Published in final edited form as:

Behav Res Ther. 2015 April ; 67: 41–44. doi:10.1016/j.brat.2015.02.002.

Evaluation of the DSM-5 Severity Indicator for Bulimia Nervosa

Dr. Carlos M. Grilo, Ph.D., Dr. Valentina Ivezaj, Ph.D., and Dr. Marney A. White, Ph.D., M.S.

Department of Psychiatry, Yale School of Medicine (Drs. Grilo, Ivezaj, and White), Department of Psychology, Yale University (Dr. Grilo), and Yale School of Public Health (Dr. White)

Abstract

This study examined the *DSM-5* severity criterion for bulimia nervosa (BN) based on the frequency of inappropriate weight compensatory behaviors. 199 community volunteers classified with BN were categorized using *DSM-5* severity levels and compared on demographic and clinical variables. 77 (39%) participants were categorized as mild, 68 (34%) as moderate, 32 (16%) as severe, and 22 (11%) as extreme. The severity groups did not differ significantly in demographic variables or body mass index. Shape and Weight concerns did not differ significantly across severity groups. Binge eating differed with the extreme group having higher frequency than the severe, moderate, and mild groups, which did not differ from each other. Restraint differed with the extreme group having significantly higher levels than the mild group. Eating concerns differed with the extreme group having higher levels than moderate and mild groups. Depression differed with the extreme group having higher levels than severe, moderate, and mild groups, which did not differ from each other. Findings from this non-clinical group provide new, albeit modest, support for *DSM-5* severity rating for BN based on frequency of inappropriate weight compensatory behaviors. Statistical findings indicate that differences in collateral clinical variables associated with the *DSM-5* severity ratings reflect small effect sizes. Further research is needed with treatment-seeking patient groups with BN to establish the validity of the *DSM-5* severity specifier and should include broader clinical and functional validators.

Keywords

bulimia nervosa; binge eating; purging; weight control; diagnosis

The diagnostic criteria for bulimia nervosa (BN) were recently modified in the *Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5; APA, 2013)* with three notable changes. First, BN is no longer subtyped by purging or non-purging. Second, the frequency stipulations for both recurrent episodes of binge eating (i.e., eating unusually

© 2015 Published by Elsevier Ltd.

Correspondence should be addressed to Carlos M. Grilo, Yale University School of Medicine, 301 Cedar Street (2nd Floor), New Haven, CT 06519. carlos.grilo@yale.edu.

Disclosure of Conflicts

The authors have no competing interests to report.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

large quantities of food accompanied by subjective feelings of loss of control) and inappropriate weight compensatory behaviors (e.g., self-induced vomiting, laxative or diuretic misuse, and excessive exercise) were changed from occurring, on average, twice-weekly to once-weekly during the past three months. This change followed research indicating that a once-weekly frequency of binge-eating signaled a clinically relevant problem (Wilfley et al., 2007; Wilson & Sysko, 2009). Third, a new “severity specifier” was added based on the frequency of inappropriate weight compensatory behaviors. The *DSM-5* proposed four severity groups defined based on the frequency of episodes of compensatory behaviors as follows: mild (1–3 episodes/week), moderate (4–7 episodes/week), severe (8–13 episodes/week), and extreme (14 or more episodes/week). Empirical support for the severity specifier for BN is needed.

Methods

Participants

Participants were 199 community volunteers drawn from a larger series of 3283 respondents to online advertisements seeking volunteers aged 18 years or older for a research study about eating, dieting, and/or health behaviors and attitudes. Participants were selected from the larger sample per criteria used to define our study group of persons with bulimia nervosa (BN). Advertisements with a link to a web survey were placed on Craigslist internet classified ads in various US cities. The participant group of N=199 consisted of 14 (7.0%) males and 185 (93.0%) females and the racial/ethnic distribution was: 76.4% (n=152) White, 9.0% (n=18) Hispanic, 6.5% (n=13) Black, 4.0% (n=8) Asian, and 4.0% (n=8) reported “other.”

Procedures and Assessments

Participants provided basic demographic information, self-reported height and current weight, and completed self-report questionnaires through SurveyMonkey, a secure online data gathering website server. Participants were required to give informed consent but provided no personal identifying information. The study was approved by the Yale IRB.

The *Questionnaire for Eating and Weight Patterns - Revised* (QEWP-R; Yanovski, 1993) assesses specific diagnostic criteria for BN. This measure, used in the *DSM-IV* field trials, has received psychometric support in diverse eating-disordered groups (Celio, Wilfley, Crow, Mitchell, & Walsh, 2004). The *Eating Disorder Examination Questionnaire* (EDE-Q; Fairburn & Beglin, 1994), which focuses on the past 28 days, assesses the frequency of objective bulimic episodes (OBEs; defined as feeling a loss of control while eating unusually large quantities of food; this definition corresponds to the *DSM-5* criteria for binge eating episodes) and inappropriate weight compensatory behaviors (purging, laxative misuse, diuretic misuse, and extreme exercise). In the present study, the EDE-Q questions regarding inappropriate weight compensatory behaviors assessed frequency of episodes (rather than days) thus allowing a direct examination of the *DSM-5* severity criteria. The EDE-Q also comprises four subscales (Dietary Restraint, Eating concern, Shape concern, and Weight concern) and a global total score. The EDE-Q has demonstrated good test-retest reliability (Reas, Grilo, & Masheb, 2006), convergence with the EDE interview (Mond,

Hay, Rodgers, & Owen, 2007), and good performance in community studies especially for purging behaviors (Mond et al., 2007). The *Beck Depression Inventory* (BDI; Beck & Steer, 1987) assesses depressive symptoms and levels; it has strong psychometric support (Beck, Steer, & Garbin, 1988) and performs well as a marker for severity and distress.

Creation of Bulimia Nervosa Severity Groups

The overall BN study group was created based on responses to the QEWP-R and EDE-Q per *DSM-5* (APA, 2013) criteria. The BN study group required a minimum frequency of once-weekly for both binge-eating and purging behaviors and shape/weight unduly influenced self-evaluation. BN severity subgroups were created using *DSM-5* severity definitions based on the frequency of inappropriate weight compensatory behaviors on the EDE-Q (which was modified in this study to assess the frequency of episodes, rather than days) as follows: mild (defined as an average of 1–3 episodes/week), moderate (4–7 episodes/week), severe (8–13 episodes/week), and extreme (14 or more episodes/week).

Statistical Analysis

General linear model (GLM) analysis of variance (ANOVA) was used to compare the severity categories for BN (i.e., *DSM-5* severity based on frequency of weight compensatory behaviors) on the demographic and clinical measures. In addition, partial η^2 , an effect size (ES) measure, was calculated; these values represent the proportion of variation in the criterion measure accounted for by group membership.

Results

Bulimia Nervosa: *DSM-5* Severity Groups

In the overall participant group of $N=199$ with BN, the following *DSM-5*-defined severity groups (based on frequency of episodes of extreme weight compensatory behaviors) were observed: 77 (39%) participants were categorized as mild, 68 (34%) as moderate, 32 (16%) as severe, and 22 (11%) as extreme. Table 1 summarizes the frequencies of the specific forms of inappropriate weight compensatory behaviors across the severity groups.

Table 2 summarizes demographic and physical variables for participants with BN based on the *DSM-5* severity groups. The severity groups did not differ significantly in age, sex, ethnicity/race, or body mass index. Table 3 summarizes descriptive statistics and statistical analyses (including effect-size measures) for participants with BN comparing *DSM-5* severity groups on the clinical measures. ANOVAs revealed that the severity groups differed significantly on frequency of inappropriate weight control behaviors (per the study method to create the severity categories). Shape concerns and Weight concerns did not differ significantly across severity groups. Binge eating differed with the extreme group having significantly higher frequency than the severe, moderate, and mild groups, which did not differ from each other. Restraint scores differed with the extreme group having significantly higher levels than the mild group; the other groups did not differ significantly. Eating concerns scores differed with the extreme group having significantly higher levels than the moderate and mild groups; the other groups did not differ significantly. Depression differed with the extreme group having significantly higher levels than the severe, moderate, and

mild groups, which did not differ significantly from each other. Partial eta squared values indicated small effect sizes between the DSM-5 severity categories.

Discussion

Findings from this non-clinical group of community participants classified with BN provide new, albeit modest, support for the *DSM-5* severity rating based on frequency of inappropriate weight compensatory behaviors. In this community sample with BN, 39% of participants were categorized as mild, 34% as moderate, 16% as severe, and 11% as extreme. Shape concerns and Weight concerns did not differ significantly across severity groups, although that is to be expected because the BN diagnosis required the presence of elevated scores on the importance of shape and weight for evaluating self-evaluation. Overall, the extreme group was characterized by significantly higher levels of eating concerns and depression than the severe, moderate, and mild groups, which did not differ significantly from each other. Although statistical testing provided some support for the *DSM-5* severity categories, effect sizes for the observed differences between categories were small.

The study's strengths and limitations are noted as context for interpreting the findings. Self-report assessments may be unreliable and cannot be used to determine formal diagnoses which would require diagnostic interviewing. The self-report measures, however, are widely used in the eating disorder field and have received psychometric support including acceptable convergence with interview methods (Celio et al., 2004; Fairburn & Beglin, 1994; Mond et al., 2007); the EDE-Q performs particularly well for assessing purging behaviors in community studies (Mond et al., 2007). It is also possible that self-report may enhance honest disclosure of sensitive material by reducing embarrassment which is well known to characterize persons with eating disorders. The study group was a sample of convenience comprising persons who volunteered for a research study on the internet. The internet provides anonymity which may have facilitated further the honest reporting of embarrassing behaviors. The internet is used widely by people for health-related issues, particularly women, those with higher education, and those aged less than 65 (Rice, 2006). The generalizability of our internet-based findings needs to be considered within this context and may not generalize to males or to certain ethnic/racial minority groups who may have been under-represented in our study (7% men and 24% non-white, respectively). Generalizability of the findings to clinical samples of patients with BN is also uncertain. The findings require replication and extension to treatment-seeking patient groups with BN to establish the validity of the *DSM-5* severity specifier. Future research should also test additional clinical and functional validators (i.e., non-eating specific variables) for the *DSM-5* severity specifier including, perhaps most importantly, prognostic significance for natural course and treatment outcome.

Acknowledgments

This study was supported, in part, by a grant from the National Institutes of Health (K24 DK070052). The National Institutes of Health were not involved in study design, collection, analysis, interpretation of data, the writing of the report, or in the decision to submit this article for publication. The article does not represent the views of the NIH. No additional funding was obtained for submitting this article.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders, Fifth edition (DSM-5). Washington, DC: American Psychiatric Association; 2013.
- Beck, AT.; Steer, R. Manual Revised Beck Depression Inventory. NY: Psychol. Corp; 1987.
- Beck AT, Steer R, Garbin M. Psychometric properties of the Beck Depression Inventory: 25 years of evaluation. *Clinical Psychology Review*. 1988; 8:77–100.
- Celio AA, Wilfley DE, Crow SJ, Mitchell J, Walsh BT. A comparison of binge eating scale, questionnaire for eating and weight patterns-revised, and EDE-questionnaire with the eating disorder examination in the assessment of binge eating disorder and its symptoms. *International Journal of Eating Disorders*. 2004; 36:434–444. [PubMed: 15558644]
- Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? *International Journal of Eating Disorders*. 1994; 16:363–370. [PubMed: 7866415]
- Mond JM, Hay PJ, Rodgers B, Owen C. Self-report versus interview assessment of purging in community sample of women. *European Eating Disorders Review*. 2007; 15:403–09. [PubMed: 17960776]
- Reas DL, Grilo CM, Masheb RM. Reliability of the Eating Disorder Examination-Questionnaire in patients with binge eating disorder. *Behaviour Research and Therapy*. 2006; 44:43–51. [PubMed: 16301013]
- Rice RE. Influences, usage, and outcomes of internet health information searching. *International Journal of Medical Informatics*. 2006; 75:8–28. [PubMed: 16125453]
- Wilfley D, Bishop ME, Wilson GT, Agras WS. Classification of eating disorders: toward *DSM-5*. *International Journal of Eating Disorders*. 2007; 40S:S123–S129. [PubMed: 17685383]
- Yanovski S. Binge eating disorder: current knowledge and future directions. *Obesity Research*. 1993; 1:306–324. [PubMed: 16350580]

HIGHLIGHTS

- Bulimia nervosa (BN) diagnosis was revised in *DSM-5*.
- Examined *DSM-5* severity specifier of inappropriate weight-compensatory behavior frequency.
- Findings provide new, albeit modest, support for *DSM-5* severity specifier for BN.
- Differences between *DSM-5* severity groups reflect small effect sizes.
- Further research is needed with treatment-seeking patient groups with BN.

Table 1

Weekly Inappropriate Weight Compensatory Behaviors across *DSM-5* severity groups.

	Mild N=77		Moderate N=68		Severe N=32		Extreme N=22		ANOVA	
	M	sd	M	sd	M	sd	M	sd	η^2	posthoc
Vomit episodes ^I	3.5	(3.8)	5.6	(8.5)	13.6	(12.2)	36.6	(32.6)	40.27***	bcdef
Laxative episodes ^I	1.7	(2.5)	4.3	(6.2)	5.4	(8.6)	13.4	(12.3)	17.96***	cef
Diuretic episodes ^I	0.7	(1.9)	3.6	(7.0)	6.8	(9.3)	13.2	(13.3)	19.31***	bcef
Exercise episodes ^I	2.2	(2.8)	7.1	(6.9)	13.7	(9.6)	16.0	(10.9)	35.54***	abcde

F-ratios: *p .0001. df (3,195)

^I Inappropriate weight compensatory behavior episodes per month assessed with the Eating Disorder Examination–Questionnaire (EDE-Q). The EDE-Q was modified to assess the frequency of episodes (rather than days).

Posthoc tests (Scheffe) for ANOVAs indicate significant group differences as follows: a: Mild vs. Moderate; b: Mild vs. Severe; c: Mild vs. Extreme; d: Moderate vs. Severe; e: Moderate vs. Extreme; f: Severe vs. Extreme; NS indicates no statistically significant differences in post-hoc tests.

Demographic and physical characteristics of participants with bulimia nervosa across *DSM-5* severity groups.

Table 2

	Mild N=77	Moderate N=68	Severe N=32	Extreme N=22	Test Statistic	P value	Effect size
Age, mean (SD)	33.1 (11.4)	34.7 (11.1)	29.7 (9.0)	30.1 (11.6)	F(3, 169)=1.71	.168	.029
Female, No (%)	71 (92.2%)	61 (89.7%)	32 (100.0%)	21 (95.5%)	χ^2 (3, n=199)=3.80	.284	.138
White, No (%)	61 (79.2%)	51 (75.0%)	22 (68.8%)	18 (81.8%)	χ^2 (3, n=199)=1.81 ^a	.613	.095
Body Mass Index	30.4 (8.3)	29.7 (10.9)	27.4 (9.1)	27.0 (10.7)	F(3, 195)=1.18	.319	.018

^a Chi-square analysis for White versus Non-White.

Table 3

Comparison of participants with bulimia nervosa across *DSM-5* severity groups.

	Mild N=77		Moderate N=68		Severe N=32		Extreme N=22		ANOVA	
	<i>M</i>	<i>sd</i>	<i>M</i>	<i>sd</i>	<i>M</i>	<i>sd</i>	<i>M</i>	<i>sd</i>	η^2	Posthoc
OBE	10.9	(11.6)	11.1	(7.5)	16.6	(10.4)	30.9	(28.8)	14.50**	.182 cef
Weight Control	8.1	(2.7)	20.6	(5.1)	39.5	(6.6)	79.1	(24.6)	375.37**	.852 ALL
Restraint	3.6	(1.6)	3.9	(1.4)	4.3	(1.3)	4.7	(1.1)	4.06*	.059 c
Eating Concern	3.8	(1.4)	4.0	(1.3)	4.4	(1.1)	5.0	(0.9)	5.73*	.081 ce
Shape Concern	5.2	(1.0)	5.1	(0.9)	5.2	(0.8)	5.6	(0.5)	1.58	.024 NS
Weight Concern	4.6	(1.0)	4.6	(1.0)	4.6	(1.0)	5.1	(0.7)	2.06	.031 NS
EDE-Q Global	4.3	(1.0)	4.4	(0.9)	4.6	(0.8)	5.1	(0.6)	5.07*	.072 ce
BDI	21.7	(11.2)	22.9	(11.1)	23.3	(11.8)	32.8	(9.4)	5.90*	.090 cef

Note.

OBE = objective bulimic episodes (i.e., binge eating) frequency per month.

Weight control = inappropriate weight compensatory behavior episodes per month assessed with the Eating Disorder Examination – Questionnaire (EDE-Q). The EDE-Q was modified to assess the frequency of episodes (rather than days). The Weight Control variable is a composite of the different forms of inappropriate weight control behaviors summarized separately in Table 1.

BDI = beck depression inventory.

F-ratios:

** p .0001,

* p .01. df ranged from (3,179) to (3,195).

Posthoc tests (Scheffe) for ANOVAs indicate significant group differences as follows: a: Mild vs. Moderate; b: Mild vs. Severe; c: Mild vs. Extreme; d: Moderate vs. Severe; e: Moderate vs. Extreme; f: Severe vs. Extreme; NS indicates no statistically significant differences in post-hoc tests. ALL indicates all groups are statistically different.