

Weight suppression as a predictor variable in the treatment of eating disorders: a systematic review

Article

Accepted Version

Jenkins, P. E. ORCID: <https://orcid.org/0000-0003-1673-2903>, Lebow, J. and Rienecke, R. D. (2018) Weight suppression as a predictor variable in the treatment of eating disorders: a systematic review. *Journal of Psychiatric and Mental Health Nursing*, 25 (5-6). pp. 297-306. ISSN 1351-0126 doi: 10.1111/jpm.12462 Available at <https://centaur.reading.ac.uk/76707/>

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/10.1111/jpm.12462>

Publisher: Wiley

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

Weight Suppression as a Predictor Variable in the Treatment of Eating Disorders: A Systematic Review

Paul E. Jenkins^a, D.Clin.Psychol., Jocelyn Lebow^b, Ph.D., & Renee D. Rienecke^{c,d}, Ph.D.

^aSchool of Psychology and Clinical Language Sciences, University of Reading, UK. RG6 6AL.

^bDepartment of Psychiatry and Psychology, Mayo Clinic College of Medicine, Rochester, MN

^cDepartment of Pediatrics, Medical University of South Carolina, 261 Calhoun St., Ste. 220, Charleston, SC, 29425, USA.

^dDepartment of Psychiatry and Behavioral Sciences, Medical University of South Carolina, 261 Calhoun St., Ste. 220, Charleston, SC, 29425, USA.

Correspondence should be addressed to: Paul E. Jenkins, D.Clin.Psychol.; School of Psychology and Clinical Language Sciences, University of Reading, UK. RG6 6AL. Tel: +44 (0)118 378 7549, E-mail: pej106@gmail.com

Acknowledgements

We would like to thank the investigators who responded to our requests for further information and Prof. Kelly Klump for comments on an earlier draft. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Weight Suppression as a Predictor Variable in the Treatment of Eating Disorders: A Systematic Review

Running head: *WEIGHT SUPPRESSION IN EATING DISORDERS*

Abstract

Introduction: Weight suppression (WS – the difference between highest body weight and current body weight) has been proposed as a predictor of treatment outcome within eating disorders (EDs), although this hypothesis has not been consistently supported.

Aim/Question: Review the association between pre-treatment WS and outcome following psychological treatment for EDs. **Method:** A comprehensive electronic database search for published and unpublished literature from 1979 to 2017. Reference lists were also inspected.

Eligibility criteria were determined according to relevant guidelines and a quality appraisal was conducted. **Results:** Twelve studies met inclusion criteria (one was subsequently excluded based on insufficient data). Greater WS was generally associated with weight gain at post-treatment although not with other treatment outcomes. **Discussion:** The existing evidence, with data from 1566 participants, is summarised according to three main post-treatment outcomes: weight change; treatment completion; and symptom abstinence. Patients with disordered eating and greater WS may need to gain more weight than others during treatment to achieve good outcomes. Recommendations for future studies are provided.

Implications for Practice: Evidence-based treatments for EDs may benefit from considering WS when planning treatment, such as further psychoeducation on weight changes. Societal interventions regarding promotion of healthy eating may also draw on these findings.

Keywords: eating disorders, weight suppression, treatment outcome, CBT

Relevance Statement

We believe that this review paper can help inform mental health professionals who work directly with eating disorders, particularly those providing psychological interventions, such as CBT. The results of the study help to clarify the role of weight suppression as a predictor of outcome in treatment, and support calls for the provision of additional support (e.g., psychoeducation) for patients high in weight suppression. Discouragement of strict dieting as a public health intervention might also be considered.

Accessible Summary:

What is known on the subject?

- Weight suppression (WS) has been suggested to predict outcome following psychological treatment for an eating disorder (ED)
- Some findings are contradictory and have not been considered systematically

What the paper adds to existing knowledge:

- The review suggests that weight gain at post-treatment is reliably predicted by pre-treatment WS, but findings regarding other outcomes (e.g., treatment drop-out) are less consistent
- Approximate effect sizes for observed relationships are identified, alongside support for biobehavioural theories of metabolic adaptation to weight loss

What are the implications for practice?

- Degree of WS at pre-treatment is associated with weight gain and is important for clinicians to consider before offering treatment to patients with EDs
- Patients high in WS might benefit from further support (e.g., psychoeducation) prior to beginning treatment

Introduction

Whilst significant progress has been made in refining and evaluating psychological treatments for eating disorders (EDs), less is known about factors that are associated with better outcome (Cooper et al., 2016). Weight suppression (WS) is a variable that may represent one such predictor (Butryn, Lowe, Safer, & Agras, 2006). Relative WS, defined as one's highest ever body weight minus current body weight, is a metric that is easy to calculate and may serve as a useful prognostic factor with importance in treatment decision-making and recommendations. However, its utility in this regard has been called into question given failure to confirm this finding in a number of samples (e.g., Dawkins, Watson, Egan, & Kane, 2013).

A review by Keel (2010) suggested that prognostic factors in anorexia nervosa (AN) might include severity and duration of illness, whereas psychiatric comorbidity and general psychiatric severity may be more important in bulimia nervosa (BN). Few reliable predictors of outcome have been found for binge-eating disorder (BED), but may include baseline psychopathology (such as frequency of binge eating and body dissatisfaction; e.g., Lammers, Vroling, Ouwens, Engels, & van Strien, 2015). However, WS is a feature that might represent an important prognostic indicator that is applicable transdiagnostically. Higher WS has been associated with maintenance of bulimic symptoms over time (e.g., Keel & Heatherton, 2010) and may also affect levels of serum hormones, such as leptin, that are associated with food intake and satiation (e.g., Keel, Bodell, Haedt-Matt, Williams, & Appelbaum, 2017).

It has been theorised that individuals with EDs who report long-term energy imbalance such that their weight is below what would be expected given their weight history (i.e., those high in WS) are caught in a biobehavioural, or psychobiological, “bind” (e.g., Butryn et al., 2006); suppressing weight may reduce some of the distress associated with an

ED (e.g., improving body image) but concomitantly drive binge eating and compensatory behaviours, including further dieting. Systemic adaptations that result from maintaining a reduced body weight (e.g., alterations in neuronal activity, hormone changes) have been proposed as an underlying mechanism involved in the difficulty sustaining low weight, whereby a weight-reduced state favours energy conservation and increased dietary intake (Rosenbaum, Kissileff, Mayer, Hirsch, & Leibel, 2010). In partial support of this, Stice, Durant, Burger, and Schoeller (2011) found that greater WS predicted future increases in BMI, although reduced resting metabolic rate did not seem to determine the strength of the relationship between WS and BMI.

Rationale for the review

While there has been some evidence in support of the biobehavioural bind theory in EDs (e.g., see Keel & Heatherton, 2010), the usefulness of WS specifically as a predictor of treatment outcome has been questioned, given failure to replicate many findings related to treatment outcome (e.g., Carter, McIntosh, Joyce, & Bulik, 2008; Dawkins et al., 2013; Zunker et al., 2011). Given observed inconsistencies, confidence in the predictive role of WS cannot be held without question and there is therefore a gap in our knowledge regarding the relationship between WS and treatment outcome.

Aims

There have been no systematic reviews of the literature on the predictive ability of WS in ED treatment outcome and, therefore, it is important to assess commonalities and differences across studies in a systematic way that aims to cover the breadth of current literature. A review would provide a summary of the existing state of the evidence regarding WS, enabling reliable and accurate conclusions to be drawn, thereby informing refinement and delivery of psychological treatments.

Objective

To conduct a systematic review that looks at the extent to which WS predicts outcome following treatment for EDs. As such, the review will cover treatment studies, but will exclude studies looking at, for example, maintenance of the illness without explicit delivery of treatment.

Review question

Review question: Does pre-treatment (baseline) WS predict outcome following psychological treatment for EDs?

Method

A systematic review was conducted, guided by the recommendations of PRISMA (Moher et al., 2009) and with reference to PICOS criteria for defining a research question. The review protocol has been registered in the PROSPERO International Prospective Register of Systematic Reviews (*Removed for blind review*).

Meta-analysis was considered but not completed for a number of reasons. Sample sizes were variable (range = 35 – 350) and reported measures of effect size varied across studies, making it impossible to derive a consistent metric without markedly sacrificing power. Typically, either odds ratios or regression equations were reported (sometimes, both). Although it is possible to obtain a measure of effect size from dichotomised outcomes (e.g., Sánchez-Meca, Marín-Martínez, & Chacón-Moscó, 2003), it was unclear from some studies whether standardised or unstandardized regression coefficients were reported and repeated imputation would have risked excessive ‘second-order sampling error’ (e.g., Hunter & Schmidt, 1990). Included articles were therefore analysed thematically, generating suggestions for reporting further studies in this area.

Study identification

Regarding ‘*Study Design*’, studies of relevant groups (i.e., excluding individual case studies) were considered for inclusion. Attempts were made to include studies not in the

‘published’ domain (e.g., grey databases, studies in progress). We conducted a literature search using electronic databases (Scopus, OVID, PsycINFO, PubMed [including MEDLINE], Google Scholar) without language restriction covering the period January 1979–November 2017. Grey literature searches included ClinicalTrials.gov and the ISRCTN Register. Vocabulary terms (MeSH) were used with the keywords of ‘weight suppression’ and ‘eating disorders’, using wildcards where necessary. To increase range, additional searches substituted ‘eating disorders’ with ‘anorexia nervosa’, ‘bulimia nervosa’, and ‘binge’ (again, using wildcards). Reference lists of included articles and others relevant to the topic were also manually searched. Based on this methodology, unmasked reports were used. Shortly after the search was completed, we became aware of one further article (Hessler et al., 2017) that would have met inclusion criteria. This was therefore included in the review.

Data collation and extraction

Two investigators (*Identifiers removed for blind review*) independently screened titles and abstracts and selected full texts for inclusion in the review, using a standardised template based on the Data Collection Form of Cochrane Reviews. Any disagreements were resolved by consensus, or through involvement of a third reviewer (*Removed*) and reasons for exclusion of full-texts were recorded. Where more information was required, an attempt was made to contact corresponding authors. Risk of bias was assessed to appraise the strength of the evidence reviewed (CRD, 2008). In the current study, the Newcastle-Ottawa Scale (NOS; Wells et al., 1999) formed a basis for assessment of the likelihood of bias; a star system is used such that studies can be graded on criteria relating to study design, including selection of cohorts (maximum of four stars), comparability of cohorts (maximum of one star), and assessment of outcome (maximum of three stars). The NOS is reviewer-rated and has been recommended by the Cochrane Non-Randomised Studies Methods Group (Reeves et al.,

2011) for use with non-randomised studies. The NOS was modified in our study to more accurately fit the review question and was ‘scored’ from zero to eight, with a score of eight representing the least amount of bias. As all studies reported WS at baseline, Item 4 (Demonstration that outcome of interest was present at the start of the study) was replaced by an item assessing reference to statistical power (Adequate reporting of statistical power). Item 8 (Adequacy of follow-up) was not included as this review focused on end-of-treatment data only. Due to critiques of the use of summary scores to determine quality / risk of bias (e.g., Stang, 2010), a combination of qualitative and quantitative assessment is reported.

Eligibility criteria

Eligible studies included any treatment study (*‘Intervention’*) published in a peer-reviewed journal and written in English that looked at pre-treatment (baseline) WS as a predictor of outcome following treatment for EDs, diagnosed according to formal criteria (*‘Outcome’*). Potential outcomes included measures of ED psychopathology, behavioural symptoms, treatment adherence, and weight change. Studies that met other criteria but did not explicitly look at WS were not included or their authors contacted for more information. Samples comprised of adolescents and adults (*‘Population’*) assessed before and after treatment (*‘Comparison’*) were included. No age limits were set, although no studies included participants below 12 years of age and the upper age reported in the studies was 65 years.

Study selection

Outputs from all databases were combined and duplicates removed. Articles meeting inclusion criteria were then reviewed and screened for inclusion in meta-analysis (i.e., whether a suitable effect size could be derived). Relevant data extracted from articles included: sample size and setting; demographics (e.g., gender, age, ethnicity); BMI; diagnostic breakdown; treatment type (e.g., CBT); duration of treatment (including number of

sessions); length of follow-up (where given); outcome (i.e., weight change, symptom abstinence, treatment completion); results used for calculation of effect sizes. We also looked at how WS was computed and how weight and height were obtained.

Data synthesis

Findings are synthesised in a narrative review, with summary tables to aid comparisons. The predictive ability of WS was evaluated relating to different outcomes (weight change, treatment completion, symptom abstinence) following treatment. The term ‘weight change’ is used to refer to the spectrum of weight change rather than weight gain specifically, although these terms have been used synonymously (e.g., Carter et al., 2008). A number of studies also reported outcomes relating to rate of weight change, but this invariably showed near-perfect correlation with weight change and therefore resulted in similar findings (e.g., Carter et al., 2015; Wildes & Marcus, 2012): we therefore chose not to report this as a separate outcome variable.

Weight change was treated as a continuous variable (although Carter et al. [2008] also considered an arbitrary division of weight change $\pm 5\text{kg}$), and treatment completion as binary. Reporting of behavioural outcomes was more variable; some studies reported both dimensional and categorical outcomes.

Results

Search results

The search, conducted in November 2017, identified 443 unique records (see Figure 1). Thirty-four full-text articles were reviewed, of which twelve met inclusion criteria. One published article (Zunker et al., 2011) included two separate studies – one on BED and one on BN. A further study (Striegel-Moore et al., 2010) concluded that WS was not a significant predictor of outcome within a CBT-based treatment of binge eating but did not provide more

detailed information. This study was therefore excluded from the review. Agreement between raters was high ($\kappa = 0.87$, lower 95% CI = 0.70).

Summary of studies and risk of bias

Table 1 summarises study characteristics; all were cohort studies and deemed to be at moderate risk of bias (Wells et al., 1999). The modified-NOS score (max. = 8) ranged from 2 – 5 (see Supplementary Material, Table S1). In all studies, WS was the primary variable of interest and comprised analysis of either an existing randomised study (Butryn et al., 2006; Carter et al., 2008; Carter et al., 2015; Zunker et al., 2011) or routinely-collected clinic data (Berner et al., 2013; Dawkins et al., 2013; Hessler et al., 2017; Lowe et al., 2006; Seetharaman et al., 2017; Wildes & Marcus, 2012). All but one study used DSM criteria to define groups (Hessler et al. used ICD criteria), although different versions of the manual were used (ranging from DSM-III-R [APA, 1987] to DSM-IV-TR [APA, 2000]).

[Insert Table 1]

The eleven studies included in the narrative review reported baseline data for 1903 participants (1864 female). Data relating to WS as a predictor of outcome were available for 1566 individuals. Pre-treatment BMI varied from 9.39 – 72.6kg/m², although not all studies reported minima and maxima.

Regarding calculation of WS, current weight was recorded either by clinicians or the research team and, where reported, in ‘street clothing’ (e.g., Lowe et al., 2006) or a hospital gown (e.g., Wildes & Marcus, 2012). Precise definition of highest weight varied slightly, with some studies explicitly stating that highest *ever* weight was used (Butryn et al., 2006) and others specifying highest *adult* weight (e.g., Zunker et al., 2011) or highest lifetime non-pregnancy weight at adult height (Berner et al., 2013). Participants were usually asked to recall their highest weight (either in clinical interview or via self-report). Assessment of eating pathology employed standardised measures with good psychometric properties; these

differed across studies, and thus the definition of “abstinence” – where assessed – varied (see Table 1). All studies used regression analyses to determine the predictive effect(s) of WS, except for Hessler et al. (2017) who grouped a sample of female inpatients with BN by degree of WS (0 – 4kg, 4 – 9kg, 9 – 15.7kg, and 15.8kg+) and compared groups via analysis of variance (ANOVA).

Duration of treatment

A range of outpatient and more intensive treatments were included (see Table 1). The CBT-based (outpatient) interventions were of similar length, with mean duration ranging from around eight weeks (Carter et al., 2008) to 19 weeks (Dawkins et al., 2013) for treatment completers. Inpatient treatment lasted considerably shorter, an average of 17.2 days for ‘normal’-weight (Lowe et al., 2006) and around 40 days for underweight samples (Berner et al., 2013; Wildes & Marcus, 2012). Hessler et al. (2017) reported a mean length of stay of 77.8 days for their sample of females with BN.

Weight status

The majority of studies included groups comprising more than one diagnosis and according to different versions of diagnostic manuals. Therefore, for the purposes of this review, studies were classified according to inclusion of predominantly underweight or non-underweight samples (see Table 1).

Relationship between Weight Suppression and outcomes

It is important to comment briefly that around half of the included studies addressed WS as a moderator (Berner et al., 2013; Dawkins et al., 2013; Hessler et al., 2017; Lowe et al., 2006). According to Kraemer et al. (2002), moderation is defined as the presence of an interaction between the proposed moderator (in this case, WS) and treatment outcome. Most studies looked at WS as a nonspecific predictor of outcome – i.e., a variable that has a main

effect on outcome but not an interactive effect (although interactive effects with treatment were often not reported).

Three main outcomes were identified at post-treatment: weight change; treatment adherence / completion; and abstinence from binge eating / purging. Not all studies looked at all variables, and results regarding individual outcomes will be discussed in turn.

Weight change

Greater WS was associated with greater weight gain at post-treatment, with six studies finding that WS independently accounted for between 1.0% (Wildes & Marcus, 2012) and 12.3% (Carter et al., 2015) of the variance in weight gain. Two studies failing to find a predictive effect of WS were both reported by Zunker et al. (2011). Although the sample size is small, there were no clear associations with weight status or diagnosis; of those studies finding a relationship, three included underweight samples (Berner et al., 2013; Carter et al., 2015; Wildes & Marcus, 2012) and three included non-underweight samples (Carter et al., 2008; Hessler et al., 2017; Lowe et al., 2006). A chi-square test confirmed no association between weight status and outcome ($\chi^2(1) = 1.806, p = .179$).

Treatment completion

This was specified in seven studies, only one of which found that completion was significantly associated with pre-treatment WS. Butryn et al. (2006) found that WS accounted for 22% of the variance in treatment completion and that for each increase of 1kg of weight suppression, odds of not completing treatment are “multiplied by a factor of 0.69” (p. 64). Completion was defined according to slightly different criteria across studies; some reports defined completion according to number of sessions attended (e.g., Carter et al., 2008), and others specifically described “nonmutual premature termination of treatment” (Dawkins et al., 2013, p. 589).

Abstinence

Studies also assessed abstinence from ED behaviours at post-treatment, specifically binge eating and purging (usually self-induced vomiting). The results were broadly consistent with those regarding treatment completion, with eight studies reporting this as an outcome and two (Butryn et al., 2006; Wildes & Marcus, 2012) finding that higher WS at pre-treatment predicted a lower likelihood of abstinence from binge eating and purging at post-treatment. Presence of ED behaviours was assessed in different ways, with most studies using data from the Eating Disorder Examination interview (EDE; Fairburn & Cooper, 1993) and others employing the self-report version (EDE-Q; Fairburn & Beglin, 1994), both of which assess behaviours over the past four weeks. Carter et al. (2008) used a structured interview which, although similar to the EDE, assessed behaviours over the previous two weeks and Lowe et al. (2006) used a different self-report questionnaire, the BULIT-R (Thelen, Farmer, Wonderlich, & Smith, 1991).

Return of Menses

One study (Seetharaman et al., 2017) examined WS as a predictor of return of menses in a transdiagnostic sample of 163 adolescents treated at an academic ED program. The authors found that, in those not previously overweight ($n = 33$), the odds of menses returning decreased by 10% for every 1kg of weight suppression. No other studies were identified in this review as addressing this outcome.

Methodology of included studies

Across studies, calculation of WS generally referred to ‘adult’ or ‘current’ height, either through self-report or recall during an interview. Similarly, ED symptoms (and thus abstinence) were assessed by semi-structured interview ($k = 6$) or self-report ($k = 3$). In two studies (Hessler et al., 2017; Seetharaman et al., 2017), psychometric measures were not reported. Most studies reported on secondary analysis of existing data (e.g., from a RCT), and thus recruitment strategies also varied.

No studies found that greater WS significantly predicted better outcomes (such as increased likelihood of treatment completion), although some failed to find hypothesised associations. It was only possible to evaluate some variables in categorical form (e.g., treatment completion) and thus relationships were typically examined using logistic regression. Where both categorical (e.g., abstinent / non-abstinent) and continuous (e.g., frequency of binge eating and purging) outcomes were possible, individual studies found consistent results across methods (e.g., logistic and linear regression).

Discussion

A number of authors have looked at the ability of pre-treatment WS – defined as the difference between highest past weight and current weight – to predict outcome following treatment of EDs. The current review describes an extensive search of existing literature to test this hypothesis. Greater WS was associated with greater total weight gain at post-treatment although this finding was not replicated in all studies. Findings lend support to the theory that individuals “highest in weight suppression [are] most susceptible to future weight gain” (Butryn et al., 2006, p. 63), consistent with theories on metabolic adaptation and energy intake (e.g., see Rosenbaum et al., 2010). However, little support was found for the association between WS and treatment outcomes other than weight gain.

Evidence of the association between WS and weight change was seen in samples comprising both underweight and non-underweight samples, although it is possible that WS may have different effects across different ED diagnoses. Of note, only one study (Zunker et al., 2011) looked at a sample with BED, and found no associations. Further studies might examine WS in BED in greater detail. Individuals with BED tend to report higher maximum weight than controls but also greater weight cycling (e.g., Borges, Jorge, Morgan, de Silveira, & Custódio, 2002). In the treatment of AN, for example, weight regain is often a treatment

target, whereas weight change in other diagnostic groups may not necessarily be a goal of treatment.

Regarding psychological and other behavioural outcomes, evidence is less conclusive. Of seven studies reporting data, only two (Butryn et al., 2006; Wildes & Marcus, 2012) found that WS predicted either treatment completion or abstinence from binge eating and vomiting (and in one of these [Wildes & Marcus, 2012], only abstinence).

Studies reporting significant associations between WS and outcomes of interest tended to cite higher average WS at baseline, suggesting that WS may need to reach a certain degree to significantly impact treatment outcome (e.g., see Hessler et al., 2017). It was not possible to test this hypothesis in the current review, but one recommendation is that future studies consider that WS may have a non-linear relationship with outcome.

Where significant associations between WS and outcomes were found, these remained when covariates were considered (e.g., Butryn et al., 2006; Lowe et al., 2006). One study included only adolescents (Seetharaman et al., 2017), with a number looking at mixed adolescent/adult samples. It may not be possible to describe WS in adolescents who may not have achieved their highest adult weight, although it has been suggested that BMI z-scores may be more appropriate as a metric among youth (Accurso et al., 2016).

Long-term follow-up was not a focus of this review, although there is some limited evidence that WS can predict longer-term outcomes. Two studies have looked at the predictive significance of WS in cohorts followed for up to five (Herzog et al., 2010) or ten years (Keel & Heatherton, 2010), finding that WS predicts maintenance of symptoms of BN over time. Keel and Heatherton (2010) also found that greater WS at baseline in those without an ED predicted onset of a bulimic syndrome at follow-up. Bodell, Racine, and Wildes (2016) looked at degree of WS at post-treatment and found that it acted as both a predictor and moderator of BMI change over time which, together with findings from a

recent study of adolescents (Swenne, Parling, & Salonen Ros, 2017), suggests that the effects reported here regarding WS on weight change may endure over time.

As noted above, a minority of studies considered whether WS has a moderator role. Lowe et al. (2006) examined pre-treatment WS and BMI, but failed to find evidence of a significant interaction, and Dawkins et al. (2013) also failed to find interactions between WS and pre-treatment BMI, parental history of overweight, self-reported childhood body shape, or the difference between highest and lowest ever adult weight. By contrast, two studies (Berner et al., 2013; Hessler et al., 2017) reported a significant interaction effect between WS and BMI when predicting weight change, suggesting that the association between BMI and weight change over treatment may depend to some extent on degree of WS. In a study not included in the current review, Bodell et al. (2016) looked at WS taken at post-treatment in a sample of females with AN and found that, not only was change in BMI over the year following treatment predicted, but that the interaction between WS and BMI predicted change in ED severity and purging frequency over the same period. Therefore, it is possible that effects of WS on psychological symptoms are not evident immediately following treatment, but begin to emerge over time, and that interaction with BMI is of principal relevance for those with underweight EDs.

The review identified twelve eligible studies, although sufficient data for review were not available from one of these (which concluded that there was not a significant predictive effect of WS on treatment outcome; Striegel-Moore et al., 2010). Cohort designs were the singular means of study and the overall risk of bias was moderate, although a modified scale was used to suit the needs of the review. As such, the criteria used to grade risk of bias were not standardised and may have represented overly strict scoring metrics. In the studies reviewed, CBT was the most common treatment, although there was some variability across approaches. Individual differences in WS may determine the effects of different treatments,

and this is one reason we advocate explicitly looking at the possible moderating effect of WS in addition to its predictive effect (see Kraemer et al., 2002). Further work is required to look at the role of gender in more detail; although men typically comprise around 1 in 10 cases of EDs (which may even be an underestimate), the number of included males in this review was only around 2%.

Strengths and limitations

The strengths of this review include adherence to best practice guidelines for the conduct of systematic reviews and evaluation of study design and risk of bias. Attempts were also made to contact corresponding authors to clarify information that was not clearly reported. A limitation lies in the designs of included studies (typically cohort designs and/or secondary analysis of data) and that data were insufficient for meta-analysis. Due to methodological constraints unmasked reviews were conducted, which may have increased the likelihood of bias (Jadad et al., 1998). Although a scale was used to evaluate risk of bias (Wells et al., 1999), adjustments were made to its content and many of the remaining items did not deal well with the designs under review; thus, the scores obtained lack external consistency (but might retain internal consistency). Generalisability of findings was limited in some respects (e.g., gender), but sampled different populations, and often analysis was of multi-site trials, which may have reduced selection bias. Furthermore, the review also considered all ED diagnoses and is therefore applicable ‘transdiagnostically’.

Implications and Recommendations

The studies reviewed are relevant to clinical practice as all focused on how WS at pre-treatment might predict response to treatment. As stated above, the evidence does not conclusively support the hypothesis that higher WS is associated with better outcome following treatment, but it does seem to be associated with weight change. As a result, it seems that some patients (i.e., those with high WS) may need to gain more weight during

treatment to achieve good outcomes, including normalisation of eating (Lowe et al., 2006). Although meta-analysis was not possible, combining results from six studies (Berner et al., 2013; Carter et al., 2008; Carter et al., 2015; Hessler et al., 2017; Lowe et al., 2006; Wildes & Marcus, 2012) suggests that around 5% of the variance in weight gain is accounted for by pre-treatment WS (weighted mean estimate; Hunter & Schmidt, 1990). Further research studies might consider this in planning stages, for example when determining sample sizes based on power analysis or considering a number of predictors of outcome.

Addressing WS has not been routinely described in evidence-based treatments for EDs, and one group has recently suggested that this be included as part of current CBT models and that treatment might be augmented by addressing WS directly (Juarascio et al., 2018). The evidence summarised here provides indirect support for psychoeducation around weight gain, preparing patients for the possibility of weight gain, discussing possible weight fluctuations during treatment, helping patients set realistic weight goals, focusing on health rather than weight, and working on healthy weight management strategies (see Bulik, Marcus, Zerwas, Levine, & La Via, 2012; Juarascio et al., 2018). Although psychoeducation is likely to confer some health benefits, a societal approach to reducing weight stigma and promoting healthy eating (e.g., see Paxton, Wertheim, Pilawski, Durkin, & Holt, 2002) may be necessary to effect lasting and widespread changes, similar to approaches used in the fields of diabetes and obesity (e.g., Glasgow et al., 1999; Lipek, Igel, Gausche, Kiess, & Grande, 2015). Such interventions might include reference, at both the individual and environmental level, to the difficulty and costs of sustaining reduced body weight (e.g., Rosenbaum et al., 2010); this might include developing public understanding of weight loss (e.g., see Puhl & Heuer, 2010) and tailoring information to those at high-risk of developing EDs (e.g., individuals engaging in extreme dieting).

At present, there is insufficient evidence to conclude that WS is a strong predictor of outcomes other than weight change. Further work in this area may be warranted, but we also suggest that authors consider the role of WS as a predictor and/or as a moderator of outcome. Similarly, further research might consider whether there is a level of WS above which treatment outcomes are significantly impacted. We recommend that future studies provide sufficient information to enable detailed secondary (e.g., meta-) analysis, such as providing a scale-bound measure of fit (e.g., R^2) in studies looking at the relation between WS and a given outcome. If regression analyses are reported, we advise including both standardised and unstandardized regression coefficients (slopes) with their respective standard errors and sample sizes (e.g., see Berner et al., 2013). In addition, mean (SD) values for different groups (e.g., WS in completers versus non-completers) would also be of use. Although the regression coefficient is equivalent to Pearson's r in a simple regression (Hayes, 2013), more information may be needed in hierarchical models. Finally, as many studies report data from secondary analysis of treatment studies, it would be advisable to ensure that studies have sufficient power to correctly interpret model estimates. Although meta-analysis was not performed, we hope that the summary of findings presented here might help in this regard.

In summary, WS represents a variable that has shown some promise in predicting weight change over the course of psychological treatment. However, more work in this area is needed to establish whether WS predicts outcomes more generally, whether a moderator effect exists, and if there is a WS threshold beyond which outcomes are consistently affected.

References

- Agras, W. S., Crow, S. J., Halmi, K. A., Mitchell, J. E., Wilson, G. T., & Kraemer, H. C. (2000). Outcome predictors for the cognitive behavioral treatment of bulimia nervosa: Data from a multisite study. *American Journal of Psychiatry*, *157*, 1302–1308. DOI: 10.1176/appi.ajp.157.8.1302.
- Accurso, E. C., Lebow, J., Murray, S. B., Kass, A. E., & Le Grange, D. (2016). The relation of weight suppression and BMIz to bulimic symptoms in youth with bulimia nervosa. *Journal of Eating Disorders*, *4*, 21. DOI: 10.1186/s40337-016-0111-5.
- American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., revised). Author: Washington, DC.
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., revised). Author: Washington, DC.
- *Berner, L. A., Shaw, J. A., Witt, A. A., & Lower, M. R. (2013). The relation of weight suppression and body mass index to symptomatology and treatment response in anorexia nervosa. *Journal of Abnormal Psychology*, *122*, 694-708. DOI: 10.1037/a0033930.
- Bodell, L. P., Racine, S. E., & Wildes, J. E. (2016). Examining weight suppression as a predictor of eating disorder symptom trajectories in anorexia nervosa. *International Journal of Eating Disorders*, *49*, 753-763. DOI: 10.1002/eat.22545.
- Borges, M. B. F., Jorge, M. R., Morgan, C. M., da Silveira, D. X., & Custódio, O. (2002). Binge-eating disorder in Brazilian women on a weight-loss program. *Obesity: A Research Journal*, *10*, 1127–1134. DOI: 10.1038/oby.2002.153.
- Bulik, C. M., Marcus, M. D., Zerwas, S., Levine, M. D., & La Via, M. (2012). The changing “weightscape” of bulimia nervosa. *American Journal of Psychiatry*, *169*, 1031-1036. DOI: 10.1176/appi.ajp.2012.12010147.

- Bulik, C. M., Sullivan, P. F., Carter, F. A., McIntosh, V. V., & Joyce, P. R. (1998). The role of exposure with response prevention in the cognitive behavioural therapy for bulimia nervosa. *Psychological Medicine*, 28, 611–623.
- *Butryn, M. L., Lowe, M. R., Safer, D. L., & Agras, W. S. (2006). Weight suppression is a robust predictor of outcome in the cognitive-behavioral treatment of bulimia nervosa. *Journal of Abnormal Psychology*, 115, 62-67. DOI: 10.1037/0021-843X.115.1.62.
- *Carter, F. A., Boden, J. M., Jordan, J., McIntosh, V. V., Bulik, C. M., & Joyce, P. R. (2015). Weight suppression predicts total weight gain and rate of weight gain in outpatients with anorexia nervosa. *International Journal of Eating Disorders*, 48, 912-918. DOI: 10.1002/eat.22425.
- *Carter, F. A., McIntosh, V. V. W., Joyce, P. R., Bulik, C. M. (2008). Weight suppression predicts weight gain over treatment but not treatment completion or outcome in bulimia nervosa. *Journal of Abnormal Psychology*, 117, 936-940. DOI: 10.1037/a0013942.
- Centre for Reviews and Dissemination (2008). *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. York, UK: York Publishing Services.
- Cooper, Z., Allen, E., Bailey-Straebl, S., Basden, S., Murphy, R., O'Connor, M. E., et al. (2016). Predictors and moderators of response to enhanced cognitive behaviour therapy and interpersonal psychotherapy for the treatment of eating disorders. *Behaviour Research and Therapy*, 84, 9-13. DOI:10.1016/j.brat.2016.07.002.
- *Dawkins, H., Watson, H. J., Egan, S. J., & Kane, R. T. (2013). Weight suppression in bulimia nervosa: relationship with cognitive behavioral therapy outcome. *International Journal of Eating Disorders*, 46, 586-593. DOI: 10.1002/eat.22137.
- Fairburn, C. G. (2008). *Cognitive behaviour therapy and eating disorders*. New York: Guilford Press.

- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: interview or self-report questionnaire? *International Journal of Eating Disorders*, 16, 363-370.
- Fairburn, C. G., & Cooper, Z. (1993). The eating disorder examination. In *Binge eating: nature, assessment, and treatment* (Eds. C. G. Fairburn and G. T. Wilson), pp. 317-360. New York: Guilford Press.
- Fairburn, C. G., Marcus, M. D., & Wilson, G. T. (1993). Cognitive-behavioral therapy for binge eating and bulimia nervosa: a comprehensive treatment manual. In *Binge eating: nature, assessment, and treatment* (Eds. C. G. Fairburn and G. T. Wilson), pp. 361-404. New York: Guilford Press.
- Glasgow, R. E., Wagner, E. H., Kaplan, R. M., Vinicor, F., Smith, L., & Norman, J. (1999). If diabetes is a public health problem, why not treat it as one? A population-based approach to chronic illness. *Annals of Behavioral Medicine*, 21, 159-170. DOI: 10.1007/BFO2908297.
- Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford Press.
- Herzog, D. B., Thomas, J. G., Kass, A. E., Eddy, K. T., Franko, D. L., & Lowe, M. R. (2010). Weight suppression predicts weight change over 5 years in bulimia nervosa. *Psychiatry Research*, 177, 330-334. DOI: 10.1016/j.psychres.2010.03.002.
- *Hessler, J. B., Diedrich, A., Greetfeld, M., Schlegl, S., Schwartz, C., & Voderholzer, U. (2017). Weight suppression but not symptom improvement predicts weight gain during inpatient treatment for bulimia nervosa. *European Eating Disorders Review*. DOI:10.1002/erv.2573.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of Meta-Analysis*. Newbury Park, CA: Sage.

- Jadad, A. R., Moher, D., & Klassen, T. P. (1998). Guides for reading and interpreting systematic reviews: II. How did the authors find the studies and assess their quality? *Archives of Pediatrics & Adolescent Medicine*, 152, 812-817.
- Juarascio, A., Lantz, E. L., Muratore, A. F., & Lowe, M. R. (in press). Addressing weight suppression to improve treatment outcome for bulimia nervosa. *Cognitive and Behavioral Practice*. DOI:10.1016/j.cbpra.2017.09.004.
- Keel, P. K., Bodell, L. P., Haedt-Matt, A. A., Williams, D. L., & Appelbaum, J. (2017). Weight suppression and bulimic syndrome maintenance: Preliminary findings for the mediating role of leptin. *International Journal of Eating Disorders*. DOI:10.1002/eat.22788.
- Keel, P. K., & Heatherton, T. F. (2010). Weight suppression predicts maintenance and onset of bulimic syndromes at 10-year follow-up. *Journal of Abnormal Psychology*, 119, 268-275. DOI: 10.1037/a0019190.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, 59, 877-883. DOI:10.1001/archpsyc.59.10.877.
- Lammers, M. W., Vroling, M. S., Ouwens, M. A., Engels, R. C. M. E., & van Strien, T. (2015). Predictors of outcome for cognitive behaviour therapy in binge eating disorder. *European Eating Disorders Review*, 23, 219-228. DOI:10.1002/erv.2356.
- Lipek, T., Igel, U., Gausche, R., Kiess, W., & Grande, G. (2015). Obesogenic environments: environmental approaches to obesity prevention. *Journal of Pediatric Endocrinology & Metabolism*, 28, 485-495. DOI: 10.1515/jpem-2015-0127.
- *Lowe, M. R., Davis, W., Lucks, D., Annunziato, R., & Butryn, M. (2006). Weight suppression predicts weight gain during inpatient treatment of bulimia nervosa. *Physiology & Behavior*, 87, 487-492. DOI: 10.1016/j.physbeh.2005.11.011.

- McIntosh, V. V., Jordan, J., Carter, F. A., Luty, S. E., McKenzie, J. M., Bulik, C. M., et al. (2005). Three psychotherapies for anorexia nervosa: A randomized-controlled trial. *American Journal of Psychiatry*, 162, 741–747. DOI:10.1176/appi.ajp.162.4.741.
- Mitchell, J. E., Crosby, R. D., Wonderlich, S. A., Crow, S., Lancaster, K., Simonich, H., et al. (2008). A randomized trial comparing the efficacy of cognitive-behavioral therapy for bulimia nervosa delivered via telemedicine versus face-to-face. *Behaviour Research & Therapy*, 46, 581–592. DOI:10.1016/j.brat.2008.02.004.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*, 6. DOI:10.1371/journal.pmed.1000097.
- Paxton, S. J., Wertheim, E. H., Pilawski, A., Durkin, S., & Holt, T. (2002). Evaluations of dieting prevention messages by adolescent girls. *Preventative Medicine*, 35, 474-491. DOI:10.1006/pmed.2002.1109
- Peterson, C. B., Mitchell, J. E., Crow, S. J., Crosby, R. D., & Wonderlich, S. A. (2009). The efficacy of self-help group treatment and therapist-led group treatment for binge eating disorder. *American Journal of Psychiatry*, 166, 1347–1354. DOI: 10.1176/appi.ajp.2009.09030345.
- Puhl, R. M., & Heuer, C. A. (2010). Obesity stigma: Important considerations for public health. *American Journal of Public Health*, 100, 1019–1028. DOI:10.2105/AJPH.2009.159491
- Reeves, B. C., Deeks, J. J., Higgins, J. P. T., & Wells, G. A. (2011). Chapter 13: Including non-randomized studies. In *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.1.0 (updated March 2011) (Eds. P. T. Higgins and S. Green). The Cochrane Collaboration. Available from www.cochrane-handbook.org, accessed 4 April 2017.

- Rosenbaum, M., Kissileff, H. R., Mayer, L. E. S., Hirsch, J., & Leibel, R. L. (2010). Energy intake in weight-reduced humans. *Brain Research*, 1350, 95-102. DOI: 10.1016/j.brainres.2010.05.062.
- Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8, 448-467. DOI: 10.1037/1082-989X.8.4.448.
- *Seetharaman, S., Golden, N. H., Halpern-Felsher, B., Peebles R., Payne, A., & Carlson, J. L. (2017). Effect of a prior history of overweight on return of menses in adolescents with eating disorders. *Journal of Adolescent Health*, 60, 469-471. DOI: 10.1016/j.jadohealth.2016.10.019.
- Stang, A. (2010). Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomised studies in meta-analysis. *European Journal of Epidemiology*, 25, 603-605. DOI: 10.1007/s10654-010-9491-z.
- Stice, E., Durant, S., Burger, K. S., & Schoeller, D. A. (2011). Weight suppression and risk of future increases in body mass: effects of suppressed resting metabolic rate and energy expenditure. *The American Journal of Clinical Nutrition*, 94, 7-11. DOI: 10.3945/ajcn.110.010025.
- Striegel-Moore, R. H., Wilson, G. T., De Bar, L., Perrin, N., Lynch, F., Rosselli, F., et al. (2010). Cognitive behavioural guided self-help for the treatment of recurrent binge eating. *Journal of Consulting and Clinical Psychology*, 78, 312-321. DOI: 10.1037/a0018915.
- Swenne, I., Parling, T., & SalonenRos, H. (2017). Family-based intervention in adolescent restrictive eating disorders: early treatment response and low weight suppression is associated with favourable one-year outcome. *BMC Psychiatry*, 17, 333. DOI:10.1186/s12888-017-1486-9.

- Thelen, M. H., Farmer, J., Wonderlich, S., & Smith, M. (1991). A revision of the Bulimia Test: the BULIT-R. *Psychological Assessment*, 3, 119-124. DOI: 10.1037/1040-3590.3.1.119.
- Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., et al. (1999). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. Ottawa (ON): Ottawa Health Research Institute.
- *Wildes, J. E., & Marcus, M. D. (2012). Weight suppression as a predictor of weight gain and response to intensive behavioural treatment in patients with anorexia nervosa. *Behaviour Research and Therapy*, 50, 266-274. DOI: 10.1016/j.brat.2012.02.006.
- *Zunker, C., Crosby, R. D., Mitchell, J. E., Wonderlich, S. A., Peterson, C. B., & Crow, S. L. (2011). Weight suppression as a predictor variable in treatment trials of bulimia nervosa and binge eating disorder. *International Journal of Eating Disorders*, 44, 727-730. DOI: 10.1002/eat.20859.

Table 1. Studies included in the review.

First author, year	Pre-treatment study characteristics						Outcome				
	N (% female)	Mean age, y (SD)	Age range, y	Weight status	Mean BMI, kg/m ²	Treatment	Mean WS, kg (SD)	Definition of abstinence	Abstinence	Weight gain	Treatment completion
Berner, 2013	350 (100)	23.9 (10.0)	NR	Underweight	15.7	Inpatient	15.82 (9.18)	NA	NR	NR	+
Butryn, 2006 (see Agras et al., 2000, for details)	188 (100)	28.1 (7.9)	NR	Normal	23.5	CBT-BN	9.4 (9.5)	No binge eating or purging, assessed through EDE (covering a 4-week period)	+	NR	+
Carter, 2008 (see Bulik et al., 1998, for details)	135 (100)	26.1 (6.1)	17 – 45	Normal	22.5	CBT + BT	7.1 (7.4)	No binge eating or purging, assessed through structured interview (covering a 2-week period)	/	+	/
Carter, 2015 (see McIntosh et al., 2005, for details)	35 (100)	24.3 (6.8)	16 – 40	Underweight	17.4	CBT, IPT, or SSCM	14.3 (7.8)	No binge eating or purging, assessed through EDE (covering a 4-week period)	+	/	/
Dawkins, 2013	117 (100)	25.5 (7.6)	16 – NR	Normal	22.6	CBT-E	9.5 (11.9)	No binge eating or purging from	/	NR	/

									EDE-Q (i.e., 4 weeks)			
Hessler, 2017	179 (100)	26.6 (8.9)	18 – 55	Normal	22.0	Inpatient	NR	NA		+	NR	NR
Lowe, 2006	146 (100)	25.4 (7.9)	14 – 53	Normal	23.2	Inpatient	12 (NR)	NA		NR	+	NR
Seetharaman, 2017	163 (100)	16.6 (2.1)	12 – 18	Underweight	NR	Outpatient	NR	NA		NR	NR	NR
Wildes, 2012	185 (97.3)	27.0 (10.2)	16 – NR	Underweight	15.7	Inpatient and Daypatient	17.1 (10.8)	No binge eating or purging, assessed through modified EDE (covering a 4- week period)		+	/	+
Zunker, 2011 (see Peterson et al., 2009, for details)	259 (87.6)	47.1 (10.4)	19 – 65	Normal	39.0	Group CBT	Median = 2.27	No binge eating or purging, assessed through EDE (covering a 4-week period)		/	/	/
Zunker, 2011 (see Mitchell et al., 2008, for details)	128 (98.4)	29.0 (10.7)	18 – 60	Normal	23.4	CBT-BN	9.2 (12.1)	No binge eating or purging, assessed through EDE (covering a 4-week period)		/	/	/

Abbreviations: AN = anorexia nervosa; BMI = body mass index; BN = bulimia nervosa; BT = behavior therapy; CBT = cognitive behavior therapy; EDE = Eating Disorder Examination; EDE-Q = Eating Disorder Examination – Questionnaire; IPT = interpersonal psychotherapy; M = mean; SSCM = specialist supportive clinical management; WS = weight suppression; NA = not applicable; NR = not reported; + = predicted; / = not predicted

Figure 1. PRISMA flow diagram

