ORIGINAL ARTICLE

Evaluating the predictive validity of purging disorder by comparison to bulimia nervosa at long-term follow-up

Katherine Jean Forney PhD[1] | Tiffany A. Brown PhD[2] | Ross D. Crosby PhD[3,4] |

Kelly M. Klein PhD[5,6] | Pamela K. Keel PhD[7]

[1]Department of Psychology, Ohio University, Athens, Ohio, USA

[2]Department of Psychological Sciences, Auburn University, Auburn, Alabama, USA

[3]Sanford Center for Biobehavioral Research, Fargo, North Dakota, USA

[4]Department of Psychiatry and Behavioral Science, University of North Dakota School of Medicine and Health Sciences, Fargo, North Dakota, USA

[5]VA Boston Healthcare System, Brockton Division, Brockton, Massachusetts, USA

[6]Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, USA

[7]Department of Psychology, Florida State University, Tallahassee, Florida, USA

Correspondence
Katherine Jean Forney, Department of Psychology, Ohio University, 22 Richland Ave, Athens, OH 45701, USA.
Email: forney@ohio.edu

Funding information
American Psychological Association Dissertation Research Award; American Psychological Foundation/COGDOP Graduate Research Scholarship; Experiment.com Crowdfunding; Florida State University Dissertation Research Award; National Institute of Mental Health, Grant/Award Numbers: F31-MH105082, R01-MH111263, R01-MH61836, R03-MH61320

Action Editor: Ruth Striegel Weissman

Abstract

Objective: The current study sought to examine the predictive validity of the purging disorder diagnosis at long-term follow-up by comparing naturalistic outcomes with bulimia nervosa.

Method: Women with purging disorder ($N = 84$) or bulimia nervosa ($N = 133$) who had completed comprehensive baseline assessments as part of one of three studies between 2000 and 2012 were sought for follow-up assessment. Nearly all (94.5%) responded to recruitment materials and 150 (69% of sought sample; 83.3% non-Hispanic white; 33.40 [7.63] years old) participated at an average of 10.59 (3.71) years follow-up. Participants completed the Eating Disorder Examination, the Structured Clinical Interview for DSM-IV, and a questionnaire battery. Diagnostic groups were compared on eating disorder (illness status, recovery status, and eating pathology) and related outcomes. Group differences in predictors of outcome were explored.

Results: There were no significant differences in eating disorder presence ($p = .70$), recovery status ($p = .87$), and level of eating pathology ($p = .17$) between diagnostic groups at follow-up. Post hoc equivalence tests indicated group differences were smaller than a medium effect size (p's ≤ .005). Groups differed in diagnosis at follow-up ($p = .002$); diagnostic stability was more likely than cross-over to bulimia nervosa for women with baseline purging disorder ($p = .004$).

Discussion: Although purging disorder and bulimia nervosa do not differ in long-term outcomes, the relative stability in clinical presentation suggests baseline group differences in clinical presentation may be useful in augmenting treatments for purging disorder.

Public Significance Statement: While purging disorder is classified as an "other specified" eating disorder, individuals who experience this disorder have comparable negative long-term outcomes as those with bulimia nervosa. This highlights the

1 | INTRODUCTION

Purging disorder and bulimia nervosa are clinically significant eating disorders characterized by recurrent purging behaviors (e.g., self-induced vomiting; American Psychiatric Association, 2013). These two disorders are distinguished by the presence (in bulimia nervosa) or absence (in purging disorder) of recurrent binge-eating episodes. In addition, purging disorder requires recurrent purging whereas bulimia nervosa may include non-purging compensatory behaviors, such as fasting or excessive exercise. Concurrent data support purging disorder as distinct from bulimia nervosa on indicators of comorbidity (Keel et al., 2005, 2007, 2008; Koch et al., 2013; Smith et al., 2017) and subjective and physiological response to eating (Dossat et al., 2015; Keel et al., 2007, 2018, 2021; Maske et al., 2020). The inclusion of purging disorder as an other specified feeding or eating disorder (OSFED) in DSM-5 reflected limited information regarding the clinical utility of identifying purging disorder as a full-threshold DSM-5 eating disorder that is distinct from established DSM-5 eating disorders. If purging disorder differs from bulimia nervosa on outcomes or predictors of outcomes, this would support the predictive validity of distinguishing between purging disorder and bulimia nervosa.

A meta-analysis suggested a better natural course in purging disorder relative to bulimia nervosa (Smith et al., 2017). This would support the predictive validity of distinguishing between the diagnoses on the basis that purging disorder may be less severe than bulimia nervosa. Collectively, the studies included in this meta-analysis were limited by small sample sizes (Allen et al., 2013; Keel et al., 2005; Stice et al., 2013), limited duration of follow-up (Keel et al., 2005), inclusion of only pregnant women (Knoph et al., 2013; Watson et al., 2013), and limited age range (up to age 20) (Allen et al., 2013; Stice et al., 2013). Previous work examining course of illness over intermediate durations of follow-up in adult women have relied upon samples drawn from tertiary care (Koch et al., 2013). In contrast to meta-analytic results (Smith et al., 2017), there were no differences in remission at 5-year follow-up between individuals with purging disorder and bulimia nervosa who previously received inpatient treatment. However, women with purging disorder had a higher crude mortality rate than women with bulimia nervosa (Koch et al., 2013). Taken together, the current literature presents conflicting results in samples that limit generalizability, with limited information on outcome at longer durations of follow-up.

We recently published results describing the naturalistic outcome of purging disorder at an average of 10-year follow-up (Forney et al., 2021). Despite experiencing a significant decrease in eating pathology, women with purging disorder reported a relatively poor outcome; 58% continued to meet criteria for an eating disorder, and only 30% met criteria for full recovery. More severe weight and shape concerns predicted meeting criteria for an eating disorder at follow-up, but no other baseline severity or comorbidity indicators predicted outcome. We are aware of only one other study that has examined predictors of outcome in purging disorder. In patients with prior inpatient treatment, depressive and somatization symptoms predicted outcome at moderate to long-term follow-up (Koch et al., 2013). The utility of these predictors did not differ for purging disorder and bulimia nervosa (Koch et al., 2013). However, low levels of general psychopathology predicted a better outcome in bulimia nervosa only (Koch et al., 2013). It is unknown if these results generalize to community-based samples. The current study explores differences in predictors of outcome between women with purging disorder and bulimia nervosa recruited from the community.

The current study sought to examine the predictive validity of the purging disorder diagnosis by comparing outcomes to bulimia nervosa at long-term, naturalistic follow-up. We tested three definitions of outcome: the presence of an eating disorder using DSM-5 criteria, full recovery (Bardone-Cone et al., 2010), and a continuous measure of eating pathology. We examined the longitudinal stability of baseline diagnosis, and we compared the two groups on clinical variables at outcome. Finally, we explored whether predictors of outcome differed between the two disorders. Based on the prior meta-analysis (Smith et al., 2017), we hypothesized that purging disorder would be associated with a more favorable outcome than bulimia nervosa. Data describing outcome in purging disorder were previously published (Forney et al., 2021); however, this is the first report of long-term outcomes in participants with a baseline bulimia nervosa diagnosis and the first comparison of long-term outcomes between purging disorder and bulimia nervosa.

2 | METHODS

2.1 | Sample and procedure

Participants were 217 women from three community-based studies (Keel et al., 2005, 2007, 2018), who met research criteria for purging disorder (n = 84) (Keel & Striegel-Moore, 2009) or DSM-5 criteria for bulimia nervosa (n = 133). Research criteria required all participants to purge at least once per week, on average, in the 12 weeks prior to the baseline interview and was established through the Eating

Disorder Examination (EDE) interview (Z. Cooper & Fairburn, 1987). While the recruitment criteria of the three parent studies varied, all studies utilized the same diagnostic interviews and questionnaires. None of the three parent studies provided psychological treatment. Baseline studies were conducted between 2000 and 2012. At baseline, participants were 22.81 (5.34) years old (range: 18–43). Participants identified as white ($n = 180$; 82.9%), Asian or Asian American ($n = 18$; 8.3%), Black or African American ($n = 9$; 4.1%), and Hispanic ($n = 9$; 4.1%), with race/ethnicity data missing for one participant. Education levels at baseline were 5.6% ($n = 12$) with a high school diploma or less, 62.2% ($n = 135$) with some college, and 32.3% ($n = 70$) with a post-secondary degree.

Between fall 2014 through 2017, women were contacted to participate in a follow-up study. Letters included a consent form and questionnaires. TLO Online Investigative Systems were used to locate participants and confirm vital status. Participants were contacted by letter up to three times, and women who did not respond were contacted by telephone, text, or e-mail. Participants were contacted in five waves, with those with the earliest baseline assessments contacted first, to increase mean duration of follow-up and minimize differences in duration of follow-up in the sample. Participants were initially offered $25 for the phone interview and $10 for questionnaires, which was increased to $35 and $15, respectively, in August 2015 to facilitate greater participation. A brief 20-min version of the phone interview was offered for $15. All interviews were completed by doctoral-level interviewers trained and supervised by the last author (PKK), and interviewers were blind to participants' baseline diagnoses. All participants provided informed consent, and all study procedures were approved by the local institutional review board.

2.2 | Assessments

2.2.1 | Interviews

Eating Disorder Examination, version 12.0
The EDE was administered at baseline and follow-up (Z. Cooper & Fairburn, 1987). The EDE produces a Global score and assesses the frequency of eating disorder behaviors (e.g., binge-eating, self-induced vomiting), over the previous 12 weeks. The EDE distinguishes between "objective" bulimic episodes (OBE), which correspond to DSM-5 binge-eating (characterized by a sense of loss of control and eating a large amount of food), and "subjective" bulimic episodes (characterized by loss of control while eating an average or small amount of food). Loss of control eating (LOC) was operationalized as the sum of objective and subjective binge-eating episodes. Purging behaviors included self-induced vomiting, laxative misuse, and diuretic misuse. Non-purging compensatory behaviors included avoidance of eating and excessive exercise. Participants who completed the short interview at follow-up ($n = 12$) answered diagnostic items only. The EDE discriminates between women with and without eating disorders (Berg et al., 2012). Internal consistency for the Global score was good (baseline alpha = .81; follow-up alpha = .88) and interrater reliability

for diagnosis was excellent (baseline Kappa = .91–1.00; follow-up Kappa = 1.00).

Structured Clinical Interview for DSM-IV
The SCID assessed lifetime mood, anxiety, and substance use disorders at baseline (First et al., 1995). Duration of illness at baseline was calculated as the difference between current age and age of eating disorder onset. At follow-up, the SCID was used to assess demographic information and current DSM-5 mood and substance use disorders. Interrater reliability for diagnosis was very good across disorders (Kappa = .76–1.00).

Body Mass Index
At baseline, participants' height and weight were measured via wall-mounted ruler and digital scale at baseline to calculate BMI (kg/m^2). At follow-up, participants' height and weight were measured via self-report. Self-reported height and weight have demonstrated acceptable agreement with objective measurement (Bowman & DeLucia, 1992).

2.2.2 | Self-report assessments

Eating Disorder Examination Questionnaire
The EDE-Q was administered at follow-up (Fairburn & Beglin, 1994). The EDE-Q has good test–retest reliability and distinguishes between individuals with and without eating disorders (Berg et al., 2012). The Global score (alpha = .94) was used to inform imputation models and subscale scores were used to determine recovery status (Bardone-Cone et al., 2010).

Body Shape Questionnaire
The BSQ assessed the severity of weight and shape concerns (P. J. Cooper et al., 1987). The BSQ has a 3-week test–retest reliability of .88 and is strongly correlated with similar body image assessments (Rosen et al., 1996). Internal consistency was excellent at baseline (alpha = .95) and follow-up (alpha = .98).

Three Factor Eating Questionnaire
The TFEQ includes three subscales (Cognitive Restraint, Disinhibition around Food, and Hunger) and was used to assess self-reported eating pathology (Stunkard & Messick, 1985). The Disinhibition and Hunger subscales are positively associated with binge eating and LOC eating (Forney et al., 2016; Lowe & Caputo, 1991), with the Disinhibition subscale uniquely associated with eating episode size (Forney et al., 2016). Additionally, these scales have consistently differentiated between purging disorder and bulimia nervosa (Keel et al., 2005, 2007, 2018; Roberto et al., 2010). Internal consistencies at baseline and follow-up were good (Restraint alpha = .83 and .86; Disinhibition alpha = .83 and .85; Hunger alpha = .82 and .80).

Clinical Impairment Assessment
The CIA was administered at follow-up to assess eating disorder impairment over the past 28 days (Bohn et al., 2008). The CIA had a

3-day test–retest reliability of .86 in women with eating disorders and is correlated with clinician impairment ratings (Bohn et al., 2008). Internal consistency was excellent (alpha = .95).

World Health Organization Quality of Life-BREF

The WHOQoL-BREF is a 26-item version of the WHOQoL that assesses quality of life across four domains and was administered at follow-up (The WHOQoL Group, 1998). Individuals with eating disorders demonstrate reduced scores in the social and psychological domains compared to individuals without an eating disorder (Mond et al., 2005). The domain scores have adequate to good test–retest reliability (.66–.87) (The WHOQoL Group, 1998). Internal consistency was good (alpha range = .82–.85).

Beck Depression Inventory

The BDI assessed depressive symptomatology at baseline (Beck et al., 1961). The BDI-II (Beck et al., 1996) was used in place of the original BDI at follow-up. The BDI has good test–retest reliability in psychiatric samples and is highly correlated with other measures of depressive symptoms (Beck et al., 1988). Internal consistency was good at baseline (alpha = .91) and excellent at follow-up (alpha = .95).

State Trait Anxiety Inventory

The STAI includes subscales that assess State and Trait anxiety (Speilberger et al., 1983). Higher test–retest reliability has been observed in the Trait compared to State scale (Keel et al., 2007). Internal consistency was excellent at baseline and follow-up for State (alpha = .94 and .96) and Trait (alpha = .93 and .95) subscales.

2.2.3 | Outcome definitions

Consistent with prior work (Forney et al., 2021), three outcome definitions were tested. An eating disorder was diagnosed as present if an individual met DSM-5 criteria for anorexia nervosa, bulimia nervosa, binge-eating disorder, purging disorder, or another OSFED. Anorexia nervosa, bulimia nervosa, and binge-eating disorder were diagnosed using the corresponding items on the EDE 12.0. Purging disorder was operationalized as a minimum of 12 purging episodes, the absence of recurrent objective binge eating episodes, the undue influence of shape or weight on self-evaluation, and a self-reported BMI ≥18.5 kg/m^2 on the EDE. An OSFED diagnosis was given if a participant did not meet criteria for the other diagnoses and endorsed a minimum of 12 behavioral episodes (e.g., objective or subjective binge-eating, self-induced vomiting, laxative or diuretic misuse, fasting, excessive exercise) over the prior 12 weeks. Within OSFED, subthreshold bulimia nervosa and purging disorder were differentiated by the presence or absence of regular objective binge-eating. The EDE-Q was not used in determining the presence or absence of an eating disorder. Recovery status was defined using the criteria put forth by Bardone-Cone et al. (2010). Participants were considered recovered if all four EDE-Q subscales were within one standard deviation of community norms for their age (Mond et al., 2006), the participant reported abstinence from binge-eating, purging, and fasting over the prior 12 weeks on the EDE, and the participant self-reported a BMI ≥18.5 kg/m^2. Global EDE score was used as a continuous index of eating pathology.

2.3 | Data analytic plan

2.3.1 | Attrition and missing data

List-wise analyses were conducted in IBM SPSS v 27 and multiple imputation analyses were conducted using R 4.1.2. Those who participated at follow-up were compared to those who did not participate to examine if attrition was biased. Variables used in inferential models (i.e., posited predictors of outcome, baseline EDE Global scores) were standardized. The R package "mice" version 3.14.0 (van Buuren, 2021) was used to multiply impute missing data 40 times. Due to possible group differences, follow-up data were imputed separately by baseline diagnosis. See the Supporting Information S1 for information about multiple imputation methods. The pattern of results was the same when using list-wise deletion. Exceptions are reported in the results.

2.3.2 | Hypothesis testing

Pooled logistic regression tested hypotheses regarding group differences in eating disorder status and recovery status in the package "mice" (van Buuren, 2021). Pooled multiple linear regression models tested group differences in EDE Global scores, adjusting for baseline scores (van Buuren, 2021). Because longer duration of follow-up is linked to better outcomes (Keel & Mitchell, 1997), these analyses were also run adjusting for duration of follow-up, calculated as the difference between interview dates. The pattern of results was the same. Pooled t-tests compared group differences in study variables using the R package "MKmisc" version 1.8 (Kohl, 2021) and pooled chi-square tests compared group differences in comorbidity using "miceadds" 3.11–6 (Robitzsch et al., 2021). SPSS v 27 was used to pool frequency estimates. Post-hoc equivalence testing was completed using two one-sided tests in TOSTER version 0.3.4 (Lakens, 2018) and used medium effect sizes (Cohen's h = .50 for eating disorder status and recovery status; Cohen's d = .50 for eating disorder severity). The alpha level was set at .01 for analyses comparing group differences in outcome (eating pathology, diagnostic stability, comorbidity, and impairment). All statistical tests were two-tailed, with the exception of post-hoc equivalence testing. Exploratory analyses comparing group differences in predictors of outcome used an alpha level of .05.

3 | RESULTS

3.1 | Participation and demographics at follow-up

Available records indicated all participants were living at the time of follow-up. Nearly all women (94.5%; n = 205) responded to

TABLE 1 Demographic information at long-term follow-up for women with a previous diagnosis of bulimia nervosa or purging disorder who completed follow-up interviews

Demographic variable	Baseline bulimia nervosa (n = 85)		Baseline purging disorder (n = 56)	
	n	%	n	%
Race				
White	82	89.1	50	86.2
Asian/Pacific Islander	6	6.5	4	6.9
Black/African American	1	1.1	3	5.2
More Than 1 Race	3	3.3	1	1.7
Hispanic	6	7.1	2	3.6
College Degree[a]	76	89.4	46	85.2
Married or Living with a Partner	51	60.0	30	53.6
Parent[b]	36	42.4	26	47.3
Receiving Governmental Financial Assistance[c]	5	5.9	1	1.9
Lifetime Mental Health Treatment[b]	76	89.4	41	73.2
Current Mental Health Treatment[d]	16	20.5	18	36.0

[a]Data were available for $n = 54$ women with baseline purging disorder.
[b]Data were available for $n = 55$ women with baseline purging disorder.
[c]Data were available for $n = 53$ women with baseline purging disorder.
[d]Data were available for $n = 78$ women with baseline bulimia nervosa and $n = 54$ women with baseline purging disorder.

recruitment materials. Of the 217 women sought, 69% ($n = 150$) chose to participate at a mean (SD) of 10.59 (3.71) years follow-up (range: 2.51–16.91 years), including 92 women with bulimia nervosa and 58 women with purging disorder at baseline (see Figure S1). Participation did not vary by parent study ($X^2[2] = 1.77$, $p = .41$), baseline recruitment site ($X^2[2] = 3.89$, $p = .14$), race/ethnicity (Likelihood Ratio(3) $= 2.76$, $p = .43$), baseline diagnosis ($X^2[1] = 0.0$, $p = .98$), age (OR $= .99$, $p = .73$), or baseline eating disorder severity (OR $= .89$, $p = .47$). Among participants, 134 women completed interviews and questionnaires, seven women completed only an interview, and nine completed only questionnaires.

Table 1 presents demographic information by diagnostic group among the women who completed interviews at follow-up ($n = 141$). Overall, women had a mean (SD) age of 33.40 (7.63) years (range: 21–55), and age did not differ by baseline diagnosis (t [139] $= -.04$, $p = .97$). Participants primarily identified as white and the sample was well educated. Over half were married or living with a partner and over 40% identified as parents. Race (Likelihood Ratio(3) $= 2.46$, $p = .48$), ethnicity (Likelihood Ratio(1) $= .81$, $p = .37$), partner status ($X^2[1] = .57$, $p = .45$), parental status ($X^2[1] = .33$, $p = .57$), and education level ($X^2[1] = .55$, $p = .46$) did not differ by baseline diagnosis. Women who had bulimia nervosa at baseline were three times more likely to receive governmental assistance compared to those who had purging disorder, but this difference was not statistically significant (Likelihood Ratio (1) $= 1.41$, $p = .24$). Women with bulimia nervosa at baseline tended to be more likely to report a lifetime history of mental health treatment relative to women with purging disorder ($X^2[1] = 5.38$, $p = .02$) but did not differ significantly in prevalence of current treatment at follow-up ($X^2[1] = 3.75$, $p = .05$).

3.2 | Eating disorder presence, recovery status, and eating pathology at follow-up

Over half of women met criteria for an eating disorder at follow-up (see Table 2). Presence of a current eating disorder did not differ between the groups (OR $= 1.15$, $p = .70$, 99% CI [0.45–2.92]). Post hoc equivalence testing indicated that this effect was reliably smaller than a medium effect size (Z $= 2.86$, $p = .002$, difference equivalence bounds [-0.23, .24]). Post hoc equivalence testing reached traditional thresholds of statistical significance using listwise deletion ($p = .04$).

Among those who met criteria for a DSM-5 eating disorder diagnosis, diagnostic distribution differed at follow-up between groups (Likelihood Ratio(4) $= 16.53$, $p = .002$; see Table 3). Diagnostic stability was observed for purging disorder. Women with baseline purging disorder, compared to women with bulimia nervosa, were significantly more likely to present with full or subthreshold purging disorder at follow-up relative to bulimia nervosa or subthreshold bulimia nervosa (OR $= 7.88$, $p = .004$, 99% CI [1.26–49.20]).

Recovery did not differ between the two disorders (OR $= .94$, $p = .87$, 99% CI [.33–2.68]). Post hoc equivalence testing indicated that this difference was reliably smaller than a medium effect size (Z $= -2.86$, $p = .002$, difference equivalence bounds [$-.24$–.19]). Post hoc equivalence testing reached traditional thresholds of statistical significance using listwise deletion ($p = .04$).

Adjusting for baseline EDE Global scores, global eating pathology did not differ between the two groups (estimate $= 0.30$, $p = .17$, 99% CI [-0.28–0.89]). Equivalence testing revealed that mean differences between groups were smaller than a medium effect size (t(158.29) $= 2.63$, $p = .005$, difference equivalence bounds [-0.78–0.78]).

TABLE 2 Pooled estimates of clinical outcomes at long-term follow-up for women with a previous diagnosis of bulimia nervosa or purging disorder

| | Baseline bulimia nervosa (n = 133) | | Baseline purging disorder (n = 84) | |
Clinical outcome	n	%	n	%
Eating disorder present	72.8	54.7	48.8	58.1
Recovered	38.3	28.8	23.2	27.6
Current mood disorder	36.4	27.4	22.7	27.0
Current substance use disorder	34.2	25.7	24.5	29.2

Note: Values represent pooled estimates from multiply imputed data.

TABLE 3 Eating disorder diagnosis among women with a previous diagnosis of bulimia nervosa or purging disorder who met criteria for an eating disorder at long-term follow-up

Diagnostic category	Baseline bulimia nervosa (n = 46)	Baseline purging disorder (n = 35)
DSM-5 diagnosis		
Anorexia nervosa	6.5% (n = 3)	0.0% (n = 0)
Bulimia nervosa	13.0% (n = 6)	11.4% (n = 4)
Purging disorder	2.2% (n = 1)	28.6% (n = 10)
Binge-eating disorder	2.2% (n = 1)	0.0% (n = 0)
OSFED	76.1% (n = 35)	60.0% (n = 21)
OSFED subtypes		
Subthreshold bulimia nervosa	42.9% (n = 15)	19.0% (n = 4)
Subthreshold purging disorder	8.6% (n = 3)	9.5% (n = 2)
Subthreshold binge-eating disorder	0.0% (n = 0)	4.8% (n = 1)
Recurrent subjective binge-eating episodes	8.6% (n = 3)	14.3% (n = 3)
Recurrent non-purging compensatory behaviors, with or without subjective binge-eating episodes	34.3% (n = 12)	52.4% (n = 11)
Chewing & spitting	5.7% (n = 2)	0.0% (n = 0)

Abbreviation: OSFED, other specified feeding or eating disorder.

3.3 | Clinical presentation at follow-up

Table 4 presents measures of eating pathology and comorbidity at follow-up, and Table S3 presents correlations among predictors at baseline. Groups did not differ significantly on measures of eating pathology, the presence or absence of specific eating disorder behaviors, eating disorder-related impairment, or quality of life.

Groups did not differ on continuous measures of depressive symptoms, state or trait anxiety (see Table 4). Likewise, groups did not differ in mood disorder ($F(1, 1036.3) = .23$, $p = .63$) or substance use disorder ($F(1, 838.03) = .26$, $p = .61$) prevalence at follow-up (see Table 2).

3.4 | Predictors of outcome: Exploratory analyses

Table 5 presents models testing predictors of eating disorder presence. Duration of illness ($p = .03$) differed as a predictor of outcome between the two groups. Among women with baseline purging disorder, duration of illness did not predict eating disorder status (OR = .74, $p = .27$, 95% CI [0.43–1.27]). In contrast, a longer duration of illness at baseline tended to be associated with a greater likelihood of having an eating disorder (OR = 1.67, $p = .06$, 95% CI [.97–2.85]) among those with a history of bulimia nervosa, although this did not reach the threshold for statistical significance. No other predictors differed in predictive utility between diagnoses ($p \geq .08$).

No variables predicted recovery status (p's $\geq .20$), nor did the effects of these variables differ by diagnosis (p's $\geq .28$; see Table S1). No predictors of follow-up EDE scores varied by diagnostic group (p's $\geq .09$; see Table S2).

4 | DISCUSSION

The current study examined the predictive validity of the purging disorder diagnosis by comparing the long-term outcome of purging disorder to that of bulimia nervosa. Diagnostic stability was observed, such that women with baseline purging disorder were more likely to present with symptoms consistent with purging disorder than bulimia nervosa. Equivalence testing indicated that the groups were indistinguishable when considering the prevalence of current eating disorders, recovery, and a dimensional measure of eating pathology. Exploratory analyses revealed that a longer baseline duration of illness was associated with a greater likelihood of having an eating disorder at follow-up for bulimia nervosa, but not purging disorder. No other predictors of illness differed between the two groups. We also did not observe differences in comorbidity at follow-up. Taken together, results refute the idea that purging disorder represents a less severe form of eating pathology than bulimia nervosa. Instead, the courses of both diagnoses are characterized by chronicity, with over half remaining ill at long-term follow-up. Although the courses of purging disorder and bulimia nervosa are characterized by symptom stability (i.e., diagnostic maintenance is more likely than cross-over to one another), the two groups exhibited comparable global outcomes at long-term follow-up. Thus, distinguishing between diagnoses of purging disorder and bulimia nervosa does not provide clinically useful information on

TABLE 4 Comparisons of eating pathology and comorbidity at 10-year follow-up among women previously diagnosed with bulimia nervosa or purging disorder

Assessment	Baseline bulimia nervosa (n = 133)		Baseline purging disorder (n = 84)		t/X^2	df	p	99% CI	Cohen's d/φ
	M/n	SD/%	M/n	SD/%					
EDE Global Score	1.92	1.45	2.12	1.67	−.85	68.10	.40	−.82 to 0.42	−.13
Presence of Purging Behaviors	46.7	35.1	38.5	45.8	1.40	1435.19	.24	-	.08
Presence of Objective Binge Episodes	54.2	40.8	20.7	24.6	3.82	1570.38	.05	-	.13
Presence of Loss of Control Eating	87.5	65.8	48.1	57.3	.93	1417.58	.34	-	.07
Body Shape Questionnaire	95.58	47.30	107.85	59.85	−1.56	72.08	.13	−33.02 to 8.50	−.23
TFEQ Restraint	11.18	6.08	13.21	6.53	−2.24	79.33	.03	−4.42 to 0.37	−.32
TFEQ Disinhibition	8.58	5.13	7.14	5.03	2.00	78.27	.05	−0.46 to 3.35	.28
TFEQ Hunger	6.17	4.29	5.22	4.48	1.50	80.15	.14	−0.72 to 2.61	.22
Clinical Impairment Assessment	13.08	12.92	13.53	15.69	−.21	70.05	.83	−6.00 to 5.10	−.03
WHOQOL Physical	76.04	20.80	75.10	18.08	.35	116.94	.73	−6.06 to 7.93	.05
WHOQOL Psychological	59.29	23.30	62.27	19.57	−1.02	121.63	.31	−10.61 to 4.67	−.14
WHOQOL Social	59.35	29.10	67.10	30.69	−1.90	99.86	.06	−18.46 to 2.96	−.26
WHOQOL Environment	76.02	21.42	76.37	20.31	−.12	121.73	.90	−7.65 to 6.95	−.02
Beck Depression Inventory	14.40	15.52	12.18	11.69	1.20	107.47	.23	−2.65 to 7.10	.16
State Anxiety	41.81	18.70	38.94	15.99	1.20	99.87	.23	−3.38 to 9.12	.16
Trait Anxiety	44.35	16.87	43.71	14.15	.31	126.57	.76	−4.75 to 6.03	.04

Abbreviations: EDE, Eating Disorder Examination; LOC, Loss of Control Eating Episode; OBE, Objective Binge Episode; TFEQ, Three Factor Eating Questionnaire; WHOQOL, World Health Organization Quality of Life BREF.

TABLE 5 Logistic regression models testing predictors of DSM-5 eating disorder presence at 10-year follow-up in women with a history of bulimia nervosa or purging disorder

Predictor variable	Intercept OR [95% CI]	Diagnosis OR [95% CI]	Predictor OR [95% CI]	Diagnosis × predictor OR [95% CI]
Severity indicators				
Illness Duration	1.03 [.36–2.95]	1.13 [0.55–2.32]	**3.74 [1.13–12.43]**	**.45 [.21–.95]**
Purging Frequency	.88 [.31–2.49]	1.32 [.64–2.73]	1.31 [.43–3.95]	1.01 [.43–2.38]
Body Shape Questionnaire	.86 [.30–2.43]	1.41 [.68–2.94]	.48 [.17–1.34]	2.06 [.93–4.59]
Loss of Control Eating Frequency	.92 [.29–2.89]	1.22 [.52–2.85]	1.77 [.52–5.99]	0.73 [.27–2.00]
Comorbidity measures				
Beck Depression Inventory	.90 [.32–2.54]	1.34 [.64–2.78]	.76 [.27–2.12]	1.40 [.64–3.07]
State Anxiety	.99 [.36–2.75]	1.20 [.60–2.43]	1.19 [.43–3.28]	.99 [.48–2.03]
Lifetime Mood Disorder	.86 [.14–5.38]	1.13 [.32–3.95]	1.23 [.13–11.74]	1.10 [.24–5.06]
Lifetime Anxiety Disorder	.87 [.23–3.33]	1.28 [.51–3.24]	1.67 [.20–13.85]	.75 [.19–2.99]
Lifetime Substance Use Disorder	.88 [.20–3.93]	1.26 [.42–3.82]	1.48 [.19–11.49]	.82 [.20–3.39]

Note: Each row represents a separate logistic regression model. Bold font indicates statistically significant parameter with a threshold of .05.
Abbreviation: OR, odds ratio.

the likelihood of having an eating disorder, the likelihood of recovering, or eating disorder severity at long-term follow-up. However, this distinction does predict whether purging presents in the presence or absence of binge eating at follow-up.

The largely similar outcomes from purging disorder and bulimia nervosa replicate prior findings from a tertiary care sample that also found no differences in remission between diagnostic groups over a

5-year follow-up (Koch et al., 2013). In contrast, a recent meta-analysis found a more favorable natural course in purging disorder relative to bulimia nervosa (Smith et al., 2017). This discrepancy may reflect the longer duration of follow-up and older age in both the current sample and prior tertiary care sample (Koch et al., 2013). Indeed, differences apparent earlier in follow-up of anorexia nervosa and bulimia nervosa (Herzog et al., 1999) also dissipated over longer durations of follow-up

(Eddy et al., 2017). The meta-analysis relied primarily on adolescent samples (Allen et al., 2013; Stice et al., 2013), which tend to have a shorter course of illness (Steinhausen, 2009) and greater diagnostic instability (Steinhausen et al., 2005) relative to adults.

Despite comparable levels of eating pathology and comorbidity, purging disorder and bulimia nervosa differed in clinical presentation at follow-up in ways that reflected clinical differences at baseline. This is similar to findings in an adolescent sample where the prevalence of diagnostic stability in both full threshold and subthreshold diagnoses was similar to or greater (23% of bulimia nervosa cases, 31% of purging disorder cases) than the likelihood of cross-over between bulimia nervosa and purging disorder over shorter durations of follow-up (17% of bulimia nervosa cases; 9% of purging disorder cases; Glazer et al., 2019). Although data from anorexia nervosa suggest that binge-eating tends to develop over time (Eddy et al., 2008), we observed the development of binge eating in only a minority of the purging disorder sample. This may reflect elevated postprandial peptide YY (PYY) response in purging disorder (Keel et al., 2018). Elevated PYY may contribute to excessive satiety and reduce risk of objective binge-eating. Consistent with this, the restricting subtype of anorexia nervosa is associated with higher fasting PYY than the binge-purge subtype (Eddy et al., 2015) and the restricting subtype has an elevated postprandial PYY response compared to healthy controls (Becker et al., 2021; Heruc et al., 2019).

Among OSFED cases at follow-up, women with baseline bulimia nervosa were most likely to present with subthreshold bulimia nervosa at follow-up (43%) whereas women with baseline purging disorder most frequently presented with recurrent non-purging compensatory behaviors (52%) rather than subthreshold purging disorder (10%). This may reflect differences in the conceptualization of purging disorder and bulimia nervosa. If a woman with bulimia nervosa switched from a purging to a non-purging compensatory behavior, her diagnosis would stay unchanged whereas a woman with purging disorder who made the same behavioral shift would be categorized as engaging in recurrent non-purging behaviors. These definitional distinctions make longitudinal comparisons between bulimia nervosa and purging disorder more complex.

A small proportion of women with bulimia nervosa, but none with baseline purging disorder, met criteria for anorexia nervosa and/or reported chewing and spitting at follow-up. Notably, the groups reported equivalent histories of anorexia nervosa at baseline (X^2[1] = .02, p = .88; 14.3% of baseline purging disorder; 15.0% of baseline bulimia nervosa). This pattern of findings is somewhat inconsistent with findings of no differences between purging disorder and bulimia nervosa in cross-over to anorexia nervosa (Koch et al., 2013). Given the small case count of both anorexia nervosa and chewing and spitting, it is too early to draw conclusions about group differences in this regard. In exploratory analyses, duration of illness predicted having an eating disorder differently between groups. Purging disorder tends to have both a longer range and later period of onset than bulimia nervosa (Glazer et al., 2019), and consistent with this, the bulimia nervosa group had a longer duration of illness than purging disorder at baseline (t[186.96] = 2.61, p = .01, Cohen's d = .36). Given differences in age of onset and the predictive utility of illness duration, future research should strive to better study the

development and maintenance of purging disorder and bulimia nervosa across the lifespan.

The study benefited from a large duration of follow-up, reasonably high retention, blinded interviewers, and use of psychometrically sound assessments. Our use of a community-based sample ensures results are not biased by treatment-seeking. Our use of multiple imputation further improves the generalizability of our results. However, this study is not without limitations, including variability in baseline parent study criteria, lack of objective measures of weight at follow-up, variability in duration of follow-up, and a relatively small sample size. Although women of different racial and ethnic backgrounds have comparable rates of bulimia nervosa as white women (Udo & Grilo, 2018) and men also experience bulimia nervosa and purging disorder (Masheb et al., 2021), our sample was restricted to women and differed in racial/ethnic and educational background from the US population. Study findings may or may not generalize fully to individuals from different demographic groups. Future work should strive to increase inclusivity and representation of these groups.

Although differences in baseline clinical presentation were not meaningful in predicting long-term illness outcome, these differences may be meaningful for treatment planning. For example, baseline group differences in fullness sensitivity (Keel et al., 2007) suggest that interoceptive exposures may be beneficial in purging disorder (Boswell et al., 2019), but may not be clinically relevant in bulimia nervosa. Transdiagnostic samples have often included purging disorder in "bulimia nervosa-spectrum illness," but more work is needed to understand how treatment may need to be tailored differently based on symptom presentation, particularly given the relative stability of diagnostic differences. Future research should also consider how purging disorder may differ from other types of OSFED, such as atypical anorexia nervosa. Baseline data were collected before the inclusion of atypical anorexia nervosa in the DSM-5 (American Psychiatric Association, 2013), and we are unable to examine this differential diagnosis. While atypical anorexia nervosa without purging is distinguishable from purging disorder (Krug et al., 2022), approximately half of individuals with purging disorder are also weight suppressed and meet both sets of criteria simultaneously (Forney et al., 2017). As similar patterns also exist in bulimia nervosa (Forney et al., 2017), and weight suppression predicts the onset of BN and PD (Stice et al., 2020), the consideration of weight history may be a more potent predictor of outcome. Future work is needed to understand how weight history predicts purging disorder treatment response, course, or outcome and to determine whether there is a meaningful boundary between purging disorder and atypical anorexia nervosa with purging. Continued refinement of nosological schemes will allow the field to provide patients with better prognostic information and tailored treatments.

CONFLICT OF INTEREST

AUTHOR CONTRIBUTIONS

DATA AVAILABILITY STATEMENT

Data are available upon request from the corresponding author.

ORCID

Katherine Jean Forney https://orcid.org/0000-0002-8215-9335
Tiffany A. Brown https://orcid.org/0000-0002-7349-7228
Pamela K. Keel https://orcid.org/0000-0001-6542-5147

REFERENCES

Allen, K. L., Byrne, S. M., Oddy, W. H., & Crosby, R. D. (2013). Early onset binge eating and purging eating disorders: Course and outcome in a population-based study of adolescents. *Journal of Abnormal Child Psychology*, 41(7), 1083–1096. https://doi.org/10.1007/s10802-013-9747-7

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Bardone-Cone, A. M., Harney, M. B., Maldonado, C. R., Lawson, M. A., Robinson, D. P., Smith, R., & Tosh, A. (2010). Defining recovery from an eating disorder: Conceptualization, validation, and examination of psychosocial functioning and psychiatric comorbidity. *Behaviour Research and Therapy*, 48(3), 194–202. https://doi.org/10.1016/j.brat.2009.11.001

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation. https://blog.naver.com/mistyeyed73/220427762670

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100. https://doi.org/10.1016/0272-7358(88)90050-5

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4(6), 561–571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Becker, K. R., Mancuso, C., Dreier, M. J., Asanza, E., Breithaupt, L., Slattery, M., Plessow, F., Micali, N., Thomas, J. J., Eddy, K. T., Misra, M., & Lawson, E. A. (2021). Ghrelin and PYY in low-weight females with avoidant/restrictive food intake disorder compared to anorexia nervosa and healthy controls. *Psychoneuroendocrinology*, 129, 105243. https://doi.org/10.1016/j.psyneuen.2021.105243

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. *International Journal of Eating Disorders*, 45(3), 428–438. https://doi.org/10.1002/eat.20931

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. *Behaviour Research and Therapy*, 46(10), 1105–1110. https://doi.org/10.1016/j.brat.2008.06.012

Boswell, J. F., Anderson, L. M., Oswald, J. M., Reilly, E. E., Gorrell, S., & Anderson, D. A. (2019). A preliminary naturalistic clinical case series study of the feasibility and impact of interoceptive exposure for eating disorders. *Behaviour Research and Therapy*, 117, 54–64. https://doi.org/10.1016/j.brat.2019.02.004

Bowman, R. L., & DeLucia, J. L. (1992). Accuracy of self-reported weight: A meta-analysis. *Behavior Therapy*, 23(4), 637–655. https://doi.org/10.1016/S0005-7894(05)80226-6

Cooper, P. J., Taylor, M. J., Cooper, Z., & Fairburn, C. G. (1987). The development and validation of the Body Shape Questionnaire. *International Journal of Eating Disorders*, 6(4), 485–494. https://doi.org/10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O

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. https://doi.org/10.1002/1098-108X(198701)6:1%3C1::AID-EAT2260060102%3E3.0.CO;2-9

Dossat, A. M., Bodell, L. P., Williams, D. L., Eckel, L. A., & Keel, P. K. (2015). Preliminary examination of glucagon-like peptide-1 levels in women with purging disorder and bulimia nervosa. *International Journal of Eating Disorders*, 48(2), 199–205. https://doi.org/10.1002/eat.22264

Eddy, K. T., Dorer, D. J., Franko, D. L., Tahilani, K., Thompson-Brenner, H., & Herzog, D. B. (2008). Diagnostic crossover in anorexia nervosa and bulimia nervosa: Implications for DSM-V. *American Journal of Psychiatry*, 165(2), 245–250. https://doi.org/10.1176/appi.ajp.2007.07060951

Eddy, K. T., Lawson, E. A., Meade, C., Meenaghan, E., Horton, S. E., Misra, M., Klibanski, A., & Miller, K. K. (2015). Appetite regulatory hormones in women with anorexia nervosa: Binge-eating/purging versus restricting type. *The Journal of Clinical Psychiatry*, 76(1), 19–24. https://doi.org/10.4088/JCP.13m08753

Eddy, K. T., Tabri, N., Thomas, J. J., Murray, H. B., Keshaviah, A., Hastings, E., Edkins, K., Krishna, M., Herzog, D. B., Keel, P. K., & Franko, D. L. (2017). Recovery from anorexia nervosa and bulimia nervosa at 22-year follow-up. *The Journal of Clinical Psychiatry*, 78(2), 184–189. https://doi.org/10.4088/JCP.15m10393

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363–370.

First, M., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). *Structured Clinical Interview for DSM-IV Axis I Disorders*. New York State Psychiatric Institute.

Forney, K. J., Bodell, L. P., Haedt-Matt, A. A., & Keel, P. K. (2016). Incremental validity of the episode size criterion in binge-eating definitions: An examination in women with purging syndromes. *International Journal of Eating Disorders*, 49(7), 651–662. https://doi.org/10.1002/eat.22508

Forney, K. J., Brown, T. A., Holland-Carter, L. A., Kennedy, G. A., & Keel, P. K. (2017). Defining "significant weight loss" in atypical anorexia nervosa. *International Journal of Eating Disorders*, 50(8), 952–962. https://doi.org/10.1002/eat.22717

Forney, K. J., Crosby, R. D., Brown, T. A., Klein, K. M., & Keel, P. K. (2021). A naturalistic, long-term follow-up of purging disorder. *Psychological Medicine*, 51(6), 1020–1027. https://doi.org/10.1017/S0033291719003982

Glazer, K. B., Sonneville, K. R., Micali, N., Swanson, S. A., Crosby, R., Horton, N. J., Eddy, K. T., & Field, A. E. (2019). The course of eating disorders involving bingeing and purging among adolescent girls: Prevalence, stability, and transitions. *Journal of Adolescent Health*, 64(2), 165–171. https://doi.org/10.1016/j.jadohealth.2018.09.023

Heruc, G. A., Little, T. J., Kohn, M., Madden, S., Clarke, S., Horowitz, M., & Feinle-Bisset, C. (2019). Appetite perceptions, gastrointestinal symptoms, ghrelin, peptide YY and state anxiety are disturbed in adolescent females with anorexia nervosa and only partially restored with short-term refeeding. *Nutrients*, 11(1), 59. https://doi.org/10.3390/nu11010059

Herzog, D. B., Dorer, D. J., Keel, P. K., Selwyn, S. E., Ekeblad, E. R., Flores, A. T., Greenwood, D. N., Burwell, R. A., & Keller, M. B. (1999). Recovery and relapse in anorexia and bulimia nervosa: A 7.5-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38(7), 829–837. https://doi.org/10.1097/00004583-199907000-00012

Keel, P. K., Eckel, L. A., Hildebrandt, B. A., Haedt-Matt, A. A., Appelbaum, J., & Jimerson, D. C. (2018). Disturbance of gut satiety

peptide in purging disorder. *International Journal of Eating Disorders*, 51(1), 53–61. https://doi.org/10.1002/eat.22806

Keel, P. K., Eckel, L. A., Hildebrandt, B. A., Haedt-Matt, A. A., Murry, D. J., Appelbaum, J., & Jimerson, D. C. (2021). Disentangling the links between gastric emptying and binge eating v. Purging in eating disorders using a case-control design. *Psychological Medicine*, 1–8. https://doi.org/10.1017/S0033291721003640

Keel, P. K., Haedt, A., & Edler, C. (2005). Purging disorder: An ominous variant of bulimia nervosa? *International Journal of Eating Disorders*, 38(3), 191–199. https://doi.org/10.1002/eat.20179

Keel, P. K., & Mitchell, J. E. (1997). Outcome in bulimia nervosa. *American Journal of Psychiatry*, 154(3), 313–321. https://doi.org/10.1176/ajp.154.3.313

Keel, P. K., & Striegel-Moore, R. H. (2009). The validity and clinical utility of purging disorder. *International Journal of Eating Disorders*, 42(8), 706–719. https://doi.org/10.1002/eat.20718

Keel, P. K., Wolfe, B. E., Gravener, J. A., & Jimerson, D. C. (2008). Co-morbidity and disorder-related distress and impairment in purging disorder. *Psychological Medicine*, 38(10), 1435–1442. https://doi.org/10.1017/S0033291707001390

Keel, P. K., Wolfe, B. E., Liddle, R. A., Young, K. P. D., & Jimerson, D. C. (2007). Clinical features and physiological response to a test meal in purging disorder and bulimia nervosa. *Archives of General Psychiatry*, 64(9), 1058–1066. https://doi.org/10.1001/archpsyc.64.9.1058

Knoph, C., Von Holle, A., Zerwas, S., Torgersen, L., Tambs, K., Stoltenberg, C., Bulik, C. M., & Reichborn-Kjennerut, T. (2013). Course and predictors of maternal eating disorders in the postpartum period. *International Journal of Eating Disorders*, 46(4), 355–368. https://doi.org/10.1002/eat.22088

Koch, S., Quadflieg, N., & Fichter, M. (2013). Purging disorder: A comparison to established eating disorders with purging behaviour. *European Eating Disorders Review*, 21(4), 265–275. https://doi.org/10.1002/erv.2231

Kohl, M. (2021). *Package "MKmisc"* (1.8) [Computer software]. https://cran.r-project.org/web/packages/MKmisc/MKmisc.pdf

Krug, I., Giles, S. E., Granero, R., Agüera, Z., Sánchez, I., Sánchez-Gonzalez, J., Jimenez-Murcia, S., & Fernandez-Aranda, F. (2022). Where does purging disorder lie on the symptomatologic and personality continuum when compared to other eating disorder subtypes? Implications for the DSM. *European Eating Disorders Review*, 30(1), 36–49. https://doi.org/10.1002/erv.2872

Lakens, D. (2018). *Package "TOSTER"* (0.3.4) [Computer software]. https://cran.r-project.org/web/packages/TOSTER/TOSTER.pdf

Lowe, M. R., & Caputo, G. C. (1991). Binge eating in obesity: Toward the specification of predictors. *International Journal of Eating Disorders*, 10(1), 49–55. https://doi.org/10.1002/1098-108X(199101)10:1<49::AID-EAT2260100106>3.0.CO;2-X

Masheb, R. M., Ramsey, C. M., Marsh, A. G., Decker, S. E., Maguen, S., Brandt, C. A., & Haskell, S. G. (2021). DSM-5 eating disorder prevalence, gender differences, and mental health associations in United States military veterans. *International Journal of Eating Disorders*, 54(7), 1171–1180. https://doi.org/10.1002/eat.23501

Maske, C. B., Williams, D. L., & Keel, P. K. (2020). Preliminary examination of insulin and amylin levels in women with purging disorder. *International Journal of Eating Disorders*, 53(6), 997–1001. https://doi.org/10.1002/eat.23230

Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2006). Eating Disorder Examination Questionnaire (EDE-Q): Norms for young adult women. *Behaviour Research and Therapy*, 44(1), 53–62. https://doi.org/10.1016/j.brat.2004.12.003

Mond, J. M., Owen, C., Hay, P. J., Rodgers, B., & Beumont, P. J. V. (2005). Assessing quality of life in eating disorder patients. *Quality of Life Research*, 14(1), 171–178. https://doi.org/10.1007/s11136-004-2657-y

Roberto, C. A., Grilo, C. M., Masheb, R. M., & White, M. A. (2010). Binge eating, purging, or both: Eating disorder psychopathology findings from an internet community survey. *International Journal of Eating Disorders*, 43(8), 724–731. https://doi.org/10.1002/eat.20770

Robitzsch, A., Grund, S., & Henke, T. (2021). *Package "miceadds"* (3.11-6) [Computer software]. https://cran.r-project.org/web/packages/miceadds/miceadds.pdf

Rosen, J. C., Jones, A., Ramirez, E., & Waxman, S. (1996). Body Shape Questionnaire: Studies of validity and reliability. *International Journal of Eating Disorders*, 20(3), 315–319. https://doi.org/10.1002/(SICI)1098-108X(199611)20:3<315::AID-EAT11>3.0.CO;2-Z

Smith, K. E., Crowther, J. H., & Lavender, J. M. (2017). A review of purging disorder through meta-analysis. *Journal of Abnormal Psychology*, 126(5), 565–592. https://doi.org/10.1037/abn0000243

Speilberger, C. D., Gorusch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *The State-Trait Anxiety Inventory*. Consulting Psychologists Press.

Steinhausen, H.-C. (2009). Outcome of eating disorders. *Child and Adolescent Psychiatric Clinics of North America*, 18(1), 225–242.

Steinhausen, H.-C., Gavez, S., & Metzke, C. W. (2005). Psychosocial correlates, outcome, and stability of abnormal adolescent eating behavior in community samples of young people. *International Journal of Eating Disorders*, 37(2), 119–126. https://doi.org/10.1002/eat.20077

Stice, E., Marti, C. N., & Rhode, 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. *Journal of Abnormal Psychology*, 122(2), 445–457. https://doi.org/10.1037/a0030679

Stice, E., Rohde, P., Shaw, H., & Desjardins, C. (2020). Weight suppression increases odds for future onset of anorexia nervosa, bulimia nervosa, and purging disorder, but not binge eating disorder. *The American Journal of Clinical Nutrition*, 112(4), 941–947. https://doi.org/10.1093/ajcn/nqaa146

Stunkard, A. J., & Messick, S. (1985). The Three-Factor Eating Questionnaire to measure dietary restraint, disinhibition and hunger. *Journal of Psychosomatic Research*, 29(1), 71–83. https://doi.org/10.1016/0022-3999(85)90010-8

The WHOQoL Group, T. W. (1998). Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment. *Psychological Medicine*, 28(3), 551–558.

Udo, T., & Grilo, C. M. (2018). Prevalence and correlates of DSM-5–defined eating disorders in a nationally representative sample of U.S. adults. *Biological Psychiatry*, 84(5), 345–354. https://doi.org/10.1016/j.biopsych.2018.03.014

van Buuren, S. (2021). *Package "mice"* (3.14.0) [Computer software]. https://cran.r-project.org/web/packages/mice/mice.pdf

Watson, H. J., Holle, A. V., Hamer, R. M., Berg, C. K., Torgersen, L., Magnus, P., Stoltenberg, C., Sullivan, P., Reichborn-Kjennerud, T., & Bulik, C. M. (2013). Remission, continuation and incidence of eating disorders during early pregnancy: A validation study in a population-based birth cohort. *Psychological Medicine*, 43(8), 1723–1734. https://doi.org/10.1017/S0033291712002516

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.