

Eighty-five per cent of what? Discrepancies in the weight cut-off for anorexia nervosa substantially affect the prevalence of underweight

J. J. Thomas^{1,2,3*}, C. A. Roberto¹ and K. D. Brownell¹

¹ Department of Psychology, Yale University, New Haven, CT, USA

² Klarman Eating Disorders Center, McLean Hospital, Harvard Medical School, MA, USA

³ Eating Disorders Clinical and Research Program, Massachusetts General Hospital, Boston, MA, USA

Background. DSM-IV cites <85% of expected body weight (EBW) as a guideline for the diagnosis of anorexia nervosa (AN) but does not require a specific method for calculating EBW. The purpose of the present study was to determine the degree to which weight cut-off calculations vary across studies, and to evaluate whether differential cut-offs lead to discrepancies in the prevalence of individuals who are eligible for the AN diagnosis.

Method. Two coders independently recorded the EBW calculation methods from 99 studies that either (a) compared individuals with AN to those with subclinical eating disorders or (b) conducted AN treatment trials. Each weight cut-off was applied to a nationally representative ($n=12001$) and treatment-seeking ($n=189$) sample to determine the impact of EBW calculation on the proportion who met the AN weight criterion.

Results. Coders identified 10 different EBW methods, each of which produced different weight cut-offs for the diagnosis of AN. Although only 0.23% of the national sample met the lowest cut-off, this number increased 43-fold to 10.10% under the highest cut-off. Similarly, only 48.1% of treatment seekers met the lowest cut-off, whereas 89.4% met the highest.

Conclusions. There is considerable variance across studies in the determination of the AN weight cut-off. Discrepancies substantially affect the proportion of individuals who are eligible for diagnosis, treatment and insurance reimbursement. However, differences may not be fully appreciated because the ubiquitous citation of the 85% criterion creates a sense of false consensus.

Received 2 October 2007; Revised 22 June 2008; Accepted 12 July 2008; First published online 8 September 2008

Key words: Anorexia nervosa, expected body weight, body mass index, DSM-IV criteria for eating disorders, Centers for Disease Control and Prevention (CDC), National Health and Nutrition Examination Survey (NHANES).

Introduction

Determining whether a patient is underweight is a crucial step in eating disorder evaluation. Indeed, 'refusal to maintain body weight at or above a minimally normal weight for age and height (e.g. weight loss leading to maintenance of body weight less than 85% of that expected)' is listed as the first diagnostic criterion for anorexia nervosa (AN) in DSM-IV (APA, 2000, p. 589). Although the 85% weight cut-off is intended to represent a 'suggested guideline' for diagnosis (APA, 2000, p. 584), investigators who enroll

eating disorder patients in clinical trials (Dare *et al.* 2001; Powers *et al.* 2002) and insurance companies that determine treatment eligibility typically adhere to this percentage when assessing underweight. The 85% criterion is also frequently used to calculate AN prevalence in epidemiological studies (Walters & Kendler, 1995; Garfinkel *et al.* 1996), which inform the perceived public health significance of the disorder. The widespread use of the 85% criterion probably reflects a desire to standardize diagnosis across diverse settings. However, because DSM-IV provides only general guidelines on expected body weight (EBW) calculation, researchers and clinicians have used several different methods to create the denominator of the 85% equation, including various versions of the Metropolitan Life Insurance Tables (Metropolitan Life Insurance Company, 1959, 1983) and the 1979 Department of Health, Education, and Welfare standards (DHEW, 1979). The degree to which these methods

* Address for correspondence: J. J. Thomas, Ph.D., Eating Disorders Clinical and Research Program, Massachusetts General Hospital, WAC-816D, 15 Parkman St, Boston, MA 02114, USA.

(Email: jjthomas@partners.org)

This paper was presented at the Association for Cognitive and Behavioral Therapies Convention on 15–18 November 2007 in Philadelphia, PA, USA.

converge remains unknown, and the field may not fully appreciate the potential impact of EBW calculation on differential diagnosis because the ubiquitous citation of the 85% criterion creates a sense of false consensus.

Data from clinical and non-clinical samples suggest that eating disorder not otherwise specified (EDNOS) is the most prevalent of DSM-IV eating disorders, and individuals who meet all criteria for AN except the weight cut-off represent a common subtype of this group (Watson & Andersen, 2003; McIntosh *et al.* 2004). A computer simulation of 193 eating-disorder treatment seekers indicated that the prevalence of AN would increase significantly if the weight criterion were relaxed from 85% to 90% of EBW (Thaw *et al.* 2001). It is therefore likely that if some clinics use more lenient methods of calculating EBW, they will diagnose a greater proportion of their patients with AN and a relatively smaller proportion of patients with EDNOS, even if they consistently apply an 85% cut-off.

The calculation of EBW plays an important role in eating disorder treatment as well as diagnosis. According to recent guidelines produced by the National Institute for Clinical Excellence (NICE, 2004), evidence-based treatments differ substantially across eating disorders, which are in part defined by degree of underweight. NICE strongly recommends individual cognitive behavioral therapy for the treatment of adult bulimia nervosa (BN) (individuals >85% EBW) and tentatively recommends family-based treatment for adolescent AN (individuals <85% EBW). However, in the absence of data, no firm recommendations could be made for the treatment of EDNOS (individuals of variable body weight). Furthermore, the American Psychiatric Association Work Group on Eating Disorders suggests that in-patient hospitalization or residential treatment should be considered for eating disorder patients who weigh <85% of healthy body weight (APA, 2006), and a major third-party health-care provider requires that patients weigh <80% of EBW in order to receive residential treatment reimbursement. The decision to move from one phase of treatment to another is also informed by percentage EBW. Howard *et al.* (1999) recommended discharging AN in-patients from the hospital at 90% of EBW, and Lock *et al.* (2001) used the same 90% threshold to determine when to shift food choices from parents to patients in their family-based AN treatment. Thus, differential EBW calculation across clinical sites could result in eating disorder patients with identical height-and-weight profiles receiving very different treatment approaches, despite attempts to adhere to evidence-based practice.

Therefore, the purpose of the present study was threefold: (1) to identify different methods investigators have used for determining whether patients meet the weight criterion for AN, (2) to evaluate the degree of discrepancy across methods, and (3) to determine the extent to which these discrepancies impact the proportion of individuals who meet the weight criterion for AN in both population-based and treatment-seeking samples. Because diagnosing underweight requires different procedures for children *versus* adults (i.e. children must be evaluated with growth charts that account for projected height and weight increases over time), the present study focused on the assessment of the AN weight criterion among individuals aged ≥ 18 years.

Method

Study population and inclusion criteria

To identify studies that provided descriptive information on how AN diagnoses are determined in clinical and research settings, we targeted two distinct empirical literatures. First, we identified studies comparing AN to subclinical eating disorders (i.e. EDNOS) because clinicians seeking to assign full *versus* partial eating disorder diagnoses must assess each diagnostic criterion. EDNOS studies were included if they applied the same diagnostic methods to assess (and thus differentiate between) AN and EDNOS subjects, and assessed current rather than past symptoms (to make it more likely that diagnostic methods would be specifically described in the article). Second, we identified treatment outcome studies of adult AN because investigators must apply the AN weight criterion to evaluate trial eligibility and treatment efficacy. As the 85% criterion was introduced in DSM-III-R, EDNOS studies and AN treatment trials were eligible for inclusion if they were published between January 1987 and February 2007 and used DSM-IV, DSM-III-R or ICD-10 criteria for AN. Only English-language reports were included.

Study search strategy

EDNOS studies

Because the EDNOS literature has not yet been comprehensively reviewed, we conducted an original literature search. Eligible studies were identified by four steps. First, five electronic databases (PsycINFO, Medline, EMBASE, PubMed and CINAHL) were searched with the terms 'EDNOS' and 'eating disorder not otherwise specified'. Four databases that feature the capability to search for adjacent words within the body of an article (PsycINFO, Medline,

EMBASE and CINHAI) were additionally queried with the terms 'eating disorders', 'anorexia', 'bulimia' and 'binge eating disorder' adjacent within five words to the terms 'atypical', 'partial', 'residual', 'subclinical', 'subthreshold', 'subsyndromal', 'continuum', 'unspecified', 'non-specified', 'NOS' or 'non-classified'. Second, all issues published between January 1987 and February 2007 of the four journals determined by the SCOPUS database to publish the greatest number of eating disorder studies were hand-searched for eligible articles. These journals included the *International Journal of Eating Disorders*, *European Eating Disorders Review*, *Eating and Weight Disorders*, and *American Journal of Psychiatry*. Third, the online database Interdisciplinary Dissertations & Theses was queried with 'anorexia nervosa', 'EDNOS' and 'eating disorders not otherwise specified' to locate unpublished studies. Fourth, the reference sections of studies retrieved through these first three methods were searched for eligible citations. At the end of this process, 88 EDNOS articles met eligibility criteria for the present study.

AN treatment trials

Because this literature has already been comprehensively reviewed, we identified a list of controlled and uncontrolled psychotherapy and medication treatment trials for AN from two recent reviews (Le Grange & Lock, 2005; Bulik *et al.* 2007). Of these, 11 articles evaluating treatments for adult AN were included in the present study. (This gave an overall total of 99 articles included in the present study, available in the online Appendix.)

Study coding

Two master's-level clinical psychology doctoral students coded the 99 studies by identifying the method investigators had used to determine whether subjects met the weight criterion for AN. Coders agreed on the methods used in 93 (94.0%) of the 99 studies and came to a mutual consensus through discussion on the remaining six (6.0%), inter-rater reliability $\kappa = 0.90$.

Database of AN weight thresholds

After ascertaining how each study calculated the AN weight cut-off, we re-created a distribution of AN weight thresholds for individuals of each height and sex. To calculate 85% of EBW, we referred to the original normative weight tables cited by study authors. When weight tables provided ranges rather than point estimates, the midpoint of the range was defined as the EBW. When tables provided different weight ranges for small, medium and large frames, the

midpoint of the medium-frame range was used because approximately 50% of the population is classified as medium frame (Metropolitan Life Insurance Company, 1983). If study authors used age-adjusted weights, we created separate sets of weight thresholds for each age group. Because some tables provided clothed and others provided unclothed weights, all weights were standardized to represent weight without clothes and height without shoes[†]. Once we adjusted for age and clothing, we multiplied the distribution of EBWs derived through each method by 0.85 to determine the respective AN weight thresholds.

Participants

To determine the impact of discrepancies in the AN weight threshold on the proportion of individuals meeting the weight criteria for AN, we applied each set of weight thresholds to a nationally representative and treatment-seeking sample.

Nationally representative sample

Participants were drawn from the publicly available National Health and Nutrition Examination Survey (NHANES) 1999–2004 database, which provides a representative sample of the non-institutionalized civilian US population (CDC, 2007). Of the 12962 participants aged 18–65 years who underwent the NHANES medical examination, 12001 (mean age = 38.4 years, s.d. = 14.60) provided height and weight data. Participants included 5651 (47.1%) males and 6350 (52.9%) females. Twenty-six per cent self-identified as Mexican American, 25.9% as non-Hispanic White, 22.3% as other Hispanic, 21.9% as non-Hispanic Black, and 4.2% as multi-racial or other race. Health technicians measured participants' weight in paper gowns using a digital scale, and participants' height using a digital stadiometer. The National Center for Health Statistics Institutional Review Board reviewed and approved the data collection, and written informed consent for NHANES was obtained from each participant.

Treatment-seeking sample

Participants comprised 189 females aged 18–65 (mean = 28.62, s.d. = 8.39) years who telephoned the Eating Disorders Research Unit at the New York State Psychiatric Institute seeking treatment for AN from January 2005 to March 2007. Participants seeking treatment exclusively for BN or binge eating disorder were excluded. Participants self-reported their height, weight and other clinical information during a

[†] The notes appear on p. 841.

telephone screen. When subjects reported a weight range, we recorded the average as a point estimate. Because of the brevity of telephone screening, ethnicity data were unavailable. The Institutional Review Board at the New York State Psychiatric Institute approved the collection of these data and their use in the present study.

Data analysis

For participants in both samples, body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared (kg/m^2). After BMI was calculated, measurements were converted to the Imperial system and heights were rounded to the nearest inch for use with normative weight tables. To assess the degree of correlation among weight cut-offs produced through each method, we calculated a series of bivariate correlation coefficients. To evaluate discrepancies across methods, we subtracted the lowest cut-off from the highest cut-off produced for each height, and calculated the mean and standard deviation of these values².

Participants whose height fell outside the range provided by a normative weight table were excluded from analyses using that particular weight table because no AN weight cut-off could be calculated. To assess the prevalence of individuals in the NHANES sample who met the weight cut-off for AN under each method, we used the SAS PROC SURVEYMEANS procedure (SAS Institute, Cary, NC, USA) to account for oversampling of minority groups, survey non-response, and other stratification factors. We calculated 6-year sampling weights for use in all statistical analyses so that results are reflective of the demographic breakdown of 2000 US Census data. To evaluate whether the proportion of participants who met the AN weight criterion differed by EBW calculation, we conducted a series of McNemar tests to assess for significant differences between dependent proportions. We set the overall α level to 0.001 to provide a Bonferroni correction for family-wise error rate across the (nine choose two) 36 unique pairwise comparisons within each sample.

Results

Methods of determining the AN weight threshold

Sixty-three (63.6%) of the 99 articles in the study population did not describe how authors calculated EBW, and a further nine studies reported using percentage of EBW but did not cite a specific table. Another three studies (Vandereycken & Pieters, 1992; Klibanski et al. 1995; Cachelin & Maher, 1998) cited the Metropolitan Life Insurance Tables but did not specify

which version, and one study (Fairburn et al. 2005) described using BMI but did not cite which value.

The remaining 23 studies (22.3%) described their methods in sufficient detail so that specific weight cut-offs could be recalculated for use in the present study. Seven of the 23 studies (Clinton & Norring, 1999; Lee et al. 2001, 2003; Solenberger, 2001; Turner & Bryant-Waugh, 2004; McIntosh et al. 2005; Abbate-Daga et al. 2007) created an absolute BMI cut-off, and another study (Strokosch et al. 2006) described using the 10th percentile BMI for gender and age based on Hebebrand et al. (1996). The other 15 studies (Lee et al. 1993; Gowers et al. 1994; Schork et al. 1994; Fullerton et al. 1995; Carlat et al. 1997; Attia et al. 1998; Schaefer et al. 1998; Mizes et al. 2000, 2004; Kaye et al. 2001; Williamson et al. 2002; Pike et al. 2003; Miller et al. 2005; Levine et al. 2007; Roberto et al., 2008) calculated 85% of EBW based on specific tables of norms, including Kemsley's (1951/2) Average Body Weights; 1959 Metropolitan Life Insurance Tables; 1983 Metropolitan Life Insurance Tables; the 1975 Fogarty Table of Desirable Weights (Bray, 1975); 1979 Department of Health, Education, and Welfare norms (DHEW, 1979); and Chiu's (1978) norms for Chinese adults. In sum, coders identified 10 distinct methods for determining whether individuals met the weight criterion for AN (see Table 1 for a brief description of each method). Because one study (Lee et al. 2003) used weight tables available only in Chinese (Chiu, 1978), the following analyses are based on nine different English-language methods cited in the recent literature as representing possible weight cut-offs for AN.

Degree of discrepancy across methods

Figs 1 and 2 depict AN weight cut-offs derived through each method across a full range of heights for females and males. Because all methods produced progressively higher cut-offs for individuals of increasing heights, cut-offs correlated positively with one another across methods for both males and females (all r 's ≥ 0.92). However, point estimates differed substantially. Table 1 provides example AN weight cut-offs for a 20-year-old female of average height (5' 4") and a 20-year-old male of average height (5' 9"). The mean difference between the lowest and the highest AN weight cut-off for each height was 15.03 (s.d. = 2.38) lb for females and 25.88 (s.d. = 5.29) lb for males.

Proportion of the nationally representative sample meeting the weight criterion for AN

Table 2 presents the percentage of individuals in the nationally representative sample who met the AN

Table 1. Nine methods used in the recent literature for calculating the AN weight criterion and the cut-offs derived from each method for a 20-year-old female and male of average height

Method type	Investigators who used this method	Method description	AN weight cut-off (lb)	
			5' 4" female	5' 9" male
Body mass index (BMI)	Lee <i>et al.</i> (2003)	BMI < 16.5	96.14	111.74
	Lee <i>et al.</i> (2001); Solenberger (2001); Turner & Bryant-Waugh (2004); McIntosh <i>et al.</i> (2005); Abbate-Daga <i>et al.</i> (2007)	BMI < or ≤ 17.5	101.96	118.52
Age-adjusted BMI at 10th percentile	Clinton & Norring (1999)	BMI < 18.0	104.88	121.90
	Strokosch <i>et al.</i> (2006)	BMI < 18.4	107.21	133.42
1951/2 Kemsley Average Body Weights	Gowers <i>et al.</i> (1994)	85% of graduated mean weight	103.70	117.30
1959 Metropolitan Life Insurance Tables	Schork <i>et al.</i> (1994); Attia <i>et al.</i> (1998); Kaye <i>et al.</i> (2001); Pike <i>et al.</i> (2003); Miller <i>et al.</i> (2005); Levine <i>et al.</i> (2007); Roberto <i>et al.</i> (2008)	85% of mean of medium-frame range	104.98	124.10
1975 Fogarty Table of Desirable Weights	Schaefer <i>et al.</i> (1998)	85% of average desirable weight	102.00	126.65
1979 Department of Health, Education, and Welfare	Mizes <i>et al.</i> (2000, 2004)	85% of weight at 50th percentile (age 18–24)	110.50	136.85
1983 Metropolitan Life Insurance Tables	Fullerton <i>et al.</i> (1995); Carlat <i>et al.</i> (1997); Williamson <i>et al.</i> (2002)	85% of mean of medium-frame range	111.35	129.20

AN, Anorexia nervosa; EBW, expected body weight; lb, pounds.

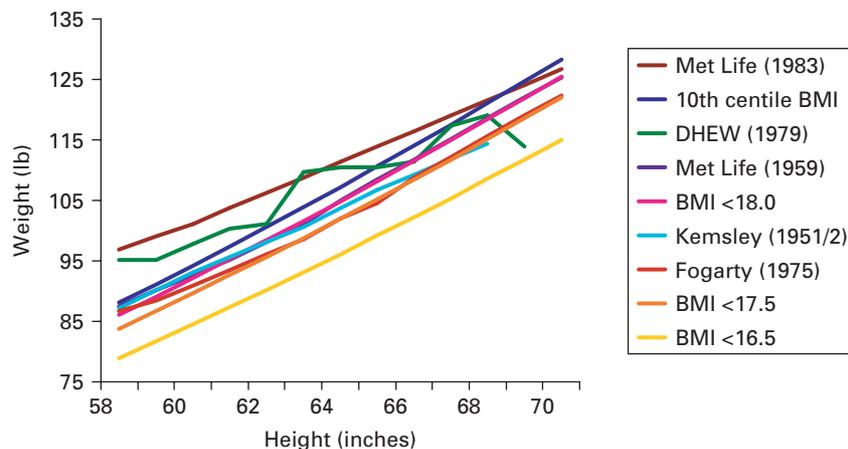


Fig. 1. Anorexia nervosa (AN) weight cut-offs (y axis) for females 4' 10" to 5' 11" (x axis) ascertained through nine different methods recently used in the empirical literature.

weight criterion under each of the nine methods. Using the lowest weight cut-off (BMI < 16.5), only 0.23% met the weight criterion, whereas the prevalence increased 43-fold to 10.10% under the highest cut-off (DHEW, 1979). McNemar tests for all 36 unique pairwise comparisons indicated that the majority of

methods produced proportions that differed significantly from one another at $\alpha = 0.001$ (see Table 2). The only methods that produced statistically indistinguishable proportions in the nationally representative sample were BMI < 18.0 and Fogarty (1975); BMI < 18.0 and Kemsley (1951/2); Kemsley (1951/2) and

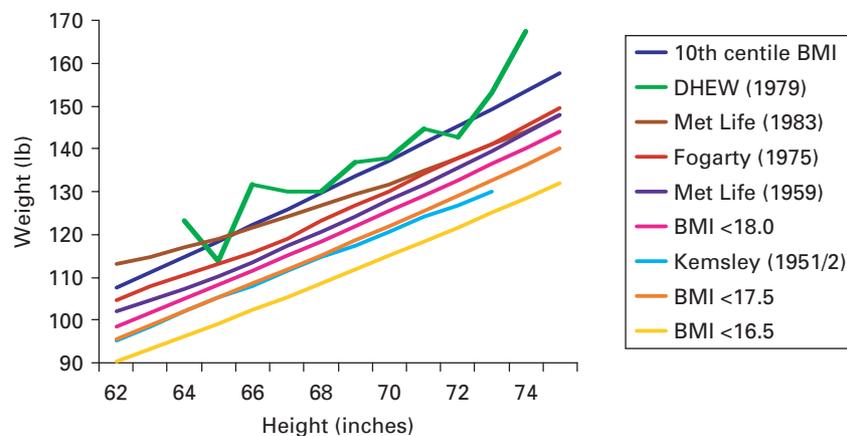


Fig. 2. Anorexia nervosa (AN) weight cut-offs (*y* axis) for males 5' 2'' to 6' 3'' (*x* axis) ascertained through nine different methods recently used in the empirical literature.

Fogarty (1975); Kemsley (1951/2) and Metropolitan Life Insurance (1959); and Fogarty (1975) and Metropolitan Life Insurance (1959) (all p 's > 0.001).

Proportion of the treatment-seeking sample meeting the weight criterion for AN

Table 2 also displays the percentage of individuals in the treatment-seeking sample who met the weight criterion for AN under each method. Forty-eight per cent met the weight criterion using the lowest cut-off (BMI < 16.5), whereas nearly twice as many (89.4%) met the criterion using the highest cut-off (DHEW, 1979). Pairwise McNemar tests demonstrated that BMI < 16.5 classified a significantly lower proportion of treatment seekers as underweight than all other methods (all p 's < 0.001). The proportion meeting BMI < 17.5 was similar to BMI < 18.0, Kemsley (1951/2) and Fogarty (1975) (all p 's > 0.001) but significantly lower than all other methods (all p 's < 0.001), with the exception of BMI < 16.5. The proportion with BMI < 18.0 did not differ from Fogarty (1975), Kemsley (1951/2) and the 1959 Metropolitan Life Insurance Tables but was significantly higher than BMI < 16.5 and BMI < 17.5 (all p 's < 0.001). The 1983 Metropolitan Life Insurance Tables, age-adjusted BMI at the 10th percentile, and DHEW (1979) produced higher proportions than the other six methods (all p 's < 0.001) but did not differ significantly from one another (all p 's > 0.001).

Discussion

There is considerable variation across studies in the determination of the weight cut-off for AN diagnosis. Most of the 99 articles focusing specifically on distinctions between eating disorder diagnostic categories

and AN treatment efficacy did not report their methods for assessing degree of underweight. Of the 23 studies that did describe calculation methods, coders identified 10 distinct methods of establishing the weight criterion. Applying nine of these methods to nationally representative and treatment-seeking samples produced large and statistically reliable differences in the proportion of individuals who were classified as underweight. Our disparate prevalences highlight substantial discrepancies in the pool of individuals who would be eligible for the AN diagnosis if other diagnostic criteria were met.

The finding that investigators use different weight criteria for AN has important implications for eating disorder diagnosis, treatment, research and insurance reimbursement. Our results raise the possibility that a patient of a particular height, weight and symptom profile could receive a diagnosis of AN at one treatment center and a diagnosis of BN or EDNOS at another, and be eligible for one investigator's AN treatment outcome study but not another. On average, discrepancies are possible within a 15-lb weight range for females and a 25-lb weight range for males, and could occur even if the assessing clinicians at each treatment center referred to the same DSM-IV criteria to assign diagnoses. If each clinician then attempted to recommend an evidence-based treatment, the patient diagnosed by the stricter weight cut-off and therefore classified as BN or EDNOS might receive out-patient therapy whereas the patient diagnosed by the more lenient weight cut-off and therefore classified as AN might receive a more intensive intervention (e.g. in-patient care) because of the perception that he or she is more underweight.

Discrepancies in the application of the weight criterion may stem in part from a well-intentioned clinical desire to account for the unique presentation of each

Table 2. Nine methods used in the recent literature for calculating the AN weight cut-off and the percentage (\pm standard error) of a population-based and treatment-seeking sample meeting the AN weight criterion under each method

Method for determining AN weight cut-off	Population-based sample ($n = 1201$)		Treatment-seeking sample ($n = 189$)	
	Percentage (s.e.) meeting AN weight criterion	Comparable methods ($p_{\text{pairwise diff}} > 0.001$)	Percentage (s.e.) meeting AN weight criterion	Comparable methods ($p_{\text{pairwise diff}} > 0.001$)
BMI < 16.5	0.23 (0.05)	None	48.1 (3.6)	None
BMI < 17.5	0.91 (0.11)	None	70.9 (3.3)	BMI < 18.0 Kemsley (1951/2) Fogarty (1975)
BMI < 18.0	1.33 (0.13)	Kemsley (1951/2) Fogarty (1975)	76.2 (3.1)	BMI < 17.5 Kemsley (1951/2) Fogarty (1975) Met Life (1959)
Weight < 85% of 1951/2 Kemsley Average Body Weights	1.50 (0.11)	BMI < 18.0 Fogarty (1975) Met Life (1959)	71.0 (3.4)	BMI < 17.5 BMI < 18.0 Fogarty (1975) Met Life (1959)
Weight < 85% of Fogarty 1975 Average Desirable Weight	1.47 (0.11)	BMI < 18.0 Kemsley (1951/2) Met Life (1959)	71.4 (3.3)	BMI < 17.5 BMI < 18.0 Kemsley (1951/2) Met Life (1959)
Weight < 85% of 1959 Metropolitan Life Insurance Tables medium-frame mean	1.61 (0.14)	Kemsley (1951/2) Fogarty (1975)	77.0 (3.1)	BMI < 18.0 Kemsley (1951/2) Fogarty (1975)
Weight < 85% of 1983 Metropolitan Life Insurance Tables medium-frame mean	3.48 (0.19)	None	87.2 (2.4)	10th centile BMI DHEW (1979)
BMI < age-adjusted BMI at 10th percentile	6.64 (0.31)	None	84.9 (2.6)	Met Life (1983) DHEW (1979)
Weight < 85% of DHEW 1979 age-adjusted 50th percentile	10.1 (0.32)	None	89.4 (2.3)	Met Life (1983) 10th centile BMI

AN, Anorexia nervosa; BMI, body mass index; DHEW, Department of Health, Education, and Welfare; Met Life, Metropolitan Life Insurance Company.

Within each row, 'comparable methods' represent methods that *did not* exhibit significant differences in proportions based on pairwise McNemar tests with α set to 0.001 to control for family-wise error.

individual case. Indeed, DSM-IV encourages clinicians to account for patient variables such as height and age in the calculation of EBW. However, our data suggest that, at present, the application of the weight cut-off for AN varies at the level of the individual study rather than the level of the individual patient. The inconsistent application of myriad weight thresholds ultimately undermines rather than enhances the ideal of idiographic assessment. Each of the nine methods explored in the present study differs in the extent to which it accounts for patient variables (i.e. gender and age), and future research is needed to elucidate which variables may be most important to consider. For example, normative weight tables present different

weight ranges for each sex whereas BMI calculation remains constant across sex.

Similarly, DHEW (1979), Hebebrand *et al.* (1996) and Kemsley (1951/2) yield higher weight recommendations for older adults whereas the other six methods do not. Because weight gain is desirable from infancy to adolescence, age-specific guidelines are important for ascertaining degree of underweight in children. However, children grow at different rates, and it is unclear whether underweight should be defined nomathetically, as < 85% of the age-adjusted 50th percentile BMI or < the age-adjusted 5th percentile BMI, or idiographically, by comparing children to their own projected growth trajectories. Furthermore,

children reach developmental milestones (e.g. puberty, growth spurts) at different velocities, creating phase differences that may temporarily make late bloomers appear underweight. Moreover, many different versions of childhood growth charts are widely used. Future work should catalogue methods for ascertaining childhood underweight, evaluate whether proposed trajectories converge, and determine which are most appropriate for juvenile AN diagnosis. Although children who fail to make anticipated weight gains are classified as underweight in DSM-IV, the desirability of weight gain throughout the adult lifespan is less clear. Because adults typically continue to gain weight after achieving full stature, at least two normative weight tables (Kemsley, 1951/2; DHEW, 1979) provide higher expected weights for successively older adult age groups. However, available data suggest that even relatively modest weight gains (i.e. 11–22 lb) after age 18 are associated with increased risk for heart disease and hypertension (Willett *et al.* 1999). Thus, weight tables that provide normative adult weights graduated by age ranges may overdiagnose underweight among older adults. In sum, we recommend that the optimal weight cut-off for AN should increase with age until early adulthood but remain constant throughout the remainder of the lifespan.

The adoption of a mutually agreed upon weight cut-off for DSM-V would enhance the diagnostic reliability of AN. If a universal criterion were adopted, several considerations should factor into its selection, including ease of calculation, applicability to individuals of wide-ranging heights, and empirical relationship to morbidity and mortality. Normative weight tables exhibit many disadvantages from the standpoint of these criteria. First, some tables are difficult to interpret because they provide weight ranges rather than point estimates, and tables that provide clothed weights are not directly comparable to those that provide unclothed weights. Furthermore, Keys (1977) and others have criticized the Metropolitan Life Insurance Tables because their creators did not measure frame size, which directly informs weight recommendations, in the reference population. A second disadvantage of normative weight tables is that they do not provide weight recommendations for individuals of all possible heights, and therefore fail to classify very tall and very short individuals. In the current study, the DHEW and 1983 Metropolitan Life Insurance Tables methods could not classify 6.9% and 1.3% of NHANES participants respectively.

A final consideration when using normative weight tables is that recommended weights have increased over time; our results indicate that the 1983 Metropolitan Life Insurance Tables and the 1979 DHEW guidelines produce higher weight cut-offs than the

1959 Metropolitan Life Insurance Tables or Kemsley's 1951/2 Average Body Weights. This upward trend reflects the continued increase in obesity at the population level (Hedley *et al.* 2004). If 'expected' body weight continues to be equated with 'average' body weight in DSM-V, then the weight criterion for AN may continue to rise. Mean weights skewed to reflect normative overweight may lead clinicians and researchers to speciously pathologize individuals whose weights fall below a new, higher average, but who do not in fact experience increased morbidity and mortality. Taking a constant percentage of increasingly heavier average body weights could create a longitudinal drift in the AN phenotype that would greatly reduce the generalizability of extant knowledge to future research and clinical practice.

BMI cut-offs circumvent many of the limitations presented by normative weight tables. Specifically, BMI can be applied to persons of any height, it can be calculated unambiguously with a single formula, and the designation of a universal BMI cut-off would be invulnerable to upward pressures emanating from increasing population body weights. Indeed, the ICD-10 (WHO, 1992–1994) sets a BMI of ≤ 17.5 as the weight criterion for AN. Unfortunately, available data do not provide definitive evidence for one BMI cut-off (i.e. 16.5, 17.5 or 18.0) over another. Low weight may be confounded with smoking status or chronic disease in large population-based studies, thus artificially elevating associated mortality rates (Willett *et al.* 1999). Therefore, there is considerably less empirical support for defining 19.0 as the lower bound of the normative BMI range than for defining 25.0 as the upper bound (Willett *et al.* 1999). Multiple classes of studies are needed to determine which BMI cut-off would be most informative for the AN diagnosis. First, cross-sectional studies could use non-linear methods to identify whether a specific BMI is associated with discontinuities in eating pathology severity, functional impairment or physical complications among individuals with heterogeneous eating disorder presentations. Second, eating disorder treatment studies could stratify groups by proposed BMI cut-offs to conduct moderator analyses identifying the BMI at which treatment becomes least effective. Third, prospective studies could determine which of the proposed BMI cut-offs best differentiates between individuals with favorable *versus* unfavorable long-term outcomes.

The present study should be interpreted in light of the following limitations. First, three of the four diagnostic criteria for AN (amenorrhea, fat phobia and body image disturbance) were not assessed in either the nationally representative or treatment-seeking samples. Furthermore, the weight criterion for AN

requires that individuals deliberately refuse to maintain a minimally normal weight for height. Because AN represents only one of many reasons for underweight in the general population, including chronic disease and genetic factors, it is likely that the influence of weight cut-off calculation method would be diminished if all other diagnostic criteria were applied. A second limitation is that not all prospective patients in our treatment-seeking sample ultimately enrolled in clinical trials. Thus, the accuracy of their self-reported weights could not be assessed. A third limitation is that, because many of the articles in our study set did not describe their EBW calculation methods, it is possible that they used methods other than the 10 that we identified. Indeed, we are aware of adolescent-focused studies that have used still other calculation methods (e.g. 85% of the 50th BMI percentile for sex and age, cf. Peebles *et al.* 2006) that did not appear in the study population. It is also possible that authors who referenced the same normative tables arrived at different EBWs. For example, the Metropolitan Life Insurance Tables provide weight ranges rather than point estimates, and it would be defensible to define the lower weight limit, mean weight, or some other in-range weight as the 'expected' number. However, this possibility provides further support for our observation of the lack of consensus in the field.

In conclusion, our data indicate that investigators interpret the AN weight criterion in myriad ways, and their differential interpretations lead to significant discrepancies in the pool of individuals who are eligible for AN diagnosis. Unresolved discrepancies in the interpretation of the weight criterion could exert even greater influence on eating disorder diagnosis in the future if recommendations to omit the amenorrhea criterion (Mitchell *et al.* 2005) come to fruition in DSM-V. Such discrepancies also render recommendations to relax the AN weight criterion (Andersen *et al.* 2001; McIntosh *et al.* 2004) difficult to evaluate empirically. Altering the numerator of the EBW equation will have indeterminate impact if the denominator fluctuates across studies. Therefore, efforts to adopt a mutually acceptable weight cut-off for AN diagnosis would not only enhance short-term diagnostic reliability and treatment disposition but also inform long-term improvements to our nosological system.

Note

Supplementary material accompanies this paper on the Journal's website (<http://journals.cambridge.org>).

Acknowledgments

This project was sponsored by National Institute of Mental Health grant 1F31 MH078394 to J.J.T.

Declaration of Interest

None.

Notes

¹ To standardize across tables, we obtained the estimated heel height and clothing weight for each sample, and then subtracted these constants from the normative heights and weights provided. For Kemsley's (1951/2) Average Body Weights, we subtracted 1 inch of height for males and 1.5 inches for females, and 10 lb of weight for males and 6 lb for females. For the 1959 Metropolitan Life Insurance Tables, we subtracted 1 inch of height for males and 2 inches for females, and 7 lb of weight for males and 4 lb for females. For the 1983 Metropolitan Life Insurance tables, we subtracted 1 inch of height for both males and females, and 5 lb of weight for males and 3 lb for females. Because subjects in the DHEW (1979) sample wore paper examination gowns and foam slippers weighing less than 1 lb, we took heights and weights directly from DHEW tables without adjustment.

² Three methods (Kemsley, 1951/2; DHEW, 1979; Hebrand *et al.* 1996) provided separate weight thresholds by age group. To ensure that the weight cut-offs we compared would represent thresholds that would apply to the same hypothetical patient, we held age constant in Table 1 by using thresholds for the youngest age group only. Later, in the prevalence analyses, we applied each threshold according to its designated age group.

References

- Abbate-Daga G, Piero A, Gramaglia C, Gandione M, Fassino S (2007). An attempt to understand the paradox of anorexia nervosa without drive for thinness. *Psychiatry Research* **149**, 215–221.
- Andersen AE, Bowers WA, Watson T (2001). A slimming program for eating disorders not otherwise specified. *Psychiatric Clinics of North America* **24**, 1–9.
- APA (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, revised. American Psychiatric Association: Washington, DC.
- APA (2006). *Practice Guideline for the Treatment of Patients with Eating Disorders*, 3rd edn. American Psychiatric Association: Washington, DC.
- Attia E, Haiman C, Walsh BT, Flater S (1998). Does fluoxetine augment the inpatient treatment of anorexia nervosa? *American Journal of Psychiatry* **155**, 548–551.
- Bray GA (1975). *Obesity in Perspective*. DHEW Publication No. 708. U.S. Government Printing Office: Washington, DC.
- Bulik CM, Berkman ND, Brownley KA, Sedway JA, Lohr KN (2007). Anorexia nervosa treatment: a systematic review of randomized controlled trials. *International Journal of Eating Disorders* **40**, 310–320.
- Cachelin FM, Maher BA (1998). Is amenorrhea a critical criterion for anorexia nervosa? *Journal of Psychosomatic Research* **44**, 435–440.

- Carlat DJ, Camargo CA, Herzog DB** (1997). Eating disorders in males: a report on 135 patients. *American Journal of Psychiatry* **154**, 1127–1132.
- CDC** (2007). *National Health and Nutrition Examination Survey Data*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention: Hyattsville, MD (www.cdc.gov/nchs/about/major/nhanes/datalink.htm). Accessed 18 February 2007.
- Chiu CH** (1978). Standard weights for Chinese adults. *Journal of the Chinese Nutrition Society (Taiwan)* **3**, 85–94.
- Clinton D, Norring C** (1999). The rating of anorexia and bulimia (RAB) interview: development and preliminary validation. *European Eating Disorders Review* **7**, 362–371.
- Dare C, Eisler I, Russell G, Treasure J, Dodge L** (2001). Psychological therapies for adults with anorexia nervosa: randomised controlled trial of out-patient treatments. *British Journal of Psychiatry* **178**, 216–221.
- DHEW** (1979). *Weight by Height and Age for Adults 18–74 Years: United States 1971–1974*. DHEW Publication No. 79-1656. Series 11, No. 280. Department of Health, Education, and Welfare: Rockville, MD.
- Fairburn CG, Cooper Z, Doll HA, Davies BA** (2005). Identifying dieters who will develop an eating disorder: a prospective, population-based study. *American Journal of Psychiatry* **162**, 2249–2255.
- Fullerton DT, Wonderlich SA, Gosnell BA** (1995). Clinical characteristics of eating disorder patients who report sexual or physical abuse. *International Journal of Eating Disorders* **17**, 243–249.
- Garkfinkel PE, Lin E, Goering P, Spegg C, Goldbloom DS, Kennedy S, Kaplan AS, Woodside DB** (1996). Should amenorrhoea be necessary for the diagnosis of anorexia nervosa? Evidence from a Canadian community sample. *British Journal of Psychiatry* **168**, 500–506.
- Gowers S, Norton K, Halek C, Crisp AH** (1994). Outcome of outpatient psychotherapy in a random allocation treatment study of anorexia nervosa. *International Journal of Eating Disorders* **15**, 165–177.
- Hebebrand J, Himmelmann GW, Hesecker H, Schafer H, Remschmidt H** (1996). Use of percentiles for the body mass index in anorexia nervosa: diagnostic, epidemiological, and therapeutic considerations. *International Journal of Eating Disorders* **19**, 359–369.
- Hedley AA, Ogden CL, Johnson CL, Carroll MD, Curtin LR, Flegal KM** (2004). Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2002. *Journal of the American Medical Association* **291**, 2847–2850.
- Howard WT, Evans K, Quintero-Howard C, Bowers WA, Andersen A** (1999). Predictors of success or failure of transition to day hospital treatment for inpatients with anorexia nervosa. *American Journal of Psychiatry* **156**, 1697–1702.
- Kaye W, Toshihiko N, Weltzin T, Hsu G, Sokol M, McConaha C, Plotnicov K, Weise J, Deep D** (2001). Double-blind placebo-controlled administration of fluoxetine in restricting- and restricting-purging-type anorexia nervosa. *Biological Psychiatry* **49**, 644–652.
- Klibanski A, Biller B, Schoenfeld D, Herzog D, Saxe V** (1995). The effects of estrogen administration on trabecular bone loss in young women with anorexia nervosa. *Journal of Clinical Endocrinology and Metabolism* **80**, 898–904.
- Kemsley WFF** (1951/2). Body weight at different ages and heights. *Annals of Eugenics* **16**, 316–334.
- Keys A** (1977). Overweight and the risk of heart attack and sudden death. In *Obesity in Perspective* (ed. G. Bray), DHEW Publication No. 708. U.S. Government Printing Office: Washington, DC.
- Le Grange D, Lock J** (2005). The dearth of psychological treatment studies for anorexia nervosa. *International Journal of Eating Disorders* **37**, 79–91.
- Lee S, Chan YY, Hsu LK** (2003). The intermediate-term outcome of Chinese patients with anorexia nervosa in Hong Kong. *American Journal of Psychiatry* **160**, 967–972.
- Lee S, Ho TP, Hsu LK** (1993). Fat phobic and non-fat phobic anorexia nervosa: a comparative study of 70 Chinese patients in Hong Kong. *Psychological Medicine* **23**, 999–1017.
- Lee S, Lee AM, Ngai E, Lee DT, Wing YK** (2001). Rationales for food refusal in Chinese patients with anorexia nervosa. *International Journal of Eating Disorders* **29**, 224–229.
- Levine J, Gur E, Loewenthal R, Vishne T, Dwolatzky T, Van Beynum IM, Sela B-A, Vered I, Yoseff I, Stein D** (2007). Plasma homocysteine levels in female patients with eating disorders. *International Journal of Eating Disorders* **40**, 277–284.
- Lock J, Le Grange D, Agras WS, Dare C** (2001). *Treatment Manual for Anorexia Nervosa: A Family-based Approach*. Guilford Press: New York.
- McIntosh V, Jordan J, Carter F, Luty S, McKenzie J, Bulik C, Frampton C, Joyce P** (2005). Three psychotherapies for anorexia nervosa: a randomized, controlled trial. *American Journal of Psychiatry* **162**, 741–747.
- McIntosh VVW, Jordan J, Carter FA, McKenzie JM, Luty SE, Bulik CM, Joyce PR** (2004). Strict versus lenient weight criterion in anorexia nervosa. *European Eating Disorders Review* **12**, 51–60.
- Metropolitan Life Insurance Company** (1959). New weight standards for men and women. *Statistical Bulletin* **40**, 1–4.
- Metropolitan Life Insurance Company** (1983). Metropolitan height and weight tables. *Statistical Bulletin* **64**, 2–9.
- Miller K, Grieco K, Klibanski A** (2005). Testosterone administration in women with anorexia nervosa. *Journal of Clinical Endocrinology and Metabolism* **90**, 1428–1433.
- Mitchell JE, Cook-Myers T, Wonderlich SA** (2005). Diagnostic criteria for anorexia nervosa: looking ahead to the DSM-V. *International Journal of Eating Disorders* **37**, S95–S97.
- Mizes JS, Christiano B, Madison J, Post G, Seime R, Varnado P** (2000). Development of the Mizes Anorectic Cognitions Questionnaire-Revised: psychometric properties and factor structure in a large sample of eating disorder patients. *International Journal of Eating Disorders* **28**, 415–421.
- Mizes JS, Heffner M, Madison JK, Varnado-Sullivan P** (2004). The validity of subjective measures of body image disturbance. *Eating Behaviors* **5**, 55–66.
- NICE** (2004). *Eating Disorders – Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders*. NICE Clinical Guideline No. 9. National Institute for Clinical Excellence

- (<http://guidance.nice.org.uk/CG9/niceguidance/pdf/English>). Accessed 24 March 2007.
- Pike KM, Walsh BT, Vitousek K, Wilson GT, Bauer J** (2003). Cognitive behavior therapy in the posthospitalization treatment of anorexia nervosa. *American Journal of Psychiatry* **160**, 2046–2049.
- Peebles R, Wilson JL, Lock JD** (2006). How do children with eating disorders differ from adolescents with eating disorders at initial evaluation? *Journal of Adolescent Health* **39**, 800–805.
- Powers PS, Santana CA, Bannon YS** (2002). Olanzapine in the treatment of anorexia nervosa: an open label trial. *International Journal of Eating Disorders* **32**, 146–154.
- Roberto CA, Steinglass J, Mayer LE, Attia E, Walsh BT** (2008). The clinical significance of amenorrhea as a diagnostic criterion for anorexia nervosa. *International Journal of Eating Disorders* **41**, 559–563.
- Schaefer WK, Maclellan RN, Yaholnitsky-Smith SA, Stove ED** (1998). Psychometric evaluation of the eating disorder inventory (EDI) in a clinical group. *Psychology and Health* **13**, 873–881.
- Schork EJ, Eckert ED, Halmi K** (1994). The relationship between psychopathology, eating disorder diagnosis, and clinical outcome at 10-year follow-up in anorexia nervosa. *Comprehensive Psychiatry* **35**, 113–123.
- Solenberger SE** (2001). Exercise and eating disorders: a 3-year inpatient hospital record analysis. *Eating Behaviors* **2**, 151–168.
- Strokosch GR, Friedman AJ, Wu S-C, Kamin M** (2006). Effects of an oral contraceptive (norgestimate/ethinyl estradiol) on bone mineral density in adolescent females with anorexia nervosa: a double-blind, placebo-controlled study. *Journal of Adolescent Health* **39**, 819–827.
- Thaw J, Williamson D, Martin C** (2001). Impact of altering DSM-IV criteria for anorexia and bulimia nervosa on the base rates of eating disorder diagnoses. *Eating and Weight Disorders* **6**, 121–129.
- Turner H, Bryant-Waugh R** (2004). Eating disorder not otherwise specified (EDNOS): profiles of clients presenting at a community eating disorder service. *European Eating Disorders Review* **12**, 18–26.
- Vandereycken W, Pieters G** (1992). A large-scale longitudinal follow-up study of patients with eating disorders: methodological issues and preliminary results. In *The Course of Eating Disorders: Long-term Follow-up Studies of Anorexia and Bulimia Nervosa* (ed. W. Herzog, H.-C. Deter and W. Vandereycken), pp. 182–197. Springer-Verlag: New York.
- Walters EE, Kendler KS** (1995). Anorexia nervosa and anorexia-like syndromes in a population-based female twin sample. *American Journal of Psychiatry* **152**, 64–71.
- Watson TL, Andersen AE** (2003). A critical examination of the amenorrhea and weight criteria for diagnosing anorexia nervosa. *Acta Psychiatrica Scandinavica* **108**, 175–182.
- WHO** (1992–1994). *International Statistical Classification of Diseases and Related Health Problems*, 10th edn. World Health Organization: Geneva.
- Willett WC, Dietz WH, Colditz GA** (1999). Guidelines for a healthy weight. *New England Journal of Medicine* **341**, 427–434.
- Williamson DA, Womble LG, Smeets MAM, Netemeyer RG, Thaw JM, Kutlesic V, Gleaves DH** (2002). Latent structure of eating disorder symptoms: a factor analytic and taxometric investigation. *American Journal of Psychiatry* **159**, 412–418.