

DSM-IV Versus DSM-5: Implementation of Proposed DSM-5 Criteria in a Large Naturalistic Database

Andreas Birgegård, PhD*
Claes Norring, PhD
David Clinton, PhD

ABSTRACT

Objective: Problems with the current DSM-IV eating disorder (ED) section have resulted in proposed changes toward the upcoming DSM-5 (<http://www.dsm5.org/ProposedRevisions/Pages/EatingDisorders.aspx>). We investigated consequences of these by implementing the proposal in a large naturalistic database.

Method: Patients were 2,584 children/adolescents and adults enrolled at specialized ED clinics in Sweden. DSM-IV diagnoses anorexia nervosa, bulimia nervosa, and “not otherwise specified” examples were compared with DSM-5 anorexia, bulimia, and binge ED, as well as atypical anorexia, subthreshold bulimia, and binge eating, purging disorder, and the residual unspecified category. Assessment methods included a semistructured

diagnostic interview and self-ratings of ED and psychiatric symptoms.

Results: We studied age-separated diagnostic distributions and explained variance in clinical variables associated with the two systems. Results showed some improvement of diagnostic specification as well as a slight increase in explained variance.

Discussion: Remaining problems with the proposed changes were also highlighted, and possible further refinement is discussed. © 2011 by Wiley Periodicals, Inc.

Keywords: diagnostics; anorexia; bulimia; EDNOS; criteria; DSM-5

(*Int J Eat Disord* 2012; 45:353–361)

Introduction

Present DSM-IV diagnostic criteria for eating disorders (EDs) have important limitations, including the fact that the residual “not otherwise specified” diagnosis (EDNOS) is often the largest category,^{1–3} that the criteria do not well capture ED symptoms among children and adolescents,⁴ and that some criteria have failed to prove sufficiently clinically relevant (e.g., amenorrhea⁵) or are overly difficult to assess relative to their diagnostic importance (e.g., weight phobia⁶). Proposed DSM-5 revisions of these criteria attempt to address many of these problems and have been published on the DSM-5 website (<http://www.dsm5.org/ProposedRevisions/Pages/EatingDisorders.aspx>), with the latest revision in October 2010.

In DSM-5, it is proposed that some disorders presently included in the DSM-IV section

“Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence,” such as pica and rumination disorder, be grouped with ED, resulting in a general diagnostic category of “feeding and eating disorders.” Additionally, two major proposed changes in DSM-5 involve a lowering of thresholds for inclusion in the formal categories of anorexia nervosa (AN) and bulimia nervosa (BN), and making binge ED (BED), presently an EDNOS subtype, a formal ED diagnosis. Other proposed changes include renaming EDNOS “feeding and eating conditions not elsewhere classified” (FECNEC), along with some specifications for subtypes (i.e., examples of what clinical phenotypes this category might include).

As yet, there is precious little empirical data supporting the validity of these proposed changes. Such data are imperative if formal changes in criteria are to result in clinically meaningful and scientifically useful diagnostic categories. This study thus aimed to test empirically the implementation of proposed DSM-5 changes to ED diagnosis. Using a large naturalistic ED database, we tested the effects of proposed changes on diagnostic distributions, as well as the relative abilities of DSM-IV and DSM-5 to capture symptom variance. Specifically, the study focused on proposed criteria for AN, BN, and BED, along with FECNEC and its suggested subtypes atypical AN, subthreshold BN, subthres-

Accepted 12 October 2011

*Correspondence to: Andreas Birgegård, KÄTS, Z8:02 Karolinska University Hospital, SE-17176 Stockholm, Sweden. E-mail: andreas.birgegard@ki.se

Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Published online 23 December 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/eat.20968

© 2011 Wiley Periodicals, Inc.

hold BED, purging disorder, and the residual category other feeding and eating condition not elsewhere classified (OFECNEC). We compared these suggested categories with the existing DSM-IV AN, BN, and EDNOS, as well as all six EDNOS examples (including BED), some of which should logically be subsumed under AN and BN in the new proposal. The EDNOS examples represent heterogeneous clinical presentations although they are found in the same formal diagnosis, and it was of interest to study how they figure into the proposed DSM-5 scheme.

Unfortunately, we did not have adequate data for night eating syndrome or the former childhood and adolescence feeding disorders, proposed to include pica, rumination disorder, and avoidant/restrictive food intake disorder (ARFID), so these were not considered.

Since the proposed changes to DSM-5 largely entail allowing more patients into the AN and BN categories by lowering thresholds and removing criteria, it might be feared that this could make the disorders less clinically distinct and informative due, for example, to a lesser average symptom severity. Thus, we wanted to investigate whether a change of diagnostic system might be associated with less explained variance in clinical variables. A recent study did not find differences between the diagnostic schemes in this regard, using an epidemiological sample and clinically rated dependent variables.⁷ To complement this research in an entirely clinical population and using self-ratings, we investigated the ability of each diagnostic system to capture variance in self-rated ED and psychiatric symptoms.

Method

Participants

Data came from the naturalistic Stepwise quality assurance database, an Internet-based data collection system for specialized ED care in Sweden.⁸ Criteria for inclusion in the database are medical or self-referral to one of the participating treatment units (30 units across Sweden at the time of data extraction, representing a mixture of inpatient and outpatient settings), a diagnosed DSM-IV ED, and intent to treat the patient at the unit in question. The database has been in use since 2005. For this study, complete data on 2,735 patients were available, of whom 2,584 patients were analyzed, with a mean age of 23.0 (SD = 8.29, range 10–67), and 83 (3.2%) participants were males.

To accurately test the proposed DSM-5 scheme, we wanted to retain as many patients in the database as possible although they may not meet the formal criteria of a DSM-IV ED diagnosis according to our operational definitions based on the structured ED interview (SEDI) diagnostic interview (see below). The SEDI did not result in a formal DSM-IV ED diagnosis in 438 cases, all of which were subsequently examined to see whether they had been erroneously registered, or if clinical ED symptoms were in fact present. In these cases, ED was judged to be present if individual SEDI items indicated that cases had at least one psychological symptom (i.e., weight phobia, self-esteem unduly affected, disturbed body experience, or denial of seriousness of low weight) in addition to at least one behavioral symptom (i.e., binge eating, loss of control, purging, or fasting/exercise) or physical symptom (i.e., anorexic weight or amenorrhea). A total of 287 such cases were found and considered to belong to the category of EDNOS Other. The remaining 151 cases were considered not to be cases of ED and were deleted from the data set. Further attesting to the clinical relevance of the retained 287 patients, in the Stepwise procedure, an “EDNOS Other” alternative is automatically prechecked in a subsequent summation form if the SEDI does not result in a diagnosis. In 211 of the 287 cases we retained, interviewers had changed this to AN, BN, or specific EDNOS subtypes (i.e., the SEDI result was still recorded, but the summation form contained the diagnosis, either confirmed or altered by the clinician). In the majority of cases then, clinicians’ judgments were specific enough to warrant choosing another category, and even in the 76 cases where EDNOS Other was retained by clinicians, our syntax did require at least two symptom domains to be analyzed. The resulting final *N* was 2,584.

Measures

*Structured ED Interview (SEDI)*⁹: ED diagnosis in Stepwise is based on the SEDI, developed specifically for the Stepwise system. Prior to February 2008, other diagnostic methods were used, but to retain methodological homogeneity, those data were excluded from the present study. The SEDI is based on DSM-IV ED criteria and comprises 20–30 questions, depending on which additional questions need to be asked. Preliminary validation against the EDE interview¹⁰ has shown a concordance of 81% concerning specific ED diagnosis (including EDNOS and BED) and Kendall’s Tau-b of $\tau = 0.69$ ($p < .0001$).⁹ Age- and sex-specific BMI cutoffs define anorexic underweight, based on published Swedish norms,¹¹ starting from BMI = 17.5 at 18 years (see Table 1). The SEDI is a semistructured interview, i.e., questions are first read verbatim, and the clinician can then probe further by asking for elaboration or examples (probes are suggested for some of the criteria). For each question, the relevant

TABLE 1. Operational definitions of the proposed DSM-5 criteria for AN, BN, and BED using SEDI items and EDEQ data

DSM-5 Criterion	SEDI Interview Item	Complementary EDEQ Data
AN Criterion A AN Criterion B ^b	Age- and sex-specific BMI cutoff ^a Fear of gaining weight, becoming fat	Restraint ≥ 3 (restrictive behaviors at least half the time)
AN Criterion C	At least one of (1) disturbed body experience, (2) lack of recognition of low weight, and (3) undue influence of weight/shape on self-esteem	
BN/BED Criterion A	Eating large amount in brief time period AND sense of loss of control	OBE frequency ≥ 1 /week
BN Criterion B ^c	Compensatory behaviors: purging or fasting/exercise	
BN Criterion C	DSM-IV frequency and duration criteria met ^d	
BN Criterion D	Undue influence of weight/shape on self-esteem	
BN Criterion E BED Criterion B	DSM-5 criteria for AN not met ≥ 3 of (1) eating rapidly, (2) eating until too full, (3) eating although not hungry, (4) eating alone due to guilt, and (5) disgust with binge eating	
BED Criterion C	Distress over binge eating	
BED Criterion D	Binge frequency ≥ 2 /week and duration ≥ 6 months	
BED Criterion E	No purging, AN, or BN	

Notes: SEDI: structured eating disorder interview; EDEQ: eating disorder examination questionnaire; AN: anorexia nervosa; BN: bulimia nervosa; BED: binge eating disorder.

^aBased on published norms,¹¹ and corresponding to BMI = 17.5 at 18 years, anorexic BMI cutoffs were mean -2 SD for girls and -2.1 SD for boys, for each respective age. Thus, for girls: 17 years BMI ≤ 17.8 ; 16 years BMI ≤ 17.1 ; 15 years BMI ≤ 16.5 ; 14 years BMI ≤ 15.9 ; 13 years BMI ≤ 15.4 ; 12 years BMI ≤ 15.0 ; 11 years BMI ≤ 14.6 ; 10 years BMI ≤ 14.3 . Boys: 17 years BMI ≤ 17.9 ; 16 years BMI ≤ 17.2 ; 15 years BMI ≤ 16.6 ; 14 years BMI ≤ 16.0 ; 13 years BMI ≤ 15.4 ; 12 years BMI ≤ 14.9 ; 11 years BMI ≤ 14.5 ; 10 years BMI ≤ 14.1 .

^bFor AN Criterion B, either the SEDI item or the EDEQ data were sufficient to fulfill the criterion.

^cFor BN Criterion B, both the SEDI and the EDEQ data were required to fulfill the criterion.

^dIn a few cases where the SEDI indicated that the DSM-IV binge eating frequency criterion was met although patients did not endorse this on the EDEQ, we decided that interview should override self-report. In these cases, the less stringent DSM-5 criterion was judged to be met.

criterion is displayed on the computer screen to help direct interviewers' attention toward the relevant symptom, rather than merely the patient's initial answer to the question. The SEDI contains "skip-rules," such that, for example, if no anorexic underweight is present, no question about denying the importance of underweight is asked. The SEDI also specifically asks about the DSM-IV research criteria for BED if binge eating is present but not compensation. Other EDNOS examples distinguishable by the SEDI are EDNOS1 (all AN criteria except amenorrhea), EDNOS2 (all AN criteria except underweight; since we did not have data for "significant weight loss," only normal-range BMI could be used to operationalize this group), EDNOS3 (all BN criteria except that binge or purge frequency are less than twice/week or of lesser duration than 3 months), EDNOS4 (regular compensatory behavior in a normal-weight person after small amounts of food), and EDNOS5 (regular chewing and spitting out food).

*EDs Examination Questionnaire (EDEQ)*¹²: The 36-item EDEQ contains four subscales: restraint, eating concern, shape concern, and weight concern, as well as a total score average of all four (ratings of key diagnostic binge/compensation behaviors are also included but are not treated further here). For patients under 18 years of age, an adolescent version is used,¹³ with simplified language and a 14- instead of 28-day time frame.

*The Clinical Impairment Assessment (CIA version 3.0)*¹⁴: It is a 16-item self-report measure of secondary psychosocial impairment resulting from ED symptoms

used for patients ≥ 18 years. Like the EDEQ, it covers a 28-day period and is designed for use immediately following administration of the EDE-Q (this is how it is administered in the Stepwise system). Items on the four-point Likert scale yield a single score; satisfactory internal consistency, test-retest reliability, as well as construct and discriminant validity have been reported.¹⁴

The Comprehensive Psychiatric Rating Scale, self-rated version of the affective subscales (CPRS-S-A), used for patients ≥ 18 years. This 19-item questionnaire is a short version of a longer measure designed to cover DSM-IV Axis I in a self-report format. Responses are given on a 0–3 scale in 0.5-increments, with nine items for depression, nine for anxiety, and eight for compulsion; some items overlap and belong to more than one scale.^{15–17}

*The Strengths and Difficulties Questionnaire (SDQ)*¹⁸: It is a 25-item screening of behavioral/psychological difficulties for children/adolescents, with the subscales emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. The first four are summed to form a total difficulties score, and this variable is analyzed in the present study, and the SDQ is used for patients < 18 years.

Procedure

Patients were assessed using Stepwise by ED specialists prior to treatment, usually within the first three visits to the unit, following, if necessary, medical assessment

TABLE 2. Operational definitions of the proposed DSM-5 criteria for FECNEC, including subtypes, using SEDI items and EDEQ data

DSM-5 Description	Operationalization	Complementary EDEQ Data
FECNEC	AN, BN, or BED criteria not met	AN, BN, or BED criteria not met
Atypical AN ^a	All criteria for AN except that BMI > anorexic cutoff	
Subthreshold BN	All criteria for BN except that the frequency criteria are not met	OBE and/or compensation frequency <1/week
Subthreshold BED	All criteria for BED except that the frequency criterion is not met	OBE frequency <1/week
Purging disorder	Purging (vomiting, laxatives, and diuretics) but not binge eating is present. Fear of gaining weight/becoming fat or undue influence of weight/shape on self-esteem	
OFECNEC	All patients with a DSM-IV diagnosis not captured in other categories	

Notes: SEDI: structured eating disorder interview; EDEQ: eating disorder examination questionnaire; AN: anorexia nervosa; BN: bulimia nervosa; BED: binge eating disorder; FECNEC: feeding and eating disorder not elsewhere classified; OFECNEC: other feeding and eating disorder not elsewhere classified.

^aThe DSM-5 proposal contains the phrase “despite significant weight loss,” which we lacked data for, so only the BMI > anorexic cutoff could be used.

by a physician. For inpatient cases with severe physical complications, assessments were performed within the first week of treatment. After a brief orienting interview lasting a few minutes SCID-I or MINIKid was conducted followed by the SEDI, clinical ratings, and collection of demographic and psychiatric history, including BMI and weighing the patient if necessary. The SEDI result is shown as a prechecked alternative in a list of ED diagnoses in a summation form (AN/BN with subtypes, EDNOS with BED and other examples, including an EDNOS Other category if none of the suggested examples fit), and interviewers can change or retain this suggestion for clinical use. In the present analyses, we used the raw criterion-based SEDI result as DSM-IV diagnosis (with one exception; see below). Interviewers were seated at a computer and recorded answers on the screen. Patients subsequently completed the self-report scales (including the EDEQ, CIA, CPRS-S-A, and SDQ) while sitting at a computer. The entire assessment took on average slightly <1.5 h for patients ≥ 18 years and around 45 min for younger patients. Informed consent for research purposes was obtained from each patient. The Stockholm regional ethical review board has approved the study.

Operationalization of DSM-5 categories was based on SPSS syntax applied to the SEDI and EDEQ; Table 1 shows how each proposed DSM-5 criterion was operationalized using this syntax. Note that frequency of binge eating and compensatory behavior was possible to ascertain separately, but not duration, since the SEDI asks about frequency and duration in the same item, and the EDE-Q only asks for the last 28/14 days.

Statistical Analysis

DSM-IV diagnoses were operationalized using SEDI interviews, and DSM-5 diagnoses were generated using SPSS syntax as described above (data were analyzed using SPSS 19.0.0 for Mac). The fact that we did not have data for weight loss as required in atypical AN (see Table 2) had the effect that the order of the SPSS syntax impacted DSM-5 distributions specifically regarding the atypical AN—purging disorder distinction. Since Atypical

AN without the weight loss requirement was defined only by psychological criteria and the absence of underweight, there was considerable conceptual overlap with purging disorder. Therefore, if the purging disorder syntax section came last, a number of patients who met purging disorder criteria were captured by that category. On the other hand, if atypical AN came last, nearly all purging disorder patients were captured in atypical AN. We present both sets of data with an emphasis on the solution that produced more purging disorder patients (i.e., the first order above), since our data could define that category faithfully but were less specific with regard to atypical AN, and over inclusion in atypical AN was therefore the greater risk. The SEDI results in one of 10 possible categories: AN/BN including subtypes, the research-criteria-defined BED, and five additional EDNOS examples. Comparisons of clinical variables between diagnostic groups in each categorization was carried out using ANOVA, and partial eta squared effect size (η^2) is reported, for each analysis. Rule-of-thumb conventions are small effect ≥ 0.01 , medium effect ≥ 0.06 , and large effect ≥ 0.14 .

Results

The diagnostic distribution of DSM-IV is presented in **Table 3** and the DSM-5 distribution in **Table 4**. Distributions according to both systems follow an expected age-related trajectory, with more EDNOS among younger patients and more BN among older patients. Frequency of AN decreased as age increased in DSM-IV, but this was less distinct in DSM-5, consistent with removing the amenorrhea criterion that chiefly concerns patients above 14 years of age; the increase in AN occurred in these groups.

Concerning the EDNOS/FECNEC categories, the overall proportion decreased from 55% to 38%. This decrease occurred, again, mainly in the older

TABLE 3. DSM-IV category distribution, according to the SEDI interview in combination with (for EDNOS Other) interviewers' decisions, divided into age groups and in total

SEDI Diagnosis (DSM-IV)	10–14 Years <i>n</i> (% Within the Age Group)	15–17 Years <i>n</i> (% Within the Age Group)	≥18 Years <i>n</i> (% Within the Age Group)	Total <i>N</i> (%)
AN	47 (24%)	106 (22%)	327 (17%)	480 (19%)
BN	4 (2%)	75 (16%)	594 (31%)	673 (26%)
EDNOS	148 (74%)	295 (62%)	998 (52%)	1,431 (55%)
EDNOS1 (% of EDNOS)	1 (<1%)	8 (3%)	44 (4%)	53 (4%)
EDNOS2 (% of EDNOS)	73 (49%)	119 (40%)	178 (18%)	370 (26%)
EDNOS3 (% of EDNOS)	2 (1%)	14 (5%)	131 (13%)	147 (10%)
EDNOS4 (% of EDNOS)	27 (18%)	73 (25%)	300 (30%)	400 (28%)
EDNOS5 (% of EDNOS)	8 (5%)	16 (5%)	41 (4%)	65 (5%)
EDNOS6/BED (% of EDNOS)	1 (<1%)	7 (2%)	101 (10%)	109 (8%)
EDNOS other (% of EDNOS)	36 (24%)	58 (20%)	193 (19%)	287 (20%)
Total (% of total)	199 (8%)	476 (18%)	1,909 (74%)	2,584 (100%)

Notes: SEDI: structured eating disorder interview; AN: anorexia nervosa; BN: bulimia nervosa; BED: binge eating disorder; EDNOS: eating disorder not otherwise specified.

TABLE 4. DSM-5 category distribution, according to the SPSS-syntax operationalization based on SEDI interview data and complementary EDEQ data, divided into age groups and in total

DSM-5 Category	10–14 Years <i>n</i> (%)	15–17 Years <i>n</i> (%)	≥18 Years <i>n</i> (%)	Total <i>N</i> (%)
AN	50 (25%)	121 (25%)	399 (21%)	570 (22%)
BN	11 (6%)	114 (24%)	806 (42%)	931 (36%)
BED	2 (1%)	6 (1%)	86 (5%)	94 (4%)
FECNEC	136 (68%)	235 (49%)	618 (32%)	989 (38%)
Atypical AN (% of FECNEC)	89 (65%)	139 (59%)	314 (51%)	542 (55%)
Subthresh. BN (% of FECNEC)	0	0	1 (<1%)	1 (<1%)
Subthresh. BED (% of FECNEC)	0	2 (1%)	31 (<1%)	33 (3%)
Purging dis. (% of FECNEC)	26 (19%)	52 (22%)	181 (29%)	259 (26%)
OFECNEC (% of FECNEC)	21 (15%)	42 (18%)	91 (15%)	154 (16%)
Total (% of total)	199 (8%)	476 (18%)	1,909 (75%)	2,584 (100%)

Notes: SEDI: structured eating disorder interview; EDEQ: eating disorder examination questionnaire; AN: anorexia nervosa; BN: bulimia nervosa; BED: binge eating disorder; FECNEC: feeding and eating disorder not elsewhere classified; OFECNEC: other feeding and eating disorder not elsewhere classified.

age groups. For the youngest patients, 68% still did not receive an AN or BN diagnosis in DSM-5. Among adults, the proportion went from over half to about a third of patients, owing mostly to an increase in patients allocated to BN and the addition of BED. Subthreshold BN and subthreshold BED captured very few patients in our data, whereas atypical AN and purging disorder together constituted well over half of the residual-category patients, with the remaining quarter ending up in OFECNEC.

As shown in **Table 5**, a total of 418 patients in DSM-IV EDNOS were classified in DSM-5 as AN, which increased from 19% to 22%, BN, which increased from 26% to 36%, or BED, which included 4% of the patients. In all, these reclassifications were responsible for the decrease of EDNOS/FECNEC from 55% to 38%. As would be expected, new cases of AN in DSM-5 came mostly from EDNOS1 (91% of this group moved to AN because of the removal of the amenorrhea criterion). New cases of BN in DSM-5 came from EDNOS3 (99% of this group moved to BN because of lowering the binge/compensation frequency threshold). Interestingly, although the elevation of

BED from informal to formal diagnostic category decreased FECNEC, the number of patients classified as BED actually decreased from 109 to 94 patients. The FECNEC subgroups atypical AN and purging disorder were mainly composed of EDNOS2 and EDNOS4, and predictably, in opposite proportions: Atypical AN captured EDNOS2, and purging disorder drew from EDNOS4. However, as noted above, if we ran the SPSS DSM-5 syntax with atypical AN and purging disorder in the opposite sequence, a different result was produced. Then, DSM-5 consisted of 734 atypical AN (instead of 531) and 56 purging disorder (instead of 259).

Next, we analyzed differences between diagnostic groups in each of the two systems to see whether more or less variance was captured using the DSM-5 instead of the DSM-IV scheme. In these analyses, the latter was treated as both a three-group (AN/BN/EDNOS) independent variable and a four-group one (AN/BN/BED/EDNOS), since the latter was more similar to the DSM-5 categories in comprising four groups and since treating BED as a separate DSM-IV category is common in research already. As shown in **Table 6**, with the exceptions of EDEQ restraint and the SDQ total score, the DSM-5

TABLE 5. Crosstabulation of DSM-IV and DSM-5, including EDNOS examples/FECNEC subtypes, N = 2,584. The shaded area shows crosstabulation of formal diagnoses, and the white area shows EDNOS/FECNEC subtypes

DSM-IV	DSM-5						FECNEC Subtypes			
	AN n = 570 (2.2%)	BN n = 931 (3.6%)	BED n = 94 (4%)	FECNEC n = 989 (38%)	Atypical AN n = 542 (21%)	Subthresh. BN n = 1 (<1%)	Subthresh. BED n = 33 (1%)	Purging Dis. n = 259 (10%)	OFECNEC n = 154 (6%)	
AN n = 480 (19%)	480 100%/84%	—	—	—	—	—	—	—	—	
BN n = 673 (26%)	—	671 100%/72%	2 <1%/2%	—	—	—	—	—	—	
EDNOS n = 1,431 (55%)	92 6%/16%	258 18%/28%	92 6%/98%	989 69%/100%	—	—	—	—	—	
EDNOS examples	n	n	n	n	n	n	n	n	n	
EDNOS1 n = 53 (2%)	48 91%/8%	3 6%/<1%	—	—	1 2%/<1%	—	1 2%/<1%	1 2%/<1%	—	
EDNOS2 n = 370 (14%)	—	51 14%/6%	6 2%/6%	—	233 63%/44%	—	77 21%/30%	—	—	
EDNOS3 n = 147 (6%)	—	146 99%/16%	1 1%/1%	—	—	—	—	—	—	
EDNOS4 n = 400 (16%)	—	48 12%/5%	—	—	183 46%/34%	—	—	149 37%/58%	27 7%/16%	
EDNOS5 n = 65 (3%)	8 12%/1%	6 9%/1%	1 2%/1%	—	30 46%/6%	—	—	6 9%/2%	14 22%/9%	
BED n = 109 (4%)	—	—	79 73%/84%	—	—	—	—	—	—	
EDNOS n = 287 (11%)	35 12%/6%	4 1%/<1%	5 2%/5%	—	95 33%/18%	—	—	26 9%/10%	123 43%/75%	

Notes: AN: anorexia nervosa; BN: bulimia nervosa; BED: binge eating disorder; EDNOS: eating disorder not otherwise specified; FECNEC: feeding and eating disorder not elsewhere classified; OFECNEC: other feeding and eating disorder not elsewhere classified.

TABLE 6. ANOVAs for clinical variables using each of the DSM-IV and DSM-5 (AN/BN/BED/FECNEC) diagnoses, with estimates of effects size (eta squared: η^2). Effect statistics are shown both for the formal DSM-IV diagnoses (AN/BN/EDNOS) and, in parentheses, for DSM-IV with BED as a separate category (AN/BN/BED/EDNOS), to allow fair effect comparisons

Variable	Diagnostic System	F	P	η^2
EDEQ restraint	DSM-IV (with BED separate)	35.03 (44.07)	<.0001 (<.0001)	0.026 (0.049)
	DSM-5	37.61	<.0001	0.042
EDEQ eating concern	DSM-IV (with BED separate)	99.50 (71.42)	<.0001 (<.0001)	0.072 (0.077)
	DSM-5	95.39	<.0001	0.100
EDEQ shape concern	DSM-IV (with BED separate)	71.19 (50.04)	<.0001 (<.0001)	0.052 (0.055)
	DSM-5	75.78	<.0001	0.081
EDEQ weight concern	DSM-IV (with BED separate)	75.52 (51.94)	<.0001 (<.0001)	0.055 (0.057)
	DSM-5	75.07	<.0001	0.080
EDEQ global score	DSM-IV (with BED separate)	79.81 (53.21)	<.0001 (<.0001)	0.058 (0.058)
	DSM-5	73.68	<.0001	0.079
CIA ^a (≥ 18 years only)	DSM-IV (with BED separate)	29.68 (21.12)	<.0001 (<.0001)	0.035 (0.038)
	DSM-5	24.48	<.0001	0.043
CPRS depression (≥ 18 years only)	DSM-IV (with BED separate)	10.98 (7.34)	<.0001 (<.0001)	0.011 (0.011)
	DSM-5	9.40	<.0001	0.015
CPRS anxiety (≥ 18 years only)	DSM-IV (with BED separate)	8.29 (5.90)	<.0001 (.001)	0.009 (0.009)
	DSM-5	6.90	<.0001	0.011
CPRS compulsion (≥ 18 years only)	DSM-IV (with BED separate)	21.13 (14.16)	<.0001 (<.0001)	0.022 (0.022)
	DSM-5	14.73	<.0001	0.023
SDQ total score (<18 years only) ^b	DSM-IV	7.68	<.0001 (<.0001)	0.022
	DSM-5	3.49	.014	0.015

Notes: AN: anorexia nervosa; BN: bulimia nervosa; BED: binge eating disorder; EDNOS: eating disorder not otherwise specified; FECNEC: feeding and eating disorder not elsewhere classified; EDEQ: eating disorder examination questionnaire; CIA: clinical impairment assessment; CPRS: comprehensive psychopathological rating scale—self-rated version of the affective subscales; SDQ: strengths and difficulties questionnaire.

^a The CIA has only been in use in the Stepwise system since August 2008, so $N = 1,629$ instead of 1,909 for this analysis.

^b There were too few BED patients <18 years to allow for a “BED separate” analysis of the SDQ.

proposal was associated with larger effect sizes in all cases, as compared with both the three-group and the four-group DSM-IV categorizations. For EDEQ subscales apart from restraint, DSM-5 was associated with an average effect size of 0.087, whereas DSM-IV averaged 0.060 for the same variables. Effect sizes were modest for psychiatric variables, but the pattern was fairly consistent, suggesting that DSM-5 more effectively differentiated groups.

Discussion

In this study, we aimed to compare the effects of implementing the proposed DSM-5 criteria to a large clinical ED sample, as well as testing the ability of DSM-IV and DSM-5 to capture statistical variance in self-rated symptom scales. We used structured assessment measures and strictly criterion-based diagnostics throughout, and findings were fairly clear. First, DSM-5 reduced the residual EDNOS/FECNEC category substantially, from being over half to two fifths of the sample, largely by including more patients in AN and BN. Second, although the residual category was reduced, this was to a negligible degree among the youngest patients. Our results suggested that proposed DSM-5 criteria chiefly improve diagnostics for later

adolescent and adult patients. Third, regardless of whether we treated BED as a separate formal DSM-IV diagnosis or not, the DSM-5 categorization of patients was better able to capture clinical symptom variance. This is an important finding since the inclusion of more patients in the AN and BN groups due to less strict criteria could have meant a dilution of clinical severity and distinctiveness of these groups. However, the opposite appears to be the case, at least when using self-rated dependent variables. An overall impression from our data set was that the increase in explanatory power was greatest for ED symptom variables (i.e., greater for EDEQ than CPRS or SDQ), suggesting that DSM-5 may more distinctly characterize ED.

A counterintuitive finding, given the lower BED binge frequency threshold in DSM-5, was that the major movement of such patients was that 30 patients were reclassified as subthreshold BED. The reason for this may be measurement method. DSM-IV criteria were determined by semistructured interview, whereas the DSM-5 frequency cutoff for BED (which is different from the DSM-IV criterion and is not independently measured by the SEDI) was determined from self-report data. The result pattern would suggest that BED patients report less binge eating in self-report than in interview, consistent with Wilfley et al.¹⁹ who found lower binge frequencies on the EDEQ than the EDE interview in a BED sample.

It is possible that our decision to include the DSM-IV EDNOS Other group was based on overly inclusive criteria (see “Statistical Analysis”), and that these patients did not in fact have clinical symptoms corresponding to an ED. As argued however, we believe it is reasonable to assume that many of these 287 patients may indeed represent a clinical reality that the DSM-5 proposal, like its predecessors, is unable to capture more specifically. Indeed, more than half of them were reclassified into AN, atypical AN, or purging disorder. The 123 (43%) that did end up in OFECNEC could represent a less clinically severe group with less typical ED features, but this issue awaits more research into the still undecided issue of what an ED actually is.

A problematic phrase in the DSM-5 specification of atypical AN (inherited from EDNOS4) is “despite significant weight loss.” The fact that we did not have data on this criterion may have resulted in inflated estimates of the frequency of atypical AN. However, it is uncertain what kind of data would suffice to rate this criterion. First, it is not clear what degree of weight loss is to be considered significant, especially since numerical guidelines have been removed from the AN criterion A concerning low weight. Second, the wording implies a temporal aspect, i.e., weight loss in relation to some unspecified reasonable amount of time. For example, weight loss two years previously cannot reasonably be intended. Some guiding definition of weight loss would be useful, such as “for example, a 10% weight reduction within the last 3 months.” Another option would be to remove the requirement, but this would result in a category minimally requiring only weight phobia and one of the three additional psychological criteria, which may be over-inclusive and not sufficiently clinically relevant. Recall that our database contains only patients in specialized care. In normal population studies, a large number of people might be classified as atypical AN without truly having clinically significant ED symptoms. The need for a clear operationalization was evident from the fact that we obtained widely discrepant findings depending on the relative placement of the syntax code lines defining atypical AN and purging disorder. If weight loss is not defined, the conceptual overlap between these categories, illustrated in our results, would be allowed greater impact.

Another issue concerns subthreshold BN, where proposed DSM-5 criteria state that all BN criteria should be met “except that the binge eating *and* inappropriate compensatory behaviors occur, on average, less than once a week *and/or* for less than

3 months” (*italics added*). Read literally, this requires that *both* the binge eating and the compensatory behaviors *either* fail to occur once per week or more *or* fail to meet the duration criterion *or* fail in both instances. We did not have separate duration data and could only use the frequency aspect, and this we chose to define such that *either* behavior, or both, could fail to meet the criterion. With the current wording, a patient who for example binges three times per week and compensates once per month would not meet subthreshold BN criteria (or indeed any diagnosis, since BN criteria would not be met either). “And/or” in both places would solve this problem.

The proposed DSM-5 criteria do not include a numeric weight or BMI cutoff for AN, but rather refer to “a significantly low body weight” as the result of persistent energy intake restriction. Although provision of a universal numeric cutoff is problematic, since it cannot apply to all individuals, the removal of any guiding definition raises problems. There is no universally accepted definition of low body weight, and for research purposes, the already great variability in definitions of anorexic weight can only be expected to increase.²⁰ A potentially fruitful alternative recently suggested by Hebebrand and Bulik²¹ is that the WHO cutoff of BMI < 18.5 be used for adults, and BMI below the 10th gender matched reference population age percentile in children and young adolescents, in conjunction with two or more symptoms of starvation (e.g., loss of fat mass, amenorrhea, bradycardia, and hypotension). In the present study, the SEDI used a cutoff of BMI ≤ 17.5, along with corresponding age- and sex-adjusted cutoffs below 18 years. We also tested BMI cutoffs of <18.5 (and corresponding cutoffs for patients under age 18), which more closely correspond to BMI below the 10th percentile. However, somewhat surprisingly, there were no substantial changes to AN prevalence in our data. Overall AN increased by less than one percentage point, and increased in the separate age groups by less than two percentage points at most (data not shown). Therefore, although we did not have data for specific starvation symptoms, our results are largely compatible with Hebebrand and Bulik’s suggestion.

An important caveat when interpreting results is that the SEDI and the EDEQ are intended to operationalize DSM-IV rather than DSM-5, and this may have skewed results toward conserving DSM-IV-like frequency distributions. For example, one and the same SEDI item asks about both binge frequency and duration, while no frequency data are recorded. The interview is designed to check each

DSM-IV criterion, and although many of our DSM-5 operationalizations were most likely accurate, in some instances, they were necessarily suboptimal. A specific problem may concern the operationalization of the AN Criterion B “persistent behavior that interferes with weight gain” in addition to weight phobia. We used data from the EDEQ restraint subscale to approximate this criterion, yet it is likely that we did not fully capture one desirable effect of the change, namely the ability to diagnose younger patients who for whatever reason do not verbalize fear of becoming fat but display behaviors consistent with such a fear. Such patients also may not report restraint in their self-ratings to a sufficient degree that our syntax included them; our FECNEC estimate among patients under 15 years of age may therefore be exaggerated. More careful study of “persistent behavior that interferes with weight gain” as regards the DSM-5 proposal in relation to younger patients would be useful. The addition of proposed ARFID criteria in future studies, which we did not have adequate data for, may also help address the problem of inability to adequately classify young patients.

Taken as a whole, our study suggests that the proposed DSM-5 criteria represent an improvement to the present state of affairs concerning EDNOS. However, the problem of a large residual diagnostic category still remains and is especially pronounced in relation to younger patients. This raises questions of the overall utility of the proposed changes in relation to distinct age groups. What’s more, although the proposed DSM-5 criteria appear to capture clinical variance in concurrent self-ratings better than DSM-IV, an important question remains concerning the prognostic validity of the DSM-5 categories. Answers to these questions await further studies.

The authors thank Dr. Timothy Walsh, head of the APA DSM-5 Eating Disorders Work group, for highly useful suggestions concerning the study.

References

1. Fairburn CG, Bohn K. Eating disorder NOS (EDNOS): An example of the troublesome “not otherwise specified” (NOS) category in DSM-IV. *Behav Res Ther* 2005;43:691–701.
2. Norring C, Palmer RL. EDNOS, Eating Disorders Not Otherwise Specified. New York: Routledge, 2005.

3. Walsh BT, Sysko R. Broad categories for the diagnosis of eating disorders (BCD-ED): An alternative system for classification. *Int J Eat Disord* 2009;42:754–764.
4. Bravender T, Bryant-Waugh R, Herzog D, Katzman D, Kriepe RD, Lask B, et al. Classification of eating disturbance in children and adolescents: Proposed changes for the DSM-V. *Eur Eat Disord Rev* 2010;18:79–89.
5. Attia E, Roberto CA. Should amenorrhea be a diagnostic criterion for anorexia nervosa? *Int J Eat Disord* 2009;42:581–589.
6. Becker AE, Eddy KT, Perloe A. Clarifying criteria for cognitive signs and symptoms for eating disorders in DSM-V. *Int J Eat Disord* 2009;42:611–619.
7. Keel PK, Brown TA, Holm-Denoma J, Bodell LP. Comparison of DSM-IV versus proposed DSM-5 diagnostic criteria for eating disorders: Reduction of eating disorder not otherwise specified and validity. *Int J Eat Disord* 2011;44:553–560.
8. Birgegard A, Bjorck C, Clinton D. Quality assurance of specialised treatment of eating disorders using large-scale Internet-based collection systems: methods, results and lessons learned from designing the Stepwise database. *Eur Eat Disord Rev* 2010;18:251–259.
9. de Man Lapidoth J, Birgegård A. Validation of the Structured Eating Disorder Interview (SEDI) Against the Eating Disorder Examination (EDE). Stockholm: Karolinska Institutet, 2010.
10. Fairburn CG, Cooper Z. The eating disorder examination. In: Fairburn CG, Wilson GT, editors. *Binge Eating: Nature, Assessment, and Treatment*, 12th ed. New York: Guilford, 1993, pp. 317–360.
11. Karlberg J, Luo ZC, Albertsson-Wikland K. Body mass index reference values (mean and SD) for Swedish children. *Acta Paediatr* 2001;90:1427–1434.
12. Fairburn CG, Beglin SJ. Assessment of eating disorders: Interview or self-report questionnaire? *Int J Eat Disord* 1994;16:363–370.
13. Carter JC, Stewart DA, Fairburn CG. Eating disorder examination questionnaire: Norms for young adolescent girls. *Behav Res Ther* 2001;39:625–632.
14. Bohn K, Doll HA, Cooper Z, O’Connor M, Palmer RL, Fairburn CG. The measurement of impairment due to eating disorder psychopathology. *Behav Res Ther* 2008;46:1105–1110.
15. Martinsen EW, Friis S, Hoffart A. Assessment of depression: Comparison between Beck Depression Inventory and subscales of Comprehensive Psychopathological Rating Scale. *Acta Psychiatr Scand* 1995;92:460–463.
16. Mattila-Evenden M, Svanborg P, Gustavsson P, Asberg M. Determinants of self-rating and expert rating concordance in psychiatric out-patients, using the affective subscales of the CPRS. *Acta Psychiatr Scand* 1996;94:386–396.
17. Svanborg P, Asberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatr Scand* 1994;89:21–28.
18. Goodman R. The strengths and difficulties questionnaire: A research note. *J Child Psychol Psychiatry* 1997;38:581–586.
19. Wilfley DE, Schwartz MB, Spurrell EB, Fairburn CG. Assessing the specific psychopathology of binge eating disorder patients: Interview or self-report? *Behav Res Ther* 1997;35:1151–1159.
20. Thomas JJ, Roberto CA, Brownell KD. Eighty-five per cent of what? Discrepancies in the weight cut-off for anorexia nervosa substantially affect the prevalence of underweight. *Psychol Med* 2009;39:833–843.
21. Hebebrand J, Bulik CM. Critical appraisal of the provisional DSM-5 criteria for anorexia nervosa and an alternative proposal. *Int J Eat Disord* 2010. [Epub ahead of print].