et al. (1988), we classify youths as obese if their BMI exceeds the historic 85th percentile, although the American Academy of Pediatrics considers that to be the threshold for overweight and the historic 95th percentile to be the threshold for obesity (Barlow and the Expert Committee, 2007).
defined using BMI rose slightly but steadily from 14.1% to 15.8% to 16.7% between NHES III (1966–1970) and NHANES I (1971–1975) and NHANES II (1976–1980).

Second, even though female adolescent obesity defined by BMI fell from NHANES I (1971–1975) to NHANES II (1976–1980), it still rose on net from NHES III (1966–1970) to NHANES II (1976–1980) for males and females combined (rise of 2.1 percentage points) and for males (2.6 percentage points). The point estimate of the rise for females is also positive (1.5 percentage points) but is not statistically significant (P = 0.14).

Third, a re-examination of the BMI data in NHANES using cohort analysis suggests that BMI began rising earlier than previously appreciated ([K]omlos et al., 2009). Finally, the long-term trend in adolescent obesity defined using BMI was clearly upward, as shown in Fig. 1.

It is possible for obesity to be rising by one measure but not by another because each index of fatness measures something different and the threshold for obesity is drawn at different points in their distributions ([N]HLB]I Expert Panel 1998; [W]orld Health Organization, 2000; [N]ational Institute of Diabetes and Digestive and Kidney Diseases, 2000).

![Image](282x77 to 503x272)

**Fig. 3.** Trend in adult obesity measured using skinfold thickness and body mass index. NHES I (1959–1962) to NHANES 2005–2006.
For example, more adults are classified as obese by skinfold thickness than by BMI (i.e. the threshold is lower for skinfold thickness than for BMI), so it is possible that between NHES III (1966–1970) and NHANES II (1976–1980) that adults in the healthy weight and overweight categories of BMI gained enough fat to make them obese by skinfold thickness but not enough fat to make them obese by BMI. In other words, the rise in obesity defined using skinfolds reflected increases in fatness among adults who remained pre-obese according to BMI. Consistent with this hypothesis, we calculate that the prevalence of skinfold-defined obesity among adults rose 20.6% among those with BMI less than 30 and rose only 9.7% among those with BMI greater than or equal to 30, between NHANES 1999-2000, NHANES 2001-02, NHANES 2003-04, NHANES 2005-06. Further evidence that fatness measures can move somewhat independently of each other is that waist circumference rose faster between 1959 and 2004 than one would expect given changes in BMI over the same period (Elobeid et al., 2007).

Our findings are also consistent with a recent re-examination of the NHES-NHANES data on BMI for youths that studies birth-year cohorts rather than measurement years; this study concludes that “it appears highly unlikely that the obesity pandemic appeared suddenly in the 1980s among American children as conventional analysis would suggest … but has rather manifested itself slowly and persistently for an extended period of time beginning at least … in the 1950s, but possibly earlier” (Komlos et al., 2009, p. 158). Komlos et al. (2009) point to many technological changes that could have contributed to this earlier increase: the spread of automobiles, radios, and televisions per capita, an increase from 24% in 1950 to 45% in 1995 in the share of food expenditures spent on food prepared away from home, and a doubling of the per capita number of fast food restaurants between 1972 and 1997. Cutler et al. (2003) note that the share of people who are employed in physically demanding jobs such as farm workers and laborers fell 19 percentage points (68–49%).
between 1910 and 1970, and only 3 more percentage points (from 45% to 42%) between 1980 and 1990. Cutler et al. (2003) also note that time spent watching television increased by 40 min between 1965 and 1975, which is double the increase between 1975 and 1995. Philipson and Posner (1999) describe the technological changes behind the long-run growth in weight and obesity, including decreases in the real price of food and decreases in the caloric expenditure involved in both market and household production. All of these changes are candidates to explain the rise in skinfold-defined obesity between the 1960s and 1980.

We also examine how the correlation between skinfold thickness and BMI changed over this period. For youths aged 12–17 years, the correlation between tricep skinfold thickness and BMI rose from 0.64 to 0.80 for boys, and from 0.76 to 0.85 for girls, between NHES III (1966–1970) and NHANES 2005–2006. For adults aged 20–74 years, the correlation percent body fat (calculated using tricep and subscapular skinfold thicknesses) and BMI rose from 0.73 to 0.78 for men and from 0.76 to 0.77 for women, between NHES I (1959–1962) and NHANES 2005–2006. An increasing correlation between skinfolds and BMI is to be expected if the population is getting fatter. (At the extreme, if everyone in society had zero fat, BMI would be uncorrelated with skinfold thickness—everyone would have minimal skinfold thickness but BMI would vary across people because of differences in muscle and other lean mass. As the population gets fatter, both skinfold thickness and BMI rise, increasing the correlation between the two.) Although the trends in obesity based on BMI and skinfold thickness diverge between NHANES I (1971–1975) and NHANES II (1976–1980) (especially for adolescents, but also for adults), there is negligible (1 percentage point or less) change in the correlation between skinfold thickness and BMI between those two surveys.

In 2001, the U.S. Surgeon General issued a “call to action” on obesity (U.S. Department of Health and Human Services, 2001) and the World Health Organization has stated that the need for public health action is “urgent” (World Health Organization, 2004). In recent years, states have increasingly taken legislative action to prevent or reduce obesity (Cawley and Liu, 2008). If more previous research had used skinfolds to monitor obesity, the rise in obesity might have been detected a decade or two sooner. If that had happened, public health and public policy responses might have been implemented, and progress in preventing obesity could have begun, earlier.

There are multiple valid measures of fatness and yet they can yield different trends in obesity. This suggests that one should monitor multiple measures of fatness to increase the probability of detecting trends early. Even more generally, because each measure of fatness has its own strengths and weaknesses, research on obesity can be enriched by greater consideration of alternative measures of fatness and body composition.

Acknowledgment

We thank John Komlos for helpful comments and suggestions.

References

Courtemanche, C., Forthcoming. Rising cigarette prices and rising obesity: coincidence or unintended consequence? J. Health Econ.
Elobeid, M.A., et al., 2007. Waist circumference values are increasing beyond those expected from bmi increases. Obesity 15 (10), 2380–2383.


Schmeiser, M.D., Forthcoming. Expanding wallets and waistlines: the impact of family income on the bmi of women and men eligible for the earned income tax credit. Health Econ.


