A test of the DSM-5 severity specifier for bulimia nervosa in adolescents: Can we anticipate clinical treatment outcomes?

Sasha Gorrell[1], **Lisa Hail**[1], **Kathryn Kinasz**[1], **Lindsey Bruett**[1], **Sarah Forsberg**[1], **Kevin Delucchi**[1], **James Lock**[2], and **Daniel Le Grange**[1,3]

[1]University of California, San Francisco

[2]Stanford University

[3]The University of Chicago (Emeritus)

Abstract

Objective: This study tested clinical utility of the DSM-5 severity specifier for bulimia nervosa (BN) in predicting treatment response among adolescents ($N = 110$) within a randomized clinical trial of two psychosocial treatments.

Method: Analyses grouped individuals meeting criteria for BN diagnosis by baseline severity, per DSM-5. Associations among baseline severity classification and BN behavior (i.e., binge eating and compensatory behavior) and Eating Disorder Examination (EDE) Global scores at end-of-treatment (EOT), 6- and 12-month follow-up were examined.

Results: Associations between severity categories with BN symptoms were not significant at EOT, or follow-up. Test for linear trend in BN behavior was significant at EOT, $F = 5.23$, $p = .02$, without demonstrating a linear pattern. Relation between severity categories with EDE Global scores was significant at 6-month follow-up, $F = 3.76$, $p = .01$. Tests for linear trend in EDE Global scores were significant at EOT, $F = 5.40$, $p = .02$, and at 6 months, $F = 10.73$, $p = .002$, with the expected linear pattern.

Discussion: Findings suggest the DSM-5 BN severity specifier holds questionable utility in anticipating outpatient treatment response in adolescents with BN. The specifier may have improved ability to predict attitudinal rather than behavioral treatment outcomes.

Keywords

Bulimia nervosa (BN) is a disabling eating disorder (ED) characterized by episodes of binge eating and compensatory behavior with concomitant overvaluation of weight and shape. BN often begins in adolescence though there is limited research evaluating treatment outcomes in this demographic (Le Grange, Loeb, Van Orman, & Jellar, 2004). To aid in improved

---

Article Correspondence: Sasha Gorrell, Department of Psychiatry, University of California, San Francisco. 401 Parnassus Avenue, San Francisco, CA 94143. Phone: 917-685-5436. Sasha.Gorrell@ucsf.edu.

diagnosis of BN, the fifth edition of the *Diagnostic and Statistical Manual for Mental Disorders* (DSM-5; American Psychiatric Association (APA), 2013) eliminated prior purging- and non-purging subtypes of BN, and also introduced severity specifiers. The severity classifications are based on the frequency of compensatory behaviors (i.e., self-induced vomiting, laxative, diuretic, medication misuse, fasting, or excessive exercise), according to average episodes per week: mild (1–3), moderate (4–7), severe (8–13), and extreme (14).

Introduction of severity ratings across diagnoses was intended to aid clinicians in identifying gradients of a disorder that may not be fully captured by strict categorical approaches (Regier, Kuhl, & Kupfer, 2013), and to provide information regarding associated risks and prognosis (Gianini et al., 2017). Initial tests of DSM-5 specifiers in adult samples with BN indicated significant differences between severity categories (Dakanalis et al., 2017) but some work has indicated small effect sizes, and only modest support for their utility in adult samples (Grilo, Ivezaj, & White, 2015). The specifiers have recently been tested in youth to determine valid differences among categories (Dakanalis et al., 2018). In this sample, support was evidenced for the severity indicator, such that individuals in the 'extreme' group reported significantly higher levels of ED pathology at initial evaluation. Given uncertainty regarding the robustness of severity specifiers in adult samples with BN, and the paucity of their investigation specifically in youth, further investigation of the specifiers is warranted.

The goal of evidence-based treatments (EBTs) for BN is to interrupt patterns of dieting, binge eating and compensatory behavior, with the ultimate goal of abstinence from these symptoms. Within BN treatment, behavioral symptoms (e.g., binge eating, purging) typically remit prior to attitudinal symptoms (e.g., undue influence of body shape and weight on self-evaluation) (Ciao, Accurso, Fitzsimmons-Craft, & Le Grange, 2015). Both mitigation of behavioral symptoms and attitudinal features are important indicators of recovery (e.g., Halmi et al., 2002), and thus represent BN treatment goals. Improved prediction of treatment response is critical to guiding intervention and in creating realistic expectations for patients, families, and providers alike. To this end, severity classification at baseline may offer clinical utility in predicting response to EBT for BN in adolescents, a relation not examined to date.

Current Study

Analyses were secondary, implemented specifically to investigate outcomes related to eating pathology. The current study examined associations between severity specifier categories and attitudinal (i.e., cognitive and psychological) symptoms of BN at end-of-treatment (EOT), and at 6- and 12-month follow-up. We also examined associations between severity specifier categories and rates of binge eating and compensatory behavior at all three time points. As secondary aims, analyses compared a) rates of abstinence from binge eating and compensatory behavior and b) rates of binge eating behavior (i.e., a composite of subjective and objective binge eating) between baseline severity categories. Of note, adolescents in the current sample received either family-based treatment for BN (FBT-BN) or cognitive behavior therapy adapted for adolescents (CBT-A). Although adolescents treated in FBT-BN were significantly more likely to achieve abstinence from BN symptoms at EOT, both

groups continued to improve over the year following EOT with no significant differences in outcome at 12-month follow-up (Le Grange, Lock, Agras, Bryson, & Jo, 2015). In this secondary data analysis of the original RCT, and testing across the combined sample, we hypothesized that those with elevated baseline severity would be more likely to demonstrate greater symptomatology at EOT and extended follow-up, both attitudinally and behaviorally.

Method

Participants and procedure

A detailed description of the study sample can be found in the main outcome report (c.f., Le Grange et al., 2015). Briefly, participants ($N = 110$; 93.6% female) aged 12–18 from two sites who met DSM-IV criteria for BN or BN-type ED not otherwise specified (APA, 2000) were randomized to one of two active treatment types, CBT-A ($n = 58$) or FBT-BN ($n = 52$). Diagnoses and ED symptom report were determined by Eating Disorder Examination (EDE; Cooper & Fairburn, 1987) interview. Institutional review boards at all participating institutions approved study protocols, and all participants provided informed consent or assent (in the case of minors) prior to participation.

Analytical Plan

Reported BN behaviors (i.e., binge eating; self-induced vomiting; misuse of laxatives, diuretics or medication; excessive exercise; fasting) over the past three months were tallied and divided by 12 to produce a weekly average, which was used to confirm DSM-5 diagnosis. Based on symptom report of compensatory behaviors from baseline EDE interview and DSM-5 classification criteria, participants were then classified into one of four categories: 'mild,' 'moderate,' 'severe,' and 'extreme'. The attitudinal outcome was determined as EDE Global score (i.e., mean of the four subscales) at EOT, 6- and 12-month follow-up. Also tested at all three time points, the behavioral outcome (i.e., composite of binge eating and compensatory behaviors) was based on report of symptoms in the prior month. To compare the two outcomes across categories, a general linear model was used to test for differences by severity. A linear contrast was estimated to test for a linear trend in each model. Linear trend estimation is a statistical technique that can be used to aid in the interpretation of categorical data. Specifically, trend estimation was used in the current study to determine if our categorical data exhibited an increasing or decreasing trend, in a manner statistically distinguished from random behavior. Secondary analyses tested relations between baseline severity categories and abstinence from binge eating and compensatory behavior, as well as relations between baseline severity categories and binge eating behavior, across all measured time points. Significance was determined with p values of $< .05$, and Cohen's f^2 (small: .1; medium: .25; large: .4) was used to benchmark effect sizes, based on variance explained by eta, and partial eta squared. SAS v9.4 was used for all analyses.

Results

Descriptive Statistics

This report focused on pre-treatment severity specifiers and associations with attitudinal (i.e., EDE Global scores) and behavioral indicators of BN recovery (i.e., binge eating and

compensatory behavior). At baseline, all participants met DSM-5 diagnosis for BN; 17 (15.45%) individuals were classified as mild, 32 (29.09%) moderate, 37 (33.64%) severe, and 24 (21.82%) extreme regarding severity specifiers.

Primary Analyses

As the outcome measure of BN behavior was skewed, a log-transformed version was calculated. Results were largely unchanged for the non-transformed data; accordingly, the original scale for both outcomes is reported for ease of interpretation (Table 1).

Behavioral outcome.

Tests of a behavioral model of outcome as a continuous variable (i.e., number of BN behaviors) did not indicate significance in the full model at EOT, or 6- and 12- month follow-up. The test for trend for the outcome was statistically significant at EOT, $F = 5.23$, $p = .02$, $\omega_p^2 = .06$, 95% CI [0, .18], with a small-medium effect size, but did not follow a linear pattern and did not yield differences as we might expect at the extremes of the distribution (Figure 1). Specifically, means of BN behavior at EOT were higher in the 'mild' category than for 'moderate.' And while means for 'moderate' were lower than for 'severe,' means for 'extreme' were lower than for 'severe.' Linear trends were not significant for either the 6- or 12-month follow-ups.

Attitudinal outcome.

Tests of the attitudinal model for outcome (i.e., EDE Global scores) did not indicate significance in the full model at EOT, but test for trend was significant, $F = 5.40$, $p = .02$, $\omega_p^2 = .06$, 95% CI [0, .17], with a medium effect size. This trend demonstrated a pattern of increasing means at EOT according to increase in severity category at baseline. The full model for 6-month follow-up was significant, $F = 3.76$, $p = .01$, $\omega_p^2 = .15$, 95% CI [.01, .27], and also demonstrated a significant trend, $F = 10.73$, $p = .002$, $\omega_p^2 = .14$, 95% CI [.02, .29], both of which evidenced large effect sizes. As with the EOT model, this trend also demonstrated a pattern of increasing means at 6-month follow-up according to increase in severity category at baseline. Neither the full model, nor the trend, demonstrated significance at 12-month follow-up.

Secondary Analyses

Abstinence from binge eating and compensatory behaviors.—Tests of the relation between baseline severity category and abstinence (i.e., cessation from a combination of binge eating and compensatory behaviors) were not significant at EOT, or 6-month follow-up. The abstinence model demonstrated significant outcomes at 12-month follow-up, $\chi^2 = .02$.

Binge eating.—Tests of the relation between severity category and binge eating behavior (i.e., combined subjective and objective binge eating episodes) did not indicate significant outcomes across measured time points.

Discussion

This study sought to examine the predictive utility of pre-treatment severity classification on both behavioral and attitudinal markers of treatment outcome in adolescents receiving outpatient treatment for BN within the context of a randomized clinical trial. Results of this study indicate that when considering attitudinal status following treatment, adolescents presenting to treatment with mild baseline severity reported fewer symptoms of ED pathology at EOT, with a linear trend indicating that with each advance in clinical severity category at baseline, we might expect a correspondent elevation in attitudinal ED pathology at EOT. In contrast, those who endorsed the lowest frequency of compensatory behavior at baseline (i.e., mild) were not necessarily more likely to demonstrate the lowest number of BN behaviors at EOT. In particular, the clinical utility of the severity specifier appeared to perform poorly at either end of the scale, in the 'mild' and 'extreme' categories. At 6-month follow-up, those who demonstrated elevated severity at baseline were significantly more likely to also endorse elevated scores on attitudinal symptoms of BN, with a trend that was significant and in the increasing pattern that we might expect (i.e., those with the greatest severity at baseline were also the most symptomatic at follow-up). Both the full model and the trend of attitudinal outcomes at 6-months demonstrated large effect sizes; this is in contrast to the medium and small-medium effect sizes demonstrated by the trends in attitudinal and behavioral outcome, respectively, at EOT. Of note, neither the full model for attitudinal or behavioral symptoms, nor tests for trend, were significant at 12-month follow-up, bringing into question the utility of the specifiers in predicating longer-term outcomes. When considering the behavioral outcome variable as categorical (i.e., abstinence), baseline severity categories only significantly predicted rates of abstinence from binge eating and compensatory behavior at 12-month follow-up. Further, the current study did not indicate differences in the severity categories in predicting rates of binge eating behavior. These models favoring categorial outcomes did not consistently bear out significant results, but depending on context, reduction in a symptom may hold more clinical meaning than simply its presence versus absence. For instance, it may be that an individual had engaged in purging only once in the prior month, as compared to daily at baseline; while standards of remission might favor abstinence from purging at EOT, this reduction would certainly constitute a clinically meaningful result of treatment.

There are several possible explanations for our findings. Large-scale treatment trials for BN have indicated that there are differences in the rate at which behavioral symptoms subside, relative to attitudinal symptoms (e.g., Ciao et al., 2015). Results from the current study confirm this general finding, and suggest that there are other factors that influence more rapid behavioral change within the context of psychosocial treatment, other than simply baseline severity. One such factor could be motivation for change. While research on motivation within the context of ED treatment does not demonstrate consistency in improved outcomes directly related to motivation (Waller, 2012), it is possible that variable motivation throughout treatment contributed to the inconsistencies in behavioral outcomes at EOT evidenced in the current study.

Other possible explanations for the current findings include the nature and size of the sample. The ability of the severity specifiers to predict behavioral outcomes at EOT was the

least effective within the extremes of the distribution, specifically for the 'mild' and 'extreme' categories. It is possible that our findings are reflective of inadequate power in a sample size that is not sufficient to detect differences between categories. It is also possible that the DSM-5 criteria for each category (e.g., ≥14 symptoms corresponds to 'extreme') allows for potentially exaggerated variability in classification, as one individual reporting 15 symptoms per week was considered categorically equivalent to another with 264. When considering the 'mild' category, one such individual may be self-inducing vomiting three times per week, whereas another might be engaging in fasting twice per week, and driven exercise once per week. While this is a hypothetical example, we might assume that the individual who is vomiting regularly at baseline may have different treatment outcomes than the individual with fasting and exercise behavior. The current study lacked power to examine specific patterns of behavior within classification categories (e.g., differences between those who self-induce vomiting versus other types of compensatory behavior), and further investigation is warranted.

The ability to use baseline presentation to predict outcome would aid clinicians, parents, and patients in understanding who might benefit most from existing EBT approaches, and potentially inform future efforts to increase precision medicine decisions in ED treatment. It should be noted that clinicians are not bound strictly to these severity cut-offs, the level of which may be increased to reflect functional disability and other clinically meaningful sequelae (APA, 2013). However, as the current study did not demonstrate the clinical utility of baseline severity in predicting behavioral outcomes following treatment, further directions might include examining how baseline severity predicts treatment outcome amongst a subset of individuals who achieve rapid symptom reduction. Significant results may be effective in planning and adapting treatment course.

Findings suggest that there may be utility to determining DSM-5 severity classification prior to the start of treatment to predict attitudinal response following treatment. Specifically, adolescents who endorse a minimum average of 14 symptomatic episodes per week in the 3-months prior to treatment, and considered 'extreme', will likely continue to report the greatest elevations in EDE Global scores at EOT, and at 6-month follow-up. Therefore, evaluating severity classification at treatment outset may aid in identifying those adolescents who are most likely to require an extended or augmented approach to existing EBTs. Correspondingly, the 'mild' severity classification may help to identify which adolescents are apt to achieve improved attitudinal outcomes after the standard course of existing EBTs. Such understanding may aid in establishing expectations for families and clinicians when planning treatment.

Limitations

Several limitations should be noted. This study included self-report of baseline behavior over the prior three months; while the EDE is a widely used assessment approach, accuracy of symptom frequency may be impaired by poor recall. Further, given the demographics of this study population, we are unable to comment on how these patterns may differ across gender. Our investigation focused specifically on ED related outcomes; future work might investigate the impact of severity on other relevant clinical features (e.g., quality of life;

treatment course and attrition). Finally, with limited sample size that rendered further divisions inadvisable for reasons of retaining statistical power, this study did not focus on outcome based on treatment type; future studies should investigate the potential differential response to different EBTs based on baseline severity classification. While the present study focuses on adolescents with BN, there may be benefit to evaluating severity classifications in other adolescent transdiagnostic samples (e.g., individuals with anorexia nervosa, who also endorse purging).

Conclusions

Severity classification based on frequency of pre-treatment behavioral symptoms predicts a linear trend in attitudinal severity but not behavioral symptom frequency at the conclusion of outpatient treatment for adolescents with BN. At 6-month follow-up, severity categories predicted attitudinal symptoms, as well as a significant trend in these characteristics, but not at later assessment. These findings suggest that specifiers may have enhanced utility in predicating attitudinal outcomes over only a limited amount of time following treatment. Adaptations to existing EBT approaches may be necessary to help some individuals presenting to treatment with extreme compensatory behavior achieve behavioral symptom remission.

Acknowledgments

References

American Psychiatric Association (2000). Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) doi:10.1176/appi.books.9780890423349

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders. Fifth Edition (DSM-5) doi:10.1176/appi.books.9780890425596

Ciao AC, Accurso EC, Fitzsimmons-Craft EE, & Le Grange D (2015). Predictors and moderators of psychological changes during the treatment of adolescent bulimia nervosa. Behaviour research and therapy, 69, 48–53. doi:10.1016/j.brat.2015.04.002 [PubMed: 25874955]

Cooper Z, & Fairburn C (1987). The Eating Disorder Examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. International Journal of Eating Disorders, 6(1), 1–8.

Dakanalis A, Bartoli F, Caslini M, Crocamo C, Zanetti MA, Riva G, ... & Carrà G (2017). Validity and clinical utility of the DSM-5 severity specifier for bulimia nervosa: results from a multisite sample of patients who received evidence-based treatment. European archives of psychiatry and clinical neuroscience, 267(8), 823–829. doi: 10.1007/s00406-016-0712-7 [PubMed: 27435722]

Dakanalis A, Colmegna F, Zanetti MA, Di Giacomo E, Riva G, & Clerici M (2018). Evaluation of the DSM-5 severity specifier for bulimia nervosa in treatment-seeking youth. Child Psychiatry & Human Development, 49(1), 137–145. doi:10.1007/s10578-017-0735-y [PubMed: 28510006]

Gianini L, Roberto CA, Attia E, Walsh BT, Thomas JJ, Eddy KT, ... & Sysko R (2017). Mild, moderate, meaningful? Examining the psychological and functioning correlates of DSM-5 eating disorder severity specifiers. International Journal of Eating Disorders, 50(8), 906–916. doi:10.1002/eat.22728 [PubMed: 28489323]

Grilo CM, Ivezaj V, & White MA (2015). Evaluation of the DSM-5 severity indicator for bulimia nervosa. Behaviour research and therapy, 67, 41–44. doi: 10.1016/j.brat.2015.02.002 [PubMed: 25744910]

Halmi KA, Agras WS, Mitchell J, Wilson GT, Crow S, Bryson SW, & Kraemer H (2002). Relapse predictors of patients with bulimia nervosa who achieved abstinence through cognitive behavioral therapy. Archives of General Psychiatry, 59(12), 1105–1109. doi.org/10.1001/archpsyc.59.12.1105 [PubMed: 12470126]

Le Grange D, Loeb KL, Van Orman S, & Jellar CC (2004). Bulimia Nervosa in Adolescents. Archives of Pediatrics & Adolescent Medicine, 158(5), 478. doi:10.1001/archpedi.158.5.478 [PubMed: 15123482]

Le Grange D, Lock J, Agras WS, Bryson SW, & Jo B (2015). Randomized clinical trial of family-based treatment and cognitive-behavioral therapy for adolescent bulimia nervosa. Journal of the American Academy of Child & Adolescent Psychiatry, 54(11), 886–894. doi.org/10.1016/j.jaac. 2015.08.008 [PubMed: 26506579]

Regier DA, Kuhl EA, & Kupfer DJ (2013). The DSM-5: Classification and criteria changes. World Psychiatry, 12(2), 92–98. doi:10.1002/wps.20050 [PubMed: 23737408]

Waller G (2012). The myths of motivation: Time for a fresh look at some received wisdom in the eating disorders? International Journal of Eating Disorders, 45(1), 1–16. doi:10.1002/eat.20900 [PubMed: 22170018]

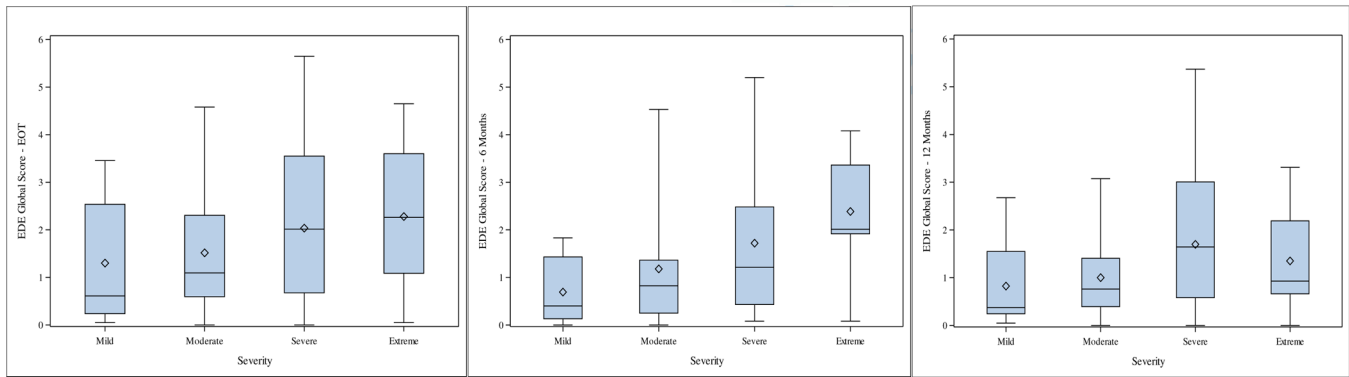**Behavioral Outcomes**
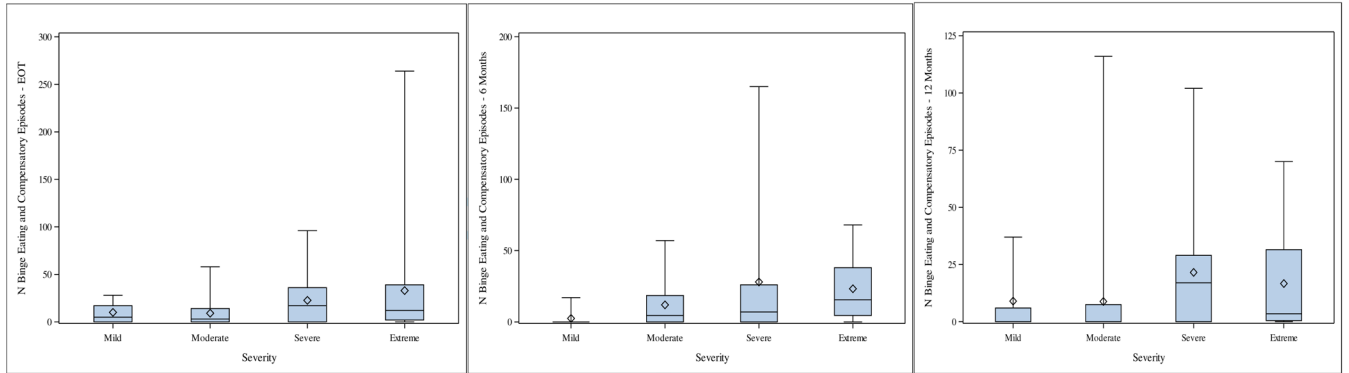
EDE Global Scores

Figure 1.
Behavioral Outcomes and EDE Global Scores According to Baseline Severity

Table 1.

General Linear Model Regression for Behavioral and Attitudinal Outcomes

Model		M(SD)	Median	F_{Full} (df)	P_{Full}	ω^2	CI_{Full}	F_{Trend} (df)	P_{Trend}	ω_p^2	CI_{Trend}
BN											
Behavior											
$N=86$	EOT	19.05 (34.79)	7.00	2.26 (3)	.09	.08	[0, .18]	5.23 (1)	**.02**	.06	[0, .18]
$N=68$	6-mth	19.06 (31.86)	6.50	2.01 (3)	.12	.09	[0, .20]	3.96 (1)	.05	.06	[0, .19]
$N=70$	12-mth	15.20 (24.91)	2.00	1.19 (3)	.32	.05	[0, .15]	1.63 (1)	.21	.02	[0, .13]
EDE											
Global											
$N=93$	EOT	1.82 (1.46)	1.50	1.95 (3)	.13	.03	[0, .15]	5.40 (1)	**.02**	.06	[0, .17]
$N=70$	6-mth	1.55 (1.39)	1.22	3.76 (3)	**.01**	.15	[.01, .27]	10.73 (1)	**.002** *	.14	[.02, .29]
$N=71$	12-mth	1.30 (1.17)	.86	2.02 (3)	.12	.08	[0, .20]	3.01 (1)	.09	.04	[0, .16]

Note: BN = Bulimia Nervosa; EOT = end-of-treatment; subscripts Full and Trend reflect values correspondent to each model. Bold values are significant at p < .05; bold values with

*
asterisk indicates p < .01.