Review

Assessment tools in obesity — Psychological measures, diet, activity, and body composition

Laura Beech {a,1}, Jennie Galpern {a,1}, Andrew Petrone {b,1}, Sai Krupa Das {a,c,*}

© 2012 Elsevier Inc. All rights reserved.

The global increase in the prevalence of obesity has led to an increased need for measurement tools for research, management and treatment of the obese person. The physical size limitations imposed by obesity, variations in body composition from that of normal weight, and a complex psychopathology all pose tremendous challenges to the assessment of an obese person. There is little published research regarding what tools can be used with confidence. This review is designed to provide researchers and clinicians with a guide to the current and emerging measurement tools specifically associated with obesity research and practice. Section 1 addresses psychological measures of well being. Sections 2, 3, and 4 focus on the assessment of food intake, activity, and body composition. All sections address basic challenges involved in the study and management of obesity, and highlight methodological issues associated with the use of common assessment tools. The best available methods for use in the obese both in research and clinical practice are recommended.
1. Introduction

The global increase in the prevalence of obesity has led to an increased need for measurement tools for research, management and treatment of the obese person. The physical size limitations imposed by obesity, variations in body composition from that of normal weight, and a complex psychopathology all pose tremendous challenges to the assessment of an obese person. The field of obesity research would benefit from having more uniform methods of assessment which would enable researchers for clinical and community-based studies, evaluation teams to assess intervention programs, and health professionals for counseling individuals. Standardized assessment methods support better comparison of health between different studies and across diverse populations. This is particularly important since the reported results are attributed value that drives policy, organization, and treatment.

2. Psychological assessment

2.1. Introduction

Psychological assessment measures are abundant in the field of obesity research, and are necessary to determine the psychological health of an obese patient before, during, and after treatment. In clinical practice, psychological assessment tools are important for determining the effectiveness of weight loss treatment. In a research setting, these tools are important for comparing the results of different weight loss programs, and understanding the connection between the physical and psychological problems associated with obesity. The field of obesity research would benefit from having more uniform methods of psychological assessment, which would allow for better comparison of psychological health between different studies and across diverse populations. The purpose of this section on psychological well-being is to address issues of tool validity in overweight and obese populations, highlight the most widely used tools in the field, and provide a reference for selecting the most appropriate method of assessment, depending on the context and purpose of the research.

2.2. Introduction to Quality of Life Assessment

The effects of obesity on quality of life (QOL) have been well studied, and the overall consensus is that obesity decreases QOL, and treatment improves QOL [1]. The main assessment tool used by researchers has been the questionnaire, and several authors have done extensive reviews on these questionnaires [2–5]. Questionnaires can be divided into general QOL questionnaires, which are not designed to examine the specific health problems associated with obesity, and obesity-specific QOL questionnaires. The questionnaires discussed in this review are the general Short Form-36, the obesity-specific Impact of Weight on Quality of Life, the Impact of Weight on Quality of Life — Lite, the Moorehead-Ardelt — II, the Weight Related Symptom Measure, the Obesity and Weight Loss Quality of Life questionnaire, and the Obesity Related Well Being questionnaire.
2.2.1. Short Form-36 (SF-36)

The Medical Outcomes Study SF-36 questionnaire is the most commonly used generic instrument for measuring QOL [4,6]. The SF-36 measures eight domains: i) physical functioning, ii) role limitations due to physical health problems, iii) bodily pain, iv) general health perceptions, v) vitality, vi) social functioning, vii) role limitations due to emotional problems, and viii) mental health. The SF-36 has excellent psychometric properties, has been validated across diverse populations with medical and psychiatric problems, and is easy to complete [4,6-8]. Although the SF-36 has been used in numerous studies with individuals who are overweight and obese, it is recommended that it be used in conjunction with an obesity-specific questionnaire [9-12]. BMI has been shown to be significantly associated with poor health related QOL using the SF-36, but this association is the strongest when measuring physical activity, not mental health, social functioning, role limitations due to emotional problems, or vitality. Also, the SF-36 is unable to distinguish between impairments due to BMI in mild and moderate physical activity versus intense physical activity [8]. Overall, the SF-36 does not measure disease-specific domains, lacks the sensitivity to detect small treatment effects, and further studies need to be done to assess the validity of each domain of the SF-36 with morbidly obese individuals [9,12]. However, the SF-36 is a very robust tool that can be used to compare QOL in obese individuals to the general population [7].

2.2.2. Impact of Weight on Quality of Life (IWQOL)/IWQOL-Lite

The IWQOL is a 74-item self-report questionnaire that was developed in a clinical sample of moderate to severe obese individuals, and assesses the affects of weight on QOL in eight areas: health, social and interpersonal life, work, mobility, self-esteem, sexual life, activities of daily living, and comfort with food [13]. The IWQOL is a psychometrically sound measure that has the ability to detect any post-treatment affects, which makes it a useful tool to use after clinical trials of antiobesity drugs, or after surgical treatments [14]. A shorter 31-item version, the IWQOL-Lite, has been developed that assesses QOL across five areas: physical function, self-esteem, sexual life, public distress, and work, and correlates well with the IWQOL, shows excellent psychometric properties, and has been validated in individuals with psychiatric disorders who are prone to obesity [9,12]. Because of its ease of use and its ability to detect changes in QOL associated with small changes in BMI, the IWQOL-Lite is the preferred method of assessment over the original questionnaire.

2.2.3. Moorehead–Ardelt Quality of Life Questionnaire — II (MA-II)

The MA-II is a one page obesity specific tool used as part of the Bariatric Analysis and Reporting Outcome System to measure postoperative outcomes in self-perceived QOL by using simple drawings to assess six areas: self-esteem, physical well-being, social relationships, work, sexual activity, and eating behavior [17-19]. The MA-II has been validated in gastric bypass patients who are morbidly obese, with a target population of morbidly and super obese individuals. It creates a standard for comparing QOL outcomes after the surgical treatment of severe obesity because it can be used for both pre and post intervention assessment [5]. The MA-II is an easy to use questionnaire that can be easily used for different cultures and populations. Specifically designed for morbidly obese patients who have undergone surgical operations, it takes into account complications that could arise from surgery and the potential for re-operation [17].

2.2.4. Weight Related Symptom Measure (WRSM) and Obesity and Weight Loss Quality of Life (OWLQOL)

Both the WRSM and the OWLQOL questionnaire were developed as culturally sensitive measures of QOL as development of the questionnaires involved qualitative input from six countries: the United States, the United Kingdom, France, Germany, Spain, and Italy [20]. The WRSM is a 20-item measure of the symptoms associated with obesity and obesity treatment, along with the degree to which each symptom “bothers the individual”. The OWLQOL is a 17-item measure of a person’s global evaluation of obesity and their effort to lose weight based on feelings that are unobservable to others. Both the WRSM and the OWLQOL are responsive to short and long-term reductions in weight loss, and are easy to complete questionnaires that are intended to be administered together and with other outcome measures [21].

2.2.5. Obesity Related Well-Being (ORWELL-97)

The ORWELL 97 questionnaire is an 18-item self-report measure that assesses QOL across three areas: symptoms, which include somatic symptoms and physical functioning; discomfort, which is defined as the effect of obesity on emotional status; and impact, which is defined as the effect of obesity on relationships and an individuals’ social network. The ORWELL 97 proposes that symptoms of similar intensity can have a different impact depending on the individual, so respondents are asked about the occurrence, severity, and relevance of each impact on the individual’s own life. The ORWELL-97 has high test–retest reliability, good internal consistency, and can be used as a clinical measure with a wide population. But, in a preliminary study, weaknesses were found when trying to correlate BMI with sub-scores of the ORWELL-97. Also, women were found to have lower QOL because of the greater impact of being obese had on psychosocial complaints [22]. This suggests that further studies need to be done in order to determine how to interpret the sub-scores of the ORWELL 97.

2.3. Introduction to hunger assessment

Hunger, dietary restraint, and overeating have been well studied in the obese and questions still exist as to the differences between normal and obese individuals when it comes to these dimensions [23]. While other scales, such as the Restraint Scale and Eating Behavior Scales exist, the Three-Factor Eating Questionnaire will be discussed because it encompasses both hunger and dietary restraint, and is commonly used in the study of the obese [1]. More subjective measures of hunger include Visual Analog Scales and what is described as Pictorial Measures of hunger.

2.3.1. Three Factor Eating Questionnaire (TFEQ)

The TFEQ is a 51-item self-report measure that was developed to assess restrained eating to control body weight by measuring three domains of the psychological patterns of eating: dietary restraint, disinhibition, and hunger [24–27]. The TFEQ is useful to predict weight loss in clinical patients, to monitor progress during treatment, has good psychometric properties, and it is one of the most widely used tools to study eating in obese individuals [25,28]. However, further analysis has shown that the original three factor structure may not be replicated in obese individuals, so two shortened forms, the TFEQ-21 and the TFEQ-18, which measure cognitive restraint, uncontrolled eating, and emotional eating have been developed in an obese population [29–32].

2.3.2. Visual Analog Scales (VAS)

A VAS is a type of question that is used to rate hunger, satiety, and individual’s own interpretation of their hunger sensations. To measure hunger, Visual Analog Scales were initially developed with six questions: i) How hungry do you feel? ii) How full do you feel? iii) How strong is your desire to eat? iv) How much to you think you could eat now? v) What is your urge to eat? and vi) What is your preoccupation with thoughts of food? Individuals answer these questions by making a single mark on a 100 mm straight line, where the two extreme answers to every question are anchored on opposite ends of the line. The use of VASs have been shown to be both reliable and valid, have been used extensively when studying
obese individuals, not influenced by prior diet, and can be used to assess the effects of drugs, diet composition, and alterations in energy intake [33–38]. VASs are sensitive to experimental manipulations, can be used as a proxy for energy intake, are simple to use and interpret, and can be used to compare across populations. However, because individual differences in interpretations of the scale may arise, it is advised that researchers predominantly use VASs in studies with a within-subjects design that compares hunger before and after treatment because in a validity study, within subjects comparisons were more accurate and sensitive than between-subjects comparisons [35,36].

Overall, VASs are a very common method for measurement on a screen. Although EARS increases the reliability of data collection, that individuals electronically mark their answers on lines presented (EARS) have been developed. These use handheld electronic devises that individuals electronically mark their answers on lines presented on a screen. Although EARS increases the reliability of data collection, initial studies have shown that EARS produces responses with less variation, so EARS and pen and paper VASs should not be used interchangeably [36]. Overall, VASs are a very common method for measuring hunger and are helpful for measuring an individual’s subjective hunger sensations.

2.3.3. Pictorial Measures

Pictorial Measures of hunger were first developed to assess the body areas associated with the sensations of hunger and the extent of these sensations [39]. Individuals are asked to outline on a drawing of a human body the area where they are experiencing hunger sensations, and the size of the outlined area should reflect the intensity of one’s hunger sensations. This is an emerging tool that needs further validation and should be used in conjunction with other subjective measures of hunger. Although this pictorial instrument was tested using obese individuals, it was developed in normal weight subjects, and the body used in the measure are of normal weight [23]. Despite this, an initial study using this tool has found that physical aspects of hunger may be distinguished from overall global aspects of hunger [39]. This tool may also be more sensitive to extreme hunger, and increases in hunger during fasting may be better measured by using a pictorial instrument [23,39]. Overall, more testing is needed, but an initial study suggests that a pictorial measure could be useful in the study of obese individuals as an instrument that complements more traditional measures of hunger.

2.4. Introduction to sleep assessment

Obesity is associated with sleep disturbances, excessive daytime sleepiness, and obstructive sleep apnea (OSA). Obesity increases a person’s risk for OSA 10-fold [40,41]. Subjective sleep assessment tools, such as polysomnography and actigraphy are commonly used as a way to quantify sleep disturbances. Questionnaires have also been commonly employed in the study of obese individuals, such as the Epworth Sleepiness Scale, St Mary’s Hospital Sleep Questionnaire, VSH Sleep Scale, and the Pittsburgh Sleep Quality Index [42–45]. These tools have not been developed in the obese, but the best to use when studying sleep disorders of obese individuals. This review will cover the Epworth Sleepness Scale, polysomnography, and actigraphy technology because of their validation and extensive use in the obese.

2.4.1. Epworth Sleepiness Scale

The Epworth Sleepiness Scale (ESS) is an eight-item measure of daytime sleepiness that asks individuals to rate on a scale how likely they would be to doze off or fall asleep in 8 situations [46]. The ESS is a simple test that has been found to be psychometrically sound among the general population, gives a retrospective report on dozing behavior, and high ESS scores have been significantly correlated with obstructive sleep apnea [47,48]. The ESS has been used extensively in obese individuals to study sleep disturbances, but may be more difficult to use in morbidly obese individuals [42,49,50]. However, the use of the ESS alone is not sufficient to diagnose obstructive sleep apnea or other sleep related disorders. It is recommended that morbidly obese individuals who have a high score on the ESS undergo polysomnography to further diagnose a sleep disorder.

2.4.2. Polysomnography

Overnight polysomnography is used to diagnose sleep-related breathing and respiratory disorders, including OSA. Before undergoing polysomnography, a full sleep history and physical examination are recommended. A full polysomnography includes an electroencephalography, electrooculography, chin electromyography, airflow, arterial oxygen saturation, respiratory effort, and electrocardiography. An anterior tibialis EMG can be used to help measure movement associated with arousal [51]. Overnight polysomnography is the gold standard for accurate diagnosis in obese individuals because of its reliability and its ability to accurately diagnose sleep disorders [51,52]. This tool can also measure sleep improvements after weight loss, and is recommended for use after substantial weight loss [51,53]. Despite these advantages, polysomnography is expensive, time consuming, often inconvenient, error could arise during instrument readings, data could be lost, and misclassification of patients could result because of night-to-night variability [41,51].

2.4.3. Actigraphy

Actigraphy is used to assess sleep/wake patterns via a movement detector, most commonly an accelerometer, which is worn on the wrist or ankle over a period of time [54–56]. Currently, there are different actigraph instruments on the market, and different algorithms used to determine sleep/wake patterns [57]. This lack of standardization poses a problem when comparing the results of sleep studies that use these different methods. Although there is a relationship between sleep duration and obesity as measured by an actigraph, actigraphy is not the gold standard for measuring sleep duration because it is not as accurate as polysomnography, it cannot distinguish the difference between different sleep disorders, and it likely overestimates sleep and underestimates wake [54,55]. For accurate readings, it is recommended that actigraphy measurements be supplemented with a sleep log. Also, actigraphy does not allow for routine diagnosis or assessment of severity of sleep disorders [58]. Despite these disadvantages, actigraphy can be used in an individual’s natural sleep environment, is feasible for use in large research studies, is cost effective, allows for study when polysomnography is not feasible, and it allows individuals to be tested for 24 h across multiple days [54,55].

2.5. Introduction to psychological well-being assessment

Individuals with obesity show a higher prevalence of psychiatric illness compared with the general population [1,59]. However, weight loss is associated with a reduction of depressive symptoms [60]. The most common method of psychological assessment is questionnaires designed for the general population. Described here are the questionnaires that are considered the gold standard for assessing depression, well-being, and self-esteem. These are the Beck Depression Inventory, the Center for Epidemiologic Studies Depression Scale, the General Well-Being Schedule, and the Rosenberg Self-Esteem Scale.

2.5.1. Beck Depression Inventory II (BDI-II)

The BDI-II is a 21-item measure that was initially developed in psychiatric patients to assess the intensity and the behavioral manifestations of depression [61,62]. The BDI-II is recommended in the study of obese individuals because of its widespread use with both obese and extremely obese populations, and because its items are not biased by obesity [4,60,63]. Advantages of the BDI-II are its ease of use, its ability to detect changes in depression over time and with treatment, and it is one of the most widely used, psychometrically
valid self-report measures of depression [61,62,64]. However, the BDI-II is not designed to diagnose different types of depression or psychiatric illness [61].

2.5.2. Center for Epidemiologic Studies Depression Scale (CES-D)

The CES-D was developed from previous questionnaires for use in large population based epidemiological studies. It is a 20-item self-report scale designed to measure the frequency and duration of the major symptoms associated with depression, including depressed mood, feelings of guilt and worthlessness, feelings of helplessness and hopelessness, psychomotor retardation, loss of appetite, and sleep disturbance [65]. The CES-D has been validated in diverse populations, so it is appropriate to use when studying obese individuals in a large epidemiological setting [66,67]. Although the CES-D is brief, a 10 item short form, the CESD-10 has been developed in healthy older adults [68]. The short form is not as widely used as the longer, 20 item scale, and only one study has used the CESD-10 in overweight individuals [69]. The CES-D is easy to complete, and has been used extensively in epidemiologic studies and in ethnically diverse samples, which allows for the comparison of scores across populations [67,70]. However, the scale is not designed for the clinical diagnosis of depression, to differentiate between different types of depression, nor to interpret individual scores [65,70].

2.5.3. General Well-Being Schedule (GWB)

The GWB is an 18-item self-administered questionnaire that has been validated in and is widely used in medical research on obesity, and is recommended for measuring subjective feelings of psychological well-being [71–73]. The GWB emphasizes an individual's inner personal state, rather than external conditions that could affect well-being. Six subscales: anxiety, depression, general health, positive well-being, self-control, and vitality have been identified, but have not been validated in all populations [71,72]. The GWB is easy to complete and avoids references to physical symptoms of emotional distress, which can lead to problems in interpretation [71]. Because of its relatively low test–retest reliability, it is recommended for use in large population studies, and not in determining individual changes in well being. Alternate forms include the 10-item Psychological Mental Health Index, and a version incorporated into the Rand Mental Health Inventory [74].

2.5.4. Rosenberg Self-Esteem Scale (RSE)

The RSE scale is a 10-item self-report psychological screening tool that measures global self-esteem by assessing whether a person has a favorable or unfavorable attitude toward oneself [75,76]. The RSE scale is widely used across a variety of populations, including the obese and morbidly obese, because of its excellent psychometric properties [59,71,77,78].

2.6. Introduction to perceived body image assessment

Williamson and O’Neil define body image as the cognitive perception of one’s body size and appearance, and the emotional response to that perception. Body image is less accurately estimated by obese individuals, and obesity is associated with a preoccupation with one’s body weight [1]. Various technological methods of body image assessment exist, including using video, computers, and distorted mirrors [79–82]. This review will cover questionnaires, because they are easier to use, are currently more commonly used, and are well validated. These include the Body Shape Questionnaire, the Multi-dimensional Body-Self Relations Questionnaire, and the Body Image Assessment for Obesity.

2.6.1. Body Shape Questionnaire (BSQ)

The BSQ is a 34-item self-report measure of body shape dissatisfaction, especially the construct of “feeling fat” by assessing distress with, and frequency of preoccupation with body shape and size [78,83,84]. It is a useful measure of weight and shape concerns in diverse clinical samples of obese and morbidly obese individuals, has good psychometric properties, and is easy to complete [78,83,85–88]. Short versions of the BSQ have been developed, but are not validated in nor used extensively in the obese [89,90].

2.6.2. Multidimensional Body-Self Relations Questionnaire (MBSRQ)

The MBSRQ is a 69-item self-report measure that is one of the most widely used tools to assess body image. It measures the evaluation of one’s appearance, health and illness, fitness, body satisfaction, weight attitude, and weight status, and assessing the cognitive, behavioral, and affective components of body image [91,92]. The two subscales frequently used in the study of obese individuals are the Appearance Evaluation (AE) subscale and the Body Areas Satisfaction (BAS) subscale [92–94]. The MBSRQ is an excellent tool for the use with obese individuals, but an analysis does show the questionnaire’s weakness in being able to compare different age and gender groups [91,92].

2.6.3. Body Image Assessment for Obesity (BIA-O)

The BIA-O is an extension of the original Body Image Assessment (BIA). The BIA defines body image dissatisfaction as the discrepancy between self-perceived and ideal body size estimates. The measure presents individuals with nine silhouettes ranging in body size and asks them to determine which most accurately depicts their current body size and ideal body size. However, the original BIA silhouettes depicting overweight individuals were not large enough to use with an obese population [95]. The BIA-O added an additional nine silhouettes, so the 18 silhouettes ranged from very thin to very obese. The developers also added an additional question concerning a reasonable body size that would be realistic to maintain over a long period of time. After completing the BIA-O, two measures of body size dissatisfaction can be determined: current body size minus ideal body size and current body size minus realistic body size. The BIA-O has been found to be a valid and reliable tool in individuals with a BMI of up to 50, can be used to determine the relative cause of body dissatisfaction, and is valid in ethnically diverse populations. Disadvantages of the BIA-O include its interpretability because it can only be used in the context of BMI, ethnicity, and gender, and the fact that male silhouettes do not distinguish between increasing size due to fat or due to musculature [95,96].

2.7. Conclusion

This review of the psychological assessment tools in the obese shows that while many objective tools and subjective questionnaires exist, few are well-validated and used extensively in obese populations. While each tool has its flaws, the tools presented in this review are recommended for researchers and clinicians going forward as more interest develops in the measurement of psychological health of obese individuals. Choosing well-validated and widely used measures allows for a better comparison of research methods and results. When choosing which method of assessment to use, researchers and clinicians should consider the population they are studying, the purpose and goals of their research, and what specific aspects of psychological health they are assessing. With the increase in prevalence of obesity, measuring the psychological health of this population will continue to be vital in determining proper treatments and their efficacy.

3. Dietary intake

3.1. Introduction

Dietary intake assessment is an influential measure in research and clinical communities. Whether employed by researchers for clinical
and community-based studies, evaluation teams to assess intervention programs, or by health professionals for counseling individuals, the reported results are attributed value that drives policy, organization, and treatment. Many resources provide descriptions and discussions on the most widely used methods [97,98]. The purpose of this section on dietary intakes assessment is to address issues of tool validity in overweight and obese populations, highlight new technologies emerging in the field, and provide an easy to reference table for selecting the most appropriate methods for a variety of contexts.

Table 1 presents the results of a literature review on dietary intake assessment tools used in overweight and obese populations. The chart follows the general evolution of the field, starting with classic methods involving written records and manual data input, to the newest automated technologies still in development. The purpose of this chart is to serve as a useful inventory by outlining the advantages and limitations specifically related to assessment in overweight and obese populations. In pursuit of the most valid measurement, many researchers have studied variations or combinations of traditional methods. A column of recommendations shares those techniques less widely used, provides suggestions for implementation, and highlights additional sources and areas of investigation. Italicized items highlight the impact of weight status on tool performance.

### 3.2. Conclusion

No measure has perfect construct validity. Selection of a dietary assessment method in any study must balance between greatest validity and feasibility [113]. The sources of bias in dietary intake assessment tools continue to be explored, and new methods promise to move the field forward. Even for tools that have been validated against the gold standard DLW, external validation in overweight and obese populations remains fragile. An unannounced multi-pass 24-hour recall (in person or over the phone) with portion size estimation aids, collected for 3–8 days, including a Sunday is a recommended method for assessing dietary intake of overweight and obese individuals. The recall should be conducted by staff highly trained in the tool methodology and interpersonal communication to encourage accurate reporting. Use of ancillary tools to screen for high risk of low-energy reporting is advised. Preemptive strategies to reduce low-energy reporting may include motivational training to increase social desirability of reporting certain foods. Statistical analyses may be used to identify and address misreporting. But for researchers working with overweight and obese populations, strategic triangulation of methods provides the greatest confidence in true reporting of dietary intake.

### 4. Physical activity

#### 4.1. Introduction

Physical activity (PA) assessment is the measurement of movement intensity, type, duration, or frequency [149]. Assessment in free-living obese individuals is important for the study of disease, weight management, and associated interventions [150]. Researchers and clinicians should consider participant interference and burden, the need for contextual data, data objectivity, and time and cost requirements when selecting a method to assess physical activity [151].

#### 4.2. Doubly labeled water (DLW)

DLW estimates total energy expenditure through measured excretion of isotope-labeled water. DLW is the most accurate and objective measurement for assessing physical activity in free-living individuals, and has been used extensively in obese populations [151]. Greater underestimation of energy expenditure has been shown in the obese than nonobese using this method [152], but the accuracy in obese populations still remains greater than other methods of PA assessment. DLW has a low participant burden, but the high cost of this method limits its use to small studies.

#### 4.3. Heart rate (HR) monitoring

Estimating energy expenditure and PA through HR monitoring is a popular alternative to more expensive methods. Minute-by-minute HR data is inexpensive, convenient, noninvasive, and versatile, and provides information on the frequency, intensity, and duration of free-living PA [150]. HR monitoring underestimated energy expenditure in a small group of obese women, but was not quite significant compared to DLW. Standard calculations of activity energy costs must be modified in obese populations to account for the increased basal metabolic rate and energy costs of moving greater mass [153]. HR calculations provide unreliable estimates of energy expenditure at the individual level, but provide an acceptable estimate of total energy expenditure and associated patterns of PA when applied to a group [54,55]. Combination with accelerometry may improve precision [151].

#### 4.4. Accelerometry

Accelerometry measures the intensity and duration of movement through sensors attached to the body. Known linear relationships between accelerometer counts and energy cost allow for the classification of PA by intensity [150]. Single unit accelerometers, usually placed on the waist, are small, non-invasive, and give minimal discomfort to subjects, including the obese [154]. Consistent and secure placement on the body is important to limit variance, which may be challenging in the extremely obese. Accelerometers are limited in ability to detect activity of the extremities, bicycling, or swimming [155]. Four days of ≥6 h wear time/day optimized reliability and sample size in a study of overweight and obese adults using a triaxial accelerometer on the hip [155]. Accelerometry, using a DLW-validated instrument, is the indicated method for the assessment of habitual frequency, intensity and duration of PA of both obese and non-obese individuals [151,154].

#### 4.5. Introduction to questionnaires

Questionnaires are the most widely used method to assess PA, but few have been studied in the obese [156]. The use of questionnaires to predict individual energy expenditure is largely dependent on subject compliance and ability to correctly estimate time spent in activities of varying intensities [157]. In general, questionnaires have low reliability and validity but are useful for ranking individuals by activity level [151]. Obesity is correlated with overestimation of daily PA in individuals [156] making this method particularly problematic in obese populations. Questionnaires vary in their measurement of activity domains, time frame of recall, and expression of result [156].

##### 4.5.1. Baecke Questionnaire

The Baecke questionnaire contains 16–items and a simple scoring system for calculation of an activity index. It is valid and reliable for assessing physical activity patterns in work, sport, and leisure in the general population [158] and has been used in studies with the obese [156]. Identification of misreporting can be difficult because the index results cannot be easily compared with energy expenditure measurements from other methods [156].

##### 4.5.2. International Physical Activity Questionnaire (IPAQ)

The IPAQ is a 31-item questionnaire available in 21 languages, in telephone or self-administered format. Domains of assessed PA include household and yard work, occupational activity, self-powered transport, leisure-time activity, and sedentary activity. The IPAQ was
Table 1
Dietary intake measurement tools for use in the overweight and obese population.

<table>
<thead>
<tr>
<th>Method and examples</th>
<th>Studied in obese</th>
<th>Advantages</th>
<th>Limitations</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. 24-hour recall</td>
<td>Yes [99–101]</td>
<td>– Appropriate for low-literacy populations [98]</td>
<td>– Underreporting of energy intake associated with higher BMI [102]</td>
<td>– May be conducted in person or over telephone [103]</td>
</tr>
<tr>
<td>Interviewer administered recall of exact food intake during the previous day</td>
<td></td>
<td>– Low participant time burden [98]</td>
<td>– Significant misreporting by obese in ethnic populations [99]</td>
<td>– Unannounced recall reduces potential for altered consumption patterns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– Relatively inexpensive [98]</td>
<td>– Requires trained staff for interview and data analysis [98]</td>
<td>– Additional financial incentive for accurate reporting does not improve intake among obese [101]</td>
</tr>
<tr>
<td>II. Food record</td>
<td>Yes [102,104–107]</td>
<td>– Validated measure of group intake among overweight and obese men and women in controlled environment [105,107]</td>
<td>– Accurate record of mean energy intake within general population, but unable to detect additional energy requirements related to obesity [102]</td>
<td>– Portion size estimation aids may help reduce misreporting</td>
</tr>
<tr>
<td>Participant recorded multi-day record of exact food intake.</td>
<td></td>
<td>– Recommended for estimating energy intake during treatment and follow-up in the obese [108]</td>
<td>– Significantly lower correlations with true energy and protein intakes in the obese than non-obese [102]</td>
<td>– First recall associated with highest misreporting [109]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– Accounts for daily intake variation if &gt; 1 day</td>
<td>– Under-reporting of energy intake independent of recording time length [110]</td>
<td>– 3 days of recall including Sunday optimal [109]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– Weight changes during recording period may identify misreporting</td>
<td>– Not appropriate in low-literacy populations [98]</td>
<td>– 8 days of recall, including Sunday, recommended for small cohort of obese participants to minimize random error [106]</td>
</tr>
<tr>
<td>III. Food Frequency Questionnaires (FFQ)</td>
<td>Yes [110,111]</td>
<td>– Greater precision estimating usual intake of individuals [98],</td>
<td>– Participant may change normal food intake due to social bias or convenience [98]</td>
<td>– Combination with food frequency questionnaire (FFQ) may improve bias with respect to obesity [102]</td>
</tr>
<tr>
<td>Interviewer or participant administered; frequency of categorized food intake over specified time period.</td>
<td></td>
<td>– Reduced portion size estimation error</td>
<td>– Potential error in quantifying portion sizes [98]</td>
<td>– Energy adjustment procedures may improve ranking of obese individuals for protein intake [102]</td>
</tr>
<tr>
<td>Portion-size data can be converted to estimate energy and nutrient intake</td>
<td></td>
<td>– Weighed records correspond most closely to biomarkers in normal weight participants [113]</td>
<td>– Data quality declines with recording time</td>
<td>– One day record per participant appropriate for determining average intake of group, if all days of week equally represented [98]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– Greatest precision estimating usual intake of individuals [98],</td>
<td>– High staff burden for data entry</td>
<td>– Motivational training and confrontation with prior underreporting resulted in reduced underreporting rates in a population of overweight and obese women, though one-third of the participants continued to underreport [111]</td>
</tr>
<tr>
<td>Weighed food record</td>
<td>Yes [112]</td>
<td>– Most practical and economical method in large epidemiologic studies [114]</td>
<td>– Not appropriate in low-literacy populations [98]</td>
<td>– Statistical correction methods based on energy expenditure and body weight can reduce measurement error [115]</td>
</tr>
<tr>
<td>All food and food waste weighed and recorded</td>
<td></td>
<td>– Appropriate for low-literacy populations if interviewer administered</td>
<td>– Participant may change normal food intake due to social bias or convenience [98]</td>
<td>– Using a picture-sort method to administer FFQ may reduce underreporting in low-literacy obese participants [116]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– Valuable for ranking individuals by usual nutrient intake [98]</td>
<td>– Limited use in large cohort studies due to high participant burden and access to weighing scales by all participants [113]</td>
<td>– Reviewing questionnaire with participant may increase validity, but cost-efficiency must be strongly considered [117]</td>
</tr>
<tr>
<td>Method and examples</td>
<td>Studied in obese</td>
<td>Advantages</td>
<td>Limitations</td>
<td>Recommendations</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
<td>------------</td>
<td>-------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>Classic tools</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. Food Frequency Questionnaires (FFQ)</td>
<td></td>
<td>– Low staff burden for data entry with electronic scan [98]</td>
<td>– Block questionnaire, Willett questionnaire, and Diet History Questionnaire comparable when adjusted for total energy intake [114]</td>
<td></td>
</tr>
<tr>
<td>Block 2005 FFQ</td>
<td>No</td>
<td>– Lower participant burden than other classic methods [98]</td>
<td>– Used in overweight and obese populations, but no studies reporting impact of weight status</td>
<td>– 30–40 min completion time [118]</td>
</tr>
<tr>
<td>110-items with pictures for portion size estimation. Available in paper-based electronic scan or web-based formats</td>
<td></td>
<td>– Food lists may be tailored to target certain nutrients or foods, or groups of people by region, gender, age, ethnicity, or language [118]</td>
<td>– Brief Block Questionnaire (70 questions) appropriate for ranking individuals by energy and nutrient intake [118]</td>
<td></td>
</tr>
<tr>
<td>Willett FFQ</td>
<td>No</td>
<td>– Provides option for open-ended additions to food list</td>
<td>– Used in overweight and obese populations, but no studies reporting impact of weight status</td>
<td>– Adjusting for total energy intake improves correlation with multiple 24-hour recalls [119]</td>
</tr>
<tr>
<td>Diet History Questionnaire (DHQ)</td>
<td>Yes [102]</td>
<td>– Most valid assessment of energy and nutrient intake compared to Block and Willett questionnaires in diverse weight status population [114]</td>
<td>– Weakest in assessing absolute intake compared to Block and DHQ in general population [114]</td>
<td>– Detected significant portion of extra energy requirements in obese women, but not in obese men [102]</td>
</tr>
<tr>
<td>IV. Direct Visual Estimation</td>
<td>No</td>
<td>– Validated method for measuring food intake [121]</td>
<td>– Requires highly trained staff</td>
<td>– Most effective in cafeteria or institutional settings [121]</td>
</tr>
<tr>
<td>Food selections and plate waste estimated by trained observers in comparison to weighed reference portions</td>
<td></td>
<td>– Low participant burden</td>
<td>– Vulnerable to participant altering eating habits</td>
<td></td>
</tr>
<tr>
<td>Digital Photography</td>
<td>No</td>
<td>– Validated method highly correlated with weighed foods [121]</td>
<td>– Requires highly trained staff</td>
<td>– Best suited for cafeteria or institutional settings [121]</td>
</tr>
<tr>
<td>Food selections and plate waste recorded with digital video camera. Computer images viewed by trained observers and compared to weighed reference portions.</td>
<td></td>
<td>– Portion size evaluation less hurried and may be validated by a second observer [121]</td>
<td>– Quality of data dependent on picture quality</td>
<td></td>
</tr>
<tr>
<td>V. Computer-based assessments</td>
<td></td>
<td>– May be less vulnerable to social desirability bias than direct visual estimation</td>
<td>– Results limited to meals of observation</td>
<td>– Overweight respondents may prefer completing assessment in own home; social desirability bias related to food and/or the interviewer may be decreased due to limited face-to-face contact [123]</td>
</tr>
</tbody>
</table>

(continued on next page)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Method and examples</th>
<th>Studied in obese</th>
<th>Advantages</th>
<th>Limitations</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information and communication technologies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USDA Automated Multi-Pass Method (AMP4M)</td>
<td>Yes [105,107,124]</td>
<td>– Assessed mean energy intake within 10% of mean actual intake among overweight and obese women [105]</td>
<td>– Inconsistent results among overweight and obese [105,107,124]</td>
<td>– Normal weight and overweight women significantly overestimated energy, protein, and carbohydrate intakes [105]</td>
</tr>
<tr>
<td>Computer-assisted, interview-administered multi-pass 24-hour recall</td>
<td>Yes [105,107,124]</td>
<td>– Accurately assessed energy, protein, carbohydrate, and fat intake in a population of men regardless of BMI [107]</td>
<td>– Requires large samples for accuracy [124]</td>
<td>– Among participants with metabolic syndrome, no relationships were found between age, gender or BMI and accuracy of reporting [125]</td>
</tr>
<tr>
<td><strong>DietAdvisor</strong></td>
<td>Yes [125]</td>
<td>– Computerized dietary assessment may encourage patients to report with less bias than in a verbal dietary assessment when compared with the literature [125]</td>
<td>– Participants were classified as mis-reporters based on predicted basal metabolic rate [125]</td>
<td></td>
</tr>
<tr>
<td>Participants self-report dietary intake on website.</td>
<td>No</td>
<td>– Respondent completes on own time</td>
<td>– No validation studies with comparison to other methods available.</td>
<td>– Recommended that readability, navigation, and cultural tailoring be improved [129]</td>
</tr>
<tr>
<td><strong>NutritionQuest</strong></td>
<td>No</td>
<td>– Provides reports on intake compared to 2005 Dietary Guidelines</td>
<td>– One validation study in group of university students compared to a 1-day diet record [128]</td>
<td>– Comparable repeatability and validity of the DHQ to other food intake measures [130]</td>
</tr>
<tr>
<td>See Block FFQ; online or offline format for self- or interviewer-administration</td>
<td>No</td>
<td>– Provides targeted nutrition education and immediate feedback [126]</td>
<td>– Researcher must manually aggregate data if used for group analysis</td>
<td>– See ‘Graphical Food Frequency System’ in Emerging Technologies section</td>
</tr>
<tr>
<td><strong>Web-Pictorial Diet History Questionnaire</strong></td>
<td>No</td>
<td>– Provides data comparable to 24-h recall and to an observed, weighed meal [134]</td>
<td>– WHI FFQ used in overweight and obese populations, but no studies reporting impact of weight status</td>
<td>– Does not appear to produce more valid data than paper-based approaches [135]</td>
</tr>
<tr>
<td>See DHQ; includes pictures for portion size estimation</td>
<td>No</td>
<td></td>
<td>– Half of error in caloric estimation attributable to portion size estimation [134]</td>
<td></td>
</tr>
<tr>
<td><strong>VioScreen/FFQ</strong></td>
<td>No</td>
<td>– Provides reports on intake compared to 2005 Dietary Guidelines</td>
<td>– One validation study in group of university students compared to a 1-day diet record [128]</td>
<td>– Researcher must manually aggregate data if used for group analysis</td>
</tr>
<tr>
<td><strong>Self-administered web-based 131-item FFQ with pictures for portion-size estimation</strong></td>
<td>No</td>
<td>– Provides optional behavioral feedback to participant [132]</td>
<td>– No usability or validation studies reported for web-based version.</td>
<td>– Comparable repeatability and validity of the DHQ to other food intake measures [130]</td>
</tr>
<tr>
<td><strong>Food Recall Checklist (FoRC)</strong></td>
<td>No</td>
<td>– Customizable system adapts for regional/ethnic food patterns, languages, font sizes, and target nutrients [132]</td>
<td>– No usability or validation studies reported for web-based version.</td>
<td>– Recently developed tool studied in small university population [133]</td>
</tr>
<tr>
<td>121-item self-administered web-based food checklist with pictures for estimating portion size</td>
<td>No</td>
<td>– Provides data comparable to 24-h recall and to an observed, weighed meal [134]</td>
<td>– Does not appear to produce more valid data than paper-based approaches [135]</td>
<td>– Half of error in caloric estimation attributable to portion size estimation [134]</td>
</tr>
<tr>
<td><strong>VI. Personal Digital Assistants (PDA)</strong></td>
<td>Yes [134]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method and examples</td>
<td>Studied in obese</td>
<td>Advantages</td>
<td>Limitations</td>
<td>Recommendations</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------</td>
<td>------------</td>
<td>-------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>Information and communication technologies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DietMatePro</td>
<td>- Meal records can be saved to facilitate entry of similar foods [134]</td>
<td>- Primary sources of error similar to classic food record methods: subjects must select the appropriate foods, record all foods consumed, and accurately estimate portion sizes accurately [134]</td>
<td>- Average 8.5 min recording time per meal [134]</td>
<td></td>
</tr>
<tr>
<td>Participants record all food intake in program on PDA with automatic analysis</td>
<td>- Can remind users to record meals [134]</td>
<td>- Suggested validity for monitoring food intake in weight management and obesity studies [134]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wellnavi PDA, camera, and mobile phone card. Participants photograph before and after meal, send photo to research staff for analysis</td>
<td>- No greater rate of underreporting by obese women compared to nonobese women [136]</td>
<td>- Significant underreporting by obese men compared to nonobese men [136]</td>
<td>- Low tool performance attributed to poor digital photo quality [136]</td>
<td></td>
</tr>
<tr>
<td>CalorieKing</td>
<td>- Provides automatic nutrient analysis, estimates target nutrient intakes, and charts progress over time [138]</td>
<td>- Prevalence of underreporting in an overweight and obese population did not improve compared to other methods of dietary intake assessment [139]</td>
<td>- Use in populations comfortable with PDA use may improve rates of underreporting [139]</td>
<td></td>
</tr>
<tr>
<td>Nutrition biomarkers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VII. Doubly Labeled Water (DLW)</td>
<td>Yes [140]</td>
<td>- Cold standard measurement of energy expenditure in free-living conditions, including overweight and obese [140]</td>
<td>- Assumes participant or group is weight stable (energy expenditure = energy intake) [98]</td>
<td>- Consideration of the dose enrichment ratio or exclusion from study must be given for participants traveling considerable distance (&gt; 500 miles) from study site during weeks prior to dose [98]</td>
</tr>
<tr>
<td>Measurement of total energy expenditure through oral dose of isotope labeled water</td>
<td>- Low participant burden [140]</td>
<td>- Requires sophisticated technology and highly trained staff [140]</td>
<td>- Body composition are important associated measurements for determination of energy balance [98]</td>
<td></td>
</tr>
<tr>
<td>VIII. Urinary nitrogen excretion</td>
<td>Yes [141]</td>
<td>- Valid indicator of dietary protein intake in free-living populations; useful for identifying misreporting [141]</td>
<td>- Very expensive [140]</td>
<td>- Eight 24-hour complete urine collections needed to estimate individual protein intake within 2X [98]</td>
</tr>
<tr>
<td>Measurement of protein intake through 24-hour urine collection</td>
<td>- Tends to classify underreporters same as DLW [98]</td>
<td>- Applies universal correction factor for nitrogen loss through skin and feces, but actual loss may have wide variability due to fiber intake and exercise [98]</td>
<td>- Must confirm urinary collection is complete using external marker, such as tablets of para-amino-benzoic acid (PABA) [141]</td>
<td></td>
</tr>
<tr>
<td>Emerging technologies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automated self-administered 24 h dietary recall (ASA24)</td>
<td>No</td>
<td>- Available at no cost to researchers, clinicians and teachers [142]</td>
<td>- Vulnerable to participant altering eating patterns</td>
<td>- Currently available for use with limited analysis capabilities of individual foods and daily total nutrients. Future plans include analysis reports in Pyramid equivalents, supplement use, a Spanish version for respondents, and optional modules to assess supplement intake, salt intake, with whom each meal was eaten, television viewing while eating, and where food was obtained [142]</td>
</tr>
</tbody>
</table>
| Internet-based multi-pass 24 h recall | - Respondent completes on own time or at time scheduled by researcher [142] | - Users must be very computer-literate and have high-speed internet access [142] | | (continued on next page)
validated in a 12-country study with reasonable measurement properties for monitoring population levels of physical activity among adults, and results at least as good as other established PA surveys [159]. The IPAQ produces higher estimates of physical activity compared to a shorter version described below [159]. No studies examining the impact of weight status on the questionnaire accuracy were identified.

4.5.3. Short 7-day IPAQ (IPAQ-S7)

The IPAQ-S7 is a 9-item questionnaire designed primarily for surveillance and comparison between populations. It is available in telephone or self-administered formats and provides results based on current recommendations for moderate and vigorous activity. The IPAQ-S7 is generally preferred by respondents and interviewers over the full-length IPAQ. There is no difference in reliability and validity between the short and long IPAQ forms [159]. The 7-day IPAQ may lead to overestimation of physical activity in obese populations, and needs further investigation before validity is established [156].

4.5.4. Physical Activity Recall Questionnaire (PAR-Q)

The PAR-Q is designed to estimate habitual PA [109], and can estimate energy expenditure using metabolic equivalent calculations [157]. Fourteen-items assess duration of sleep, moderate, hard, and very-hard intensity PA. Though studied in general populations with obese individuals [157], the impact of weight status has not been reported. In a small, but general population, the PAR-Q significantly overestimated energy expenditure compared to DLW. Awarding a lower intensity to hard- and very-hard activity may reduce overestimation with this tool. The PAR-Q is therefore not recommended for estimation of individual or small group energy expenditure, but may be appropriate for large epidemiological studies.

4.6. Behavioral observation

Direct behavioral observation by trained observers is a possible method for small samples of short durations, when contextual information is particularly important. The many disadvantages make the use of this method now rare. These include an extensive time requirement, potential bias with presence of observer, and subjective
classification of activity and intensity. No validation studies with DLW have been completed [151].

5. Body composition

5.1. Introduction

The human body is composed of fat and fat-free compartments and body composition assessment involves the accurate measurement of one or many of these compartments. Body composition can be assessed at the molecular, cellular, and tissue levels [160] using several different methods. Evaluating body composition of obese individuals is necessary both in research and clinical practice [161] to determine health as well as disease risk. It is well known that high amounts of body fat are associated with a greater risk of developing type 2 diabetes, cardiovascular disease, cancer, and renal failure [162]. However, assessing body composition in the obese is challenging because obesity is marked by an increase in body fat and changes in body composition different from that of a non-obese person. There is an increase in total body hydration and a relative expansion of the extracellular water (ECW) component compared to intracellular water. Due to these physiological changes the assumptions used to assess body composition in normal weight individuals, including density of tissues, concentrations of water and electrolytes, biological interrelationships between body tissues and distributions do not apply for obese persons [162] and can affect the accuracy of body composition tools [163]. The goal of this section is to review what methods are available to researchers and clinicians and to identify which are the best options to assess body composition in the obese population.

5.2. Body Mass Index (BMI)

BMI is a proportion of height to weight (weight in kilograms/height in meters squared) and is the most widely used measure for determining the prevalence of obesity [164]. A BMI above 30 kg/m² is classified as obese, with sub-classifications of Class I 30–34.9, Class II 35–39.9, and Class III ≥ 40 [165]. There is extensive national reference data for BMI [164], and because it is thought to correlate highly with percent body fat (%BF), BMI is considered an accurate indicator of body composition [166,167]. However, in obesity BMI and %BF do not have a strong correlation. Obesity is identified as >25% BF in men, and >30% BF obese in women [168]. When BMI is compared with these parameters, more than 50% of individuals who have body fat outside the %BF cutoffs do not have a BMI >30 [164]; BMI is an acceptable tool for screening for obesity and tracking weight over time. However, since it does not separate body compartments into fat-free mass and fat mass [169] or identify the distribution of fat [170] it should not be used to further assess body composition in the obese beyond classifying the level of obesity.

5.3. Anthropometrics

Anthropometry is the most basic method for assessing body composition and is used to determine body mass, size, shape, and level of fatness [160]. Measurements include height, weight and circumferences of the waist, hip, head, and neck measured with a flexible quilting tape. Body composition is assessed using these variables in standardized regression equations [171]. Anthropometric measurements are considered easy, safe, and inexpensive for assessing obesity [171]. However accuracy is dependent on the skills and training of the person taking the measurements [172] and can vary from observer to observer [173]. Specific limitations of anthropometry in the obese include the inability to distinguish subcutaneous fat from visceral adipose tissue, which is helpful to assess disease risk [160,174] and accuracy may be lowered in the severely obese, due to difficulties finding the actual waistline or drooping abdominal fat can interfere with hip measurement [161].

5.4. Skin fold thickness (SKF)

SKF is a tool used to assess body fat stores by measuring subcutaneous fat in specific locations on the body. The most common sites of measurement are the bicep, tricep, subscapular, and supra-iliac [175,176]. SKF measurements are incorporated into regression equations to predict total body fat [169,177]. SKF has limits in the general population including observer variability, elasticity of fat and skin tissue (which vary with age and between individuals), and discomfort the participant may feel during measurements [178]. All of these limitations are applicable to the obese population. Furthermore, the thickness of adipose tissue in obese participants may make it difficult to raise a skinfold that will provide an accurate measure [177]. The calipers used to measure SKF are often too small [177], especially when used on the abdominal and thigh folds of obese participants [177]. Larger calipers are available [162] but are much more difficult for the researcher to use due to their width, and consequently have a greater risk of error [162]. If a participant has edema, it further complicates the accuracy of SKF, because the degree of compression of the caliper can differ from location to location, resulting in uneven and inaccurate measurements [178]. Finally, SKF accuracy in the obese is related to the prediction equation used. Most of the current equations were developed in normal weight individuals, and have not been validated in the obese [161].

5.5. Bioelectrical impedance analysis (BIA)

BIA measures the body impedance using electrodes that are connected from one leg to the other, or to the arm, to form a circuit for the current to pass through. The impedance measure is used to predict total body water (TBW) and fat-free mass (FFM) and fat mass is calculated from the difference between weight and FFM. Different tissues offer varying resistance, with adipose tissue a poor conductor of the current because of its low water content [179]. Muscle tissue, which has higher water content, offers less resistance and is able to better conduct the current [179]. Body composition assessment by BIA is attractive because it requires little equipment, is inexpensive, noninvasive [180], and has no weight or height restrictions [181]. The method is safe, and there is no risk from frequent measurements [163]. Single frequency BIA (SF-BIA) should not be used for body composition assessment in the obese because the theory that the human body is a single cylinder with constant resistivity cannot be applied to the obese [181]. In addition, the frequency of the current applied (50 kHz) in SF-BIA is not high enough to penetrate all tissues [182]. Segmental BIA (teta- and eight-polar-BIA) recognizes the human body is complex in shape and combines several impedance measures together for a more accurate assessment [181]. However, segmental-BIA has been found to significantly overestimate %BF in obese adults [181]. Multi frequency-BIA (MF-BIA) allows multiple frequencies to assess fluid distribution [181]; low electric frequencies (e.g. 1 or 5 kHz) measure ECW and high frequencies (e.g. 100, 200, or 500 kHz) measure TBW [183]. MF-BIA has been found to overestimate %BF in the overweight and obese groups [181], significantly underestimate both total and truncal fat in obese women [184], and offer accurate estimates of TBW and ECW in women with a BMI up to 48.2 kg/m² [185]. More research needs to be conducted to determine an agreement on the use of MF-BIA in the obese before recommendations can be made.

Prediction equations developed in normal-weight subjects for BIA are based on the assumption that the hydration of FFM is a constant factor of 73.2% [182]. However, in obesity this hydration of FFM has found to be higher (approximately 77.5%) [161,186–188]. The different body build of obese also affects the accuracy of BIA, as obese subject typically have a trunk that is short, and large in diameter, leading to a different body water distribution form lean individuals [182]. As a result of these variances found in obesity, when prediction equations developed in non-obese populations are used to assess body composition in obese participants, they underestimate body composition.
fat [177,182,183,189,190]. Fatness-specific BIA equations, developed by Segal et al. have been validated for use in the obese [191] and more recently developed prediction equations specifically for the obese population are more accurate for prediction of body fat [192] and have been discussed in detail elsewhere [192–195].

5.6. Bioimpedance spectroscopy (BIS)

BIS assesses body composition by measuring TBW, differentiating between ECW and intracellular water (ICW) [196] by using a range of frequencies (5–1000 kHz) [183]. The method is noninvasive and inexpensive [197]. For an accurate BIS measurement, the participants' limbs must be completely away from the body such that they are not touching the trunk. This can be difficult in extreme obesity, leading to an overestimation of fluid volumes and BIS has been found to overestimate TBW and ECW [198]. At this time, the method has proven to be inaccurate when used to assess body composition in the obese [88,101].

5.7. Dilution technique

In the dilution technique, deuterium labeled water ($^2$H$_2$O) is used to obtain a measure of TBW to calculate an individual's FFM [161,193]. A known dose of isotope, based on the individual's weight, is provided for the participant to drink [199]. TBW volume can be assessed from the dilution spaces of the $^2$H$_2$O and then converted to kilograms using the conversion factor of 0.99336 (density of water at normal body temperature) [161]. The technique is limited to use in research because of its cost, time needed [98] required specialized equipment, and highly trained personnel [161], which make it unrealistic for routine clinical use. When the technique is used in an obese participant, there is an underestimation of FFM and overestimation of fatness [161] because of the high hydration found in obesity. For normal-weight individuals the average proportion of TBW in FFM is 73%, but in the obese it may be as high as 80%. This percentage increases with an increase in adiposity leading to more inaccurate measures. Population specific values for FFM hydration should be determined to further enhance the accuracy of the dilution technique.

The addition of Intravenous Sodium Bromide with the deuterium dilution technique to determine ECW provides a more accurate measure of ECW. ECW is calculated from the increase in bromide concentration between baseline and mean post dose blood samples and the amount of bromide injected after applying appropriate correction factors [161]. Studies have shown that bromide is able to equilibrate within a four hour period in the extracellular space in severe obesity [161].

5.8. Total body potassium (TBP)

Measuring potassium, the most abundant intracellular ion, with a whole-body counter [200] allows researchers to calculate body cell mass. Body cell mass is the metabolically active portion of the human body; quantifying it allows researchers to assess FFM and metabolism [160,200]. Measuring TBP over time in the obese has been found to be an acceptable method to monitor weight [200], but potassium content of FFM may be affected by hydration related changes, specifically in severe obesity [161]. This method is not recommended for assessment of body composition because of high price of the counter [200] and lack of strong support for accuracy.

5.9. Hydrostatic weighing

Hydrostatic weighing, also known as hydrodensitometry, estimates body composition by combining body weight, body volume, and residual lung volume. The original hydrodensitometry method requires complete submersion. However hydrostatic weighing without head submersion has also been developed [202] with comparable accuracy [189]. Hydrostatic weighing is impractical in clinical settings [189]. The test is time consuming [203], labor intensive [160], and often involves difficult maneuvers such as holding breath underwater and is highly reliant on participant performance [192]. Although hydrostatic weighing is done in the obese, the test may also cause discomfort and apprehension for some individuals [161] due to physical and technical constraints. Given the many disadvantages of this method hydrodensitometry, is not a widely recommended method for measuring body composition in obese participants.

5.10. Air Displacement Plethysmography (ADP)

ADP measures body volume by measuring air displacement. The machine used is a dual-chamber unit with a testing chamber for the participant to sit in and a reference chamber which holds the breathing circuit, electronics, and pressure transducers [181,204,205]. The tool is highly sensitive to changes in body volume, is valuable for trending small changes in body composition, is quick to perform, has low participant burden [206], and is noninvasive [181]. The ADP method is validated in the obese, including the extremely obese patients with BMI over 40 [161,206,207]. Obese participants are able to easily learn and perform the correct breathing techniques needed for accurate measurement [206]. For measurement, participants must wear minimal, tight fitting clothing (ideally a swimming suit) and a swimming cap to compact the participants' hair [192]. The clothing requirement for the ADP may limit its use in the moderate to severely obese and in certain ethnic groups. However, its ease and speed make ADP a favorable option for measuring body composition in the obese.

5.11. Dual Energy X-ray Absorptiometry (DEXA)

DEXA is a scanning technique that measures bone mineral, fat tissue, and fat-free soft tissue. Participants must lie completely still on the DEXA machine platform while X-rays at a high and low energy levels are passed over the body [201,208]. DEXA can be used to determine abdominal obesity [209] and is useful in predicting intra-abdominal fat in obese men and women [210]. DEXA also provides assessment of regional body composition in allowing for the identification of gynoid or android obesity [211]. Limitations of DEXA include its high cost, need for trained technicians, and dedicated facilities [212]. In obese participants, the DEXA scan is sensitive to difference in body thickness resulting in an overestimation of body fat [201]. As the tissue gets thicker, especially over 20 cm, there is an increased degree of beam hardening, which involves the preferential attenuation of the lower energy X-rays [208]. The instrument itself may also limit its use for measuring body composition of obese individuals, Traditional DEXA scan tables can only hold up to 300 lb and the width of the scanning area, average of 60 cm, does not accommodate the obese or severely obese [160,161,213]. It has been demonstrated that an accurate whole-body composition assessment can be predicted from a half-body scan for participants who are too large to fit on the traditional DEXA scan table [214]. A recently developed iDXA half body analysis, that is able to hold up to 400 lb with an increased scanning width of 66 cm, and scanning height of 46 cm, has been shown to provide body composition analysis for fat-mass, non-bone lean mass, and percent fat that is comparable to whole-body analysis [213].

5.12. Computerized Tomography (CT) scan

CT scan uses an X-ray beam to produce cross sectional images of the body, allowing differentiation between measured muscle mass, visceral organ volumes [176], and measures of visceral adipose tissue in overweight and obese patients [160,195]. CT scans at the whole body level involve high radiation exposure [160]. Single cross
sectional images taken at specific abdominal locations can be used to assess total body adiposity, visceral adipose tissue as well as skeletal muscle mass in healthy adults and are a more cost effective option and reduce radiation exposure [222,223]. However, when compared to multi-slice imaging, the reference method [210], the single slice method is not as accurate in detecting small changes in abdominal adiposity [224], because fat loss in the abdominal region is not uniform and should not be used to assess total abdominal fat loss [223]. It is also important to note that, single abdominal slice images provide good estimates of total body adiposity, visceral adiposity and skeletal muscle in group studies [224,225], but have limited applicability at the individual level due to individual variation [210]. A computer aided non-contrast CT can detect pericardial fat and thoracic fat, a risk factor for atherosclerosis [226], CT can also be used to quantify fat content in skeletal muscle, but in the obese this can be more difficult due to higher levels of adipose tissue surrounding muscle [227]. Although CT and magnetic resonance imaging (MRI) are currently the best methods for analyzing regional adiposity these machines are expensive [195] and are usually limited to the hospital setting [176]. The high risk, combined with high cost, make the CT scan an unattractive option for routine clinical use for assessing body composition [169] in all participant populations.

5.13. Magnetic Resonance Imaging (MRI)

MRI is a technique of generating images from interactions between the nuclei of hydrogen atoms in the body and magnetic fields generated by the MRI machine. Protons from the various tissues in the body resonate differently. The MRI recognizes these differences, generating an image of the tissues. The generated image can be used to measure body composition [176] and to examine regional fat distribution. MRIs can accurately differentiate between fat and muscle in all populations [228], measure intramyocellular lipids in skeletal muscle [215], and quantify total body lipid [176]. MRI and MRS (magnetic resonance spectroscopy) are used clinically for detection and quantification of hepatic fat, [215,216], helping to diagnose fatty liver disease [217] and type 2 diabetes [218]. More recent studies have found that single slice images at a predetermined area of the abdomen allow for a fast and reliable estimation of visceral and total adipose tissue [219]. This is particularly important in the context of assessing risk factors for diabetes. However, while single slices may be useful for cross sectional estimation of volumes of relevant fat tissue compartments it is important to note that single slice imaging may not be sensitive or accurate in detecting small changes in abdominal adiposity [223]. Additionally, visceral adipose tissue is composed of subcompartments that are largely different both in metabolic and functional properties and the traditional CT and MRI protocols are not capable of separating all of the compartments, specifically intraperitoneal from intraabdominal [220]. While technical advances are clearly needed in this area the interpretations of current scans should come with a clear definition of the type of visceral adipose tissue.

There are no known long-term side effects from MRIs so they can be used for large coverage and repeated tests [221]. Use in the obese was previously limited by the size of the MRI machines, which were not able to accommodate large body sizes [160]. The developments of open-configuration MRI scanners have helped resolve this problem [161]. MRI is a good option for assessing body composition in the obese.

5.14. Near-Infrared Interactance (NII)

NII measures body fat by assessing the absorption of infrared light. The amount of absorption and reflection of the infrared light is related to the composition of the underlying tissue [186]. A signal penetrates underlying body tissue up to 1 cm, usually on the bicep and total body fat can be calculated by a prediction equation [201]. Error from using only one site on the body to measure body fat is likely [201]. The prediction equations used for this method have been found to underestimate body fat with increasing adiposity [201] leading to an underestimation of %BF in the obese [186]. This underestimation may be a result of the NII beam being affected by the irregularities in the fat-muscle junction and fat layering with increased adiposity [186]. NII has not been validated in the obese.

5.15. Three-Dimensional Photonic Scanning (3DPS)

In 3DPS a scanner captures body surface topography [229] measuring surface geometry using digital techniques [230]. The different 3DPS techniques that have been successfully developed for assessing body composition include photogrammetric technique, laser technique, and stereovision technique [229,231,232]. A scanner generates millions of points over a scan field, and then software connects the dots creating a 3-D body image including values on total and regional body volumes [231]. Measurements of waist and hip circumference, sagittal abdominal diameter, segmental volumes, and body surface area [232] are generated in seconds [231]. %BF can be calculated using a prediction equation [231]. 3DPS accurately measures body shape, including in the obese [229,231], and is attractive for use because it is safe to use frequently, requires no special conditions, and does not require intensive technical support [229]. Participants must wear tight fitting clothes and stand still for 10 s [231]. Monitoring the body shape measurements of obese individuals over time can help track patients weight gain or loss overtime, and the photonic scanner is able to display within person change over time [229]. 3DPS is a good option to use for both clinical and research assessment of body composition [229], with practical use in public health as well [232].

5.16. Quantitative Magnetic Resonance (QMR)

QMR is an emerging method for body composition assessment. The tool measures differences in nuclear magnetic resonance properties of hydrogen atoms to divide signals originating from fat, lean tissue, and free water [233]. The test, originally developed and tested on mice [234], has recently been adapted and scaled for human use. QMR is able to detect small changes in fat mass superior to DEXA and four-compartment models [233,235] but when quantifying total fat mass, there is some discrepancy when compared to the four-compartment model [233]. The method it is quick (less than 3 min) [233] and has been shown to have promise for body composition assessment in the obese.

5.17. Multi-compartment methods

Multi-compartment models account for the fact that the human body is composed of different compartments including fat mass and fat free mass (water, muscle, protein, bone, minerals). Combinations of two, three, four, or more of the previously discussed methods to measure body composition are often used in a multi-compartment model. They are considered the reference for body composition, and therefore must avoid major assumptions and have maximal precision [236]. Many single body composition assessments are based on assumptions, like the assumption of standard hydration or FFM density and the assumption of constant hydration in fat-free soft tissues in DEXA [236]. The most basic two-compartment model (2-C) methods are based on major assumptions like the water or potassium constancy in FFM [236]. A three-compartment model allows for improvement over a two-compartment model, but it does not rely on these assumptions of standard hydration or FFM density [237]. Three-compartment models may combine ADP, BIA and TBW and has been developed in moderate to severely obese [161]. Four-compartment methods include measurements of fat, water, mineral, and protein, for
example combining the measurements from ADP, deuterium oxide (TBW), and DEXA (bone mineral mass) [238]. Multi-compartment models are useful for measuring body fat in the obese. However, multi-compartment methods rely on the accuracy of the different measurements that are combined and an error in one of the measurements will result in an inaccurate body composition assessment.

5.18. Conclusion

This review of the different tools shows that there are several options for assessing body composition in the obese. Many of the tools and methods reviewed have their limitations and should be used with caution. Emerging methods have more promise for accurate assessment and need to be validated for use in the obese. When choosing a method for assessing body composition, researchers or clinicians should consider what resources they have available, their budget, and the goals of their assessment. With the increasingly large number of obese people in the world, body composition assessment will continue to be important in both in identifying the most effective treatments of obesity and in the evaluation of patients’ health.

Acknowledgement

This article was produced as part of the First Tufts University Seminar on the Obesity Epidemic and Food Economics that was organized by Drs. Emmanuel N. Pothos (Chair), Robin B. Kanarek and Susan B. Roberts in Boston and Medford, MA, USA.

References


The obesity epidemic poses a global concern because of its related health problems. Obesity increases the risk for diabetes mellitus (DM), hypertension, coronary heart disease, dyslipidemia, certain cancers, and obstructive sleep apnea, among other comorbidities. Although overweight and obesity are associated with an increased risk for DM and cardiovascular disease, the pattern of fat distribution is also important in risk stratification. Intra-abdominal (visceral) fat deposition is linked to an increased risk for developing DM and cardiovascular disease compared with a more peripheral (subcutaneous) fat distribution. The accurate measurement of visceral fat requires the use of advanced methods not available in clinical practice such as computed tomography (CT) or magnetic resonance imaging (MRI). On the other hand, a simple measurement of waist circumference (WC) is strongly correlated with risk for DM and cardiovascular disease. Increased WC is one of the diagnostic criteria for the metabolic syndrome, according to the National Cholesterol Education Program (NCEP) Adult Treatment Panel III (ATP III), and is essential for the diagnosis of the metabolic syndrome, according to the International Diabetes Federation (IDF).

This review summarizes the scientific data linking an increased WC to increased cardiovascular, DM, and mortality risk and the importance of its assessment in clinical practice.

**Measurement of Body Adiposity**

Body fat is composed of cells that contain triglyceride stores. Body fat can be divided into 2 main compartments: subcutaneous and intra-abdominal. Fat may also accumulate at ectopic sites such as muscle, including the heart, pancreas, and liver.

There are many ways to measure total body fat; traditionally, the gold standard for body fat estimation has been hydrodensitometry (underwater weighing), based on the fact that fat tissue is less dense than muscle and bone. Dual-energy X-ray absorptiometry (DXA) is another method for evaluating body composition. These 2 methods are used mainly in research and are not readily available for clinical practice or epidemiological studies.

Imaging studies such as CT, MRI, and DXA may evaluate fat partitioning, normally with most of the fat located in the subcutaneous space and less in the visceral space and most of the visceral fat located in the intra-abdominal space (intraperitoneal and retroperitoneal). The amount of visceral fat is influenced by gender, age,
race/ethnicity, physical activity, and total adiposity. Computed tomography and MRI are now the gold standards for the evaluation of the abdominal subcutaneous and visceral adipose tissue. A single slice at the L4-L5 intervertebral level is used to measure the amount of fat, expressed in cubic cm. It is possible that the intramuscular fat is correlated more to the visceral fat than the subcutaneous fat. The deleterious effect of fat deposition in the liver is well known. Nonalcoholic steatohepatitis (NASH) is a common disorder known to correlate with insulin resistance, increased triglyceride production leading to an atherogenic dyslipidemia, type 2 DM, and cirrhosis. Intrahepatic fat deposition can be evaluated easily by ultrasound (US) of the liver, but US is not very sensitive in the earlier stages of steatosis, and it does not quantify lipid content. Computed tomography is used to measure hepatic steatosis, whereas the most accurate tool for quantifying lipid storage in the hepatocytes is achieved by nuclear magnetic resonance (NMR). Pancreatic fat infiltration is also associated with decreased insulin secretion and β-cell apoptosis. 

**Anthropometric Measurements**

**Body Mass Index**

Body mass index (BMI), as weight in kg divided by the square of the height in meters (or weight in pounds multiplied by 703, then divided by height in inches), is an easy way to calculate body fatness and is widely used in clinical practice. Both the World Health Organization (WHO) task force and the National Institutes of Health (NIH) guidelines stipulate normal, overweight, and obesity cutoffs according to BMI. A BMI between 18.5 and 24.9 kg/m² is considered normal, between 25 and 29.9 kg/m² is overweight, and >30 kg/m² is obese (Table 1). 

Because these cutoffs are arbitrary, there is some dispute whether normal weight should be defined as a BMI of < 25 kg/m². There has been an increase in the prevalence of DM and mortality with increasing BMI in patients within normal range. This trend may be explained by the pattern of fat distribution, with a more central and particularly visceral fat accumulation being associated with increased risk for DM and cardiovascular disease; when added to BMI in patients with a BMI between 25 and 35 kg/m², this improves health risk assessment. Recent data failed to show an increase in mortality in overweight individuals compared with individuals of normal weight, possibly due to improved cardiovascular risk management, whereas obese and underweight individuals had higher mortality rates. The association between increased gluteofemoral adipose tissue and cardiovascular risk is much weaker one. Another pitfall of relying on BMI alone is that muscular patients may have relatively high BMIs but low fat mass, and in the elderly, BMI may underestimate fat mass due to a decrease in lean body mass.

**Waist Circumference**

As mentioned earlier, body fat distribution is an important risk factor for obesity-related disease. Increased visceral adipose tissue (also called central or abdominal fat) is associated with increased risk for cardiometabolic disease.

Precise measurement of the visceral fat component is expensive and not feasible in clinical practice. Therefore, WC is used as a surrogate marker of abdominal fat mass. WC correlates with subcutaneous and visceral fat mass and is related to increased cardiometabolic risk. A WC >40 inches (102 cm) in men and >35 inches (88 cm) in women is considered to confer an increased cardiometabolic risk. These WC cutoffs were derived from Caucasian men and women living in the Netherlands to correlate with WC values associated with a BMI ≥30 kg/m², therefore identifying subjects who should be advised to lose weight. The National Heart, Lung, and Blood Institute (NHLBI) recommended measuring WC along with BMI to assess patient risk stratification in subjects with a BMI between 25 and 35 kg/m². Although not incorporated in the guidelines, a comment recommends measuring WC even in patients with normal BMI to improve risk stratification. The cutoffs for normal WC chosen by the NHLBI are as mentioned earlier: WC >40 inches (102 cm) in men and >35 inches (88 cm) in women, independent of BMI or ethnicity. According to the NHANES III, 1% of men and 14% of women with a normal BMI had a high WC, and in the overweight range (BMI between 25 and 29.9 kg/m²), 25% of men and almost 70% of women had a high WC. Shen et al evaluated BMI, WC, and percent fat measured by DXA in 1010 healthy white and African American men and women. WC had the strongest association with health risk, followed by BMI and percent fat.

<table>
<thead>
<tr>
<th>Classification for Body Mass Index (BMI)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Underweight</td>
<td>&lt;18.5 kg/m²</td>
</tr>
<tr>
<td>Normal weight</td>
<td>18.5-24.9 kg/m²</td>
</tr>
<tr>
<td>Overweight</td>
<td>25-29.9 kg/m²</td>
</tr>
<tr>
<td>Obesity (class 1)</td>
<td>30-34.9 kg/m²</td>
</tr>
<tr>
<td>Obesity (class 2)</td>
<td>35-39.9 kg/m²</td>
</tr>
<tr>
<td>Extreme obesity (class 3)</td>
<td>&gt;40 kg/m²</td>
</tr>
</tbody>
</table>

Waist-to-Hip Ratio

Waist-to-hip ratio (WHR) is a common anthropometric index used to assess abdominal obesity. The WHO included a high WHR defined as a ratio >0.9 in men and >0.85 in women as a criterion for diagnosing the metabolic syndrome. A high WHR has been associated in epidemiological studies with increased cardiovascular and DM risk. Because both WHR and WC are measures of abdominal adiposity, they both correlate with increased risk of cardiovascular events. Some studies have found that WHR is a more accurate tool to diagnose patients with a higher cardiovascular risk compared with WC and BMI, whereas other studies have found a better correlation of cardiometabolic risk with WC compared with WHR, particularly in women. The Interheart Study, an epidemiological, retrospective case control study in different populations, included more than 27,000 people from 52 countries, with the study variable being a history of myocardial infarction (MI). In this study, BMI, WC, and WHR were all highly significantly associated with the risk for MI, but WHR was the strongest anthropometric predictor of MI in men and women, independent of other risk factors or BMI category.

How Should Waist Circumference Be Measured?

Different anatomic landmarks used to measure WC include the midpoint between the lowest rib and the iliac crest, the umbilicus, the narrowest or widest WC, just below the lowest rib, and just above the iliac crest. The NHLBI practical guide recommends measuring just above the iliac crest. WC measurement should be done with the patient standing, bare midriff, after the patient exhales, with both feet touching and arms hanging freely. The measuring tape should be made of a material that is not easily stretched; it should be placed perpendicular to the long axis of the body and horizontal to the floor and applied with tension, but it should not put pressure on the abdominal wall (Figure 1).

A recent systematic review by Ross et al evaluated whether the site of WC measurement influenced its relationship to morbidity and mortality. The most commonly used landmarks were the minimal waist (33%), the midpoint between the lowest rib and the iliac crest (26%), and the umbilicus (27%). According to their review, WC protocol did not influence outcome concerning all-cause and cardiovascular disease (CVD) mortality, CVD, and DM. Therefore, it seems that the site of measurement does not influence clinical outcome.

Visceral Adipose Tissue

Adipose tissue is now viewed as an active endocrine organ and not merely as an energy depot. The visceral adipose tissue (VAT) accounts for approximately 15% of total body fat in the lean subject and includes the intraperitoneal (mesenteric and omental) fat, which drains into the portal circulation, and retroperitoneal fat, which drains into the systemic circulation. There is a correlation between BMI and WC with total adipose tissue mass, but WC has a higher correlation with VAT than BMI alone.

Because the WC measures the subcutaneous abdominal tissue (SAT) and the VAT, it is not possible to discern between the 2 compartments when measuring WC (Figure 2). The etiology of increased cardiometabolic risk in patients with increased VAT is not known, but there are many theories on the subject. Increased levels of free fatty acids (FFAs) due to obesity and to increased VAT worsen insulin resistance. The VAT is more insulin resistant, fueling the hyperlipolytic state and worsening insulin resistance. Although the VAT drains into the portal vein, most of the FFA drained to the liver is derived from upper body non-VAT fat. The increase in VAT mass possibly reflects the lack of capacity of the subcutaneous tissue to store the energy excess that accumulates in the liver, muscle, and pancreas, worsening insulin resistance. Numerous cytokines are secreted from the adipose tissue, including proinflammatory molecules such as interleukin-6 (IL-6) and tumor necrosis factor-α (TNF-α). Plasma levels of C-reactive protein...
CRP), an inflammatory factor produced by the liver and known to correlate with atherosclerosis, are increased in patients with visceral obesity.\textsuperscript{41,42} It has been recently shown that the adipose tissue is infiltrated by macrophages, contributing to the inflammatory process.\textsuperscript{43,44} Adiponectin, a protein derived from the adipose tissue, is known to improve insulin sensitivity and may protect against atherosclerosis.\textsuperscript{45,46} Adiponectin levels are decreased in obesity and in viscerally obese patients.\textsuperscript{46,47} Another contributing factor may be local generation of cortisol from cortisone due to increased activity of 11β-hydroxysteroid dehydrogenase in the VAT, which may further increase fat deposition, worsen insulin resistance, and increase cardiometabolic risk.\textsuperscript{48}

The Metabolic Syndrome

The metabolic syndrome consists of a cluster of risk factors strongly associated with an increased risk for atherosclerotic cardiovascular disease (ASCVD) and type 2 DM. The metabolic risk factors consist of a specific pattern of hyperlipidemia (elevated serum levels of triglycerides and apolipoprotein B, small low-density lipoprotein [LDL] particles, and low serum high-density lipoprotein [HDL] cholesterol levels), elevated blood pressure, elevated plasma glucose concentration, a proinflammatory state, and a prothrombotic state.\textsuperscript{5,49} Possibly the most important risk factors for the development of the syndrome are abdominal obesity and insulin resistance.

The NCEP ATP III held in 2001 provided diagnostic criteria for the metabolic syndrome that could be easily implemented in clinical practice. These criteria include WC, blood pressure measurements, and serum levels of triglycerides, HDL-C, and fasting glucose. The criteria were further updated in 2005 by an executive summary by the American Heart Association and NHLBI. One change in the diagnostic criteria is a decrease in normal serum glucose values to < 100 mg/dL instead of < 110 mg/dL, and although WC cutoffs were not changed, there is a recommendation that some white, black, and Hispanic adults with marginally increased WC (e.g., 94-101 cm [37-39 in] in men and 80-87 cm [31-34 in] in women) may benefit from changes in lifestyle, similar to subjects with higher WC and that a lower cutoff for WC (≥ 90 cm [35 in] in men and ≥ 80 cm [31 in] in women) appears to be appropriate for Asian Americans because this population is more prone to insulin resistance.

The IDF has adopted the same diagnostic criteria for the metabolic syndrome but proposed an adaptation to WC according to ethnic group, due to different risk in different populations. According to the IDF criteria, an increased WC is essential for the diagnosis of the metabolic syndrome because it incorporates both insulin resistance and abdominal obesity. The presence of 2 other criteria is required for the diagnosis of the metabolic syndrome: elevated blood pressure, elevated plasma glucose concentration, or elevated triglyceride levels.\textsuperscript{7,50} The WHO criteria require a diagnosis of insulin resistance (glucose intolerance, impaired glucose tolerance [IGT], or DM and/or insulin resistance measured under hyperinsulinemic euglycemic conditions). Table 2 compared the different criteria required for the diagnosis according to the 3 different groups.

There has been some dispute whether the metabolic syndrome can be called a syndrome or if it adds to risk stratification of patients beyond the regular risk factors (eg, hyperlipidemia, hypertension, DM).\textsuperscript{51} Because obesity is the main culprit for the epidemic proportion of the metabolic syndrome, those in favor of the definition point out that it is more important to focus on lifestyle changes that may improve all components of the syndrome instead of on each individual risk factor.\textsuperscript{52}

Waist Circumference and Type 2 Diabetes

The diabetes epidemic is closely linked to the obesity epidemic.\textsuperscript{53} In the Nurses’ Health Study, increases in BMI within normal range were associated with a 4-fold risk of developing type 2 DM (women with BMI <22 kg/m\textsuperscript{2} vs women with BMI 23-25 kg/m\textsuperscript{2}).\textsuperscript{18} The relationship between WC and DM is very strong. WC is a stronger predictor of DM than is BMI, and it is independent of traditional risk factors such as hypertension, lipoproteins, and glucose levels.\textsuperscript{11} In the International Day for the Evaluation of Abdominal Obesity (IDEA) study, which included 69,409 subjects from 63 countries who consulted a primary care physician, BMI and particularly WC were strongly linked to DM.\textsuperscript{54}
Waist Circumference, Cardiovascular Disease, and Mortality

The correlation between CVD risk and an increased WC is well known, although it seems to be less than the correlation between WC and DM. Values for WC are related to the risk for developing CVD and remain statistically significant after adjusting for BMI and other cardiovascular risk factors, consistent in men and women.11,36,55-57

All-cause mortality is also related to WC, even when adjusted for BMI. For any given BMI, an increased WC confers a higher mortality risk. A recent study evaluated the association between WC and mortality among 154,776 men and 90,757 women aged 51 to 72 years at baseline and followed for 9 years. All-cause mortality was higher in patients with increasing WC, even when adjusted for BMI and other covariates. Men with increased WC had a 22% increased mortality risk and women a 28% increased mortality risk. This finding was consistent in all races and among smokers and nonsmokers.60

Table 2. Metabolic Syndrome Definitions

<table>
<thead>
<tr>
<th>Study Group</th>
<th>WHO&lt;sup&gt;3b&lt;/sup&gt;</th>
<th>NCEP ATP III&lt;sup&gt;b&lt;/sup&gt;</th>
<th>IDF&lt;sup&gt;7,46&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic criteria</td>
<td>Glucose intolerance, IGT, or DM and/or insulin resistance together with</td>
<td>≥3 of the following risk factors:</td>
<td>Central obesity (waist circumference) ≥9 cm for Europid men ≥80 cm for Europid women Ethnicity-specific values for other groups&lt;sup&gt;a&lt;/sup&gt; Plus presence of ≥2 of the following:</td>
</tr>
<tr>
<td></td>
<td>≥2 of the following:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting plasma glucose</td>
<td>≥100 mg/dL</td>
<td></td>
<td>100 mg/dL (or diagnosed DM)</td>
</tr>
<tr>
<td>Serum triglycerides</td>
<td>≥150 mg/dL</td>
<td>≥150 mg/dL</td>
<td>≥150 mg/dL (or diagnosed DM)</td>
</tr>
<tr>
<td>Serum HDL-cholesterol</td>
<td>Men: &lt;35 mg/dL</td>
<td>Men: &lt;40 mg/dL</td>
<td>Men: &lt;40 mg/dL</td>
</tr>
<tr>
<td></td>
<td>Women: &lt;39 mg/dL</td>
<td>Women: &lt;50 mg/dL</td>
<td>Women: &lt;50 mg/dL</td>
</tr>
<tr>
<td>Hypertension</td>
<td>≥140/≥90 mm Hg</td>
<td>≥130/≥85 mm Hg</td>
<td>≥130/≥85 mm Hg</td>
</tr>
<tr>
<td>Obesity</td>
<td>Men: waist-hip ratio &gt;0.9</td>
<td>Men: waist circumference &gt;40 in (102 cm)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women: waist-hip ratio &gt;0.85 and/or BMI ≥30 kg/m&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Women: waist circumference &gt;35 in (88 cm)</td>
<td></td>
</tr>
<tr>
<td>Microalbuminuria</td>
<td>Albumin/creatinine &gt;30 mg/g or urinary albumin excretion rate ≥20 μg/min</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DM, diabetes mellitus; HDL, high-density lipoprotein; IDF, International Diabetes Federation; IGT, impaired glucose tolerance; NCEP ATP III, National Cholesterol Education Program, Adult Treatment Panel III; WHO, World Health Organization.

Ethnic specific cutoffs: South Asians men ≥90 cm, women ≥80 cm; Chinese men ≥90 cm, women ≥80 cm; Japanese men ≥85 cm, women ≥90 cm (there is no agreement concerning the appropriate cutoff for the Japanese population).

Waist Circumference, Cardiovascular Disease, and Mortality

Pediatric obesity is a serious medical problem worldwide. In the U.S., 19% of children between 6 and 11 years old are overweight, with BMI greater than the 95th percentile for their age and sex.61 Cardiometabolic risk factors are already present in this population (hypertension, dyslipidemia, and impaired glucose tolerance). The use of age- and gender-specific cutoffs (≥90th percentile) for WC seems to be a reliable marker for the detection of high-risk overweight pediatric patients.62 In this population, the use of WHR (normal ≤0.5) may be easier in clinical practice because it does not require age-specific tables.

Ethnicity

Ethnicity is known to have an impact on CVD and DM risk, and because abdominal obesity is a better predictor of cardiovascular disease than is BMI, the use of ethnic-specific WC seems appropriate. For example, it is known
that Asians have a higher risk for DM and CVD with lower BMIs and WC, corresponding to higher VAT for a given BMI. African Americans, on the other hand, have lower VAT compared with Caucasians of the same BMI and WC.

The IDF has adopted a lower cutoff WC for Caucasians compared with the NCEP ATP III, corresponding to a BMI of 25 kg/m², and includes ethnic- and gender-specific cutoffs (Table 2). There are still populations (sub-Saharan Africans, ethnic South and Central Americans, Eastern Mediterranean, and Middle East populations) for which ethnic-specific cutoffs for WC have not been specified; these should adopt the Caucasian cutoffs for the time being.7

Weight Loss

Weight loss is known to improve insulin resistance, DM control, hypertension, and lipid profile. Lifestyle changes, including weight loss and physical activity, have been shown to prevent or delay type 2 DM in high-risk patients. In the Diabetes Prevention Program study, lifestyle intervention was associated with a marked reduction in visceral fat (measured by CT during the first year of the study) and a smaller decrease in subcutaneous fat, body weight, BMI, and WC in both men and women. Pare et al73 evaluated the change in visceral fat, subcutaneous fat, and waist girth followed by weight loss in obese men. Their results showed a greater relative reduction in visceral fat than subcutaneous fat, suggesting greater mobilization of visceral fat following weight loss in viscerally obese men.

Weight loss achieved through diet and pharmacotherapy with sibutramine, orlistat, and rimonabant has shown a decrease in WC and improvement in cardiometabolic risk. Rimonabant therapy has shown a higher decrease in WC than expected for weight loss, possibly due to its effect on the cannabinoid receptor 1 (CB1) in the VAT.

Summary

WC reflects the magnitude of the abdominal fat tissue, including subcutaneous and visceral adipose tissue. Although it is a simple and crude measurement, it correlates with VAT and, when increased, has been shown to confer an increased risk for DM, CVD, and mortality, independent of BMI. It is now clear that different ethnic groups should be assessed according to their race-specific cutoffs to improve risk evaluation. Further study is needed to specify ethnic-specific WC values. WC measurement should be implemented as a vital sign to improve cardiometabolic risk assessment.

References


76. Pi-Sunyer FX, Aronne LJ, Heshmati HM, Devin J, Rosenstock J, for the RIONASG: Effect of rimonabant, a cannabinoid-1 receptor blocker, on weight and cardiometabolic risk factors in overweight or obese patients: RIO-North America: a randomized controlled trial. *JAMA*. 2006;295:761-775.