

How useful is the DSM-5 severity indicator in bulimia nervosa? A clinical study including a measure of impairment

Article

Accepted Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Jenkins, P. E. ORCID: <https://orcid.org/0000-0003-1673-2903>, Luck, A., Cardy, J. and Staniford, J. (2016) How useful is the DSM-5 severity indicator in bulimia nervosa? A clinical study including a measure of impairment. *Psychiatry Research*, 246. pp. 366-369. ISSN 0165-1781 doi: 10.1016/j.psychres.2016.10.011 Available at <https://centaur.reading.ac.uk/76079/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1016/j.psychres.2016.10.011>

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

How useful is the DSM-5 severity indicator in bulimia nervosa? A clinical study including a measure of impairment

Abstract

The severity criterion used in DSM-5 for bulimia nervosa (BN) was investigated in 214 individuals referred for treatment at a regional eating disorders service in the UK. In addition to comparing eating disorder symptoms, impairment secondary to these symptoms was also assessed. According to guidance in DSM-5, 94 individuals were classified as mild (43.9%), 70 as moderate (32.7%), 32 as severe (15.0%), and 8 as extreme (3.7%) levels of BN severity. Due to small numbers in the latter two groups, it was necessary to combine these to form one 'severe/extreme' group. Analyses on these three groups suggested no group effect on demographic variables but differences were seen on measures of eating pathology, psychological distress, and psychosocial impairment between the mild group and other groups. Individuals in the moderate and severe/extreme groups scored comparably on most measures of pathology and impairment. The results are broadly consistent with past studies on community samples although together question the demarcation between moderate and more severe groups of individuals with BN.

Keywords: classification, diagnoses, DSM, DSM-5, bulimia nervosa

1. Introduction

The most recent, fifth, edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013) introduced a ‘severity specifier’ for all eating disorders (EDs). For individuals meeting diagnostic criteria for bulimia nervosa (BN), level of severity is based on the frequency of inappropriate compensatory behaviours (e.g., self-induced vomiting, laxative misuse), but also takes into account “other symptoms and the degree of functional disability” (APA, 2013, p. 345). Regarding frequency of these behaviours, four groups are defined: mild (an average of 1 – 3 per week); moderate (an average of 4 – 7 per week); severe (an average of 8 – 13 per week); and extreme (an average of 14 or more per week).

There has been limited research investigating the utility and correlates of the DSM-5 severity criteria for BN, nor the approximate distribution of these subgroups in adult, clinical samples. Grilo et al. (2015b) examined the severity criterion in a community sample of 14 men and 185 women. Participants completed self-report measures assessing eating pathology and depression online, and were subsequently classified using the severity specifier. Groups did not differ on demographic variables but (aside from the expected difference according to frequency of inappropriate compensatory behaviours) some differences were noted on frequency of binge eating, some subscales of the Eating Disorder Examination – Questionnaire (EDE-Q; Fairburn and Beglin, 1994), and depression symptoms. The authors argue that the results provided “new, albeit modest, support for the DSM-5 severity rating” for BN (p. 43), although numbers in the severe and extreme groups were limited, with overall effect sizes for the group effect of severity small to medium. Similar studies have recruited relatively small samples of individuals diagnosed with BN (e.g., Smink et al., 2014, $n = 8$), making wider inferences difficult and there have been no studies of BN recruited from clinical (treatment-seeking) samples.

Although other studies have found positive associations between symptoms and impairment (e.g., Hovrud and De Young, 2015), the assumption that symptom frequency is a good indicator of ED severity has been questioned. For example, MacDonald et al. (2014) have questioned where the most appropriate symptom (binge eating and compensatory behaviours) ‘threshold’ for a diagnosis of BN should lie and other authors (e.g., Keel et al., 2013) have suggested use of cognitive measures or symptom history to inform severity (see also Grilo et al., 2008). Furthermore, there have been no published studies that have looked explicitly at degree of impairment across the severity range of BN, with generally little research into

functional impairment in DSM-5 diagnoses (Stice et al., 2013). The current study therefore has two main aims. The first is to describe the range of severity, according to DSM-5 criteria, that is seen in a treatment-seeking BN sample. A second aim is to look at association of severity with psychosocial impairment and psychological distress, and to investigate relationships with (diagnostic) severity and variance across groups. It was hypothesised that groups would differ on inappropriate compensatory behaviours, binge eating, and subsequent impairment, but would be similar on demographic measures (e.g., age) and core ED symptoms (e.g., overvaluation of weight and shape).

2. Methods

Participants were drawn from consecutive referrals to two specialist eating disorders units based in Oxfordshire and Buckinghamshire, UK; these units are governed by the same NHS Trust and cover a population of approximately one million adults. Data were collected between February 2012 and November 2015 from consecutive assessments although it is possible that a small proportion of those assessed did not provide data for the study.

At assessment, individuals met with a qualified clinician (e.g., nurse specialist, clinical psychologist) who confirmed diagnoses following semi-structured interview. Weight and height were recorded during this interview using calibrated scales. Self-report questionnaires (see Measures) were sent to individuals in advance and collected at interview.

2.1. Participants

Two hundred and fourteen adults ($n = 204$ females; 95.3%) with full-syndrome BN, according to DSM-5 criteria, were included in the study. Of 204 who provided data, the majority ($n = 193$; 94.6%) were from a 'white' ethnic background. Of 198 who provided data, 106 (53.5%) were employed at the time of assessment, 58 (29.3%) were in full-time education, 27 (13.6%) unemployed, and seven (3.5%) reported employment status as 'other'. As this was a retrospective review of routinely collected data, the local NHS Trust Research and Development department approved the study as an audit and thus further ethical review was not sought.

2.2. Measures

The Eating Disorder Examination – Questionnaire (EDE-Q; Fairburn and Beglin, 1994) is a self-report questionnaire assessing behavioural and cognitive symptoms of an ED over the previous 28 days and has good psychometric properties (see Berg et al., 2012). It asks respondents to indicate how frequently certain behaviours have occurred (such as the frequency of self-induced vomiting and objective binge-eating episodes; OBEs, defined as eating a large amount of food with an associated loss of control; see APA, 2013). A Global score can be calculated from the cognitive items, which provides a general index of ED pathology (e.g., Friberg et al., 2013). In addition, an index of the core ED symptoms of ‘overvaluation of weight and shape’ (OWS) can be computed, using an average of items 22 and 23 (e.g., Goldfein et al., 2000; Goldschmidt et al., 2010); this was used in the current study as an additional measure to compare groups. The measure has been widely used, and a large German sample has recently provided normative data (Hilbert, de Zwaan, & Braehler, 2012), finding that around 6% of women score above a cut-off of 2.30 on the Global score (see Mond et al., 2004).

The Clinical Impairment Assessment questionnaire (CIA; Bohn and Fairburn, 2008) assesses psychosocial impairment resulting from ED symptoms over the last 28 days. Good psychometric properties have been reported alongside prevalence data from a clinical sample (Jenkins, 2013). The measure asks participants to consider how issues related to their eating have affected their lives, rated on a 0 – 3 Likert scale (where 0 = ‘Not at all’, 3 = ‘A lot’) and a cut-off of 16 has been suggested to indicate clinical severity (Bohn et al., 2008).

The Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM; Barkham et al., 2001) is a 34-item measure of general psychological distress concerning symptoms experienced over the previous week. Items are scored from 0 – 4 (0 = ‘Not at all’, 4 = ‘All the time’) and a Total score is calculated as a mean of all items, subsequently multiplied by 10 to aid interpretation (e.g., Connell et al., 2007). It has been suggested that a cut-off score of 10 can distinguish between clinical and general population samples (Connell et al., 2007).

2.3. Procedures

Severity groups were created according to responses on the EDE-Q (Fairburn and Beglin, 1994, 2008), as per the method of Grilo et al. (2015b). Specifically, responses to questions assessing the frequency of inappropriate compensatory behaviours (i.e. self-induced

vomiting, laxative use, and “driven” exercise [Fairburn and Beglin, 2008]), were used to assign a DSM-5 severity rating (mild, moderate, severe, extreme). However, due to the small number of individuals in the ‘extreme’ group, this was combined to form a ‘severe/extreme’ group (see below), similar to Grilo et al. (2015c) in a sample of individuals with binge-eating disorder. One outlier was removed for statistical analysis due to potentially unreliable data (a z -score of 6.79 regarding total inappropriate compensatory behaviours and extremely low scores on other measures).

2.4. Statistical analysis

Variables were analysed by general linear model analysis of variance (ANOVA) using the Brown-Forsythe F -ratio to correct for heterogeneity of variances and control Type 1 error rates. Where a significant group effect was observed, Dunnett T3 post hoc tests were conducted to examine this in more detail. As an estimate of effect size, partial η^2 (η_p^2) was calculated, with values of approximately 0.01, 0.06, and 0.14 corresponding to small, medium, and large effect sizes respectively (Cohen, 1988). An estimate of sample size was made with G*Power (Faul et al., 2007), which suggested a minimum N of 207 for three groups. This was based on an anticipated effect size of $f = 0.25$ ($\eta_p^2 \approx 0.06$; Cohen, 1988), to achieve power of 0.9 at $\alpha = 0.05$.

3. Results

Table 1 shows demographic data for both sites, which were considered as one sample due to similarities across dependent variables and service specifications.

Table 1. Demographic data from the sample. Means and SD are reported

Variable, mean (SD)	Site 1 (n = 140)	Site 2 (n = 74)	Total sample [†]	p (2-tailed t -test)
Age, years	27.7 (9.9)	27.8 (8.8)	27.8 (9.5)	0.94
Body mass index, kg/m ²	23.3 (5.3)	23.7 (5.0)	23.4 (5.2)	0.54
[‡] Episodes of				
Vomiting	22.8 (27.6)	28.7 (39.1)	24.8 (32.1)	0.26
Laxative use	4.2 (9.5)	5.0 (10.1)	4.5 (9.7)	0.62
Driven Exercise	8.0 (9.1)	7.9 (8.7)	8.0 (9.0)	0.96
OBEs	18.3 (18.3)	19.8 (13.2)	18.8 (16.7)	0.52

EDE-Q Global	4.4 (1.0)	4.5 (1.0)	4.5 (1.0)	0.64
OWS	5.0 (1.3)	5.2 (1.3)	5.1 (1.3)	0.24
CIA Total	33.8 (8.9)	34.5 (9.6)	34.1 (9.1)	0.59
CORE-OM Total	19.9 (6.8)	20.6 (7.8)	20.1 (7.1)	0.52

[†]Exact numbers per group vary slightly as complete information was not available on all individuals [‡]Indicates frequency of behaviours over the last 28 days

Note. CIA = Clinical Impairment Assessment questionnaire; CORE-OM = Clinical Outcomes in Routine Evaluation – Outcome Measure; EDE-Q = Eating Disorder Examination – Questionnaire; OBEs = objective binge-eating episodes; OWS = overvaluation of weight and shape

Within the entire sample, 94 individuals (43.9%) comprised the mild group, 70 (32.7%) the moderate group, 32 (15.0%) severe, and 8 (3.7%) extreme (thus, 40 [18.7%] in the severe/extreme group). In line with guidance (APA, 2013), group membership was assigned based on frequency of inappropriate compensatory behaviours. However, approximately 4.7% of the sample were missing this information from their questionnaires (i.e., they did not report frequencies of inappropriate compensatory behaviours) and were therefore not classified.

Table 2 summarises group differences for individuals with BN, based on DSM-5 severity. As expected, no differences in age or BMI were observed and there was no main effect of group on the OWS criterion. In accordance with the methodology, all groups differed on frequency of self-induced vomiting, with the mild group engaging in this less frequently, and the severe/extreme group reporting more frequent symptoms. For both laxative use and driven exercise, the mild group reported significantly less frequent symptoms than the other two groups, which were equivalent (although the *p*-value for laxative use approached significance, at 0.06). Although these differences support valid application of DSM-5 severity criteria, a possible confound to this method of investigation concerns co-occurrence of different compensatory behaviours. Among those for whom data were available on inappropriate compensatory behaviours (*N* = 198), 42 individuals (21.2%) reported presence of all three behaviours studied (self-induced vomiting, laxative use, driven exercise), 47 (23.7%) reported only vomiting, two (1.0%) only using laxatives, and 13 (6.6%) only engaging in driven exercise. Eighty-five individuals (42.9%) reported using two of these three compensatory behaviours.

There were differences in frequency of OBEs and scores on the EDE-Q Global, CORE-OM Total, and CIA Total. For all of these measures, post hoc tests suggested that the mild group scored significantly lower than other groups but that scores in the severe/extreme group were statistically similar to the moderate group. Effect size estimates indicate that significant group differences seen were of at least medium size.

Table 2. Frequency of inappropriate compensatory behaviours (per week over the last 28 days) and measures of ED psychopathology, psychosocial impairment, and psychological distress across DSM-5 severity groups. Means and SD are reported

	Mean (SD)			Test statistic	η_p^2	Post hoc
	Mild	Moderate	Severe / Extreme			
Age	26.9 (9.9)	28.7 (9.5)	27.2 (8.2)	F(2,174) = 0.80	0.009	ns
BMI	24.0 (6.4)	22.7 (3.7)	23.2 (3.7)	F(2,172) = 1.67	0.019	ns
Episodes of						
Vomiting	2.2 (2.1)	6.0 (3.2)	14.9 (9.4)	F(2,48) = 56.30**	0.702	ALL
Laxative use	0.3 (0.8)	1.3 (2.3)	2.9 (4.0)	F(2,61) = 11.43**	0.272	a,b
Driven Exercise	1.2 (1.7)	2.5 (2.4)	3.1 (2.5)	F(2,120) = 11.21**	0.158	a,b
OBEs	3.5 (2.6)	4.8 (3.1)	7.3 (7.0)	F(2,59) = 8.17**	0.218	a,b
EDE-Q Global	4.2 (1.1)	4.6 (0.8)	4.7 (1.1)	F(2,136) = 5.37*	0.073	a,b
OWS	4.9 (1.4)	5.3 (1.1)	5.2 (1.3)	F(2,149) = 2.16	0.028	ns
CIA Total	31.5 (9.6)	36.3 (7.9)	36.4 (9.2)	F(2,146) = 7.40**	0.092	a,b
CORE-OM	18.1 (6.6)	21.0 (7.1)	23.6 (7.4)	F(2,138) = 9.01**	0.115	a,b
Total						

Note. CIA = Clinical Impairment Assessment questionnaire; CORE-OM = Clinical Outcomes in Routine Evaluation – Outcome Measure; EDE-Q = Eating Disorder Examination – Questionnaire; OBEs = objective binge-eating episodes; OWS = overvaluation of weight and shape

** $p \leq 0.001$, * $p \leq 0.01$, ns = nonsignificant. Post hoc tests indicate significant group differences ($p < 0.05$) as follows: a: Mild vs. Moderate; b: Mild vs. Severe/Extreme; c: Moderate vs. Severe/Extreme. ALL indicates all groups are statistically different.

4. Discussion

Alongside work from non-clinical samples (e.g., Grilo et al., 2015b), results suggest that there is some empirical validity to the severity criterion for BN adopted in DSM-5. As expected, groups did not differ on demographics or on the core ED symptom of overvaluation of weight and shape (see also Grilo et al., 2015b). An unexpected finding was that the distribution of individuals grouped by severity in this clinical sample was very similar to the community volunteers recruited by Grilo et al., although all met criteria for BN – the majority of individuals (over 80% in the current study and nearly 75% of those in Grilo et al.) comprised the mild or moderate groups. Similar findings regarding the comparability of clinical groups and individuals recruited from the community have been reported by Grilo and colleagues in studies of binge-eating disorder (see Grilo et al., 2015a,c).

Aside from the relative distribution of severity groups, effect sizes were of comparable magnitude in the current study and that of Grilo et al. (2015b), a study with a similar methodology. Although a degree of caution should be exercised when comparing effect sizes across studies, particularly those with different population characteristics (Olejnik and Algina, 2003), this result suggests that severity accounts for around 7 – 11% of the variance in eating pathology, psychosocial impairment, and psychological distress respectively. However, in both the current sample and that of Grilo et al., not all group comparisons revealed significant differences in symptoms. For example, although frequency of self-induced vomiting differed across groups, as expected, this was not the case for other inappropriate compensatory behaviours (laxative use and driven exercise). As has been previously shown (e.g., Favaro & Santonastaso, 1996), self-induced vomiting was the most common compensatory behaviour in the current sample. Those who used other methods (laxative use, driven exercise) commonly reported also using vomiting, suggesting that variability in findings may be due to significant co-occurrence of such behaviours, although more detailed analyses were not conducted. A separate study by the authors is currently underway that aims to assess which features of BN contribute to impairment in order to understand this relationship in more detail.

Although the mild group scored consistently lower than other groups, there were no differences between individuals classified as either moderate or severe/extreme on measures of general eating pathology, psychosocial impairment, and psychological distress. In the study of Grilo et al., however, differences more consistently occurred between the extreme

group and others. This finding may be attributable to slightly different methodologies between the two studies (methods of recruitment, grouping of severity) or may reflect impaired validity of the severity criterion. Results may have also been confounded by small group sizes in the extreme group in particular (in both study samples). The results are interesting in light of the changes made to the pre-existing DSM (APA, 2000), which specified that binge eating and compensatory behaviours must be present “on average, at least twice a week for 3 months” (p. 594). This has since been amended to specify that the behaviours must occur, on average, *once* a week (APA, 2013). Although the current study did not provide a direct test of the revised frequency criterion (see MacDonald et al., 2014), results suggest that significant impairment is seen in those classified as ‘mild’ BN severity, with compensatory behaviours occurring once a week on average. However, differences between ‘moderate’ (4 – 7 inappropriate compensatory behaviours per week) and more severe (7 or more inappropriate compensatory behaviours per week) are limited.

Taken alongside prior studies, the results suggest that individuals classified as ‘extreme’ in DSM-5 are rare, even in individuals seeking treatment for BN. Due to unequal group sizes, comparisons with the extreme group were deemed unsuitable and thus a composite ‘severe/extreme’ group was created that better matched the remaining groups in terms of size. It is notable that Grilo et al. (2015c) used the same method in a clinical sample of individuals with binge-eating disorder, where fewer than 13% of individuals were classified as severe or extreme. Further work might seek to investigate where the suitable ‘cut-off point’ for severity might lie. In contrast, other authors have suggested that classification of BN and similar disorders is more accurate when considered through presence of other symptoms, such as weight phobia, raising the question of how best to gauge ‘severity’ in eating disorders (e.g., Keel et al., 2013).

The current study goes beyond previous work in a number of ways. Firstly, existing studies in BN have only recruited community samples. Secondly, the study also looks at impairment secondary to ED features, finding that those in the mild group reported significantly lower levels of impairment than other groups, although moderate and severe/extreme groups scored similarly. Thirdly, a measure of psychological distress was included, finding analogous results to a measure of impairment.

There were few exclusion criteria for the study so, while results are from a large sample of individuals with BN presenting to a specialist eating disorders service, they may also be

confounded by other factors (e.g., comorbidity). Use of a measure of impairment relating specifically to the effects of ED symptoms may have improved validity (as opposed to using a more generic measure of impairment), although it is not possible to make definitive conclusions. A further limitation was that participants were sent questionnaire packs in advance of their appointment. Given that some problems with the EDE-Q in the assessment of binge eating have been highlighted (Mond et al., 2004; see also Grilo et al., 2001) there may have been some discrepancies in the frequency of symptoms reported between self-report and interview. As in the study of Grilo et al. (2015b), men were under-represented, reflecting the uneven gender distribution seen in EDs (e.g., see Jones and Morgan, 2010). Similarly, using recent census data (ONS, 2011), the population of Oxfordshire and Buckinghamshire is approximately 89% White, suggesting that ethnic minorities were also under-represented in this sample (e.g., see Marques et al., 2011). Future work might seek to establish the utility of the severity specifier as a longitudinal predictor of outcome, for example regarding treatment response, and the findings summarised here would need to be replicated in a larger and more diverse sample.

The results presented here are similar to those found in a US community sample (Grilo et al., 2015b), and therefore the generalisability appears to be high. If the severity criterion is to be continued, future studies might look to refine how this is measured in EDs, adding to the evidence for its reliability and validity.

Conflicts of Interest and Source of Funding: This work was supported by resources and use of facilities within Oxford Health NHS Foundation Trust. The authors report no competing interests.

References

- APA (American Psychiatric Association), 2000. Diagnostic and Statistical Manual of Mental Disorders–Text Revision (DSM-IV-TR). American Psychiatric Association, Washington, DC.
- APA (American Psychiatric Association), 2013. Diagnostic and Statistical Manual of Mental Disorders: DSM-5, 5thed. American Psychiatric Association, Washington, DC.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., McGrath, G., 2001. Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based evidence in the psychological therapies. *J. Consult. Clin. Psychol.* 69(2), 184-196.
- Berg, K.C., Peterson, C.B., Frazier, P., Crow, S.J., 2012. Psychometric evaluation of the Eating Disorder Examination and Eating Disorder Examination-Questionnaire: a systematic review of the literature. *Int. J. Eat. Disord.* 45(3), 428-438.
- Bohn, K, Doll, H. A., Cooper, Z., O'Connor, M., Palmer R. L., Fairburn, C. G., 2008. The measurement of impairment due to eating disorder psychopathology. *Behav. Res. Ther.* 46, 1105-1110.
- Bohn, K., Fairburn, C.G., 2008. The Clinical Impairment Assessment questionnaire (CIA 3.0), in: Fairburn, C.G. (Ed.), *Cognitive Behaviour Therapy and Eating Disorders*. Guilford Press, New York, pp. 351–318.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*, second ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, O., Miles, J.N.V., 2007. Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *Br. J. Psychiatry* 190, 69-74.
- Fairburn, C.G., Beglin, S.J., 1994. Assessment of eating disorders: interview or self-report questionnaire? *Int. J. Eat. Disord.* 16, 363-370.
- Fairburn, C.G., Beglin, S.J., 2008. Eating Disorder Examination Questionnaire (EDE-Q 6.0), in: Fairburn, C. G. (Ed.), *Cognitive Behaviour Therapy and Eating Disorders*. Guilford Press, New York, pp. 309-313.

- Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175-191.
- Favaro, A., Santonastaso, P., 1996. Purging behaviors, suicide attempts, and psychiatric symptoms in 398 eating disordered subjects. *Int. J. Eat. Disord.* 20, 99-103.
- Friborg, O., Reas, D.L., Rosenvinge, J.H., Rø, Ø., 2013. Core pathology of eating disorders as measured by the Eating Disorder Examination Questionnaire (EDE-Q): the predictive role of a nested general (g) and primary factors. *Int. J. Methods Psychiatr. Res.* 22(3), 195-203.
- Goldfein, J.A., Walsh, B.T., Midlarsky, E., 2000. Influence of shape and weight on self-evaluation in bulimia nervosa. *Int. J. Eat. Disord.* 27, 435-445.
- Goldschmidt, A.B., Hilbert, A., Manwaring, J.L., Wilfley, D.E., Pike, K.M., Fairburn, C.G., Striegel-Moore, R.H., 2010. The significance of overvaluation of shape and weight in binge eating disorder. *Behav. Res. Ther.* 48(3), 187-193.
- Grilo, C. M., Hrabosky, J. I., White, M. A., Allison, K. C., Stunkard, A. J., Masheb, R M., 2008. Overvaluation of shape and weight in binge eating disorder and overweight controls: refinement of a diagnostic construct. *J. Abnorm. Psychol.* 117, 414-419.
- Grilo, C.M., Ivezaj, V., White, M.A., 2015a. Evaluation of the DSM-5 severity indicator for binge eating disorder in a community sample. *Behav. Res. Ther.* 66, 72-76.
- Grilo, C.M., Ivezaj, V., White, M.A., 2015b. Evaluation of the DSM-5 severity indicator for bulimia nervosa. *Behav. Res. Ther.* 67, 41-44.
- Grilo, C.M., Ivezaj, V., White, M.A., 2015c. Evaluation of the DSM-5 severity indicator for binge eating disorder in a clinical sample. *Behav. Res. Ther.* 71, 110-114.
- Grilo, C.M., Masheb, R.M., Wilson, G.T., 2001. Different methods for assessing the features of eating disorders in patients with binge eating disorder: A replication. *Obes. Res.* 9(7), 418-422.
- Hilbert, A., de Zwaan, M., Braehler, E., 2012. How frequent are eating disturbances in the population? Norms of the Eating Disorder Examination-Questionnaire. *PLoS ONE* 7, e29125. doi:10.1371/journal.pone.0029125.

- Hovrud, L., De Young, K.P., 2015. Unique contributions of individual eating disorder symptoms to eating disorder-related impairment. *Eat. Behav.* 18, 103-106.
- Jenkins, P.E., 2013. Psychometric validation of the Clinical Impairment Assessment in a UK eating disorder service. *Eat. Behav.* 14, 241-243.
- Jones, W.R., Morgan, J.F., 2010. Eating disorders in men: a review of the literature. *J. Public Ment. Health* 9(2), 23-31.
- Keel, P.K., Crosby, R.D., Hildebrandt, T.B., Haedt-Matt, A.A., Gravener, J.A., 2013. Evaluating new severity dimensions in the DSM-5 for bulimic syndromes using mixture modelling. *Int. J. Eat. Disord.* 46, 108-118.
- MacDonald, D.E., McFarlane, T.L., Olmsted, M.P., 2014. "Diagnostic shift" from eating disorder not otherwise specified to bulimia nervosa using DSM-5 criteria: a clinical comparison with DSM-IV bulimia. *Eat. Behav.* 15, 60-62.
- Marques L., Alegria, M., Becker, A.E., Chen, C.-n., Fang, A., Chosak, A., Diniz, J.B., 2011. Comparative prevalence, correlates of impairment, and service utilization for eating disorders across U.S. ethnic groups: implications for reducing ethnic disparities in health care access for eating disorders. *Int. J. Eat. Disord.* 44, 412-420.
- Mond, J.M., Hay, P.J., Rodgers, B., Owen, C., Beumont, P.J.V., 2004. Validity of the Eating Disorders Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behav. Res. Ther.* 42(5), 551-567.
- Office for National Statistics (2012). 2011 Census: Key Statistics for Local Authorities in England and Wales (Table KS201EW). Author, Newport, UK.
- Olejnik, S., Algina, J., 2003. Generalised eta and omega squared statistics: measures of effect size for some common research designs. *Psychol. Methods* 8, 434-447.
- Smink, F.R.E., van Hoeken, D., Oldehinkel, A.J., Hoek, H.W., 2014. Prevalence and severity of DSM-5 eating disorders in a community cohort of adolescents. *Int. J. Eat. Disord.* 47, 610-619.
- Stice, E., Marti, C.N., Rohde, P., 2013. Prevalence, incidence, impairment, and course of the proposed DSM-5 eating disorder diagnoses in an 8-year prospective community study of young women. *J. Abnorm. Psychol.* 122, 445-457.