The Effectiveness of Screening for Obesity in Primary Care: Weighing the Evidence

Amy R. Wilson
Donna D. McAlpine
University of Minnesota

In response to growing alarm about the increase in the prevalence of obesity in the United States, several organizations have recommended that physicians screen their adult patients for this condition and initiate treatment. Screening can be an effective intervention when the condition is grave and prevalent, when an accurate test exists, when effective treatment exists, when the screening program itself does not pose undue risks, and when early detection and treatment improve outcomes. This article critically reviews the evidence supporting these criteria in the case of obesity in adults. It extends previous reviews by assessing the potential impact that uncertainties in the evidence base may have on the effectiveness of a screening program. It also examines the feasibility of such a program. We conclude that following the recommendation to screen all adults for obesity is unlikely to improve outcomes.

Keywords: obesity; overweight; mass screening; body mass index

Obesity increasingly is viewed as a significant health issue. Thus, it is not surprising that medical professionals have been asked to take greater responsibility in helping their patients lose weight. Numerous leading medical organizations, including the United States Preventive Services Task Force (USPSTF), have recommended that identification of obese patients through more intensive screening efforts should be implemented in clinical practice (Lyznicki et al. 2001; National Heart, Lung, and Blood Institute Obesity Education Initiative 1998; Nawaz and Katz 2001; USPSTF 2003).
NEW CONTRIBUTION

Implementing mass screening for obesity using a body mass index (BMI)–based criterion and the subsequent intensive treatment for every person meeting that criterion would constitute significant change in current practice and incur significant cost. While many organizations have issued obesity-screening guidelines promoting increased detection and treatment of obesity, no evaluations of such interventions have been identified. In the absence of even a single such evaluation, it is critical to examine the obesity-research literature specifically in the context of screening. This article seeks to evaluate the potential for obesity screening in an adult population by applying the criteria critical to sound screening policies. Increased detection is necessary but not sufficient for a screening program to improve outcomes. Program effectiveness requires that the consequences of the condition in question be significantly grave, that detection be followed by effective treatment and improved health outcomes, and that the program be feasible. This article extends previous reviews by outlining the uncertainties in the evidence base and how those uncertainties may translate to uncertain outcomes in a screening program. The article also addresses the feasibility of implementing a mass screening program for obesity. We focus in more depth than previous reviews on the treatment of obesity, looking at the evidence in each of five areas: Can weight loss be achieved? Do outcomes improve when weight loss is achieved? Can weight loss be maintained? What are the health consequences of failing to maintain a weight loss? Is treatment available? The central question of whether one could expect a measurable reduction in the health consequences of obesity if physicians adopted current screening recommendations is addressed. We examine the potential for obesity screening both to improve outcomes and to leave them unchanged or worse.

CONCEPTUAL FRAMEWORK

As shown in Figure 1, the rationale for screening for any condition is simple. For obesity, screening consists of the “conscious measurement of weight status” (McTigue et al. 2003, 933) and is expected to lead to increased detection. Detection is assumed to lead to increased treatment, which in turn improves health outcomes including morbidity, mortality, and quality of life.

While the rationale is simple, assessing whether a program will improve outcomes in practice is complex. The criteria for an effective screening program

This article, submitted to Medical Care Research and Review on July 29, 2004, was revised and accepted for publication on December 31, 2005.
have been well established (Morrison 1992; Russell 1994) and include six factors:
(1) the prevalence of the condition, (2) its effect on morbidity and mortality,
(3) availability of a screening test that accurately distinguishes between those
with and without the condition, (4) effects of the screening test, (5) availabil-
ity of effective treatment, and (6) increased effectiveness of early treatment.

Test accuracy and prevalence determine the percentage of the population
that will be referred for treatment, and thus, must result in a reliable distinc-
tion between those needing treatment and those not. Prevalence itself is
important because screening for a rare condition will not constitute a good
use of resources. The effects on morbidity and mortality of the screening test
itself and any follow-up diagnostic tests must be insignificant. The effects of
the condition on morbidity and mortality must be great enough to merit con-
sideration and to warrant the negative effects of the treatment. The treatment
must be available and effective, and the effects of treating in the case of a
false positive must be considered. Finally, early detection and treatment of
the condition must offer advantages over usual care.

We can make judgments about the potential for screening to improve out-
comes by weighing the evidence surrounding these factors. This article seeks
to do so by reviewing the literature in each of these areas. Moreover, we
argue that even if obesity screening were effective, there are issues around
the feasibility of such a program that must be addressed.

**METHODOLOGY**

There are excellent reviews of the literature in some of the areas needed
to address the utility of screening programs for obesity. The USPSTF has
completed two systematic reviews in areas specifically relevant to screening
(USPSTF 1996, 2003), and other reviews are available from leading health
care organizations such as the National Heart, Lung, and Blood Institute
(NHLBI; National Heart, Lung, and Blood Institute Obesity Education
Initiative 1998). It is not our intent to replicate these reviews. Instead, this
article seeks to weigh the evidence presented in these reviews and in more
recent literature to highlight where the evidence either supports or is
contrary to the criteria for effective screening programs. Where contrary or inconclusive evidence exists, we look more carefully at the impact of that uncertainty on the efficacy of a screening program. We also extend the USPSTF reviews by focusing in more detail on treatment and by examining the feasibility of implementing screening in physicians’ offices.

We updated the literature review provided by USPSTF by searching MEDLINE and the Cochrane and Health Technology Assessment databases from 2001 to 2005 for studies on humans, written in English, with the keywords obesity or overweight and appropriate terms for each subarea (prevalence, mortality, weight loss, etc). Wherever possible, we relied on systematic reviews, identifying fourteen published since the USPSTF review (Anderson, Luan, and Hoie 2004; Arterburn, Crane, and Veenstra 2004; Aucott et al. 2004; Avenell et al. 2004; Buchwald et al. 2004; Carmichael and Bates 2004; Chapman et al. 2004; Clegg et al. 2003; Colquitt et al. 2003; Harvey et al. 2002; O’Meara et al. 2002, 2004; Padwal, Li, and Lau 2003b; Poobalan et al. 2004). The quality of the articles was judged consistent with USPSTF guidelines.

MEASURING OBESITY

Definitions of overweight or obesity are somewhat arbitrary. Obesity is defined as excess body fat or adiposity, but what constitutes excess is not clear except to the degree that it is sufficient to elevate risk for morbidity and mortality. Common definitions of obesity (and overweight) based on different measures are described in Table 1. Of these, the most widely used is that proposed by the World Health Organization (WHO) and is based on the BMI, which is defined as \[
\text{BMI} = \frac{\text{weight (kg)}}{\text{height (m)}^2}
\] (World Health Organization 2000). Overweight has been defined as \(25 \leq \text{BMI} < 30\) and obesity as \(\text{BMI} \geq 30\). This definition is stricter than its predecessor, and on its adoption by the National Institutes of Health (NIH) in 1998, a large number of Americans suddenly met the criterion for obesity.

BMI is based solely on height and weight. Other measures, such as waist circumference or waist-to-hip ratio, account for the distribution of body fat. These measures may provide useful information given the potential significance of that distribution. Two large prospective studies suggest that body fat accumulated in the abdomen is associated with increased morbidity and mortality (Folsom et al. 2000; Seidell et al. 1994). In addition to the obesity measures described in Table 1, laboratory-based methods (e.g., densitometry, hydrometry, and dual energy X-ray absorptiometry) directly measure body-fat levels. While highly accurate, these methods require costly equipment, and thus, they are unlikely to be appropriate for mass screening.
All of the leading organizations that suggest physicians screen for obesity recommend the BMI as the primary instrument. BMI is simple to obtain, requiring only height and weight measurement and a simple computation. Because of this method’s popularity, the balance of this article investigates the potential impact of an intervention that uses BMI as the screening instrument.

### CURRENT DETECTION AND TREATMENT OF OBESITY IN PRIMARY CARE

The first assumption behind a screening program is that it will increase detection of hidden disease. If a condition already is detected adequately in usual care, there is no need to develop a screening tool. The suggestion that obesity is underdetected in clinical settings may come from studies that ask community respondents how often their physicians have counseled them to lose weight. For example, using data from the Behavioral Risk Factor Surveillance System (BRFSS), Galuska and colleagues (1999) found that only 42 percent of obese persons who had visited a doctor in the past twelve months reported being advised to lose weight. An analysis of data from the National Ambulatory Care Survey in the mid-1990s estimated that physicians identify obesity in only 38 percent of their obese patients (Stafford et al. 2000). However, as evidence of underdetection, such studies should be

### Table 1  Common Measures of Obesity

<table>
<thead>
<tr>
<th>Obesity Measure</th>
<th>Criteria for Obese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body Mass Index (BMI)</td>
<td></td>
</tr>
<tr>
<td>Weight (kg) / [height (m)]²</td>
<td>Overweight: 25.0–29.9</td>
</tr>
<tr>
<td></td>
<td>Obesity: 20–34.9 (Class I)</td>
</tr>
<tr>
<td></td>
<td>35–39.9 (Class II)</td>
</tr>
<tr>
<td></td>
<td>Extreme Obesity: ≥ 40 (Class III)</td>
</tr>
<tr>
<td>Waist Circumference (cm)</td>
<td>Men: &gt; 102</td>
</tr>
<tr>
<td></td>
<td>Women: &gt; 88</td>
</tr>
<tr>
<td>Waist-to-hip ratio</td>
<td>≥ 1.0</td>
</tr>
<tr>
<td>Conicity index (CI)</td>
<td>Not established</td>
</tr>
<tr>
<td><em>Waist(m)</em></td>
<td></td>
</tr>
<tr>
<td><em>0.109 √weight(kg)/height(m)</em></td>
<td></td>
</tr>
<tr>
<td>Neck circumference (cm)</td>
<td>Men: ≥ 37</td>
</tr>
<tr>
<td></td>
<td>Women: ≥ 34</td>
</tr>
</tbody>
</table>

All of the leading organizations that suggest physicians screen for obesity recommend the BMI as the primary instrument. BMI is simple to obtain, requiring only height and weight measurement and a simple computation. Because of this method’s popularity, the balance of this article investigates the potential impact of an intervention that uses BMI as the screening instrument.
treated cautiously given that surveys of physicians suggest that the failure to
counsel patients may better reflect physicians’ beliefs about the efficacy of
counseling than their detection of obesity (Kushner 1995; Yeager et al. 1996).
With respect to patients’ detection of their own obesity, Paeratakul and col-
leagues (2002), using data from two national surveys, found that 87 percent
of persons who met the BMI threshold for obesity also self-identified as
overweight.
Very few studies directly have assessed how often physicians correctly
identify obesity in usual practice in the absence of a formal screening pro-
gram. A study set in a primary care setting, however, estimated that physi-
cians detected 75 percent of cases of overweight in their patients, suggesting
that the detection rate for obesity would be even higher (Caccamese,
Kolodner, and Wright 2002).
These data suggest that improved detection by physicians may be
needed, thus warranting a screening program for obesity. However, it
remains uncertain that such a program would be effective. The remainder of
this article reviews the evidence supporting the six criteria for an effective
screening program and discusses the feasibility of obesity screening.

PREVALENCE OF OBESITY

Using the BMI criterion, there is no question that overweight and obesity
are prevalent. According to data from the National Health and Nutrition
Examination Survey (NHANES), in 2001–2002, approximately 36 percent
of the adult population met the criterion for obesity, and another 30 percent was
overweight (Hedley et al. 2004). The prevalence of obesity also is increasing.
A study based on the BRFSS showed a 50-percent increase in obesity preva-
lence from 1991 to 1998 (Mokdad et al. 1999). These estimates are based on
self-reported values for height and weight, and thus, they may underestimate
or overestimate the prevalence of obesity, depending on age and gender
(Engstrom et al. 2003; Kuczmarski, Kuczmarski, and Najjar 2001).
All else being equal, mass screening is less effective and does not constitute
a wise use of resources if the condition is so rare as to result in the identifica-
tion of very few cases. Clearly, this is not the situation with obesity. Instead,
the high prevalence of obesity lends support to a screening program.

ACCURACY OF THE SCREENING TEST

An accurate screening test discriminates between individuals who do and
do not have the condition in question. Accuracy is a function of test technol-
ogy and disease prevalence and usually is measured by sensitivity and
specificity. If one accepts the current BMI-based definition of obesity, given accurate measures of height and weight, BMI constitutes a perfect measure. There exists uncertainty, however, surrounding the ability of BMI to measure excess fat and to identify obesity in a diverse population.

Current guidelines recommend applying the BMI criterion uniformly to the adult population. In other words, every adult with a BMI exceeding 30 is obese; every obese person should lose weight. The evidence, however, does not support this recommendation. While the BMI is correlated highly with body fat in adults (USPSTF 1996), a meta-analysis of 32 studies suggests the strength of the correlation varies substantially with ethnicity (Deurenberg, Yap, and van Staveren 1998). One study compared procedures for assessing obesity in a sample of obese white and African American women. This research found that while body-fat percentages (as measured by dual energy X-ray absorptiometry) did not vary in the two groups, BMI in the African American women exceeded that in the white women (Brandon, Elliott-Loyd, and Calloway 2002). In other words, a BMI-based classification of obesity applied in these two groups would classify more African American women than white women as obese, even if they had the same body-fat percentage.

A potential explanation for variation in BMI performance across ethnicities is that BMI does not reflect the distribution of body fat—in particular, abdominal fat. Several prospective studies have found that abdominal fat varies significantly by ethnicity (Casas et al. 2001; Hill et al. 1999; Lovejoy et al. 1996; Park et al. 2001). To complicate the issue further, both a large retrospective study based on NHANES III and a smaller prospective study suggest that the health risks of abdominal obesity vary by ethnicity (Lear et al. 2002; Okosun et al. 2001).

The NHLBI guidelines attempt to address the importance of abdominal fat by recommending measurement of waist circumference in addition to BMI. An evaluation of that algorithm in an ethnically diverse population from the NHANES III, however, found that waist circumference provided no independent information when making treatment recommendations (Kiernan and Winkleby 2000).

Finally, BMI cannot account for fitness levels. A study of cardiorespiratory fitness (as measured by maximal oxygen uptake) and body fat (as measured by skin folds) in more than seven thousand persons showed that persons with high cardiorespiratory fitness levels have less fat, both overall and abdominal, than persons with low levels, controlling for BMI (Ross and Katzmarzyk 2003). Thus, variation in fitness levels may mean that BMI is not a good predictor of overall or abdominal fat.

The limitations of BMI as a measure of obesity have important implications for a screening program. If we rely on the BMI as a screening tool,
many persons will be misclassified as obese and subsequently either be advised to or attempt to lose weight. This is both a waste of resources and may (as will be discussed later) have negative consequences to the degree that weight cycling is associated with poor health outcomes.

**EFFECTS OF OBESITY ON MORTALITY AND MORBIDITY**

Obesity is correlated with increased risk for coronary heart disease, hypertension and stroke, type 2 diabetes, sleep apnea, gall bladder disease, colon cancer, and certain musculoskeletal disorders (Field et al. 2001; Giovannucci 2003; Pi-Sunyer 1999, 2002). In women, obesity has been associated with an increase in incidence of polycystic ovary syndrome and of breast and endometrial cancer (Carmichael and Bates 2004; Hu 2003). There exists controversy about the shape of the relationship between BMI and mortality. Several large studies and meta-analyses have suggested a U- or J-shaped relationship (Calle et al. 1999; Katzmarzyk, Craig, and Bouchard 2001; Singh, Lindsted, and Fraser 1999; Troiano et al. 1996; Visscher et al. 2000). Others have argued that the increased risks purportedly associated with excess leanness should be discounted because these studies did not control properly for smoking status and other illness (Manson et al. 1987; Manson et al. 1995; Seidell, Visscher, and Hoogeveen 1999). An analysis based on the Nurses’ Health Study, with more than 115,000 women, that attempted to control for these factors found no excess mortality in lean women (Manson et al. 1995). A study of more than six thousand obese individuals suggests that the increased risk associated with gross obesity is not as great as previously thought (Bender et al. 1998). Other researchers have questioned the use of BMI as a measure of obesity in mortality studies, citing its nonspecificity, and have proposed an alternative model of the relationship between body composition and mortality that would help to explain U- and J-shaped BMI-mortality curves (Allison et al. 1997). Research based on NHANES 1 examined the BMI level associated with minimum all-cause mortality (adjusting for age and smoking status) and found that (1) the minimizing value of BMI varied by ethnicity and gender and (2) for each group, a wide range of BMIs was associated with minimum mortality (Durazo-Arvizu et al. 1998).

Uncertainty about the true relationship between BMI and mortality was heightened after the recent publication of research based on NHANES data that estimated the association between weight and mortality in the United States. The authors concluded that obesity accounted for about 111,909 excess deaths in 2000, a figure much lower than prior estimates (Flegal et al. 2005). Moreover, the researchers also found that overweight (BMI 25 to 30)
was associated with fewer excess deaths than was normal weight, a finding contrary to much of the current literature. The proper interpretation of these data remains the subject of significant debate (Ding 2005; Gibbs 2005; Greenberg 2005; Strickler et al. 2005; Willett et al. 2005) in both the scientific and popular press; nonetheless, the vigor of the debate points out that scientists have yet to reach consensus about the nature of the basic relationship between mortality and body weight.

It also appears that the threshold of a BMI of 30 does not predict increased mortality equally well across ethnicities. One prospective study of more than one million African American and white adults in the United States explored the relationship between BMI and all-cause mortality in adults who never had smoked and had no history of disease (Calle et al. 1999). The study found that the relative risk of all-cause mortality indeed did increase with BMI, but it increased less significantly in the African American population than in the white. A study based on multiple national data sets of years of life lost attributable to obesity found that for all age groups, African American men did not experience a reduction in life expectancy for BMIs less than 32, nor did African American women for BMIs less than 37 (Fontaine et al. 2003). By contrast, reductions were estimated for young white men at BMIs below the threshold of 30. For young women, mortality reductions started at a BMI of 25, and for young white women, at a BMI of 27. One study used data from two national studies to explicitly address the question of the choice of a BMI cutoff value for white and African American women. This analysis looked at outcomes including hypertension, hypertriglyceridemia, and mortality for different levels of BMI in white and African American women (Stevens, Juhaeri, et al. 2002). BMI was not associated significantly with mortality for African American women.

There is some evidence suggesting that obesity and overweight have a lesser effect on mortality in older persons. A review of thirteen studies of the relationship between BMI and mortality in individuals sixty-five years or older found that, with the exception of one study, the BMI associated with minimal risk increased with age and ranged from 27 to 30 (Heiat, Vaccarino, and Krumholz 2001). What is unclear from these studies is whether obesity poses a lesser risk as a person ages or whether BMI does not predict body-fat percentage equally well across different age ranges. Some research suggests the latter and that the accuracy of BMI in predicting fat percentage depends on age and gender (Deurenberg, Weststrate, and Seidell 1991). These uncertainties clearly have implications for the guidance that should be given to patients of various ages and ethnicities regarding their weight-maintenance or weight-loss goals.

There are also several areas in which obesity diminishes quality of life, including stigma and decreased mobility and functioning (Fontaine and
Barofsky 2001; Hassan et al. 2003). The stigma associated with being overweight is manifested through negative comments from others, job discrimination, and even physical attacks (Myers and Rosen 1999). As this research points out, stigma stems from the judgment of both an obese person’s appearance and perceived character. Obese persons also may face stigma in the healthcare setting (Hebl, Xu, and Mason 2003), and biases against obese persons clearly are present even among obesity-treatment specialists (Schwartz et al. 2003; Teachman and Brownell 2001).

Finally, there exists uncertainty as to whether obesity, as measured by BMI, constitutes a sufficient screening criterion. As Flegal and colleagues (2005) acknowledge, it is possible that other factors account for observed relationships between obesity and mortality. For example, BMI does not measure another important predictor of increased morbidity and mortality: fitness. Data from the Lipid Research Clinics study found that BMI was correlated only modestly with fitness (measured with a standardized treadmill test). The correlation was –.10 in men and –.21 in women (Stevens, Cai, et al. 2002). Several large studies have suggested that, controlling for BMI, cardiorespiratory fitness is associated with mortality (Farrell et al. 2002; Lee, Jackson, and Blair 1998; Wei et al. 1999). A systematic review of 24 studies concluded that fitness protected against all-cause mortality, cardiovascular disease (CVD) mortality, coronary heart disease (CHD), hypertension, and type 2 diabetes mellitus (Blair and Brodney 1999). The CARDIA (Coronary Artery Risk Development in Young Adults) study, which involved a longitudinal cohort of young adults, found that low cardiorespiratory fitness (controlling for baseline BMI) increased the risk of developing hypertension, diabetes, and metabolic syndrome (Carnethon et al. 2003). Men who are obese and fit face lower all-cause and CVD mortality risk than men who are normal weight and unfit, although fitness does not completely negate the ill effects of obesity (Stevens, Cai, et al. 2002). Two recent studies reached different conclusions about the predictive power of BMI versus fitness with respect to important health outcomes (Weinstein et al. 2004; Wessel et al. 2004). The inconclusiveness of the studies reviewed in this section suggests that research has not yet established whether obesity is the best predictor of poor health outcomes, and thus, the appropriate candidate for a screening criterion and intervention.

There is no question that the evidence relating obesity, as measured by BMI, to adverse outcomes is substantial. The evidence related to obesity and age, ethnicity, and fitness, however, implies that we do not understand fully exactly how obesity contributes to health. This uncertainty suggests that we cannot yet know what recommendations we should make to a population that varies by age, fitness, and ethnicity. Thus, screening persons, identifying
those with a BMI $\geq 30$ as obese, and suggesting they lose weight may not improve outcomes or constitute a judicious use of health care resources.

**EFFECTS OF SCREENING TESTS**

Any screening program has the potential for negative effects. Screening tests may be uncomfortable or invasive, as may be the follow-up tests or treatment. There is always the risk of unnecessary treatment. The potential exists for all of these factors to come into play in the context of obesity screening. Being weighed and having one’s weight measured may not be as physically uncomfortable as other commonly performed screening tests. Patients who already feel stigmatized by their weight, however, may face further stigmatization when faced with a diagnosis of obesity. We already have discussed the potential for the BMI to identify persons who may not need to lose weight as obese, thus raising the potential for unnecessary treatment. Attempting to lose weight is a burden psychologically and economically, and as with any other intervention, should not be undertaken unnecessarily. Given these uncertainties and those outlined below with respect to the effectiveness of treatment, it is unclear whether the potential benefits of the screening test outweigh the potential harms.

**EFFECTS AND EFFECTIVENESS OF TREATMENT**

Existing guidelines call for clinicians to promote sustained weight loss in their obese adult patients. The USPSTF guidelines state that “high-intensity counseling—about diet, exercise, or both—together with behavioral interventions aimed at skill development, motivation, and support strategies produces modest, sustained weight loss (typically 3 to 5 kg for $\geq 1$ year)” (USPSTF 2003, 930). The NIH guidelines suggest an initial goal of a loss of 10 percent of body weight in six months (National Heart, Lung, and Blood Institute Obesity Education Initiative 1998). These guidelines recommend dietary and behavior therapy combined with physical activity and that pharmacotherapy be considered if these methods have not produced results in a six-month period. When these interventions have not succeeded, weight-loss surgery (WLS) is to be considered for those patients with BMI $\geq 40$ or those patients with BMI $\geq 35$ with comorbid conditions.

Assessing the effectiveness of obesity treatment is not straightforward. We must consider five questions: Does treatment cause weight loss? Does weight loss improve outcomes? Can weight loss be maintained? What are the health implications of failing to maintain a loss? Is treatment available?
 DOES TREATMENT CAUSE WEIGHT LOSS?

The myriad weight-loss approaches fall into three categories: lifestyle change, medication, and surgery. While a detailed review of the effectiveness of these options is beyond the scope of this article, we can outline some important findings from the literature.

Lifestyle-change approaches to weight loss generally include dietary and/or physical activity components. The numerous options in this vein range from casual and self-directed approaches to commercial programs to clinical programs. There is continuing debate in the popular and scientific press about the merits of particular diets. Official recommendations about physical-activity levels also are revisited frequently. The plethora of options may reflect scientific uncertainty about what works best. It is also possible, however, that a wide range of approaches is necessary in a population of persons with diverse needs and behaviors. This diversity has not been represented well in the research literature. The bulk of research on the effectiveness of weight-loss treatment has been conducted in very homogeneous populations (primarily white women).

Pharmaceutical therapy for weight loss is generally reserved for those with a substantial amount of weight to lose. The Cochrane Collaboration found that of eight medications reviewed, the evidence for only two met the criteria for inclusion in their systematic review: sibutramine and orlistat (Padwal, Li, and Lau 2003a). Orlistat inhibits fat absorption and sibutramine depresses appetite. The majority of studies of these medications have focused on short-term outcomes related to weight loss and have not included longer term outcomes such as obesity-related morbidity and mortality (Arterburn, Crane, and Veenstra 2004; Padwal, Li, and Lau 2003a). Weight losses range from 2 to 4 kg (Haddock et al. 2002; Padwal, Li, and Lau 2003a). While generally effective, the drugs can have significant negative side effects such as insomnia, gastrointestinal problems, hypertension, dizziness, and more rarely, both higher blood pressure and heart rate (Buchwald et al. 2004; Davidson et al. 1999; Monteforte and Turkelson 2000; O’Meara et al. 2004).

WLS is reserved for the most severely obese persons. Procedures fall into two types: gastric restrictive and malabsorptive. Meta-analyses have estimated that morbidly obese persons undergoing WLS will lose, on average, 60 percent of their excess weight (Buchwald et al. 2004), or 23 to 28 kg more (at a two-year follow-up) than persons using conventional methods (Colquitt et al. 2003). Systematic reviews have concluded that weight-loss surgery is a cost-effective intervention (Avenell et al. 2004; Clegg et al. 2003). As with any surgical intervention, however, there exists significant potential for negative consequences, including perioperative and postoperative morbidity and
mortality. Finally, there exists uncertainty as to which procedure is best. A systematic review comparing procedures concluded that no single strategy was both the safest and the most effective (Chapman et al. 2004).

In sum, research definitely supports the idea that persons can lose weight using a variety of approaches that range in invasiveness from minimal to maximal. With the exception of surgery, however, most weight loss is modest, and thus, even the most thorough and accurate screening program may have only a limited impact on outcomes.

**DOES WEIGHT LOSS IMPROVE OUTCOMES?**

The evidence addressing the question of whether persons who lose weight have improved outcomes is more conclusive for short-term than for long-term benefits. Short-term benefits include lower blood pressure, improved lipid profiles, and improved insulin action (Anderson and Konz 2001). Obese persons who lose weight, however, face a transitory increased risk of gallstone formation (Erlinger 2000; Everhart 1993). A meta-analysis of twenty-five randomized controlled trials concluded that long-term weight loss is associated with reduced risk for hypertension (Neter et al. 2003). A systematic review of the research published from 1966 to 2000 that followed subjects for at least five years also concluded that weight loss may be associated with improved lipid profiles (Poobalan et al. 2004).

The evidence suggests that long-term weight loss does appear to improve some health outcomes, although there exist numerous methodological problems with these studies, making it necessary to be cautious in interpreting results (Aucott et al. 2004; Avenell et al. 2004). There also remains substantial controversy about the link between weight loss and mortality. A review of early studies reporting decreased mortality with weight loss questioned their conclusions (Williamson and Pamuk 1993). Reviews of studies reporting increased mortality with weight loss conclude that each study failed to account for at least one important factor: intentionality of weight loss, baseline health of subjects, alcohol or tobacco use, or pregnancy (Andres, Muller, and Sorkin 1993; Lee and Paffenbarger 1996). An analysis based on a national data set of overweight women who never had smoked and who intended to lose weight found that changes in mortality depended on whether the women had obesity-related illness (Williamson et al. 1995). In healthy women, the authors found little difference in all-cause, CVD, and diabetes-related mortality between women who lost weight and those who did not. For women with obesity-related disease, mortality rates decreased for women who lost weight. This decrease was attributed primarily to obesity-related cancers.
Others have found that weight loss actually may be associated with increased mortality. For example, using data from the Finnish Twin Cohort, Sorensen and colleagues (2005) examine mortality both for healthy persons who intended to lose weight and for healthy persons who did not. Participants who either lost or gained weight had higher mortality than those with stable weight. While this study has been criticized for methodological weaknesses such as not controlling for past smoking behavior or adequately measuring intent to lose weight (Stampfer 2005), the findings do point out the significant uncertainty that remains about weight loss and mortality outcomes.

Recent systematic reviews cite lack of standard methodology and conclude that we still do not have convincing evidence about the effects of weight loss, and significant disagreement still exists among obesity researchers about the effects of intentional weight loss on mortality (Sorensen 2003; Yang et al. 2003). A possible explanation for the inconclusiveness of this evidence involves fitness. Few studies of weight loss assess whether it was accompanied by an increase in fitness. (In fact, few studies assess the method of weight loss.) It is possible to lose weight without becoming more fit. If, as discussed earlier, fitness is the fundamental risk, weight loss without increase in fitness is less likely to improve outcomes.

CAN WEIGHT LOSS BE MAINTAINED?

A proposed definition of long-term weight-loss success is “intentionally losing at least 10% of initial body weight and keeping it off for at least one year” (Wing and Hill 2001, 323). It is commonly believed that the majority of persons who lose weight will not be able to maintain that loss, but the research evidence is mixed. A study of an intensive university-based weight-loss program found that between 60 percent and 80 percent of participants had maintained a loss of at least 10 percent of their initial weight at the end of one year, and approximately 25 percent had done so at the end of seven years (Anderson et al. 1999). A meta-analysis of longitudinal studies of twenty-nine structured weight-loss programs estimated that at a five-year follow-up, participants maintained a loss of 3.2 percent of their original body weight (Anderson et al. 2001). This study found that exercise appeared to play an important role in maintenance and that those who either had used a very low-energy diet or lost a significant amount of weight initially (≥ 20 kg) were more likely to maintain losses successfully.

Most studies of weight-loss program effectiveness are set in hospital-based clinics. These programs are most likely to attract persons for whom previous efforts at weight loss have failed. This suggests that estimates of
successful maintenance may be overly pessimistic. A systematic review attempting to establish the prevalence of self-cured obesity in the general population concluded that the quality of evidence was insufficient to draw any conclusions (Bartlett et al. 1999). In a sample of adults who expressed interest in avoiding a weight gain, a three-year follow-up revealed that only one-fourth of the sample succeeded (Crawford, Jeffery, and French 2000). Only 5 percent of the sample successfully maintained a loss. Of those who achieved a significant loss, however, 40 percent maintained it through the end of the follow-up period. A survey of a nationally representative population found that 25 percent of persons who intentionally had lost more than 10 percent of their body weight had maintained that loss for five years or more (McGuire, Wing, and Hill 1999). One study of persons who had maintained a loss of at least thirty pounds for at least one year suggests that as the length of time that one maintains a loss increases, the cognitive burden associated with maintaining that loss decreases (Klem et al. 2000). The evaluation of the intervention by Anderson et al. (1999) mentioned above may support this finding. They found that in the first two to three years of follow-up, participants showed a monthly gain of weight, followed by a period of stable weight for the remainder of the follow-up period.

In addition to psychological issues, there may be physiological reasons why it is difficult to maintain weight loss. A meta-analysis that compared the resting metabolic rate for persons who had been obese (defined as BMI ≥ 27) to that for persons who never had been obese found that the rate for the formerly obese was 3 to 5 percent lower (Astrup et al. 1999).

In sum, the majority of existing research focuses on hospital-based clinics and suggests that some degree of maintenance is possible. The evidence for other types of programs and self-initiated efforts is less conclusive. In the absence of an endorsement by the USPSTF of a particular treatment program associated with the screening program, one cannot say with certainty that weight loss achieved will be maintained.

WHAT ARE THE HEALTH IMPLICATIONS OF FAILING TO MAINTAIN A LOSS?

Given that weight-loss maintenance may be difficult to achieve and that persons may make multiple attempts during a lifetime to lose weight, we must consider whether weight cycling is detrimental to health. As with so many obesity-related issues, the evidence is inconclusive, in part because of methodological problems. Several studies do not distinguish between voluntary and involuntary weight loss. Not all studies are conducted in obese populations. Some studies look at risk factors but not mortality. Other
methodological challenges include the definition and quantification of weight cycling. In light of this, a 1994 review by the National Task Force on the Prevention and Treatment of Obesity concluded that obese persons should not let concerns about the deleterious effects of weight cycling deter them from efforts to lose weight (Weight Cycling 1994). Some more recent research assessing the link between voluntary weight cycling and cardiovascular risk suggests no adverse affects (Graci et al. 2004; Wing, Jeffery, and Hellerstedt 1995). A study of women with coronary risk factors, however, found an association between voluntary cycling and lower levels of high-density lipoprotein cholesterol but no association with coronary artery disease (Olson et al. 2000). A recent study of frequent weight cycling in women suggests that it negatively affects immunocompetence (Shade et al. 2004).

In a study of weight cycling and mortality based on NHANES data, Diaz, Mainous, and Everett (2005) are able to examine changes in weight (measured at five time periods) during a period of approximately twenty years and all-cause and cardiovascular mortality. Compared to persons with stable weight, persons with weight fluctuation (controlling for body weight) had a hazard ratio of 1.81 times greater risk for all-cause mortality and 1.48 times greater risk for cardiovascular-related mortality.

The uncertainty in the effects of weight cycling becomes important if weight-loss treatment results in a loss of weight, if that loss cannot be maintained, and if treatment is reinitiated. Thus, a lesser weight should be accorded to this inconclusiveness when evaluating the potential of a screening program to improve outcomes.

IS TREATMENT AVAILABLE?

A screening intervention cannot improve outcomes if it is not possible or not encouraged for a physician to initiate treatment. That initiation will depend both on the healthcare infrastructure and on physician characteristics. As mentioned earlier, a 1996 study found that fewer than half of obese persons who had had a visit with a clinician in the previous twelve months reported being advised to lose weight (Galuska et al. 1999). Reimbursement, limited office time, lack of training, and low confidence in ability to counsel (Kristeller and Hoerr 1997) or change behavior (Kushner 1995; Yeager et al. 1996) all have been cited as barriers to physician counseling. Reimbursement issues may decrease as programs such as Medicare decide to treat obesity as a disease and cover treatment (U.S. Department of Health & Human Services 2004).

One reason why screening for obesity may not result in persons receiving treatment is that obesity is associated with low socioeconomic status (Everson et al. 2002; James et al. 1997). While this observation may be confounded by
the measurement issues discussed earlier, the relationship appears very strong. Persons who are socioeconomically disadvantaged are less likely to have insurance and to have a relationship with a health care provider (National Center for Health Statistics 2003), and thus, they are less likely to benefit from a screening intervention set in a doctor’s office.

In sum, there exists uncertainty in each of the five weight-loss effectiveness criteria outlined here. While persons can lose weight, it is not clear what approach should be taken, if treatment will be available, if the loss can be maintained and outcomes will improve, or if outcomes will be compromised by cycling.

### INCREASED EFFECTIVENESS OF EARLY DETECTION AND TREATMENT

The meaning of early detection and treatment is unclear in the context of obesity. Clearly, obesity has no asymptomatic phase. While it is possible that those who are overweight are in the early stages of obesity, we do not know much about who will and will not gain weight. Some might argue that early detection and treatment suggest focusing on obesity in children. Even given the methodological and measurement challenges of estimating childhood obesity (Troiano and Flegal 1999), it appears that its prevalence is increasing dramatically (Ball and McCargar 2003; Strauss and Pollack 2001). Furthermore, obese older children and young adults are likely to become obese adults (McTigue, Garrett, and Popkin 2002; Whitaker et al. 1997). As with adult obesity, however, there is still considerable uncertainty as to the best treatment approach (Reilly and McDowell 2003; Summerbell et al. 2003).

### THE FEASIBILITY OF IMPLEMENTING SCREENING

Even under the most optimistic assumptions about the potential for an obesity-screening program to meet the effectiveness criteria outlined above, there remain important questions about feasibility. First, the USPSTF guidelines make no recommendation about frequency of screening. Certainly, a recommendation for screening interval is necessary for physicians. Should patients be screened and counseled once a year? At every visit? Every other year? Screening and follow-up may constitute a significant portion of the already limited time that patients spend with their primary care provider. Given that physicians generally perceive the amount of time they can spend with a patient to be fixed, time spent on screening and counseling for weight loss cannot be spent on other aspects of patient care and constitutes a significant opportunity cost.
The feasibility of screening also requires that effective programs are available for persons identified as obese. Physicians face a wide array of choices about which there is little consensus as to what programs work best for whom. A survey of weight-loss experts asked them to consider eleven weight-loss programs (e.g., commercial programs, residential programs) and to identify the patient characteristics that might indicate or contraindicate a given program (Schwartz and Brownell 1995). There was significant disagreement among the group. Degree of obesity was the only patient characteristic that pertained to every treatment, and no consensus was reached as to decision rules. This uncertainty may reflect the lack of consensus in the research literature as to the best approaches to weight loss.

Finally, if medical professionals are to take increased responsibility for identifying (and subsequently treating) obesity, they must feel comfortable with the issue and believe that treatment works. Research suggests that this is not the case. As mentioned earlier, studies have found that health care professionals exhibit bias against obese patients and even are inclined to treat them differently than their nonobese patients (Hebl, Xu, and Mason 2003). Physicians have expressed limited confidence in their ability to counsel or change behavior (Kushner 1995; Yeager et al. 1996). They also may be aware that for many patients, weight is an emotionally laden topic, and even the words used to discuss it can have an effect on a patient’s reaction (Wadden and Didie 2003). This is not to suggest that this problem could not be overcome, but strategies for easing discomfort should be addressed. However, to date, there are no trials of interventions to change physicians’ attitudes toward obese or overweight patients, and there is no evidence that existing interventions can improve the ability of health professionals to manage obesity (Harvey et al. 2002). In sum, research suggests that many obese patients go uncounseled about their weight, that physicians lack confidence in their ability to counsel about weight, and that interventions to improve that ability have not been identified.

CONCLUSIONS

In the current climate characterized by ever more alarming statistics about the prevalence of obesity and associated health problems in the United States (and around the world), it is not surprising that we turn to the medical sector for solutions. The suggestion that physicians must improve identification of obese patients and initiation of treatment seems at first glance to be a simple and appealing initial step. Given that height and weight are measured easily and BMI is simple to calculate, one might ask, what is the harm of asking physicians to take these measurements and talk to their patients
about obesity? That simple measurement-and-discussion framework is not, however, what is suggested by the USPSTF and is not the focus of our review. The task force’s recommendation is that all adult patients be screened and that all those classified as obese begin intensive treatment. In this review, we have highlighted the limitations of such a recommendation, arguing that if universal screening in clinical settings were to be implemented, there is considerable uncertainty as to whether either obesity or its related health problems would be reduced.

First, if physicians were to conform strictly to BMI guidelines, they would initiate treatment for all patients who meet the current guidelines for obesity (or overweight). In reality, not all persons who meet such guidelines need to lose weight for health reasons. While physicians may try to improve the measurement of excess fat by taking other measurements such as waist circumference, not all patients who meet the thresholds for those measurements would benefit from weight loss.

Second, if physicians are to offer their patients help, they must provide treatments that are more effective than what patients currently are trying. Observational studies suggest that many persons already are trying to treat themselves. A 1996 telephone survey found that 60 percent of obese men and 70 percent of obese women reported trying to lose weight, although most were not engaging in the recommended combinations of dietary and physical-activity changes (Serdula et al. 1999). As this review points out, identifying these treatments will be difficult for physicians, and the research community has yet to build the evidence base to support such treatment decisions.

Indeed, one of the major weaknesses in arguing that increased detection of obesity will lead to improved outcomes is that research has not demonstrated the effectiveness of treatment fully. As attention increasingly focuses on the health consequences of obesity, however, we expect that there will be increased efforts to identify effective programs. Some researchers advocate a chronic-care model of the treatment of obesity (Nonas 1998), a model demonstrated effective in other areas such as diabetes (Wagner et al. 2001). However, before such models can be implemented for obesity, there is much to be resolved about design. For example, questions remain as to which types of providers would constitute an effective team (e.g., primary care provider, nutritionist, fitness instructor), who would pay, what key self-management skills are needed, and what types of treatments could be recommended as effective. This review is not meant to suggest that physicians should not discuss weight and weight-related issues with patients. Rather, we focus on the very specific question of whether the research evidence exists to suggest that mass screening for and ensuing treatment of obesity in adults can be expected to improve health outcomes.
Calls for obesity to be treated within the health care sector represent the condition’s further medicalization. In their review of the conceptualization of obesity in a leading medical textbook since 1927, Chang and Christakis (2002) find that through time, the causes of obesity are less likely to be located in what people do and more likely to be conceptualized as sickness. While there may be some advantages to medicalizing obesity, there also may be substantial unintended negative consequences. A major advantage to medicalizing any social problem is that it opens up the health care sector for individuals who may benefit from treatment. Just as Medicare recently has declared obesity to be a disease and now will cover services, many other insurers may follow suit. This may make previously expensive treatments available to patients. However, the view that obesity is a disease for which we seek a cure locates the source of the problem in the individual who is perceived as sick. Such a view inevitably classifies some persons as diseased who never will suffer any ill effects of being obese. It also distracts attention from the environmental and social causes of obesity. As has been pointed out by other writers, however, even environmental approaches to obesity have not been based on evidence (Jain 2004). Outlining uncertainties about the efficacy of screening for obesity is not tantamount to recommending inaction. Rather, we recommend action in the direction of strengthening the evidence base for both the physiology and measurement of obesity and for the long-term effectiveness of treatment.

REFERENCES


