Is physical exercise and dietary therapy a feasible alternative to cognitive behavior therapy in treatment of eating disorders? A randomized controlled trial of two group therapies

Therese Fostervold Mathisen PhD1 | Jan H. Rosenvinge PhD2 | Oddgeir Friborg PhD2 | KariAnne Vrabel PhD3 | Solfrid Bratland-Sanda PhD4 | Gunn Pettersen PhD5 | Jorunn Sundgot-Borgen PhD1

1Department of Sports Medicine, Norwegian School of Sport Sciences, Oslo, Norway
2UiT—The Arctic University of Norway, Department of Psychology, Faculty of Health Sciences, Tromsø, Norway
3Research Institute of Modum Bad, Vikersund, Norway
4Department of Outdoor Studies, Sports and Physical Education, University College of Southeast Norway, Bø, Norway
5Department of Health and Caring Sciences, Faculty of Health Sciences, UiT—The Arctic University of Norway, Tromsø, Norway

Correspondence
Therese F. Mathisen, Østfold University College, Faculty of Health and Welfare, Kobberslagervest 5, 1671 Kråkerøy, Norway.
Email: therese.f.mathisen@hiof.no

Funding information
Norwegian School of Sport Sciences; The Norwegian Women’s public health association; University of Tromsø - The arctic University of Norway

Abstract

Objective: To compare effects of physical exercise and dietary therapy (PED-t) to cognitive behavioral therapy (CBT) in treatment of bulimia nervosa (BN) and binge-eating disorder (BED).

Method: The active sample (18–40 years of age) consisted of 76 women in the PED-t condition and 73 in the CBT condition. Participants who chose not to initiate treatment immediately (n = 23) were put on a waiting list. Outcome measures were the eating disorder examination questionnaire (EDE-Q), Clinical Impairment Assessment (CIA), Satisfaction with Life Scale (SWLS), Beck Depression Inventory (BDI), and numbers in remission at posttreatment, and at 6-, 12-, and 24-months follow-up.

Results: Both treatment conditions produced medium to strong significant improvements on all outcomes with long-term effect. The PED-t produced a faster improvement in EDE-Q and CIA, but these differences vanished at follow-ups. Only PED-t provided improvements in BDI, still with no between-group difference. Totally, 30–50% of participants responded favorable to treatments, with no statistical between-group difference.

Discussion: Both treatments shared a focus on normalizing eating patterns, correcting basic self-regulatory processes and reducing idealized aesthetic evaluations of self-worth. The results point to the PED-t as an alternative to CBT for BN and BED, although results are limited due to compliance and dropout rates. Replications are needed by independent research groups as well as in more clinical settings.

Keywords
affect regulation, binge-eating disorder, bulimia nervosa, dietary therapy, exercise therapy, physical activity, self-regulation
1 | INTRODUCTION

Sustained recovery from bulimia nervosa (BN) or binge-eating disorder (BED) requires that a treatment focuses on the overvaluation of controlling body weight and food intake, and the affective dysregulation represented by binge eating with or without purging (Fairburn et al., 1995). Cognitive behavior therapy (CBT) is empirically supported as effective in alleviating behavioral and cognitive maintaining factors (Hay, 2013; Linardon, Fairburn, Fitzsimmons-Craft, Willfrey, & Brennan, 2017), and is currently held as the preferred therapy for BN and BED (Hay, 2013; Linardon et al., 2017; Linardon & Wade, 2018). Still, more than 60% do not fully abstain from symptoms after CBT (Linardon & Wade, 2018). High rates of premature treatment termination (Linardon, Hindle, & Brennan, 2018) and relapses (Södersten, Bergh, Leon, Brodin, & Zandian, 2017) may be due to insufficient delivery (Mulkens, de Vos, de Graaff, & Waller, 2018; Waller, Stringer, & Meyer, 2012), and patient's clinical heterogeneity beyond what is captured by the diagnostic criteria. Hence, the efficacy of new treatment options needs to be explored to reach out to persons with BN or BED who, for some reason, do not benefit from CBT or other kind of psychological treatments.

Physical activity is found to improve cognitive function, self-regulation and positive affect (Buckley, Cohen, Kramer, McAuley, & Mullen, 2014; Lambourne & Tomporowski, 2010; Ludwig & Rauch, 2018; Mullen & Hall, 2015; Oaten & Cheng, 2006) as well as self-esteem, self-perception, self-efficacy, global self-worth and overall quality of life (Dupuy et al., 2015; Fox, 1999; Haugen, Säfvenbom, & Ommundsen, 2011; Oaten & Cheng, 2006; Vancampfort et al., 2014). In addition, regular physical activity is found to alleviate symptoms of anxiety and depression (Fox, 1999; Rosenbaum, Tiedemann, Sherrington, Curtis, & Ward, 2014; Stubbs et al., 2018). Such effects may be explained both neurobiological, psychosocial, and behavioral changes (Ashdown-Franks et al., 2019; Lubans et al., 2016; Mullen & Hall, 2015). If physical activity is included in the treatment of BN and BED, such changes may be relevant to expect as both disorders are acknowledged as cognitive and neurobiological in nature (Donnelly et al., 2018; Grant & Chamberlain, 2019). Both immediate and longitudinal effects from moderate intensity physical activity have shown positive cognitive effect by improving eating behavior and self-regulation in healthy students (Lowe, Hall, Vincent, & Luu, 2014; Oaten & Cheng, 2006). Noncompensatory physical activity is also associated with less psychopathology in persons with BED or BN (Kerrigan, Lydecker, & Grilo, 2019). However, physical activity is rarely incorporated in the treatment of BN (Bratland-Sanda et al., 2009; Quensel et al., 2018). One possible reason may be the fact that dysfunctional exercise affects approximately 20–40% of persons with BN (Dalle Grave, Calugi, & Marchesini, 2008), a behavior explained by the need for affect regulation, or in an urge to control body weight and shape (Meyer & Taranis, 2011). Clinicians may deem it counterintuitive to increase behaviors that are understood as BN-symptoms, hence the reluctance in incorporating physical activity in treatment of eating disorders. However, such opinions and practices lack empirical evidence. On the contrary, in a previous randomized controlled trial (RCT) from our research group, we found that therapist guided group-based physical activity was equal to group-CBT in reducing binge eating, purging and pursuit of thinness in women with BN at the end of treatment, and even superior to CBT at 18 months follow-up (Sundgot-Borgen, Rosenvinge, Bahr, & Schneider, 2002). Moreover, uncontrolled studies indicate that physical activity reduces binge eating and body weight (Vancampfort et al., 2013), and increases self-efficacy in treatment of BED (Vancampfort et al., 2014). Additionally, a recent systematic review identified beneficial effects from physical activity on treatment of BED (Blanchet et al., 2018). The weak empirical support for dietary counseling as a single treatment option for BN (Laessle et al., 1991; Sundgot-Borgen et al., 2002) and BED (Hay, 2013) does not exclude the possibility that such counseling may amplify other treatment approaches such as CBT (Hsu et al., 2001). We argue that to combine dietary therapy with guided physical activity (physical exercise and dietary therapy, PED-t) may represent an effective, alternative treatment option for persons with BN and BED. This combination has to our knowledge, never been tested against a well-established psychological treatment using a RCT design.

We thus hypothesized that PED-t is comparable to CBT in alleviating BN and BED - as well as depressive symptoms, and in increasing subjective well-being and improving psychosocial impairment (Mathisen et al., 2017). We also hypothesized that the PED-t would result in less drop out, and produce a more rapid treatment response than CBT.

2 | METHODS

2.1 | Study design, sample, and recruitment procedures

This RCT recruited women with BN or BED for outpatient treatment over 16 weeks arranged at the Norwegian School of Sport Sciences, in Oslo (Norway) during 2014–2017. All outcomes were measured by self-reports at baseline (T1), after treatment (T2), and at 6- (T3), 12- (T4) and 24-months (T5) posttreatment. Additionally, an experienced clinician assessed by interview the diagnostic criteria for BN and BED according to the definition by DSM-5 (American Psychiatric Association, 2013); that is, an overvaluation of controlling body weight, –shape and food intake, and with frequency of symptomatology (binge eating and/or purging) at least once per week.

Women 18–40 years of age and with a DSM-5 diagnosis of BN or BED (all categories of severity), and a BMI between 17.5 and 35 kg m⁻² were eligible. The gender restriction and age range reflected the intention to increase treatment group homogeneity, while the upper BMI value was chosen because the manual-based exercise program did not allow individual adjustments for reasons related to severe obesity. Considering the need to comply with a weekly, repetitive exercise program with high impact loading (i.e., interval running and weightlifting exercises like squats and walking lunges), being the manual, which PED-t relies on, we did not find it safe to include persons with a BMI above 35. A comorbid axis I psychiatric disorder that precluded the eating disorder focused treatment
and needed more extensive care was excluded. More thorough description of inclusion and exclusion criteria is provided in the open-access published protocol (Mathisen et al., 2017).

Participants were recruited through general practitioners (GP’s), magazines, newspapers, national TV, social media, posters, handouts, and through webpages of the national patient organizations for EDs. We interviewed participants by phone 2–6 weeks prior to the baseline measures and randomization, and all participants and their GP’s had to sign an informed consent before final inclusion.

2.2 | Randomization and masking

Two researchers not affiliated with the study produced the randomization list (by www.randomizer.org) with block size of eight, and concealed allocation to either group CBT or group PED-t, and they performed allocation by listing the randomization results by id-number. The project leader conveyed the allocation to the participants after they had completed the baseline measures. Participants recruited during ongoing treatment or who for some reason were unable to enter therapy immediately were put on a waitlist for 16 weeks and thereafter randomized to one of the two groups. This group served as a reference group for the two therapies at pre- and post-treatment. All outcomes, anonymously coded by ID-number, were transferred to SPSS files by blinded statistical staff before the final statistical analyses.

2.3 | Procedure

The 16-weeks treatments consisted of 20 group sessions, each of 90–120 min duration with 5–8 participants in each group.

The CBT manual rests on a transdiagnostic model positing generic core ED-characteristics across ED-diagnoses (Fairburn et al., 2009). The CBT-therapy was based on Fairburn’s individual CBT-E therapy, but adapted for groups as described previously (Fairburn, 2008). This means that the treatment followed the CBT-E protocol and it was designed to mirror the 20-week treatment. The group meetings followed a similar format to that used in individual treatment and covered the exact same material. In advance of each meeting the patient met the therapist to be weighed and for a brief review of their monitoring records. The therapy consisted of 1–2 weekly therapy sessions of 90 minutes each, and the manual outlined four stages: (a) engagement and behavioral change, (b) monitoring and evaluating progress, (c) addressing the core ED pathology, and (c) relapse prevention. Homework comprised registration of meals, thoughts and emotions during mealtimes, and behavioral experiments. Clinical psychologists certified in CBT, with at least 10 years of experience, were conducting this therapy. The therapists were instructed and led by one senior therapist (Ph.D., clinical psychologist and scientist) having more than 20 years of experience in CBT and in treatment of eating disorders.

The PED-t manual combines progressive resistance training and high intensity interval running, with dietary therapy; that is, education, group discussions and exploration of behavioral and practical skills (Mathisen et al., 2017). The PED-t follows the same structure as the CBT-E; hence, 20 therapy sessions in which 60 min dietary therapy directly followed 45 min with supervised resistance exercise. The dietary part of the PED-t was arranged as focused modules, that is: (a) meals frequency, portion sizes and the eating context, (b) detailed knowledge about energy and nutrient needs, and exercise physiology theory, and (c) presentation of a personal plan for the future. Homework in the PED-t therapy consisted of two unsupervised exercise sessions per week (one resistance exercise and one interval running session, both following a prescribed, progressive program), and dietary tasks related to weekly topics (Mathisen et al., 2017). Physical exercise therapists and dietitians (M.Sc.) were responsible for this therapy, and they received supervision from an experienced therapist (Ph.D., dietitian, and exercise therapist).

A detailed account of the procedures, safety agenda and the treatment contents has been described previously (Mathisen et al., 2017). Previously described are also results with respect to effects on physical health and dysfunctional exercise routines (Mathisen et al., 2018; Mathisen, Sundgot-Borgen, Rosenvinge, & Bratland-Sanda, 2018), as well as feasibility and acceptability (Bakland et al., 2018; Bakland et al., 2019; Pettersen et al., 2017).

2.4 | Outcomes

2.4.1 | Eating disorder examination questionnaire (EDE-Q)—6.0

The EDE-Q 6.0 (Cronbach’s α = .87, T1 current study) validly assesses the frequency and severity of ED features to produce ED diagnoses according to the DSM-5 (Fairburn, 2008). Apart from a global score, four subscales measure eating restraint (ER), eating concern (EC), shape concern (SC), and weight concern (WC). Mean (SD) global score for a Norwegian cohort of healthy female controls is 1.25 (1.10), while corresponding national clinical cutoff for probable BN and BED are 2.62, and 2.63, respectively (Rø, Reas, & Stedal, 2015). Reported binge-eating episodes are based on scores from EDE-Q item 14 (reported binge eating with loss of control). Reported purging episodes are based on scores from EDE-Q item 16–18 (i.e., purging by self-induced vomiting, use of laxatives, and driven/compulsive exercise).

2.4.2 | Remission

Calculation of numbers in full remission is based on the EDE-Q global score of the national cohort sample's normative value +1 SD (i.e., ±2.35) (Rø et al., 2015), with concurrently abstinence from binge eating and weight compulsatory behavior (i.e., 0 episodes last 28 days, as reported in EDE-Q item 13–18). Calculation of numbers in partial remission is based on the EDE-Q global score below the diagnostic specific cutoff (i.e., 2.62 in BN and 2.63 in BED) (Rø et al.,
2.4.3  |  Clinical impairment assessment (CIA 3.0)

The CIA 3.0 (Cronbach's $\alpha = .90$, T1 current study) measures ED-elicited personal, social, and cognitive impairment (Fairburn, 2008). The CIA consists of 16 items scored on a 4-point Likert scale ranging from 0 (not at all) to 3 (a lot) during the past 28 days, with total scores ranging 0–48 (Bohn et al., 2008).

2.4.4  |  Beck depression inventory (BDI-Ia)

The BDI-Ia (Cronbach's $\alpha = .86$, based on T1) measures current (past 2 weeks) self-reported symptoms of depression (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). It consists of 21 items scored with a 4-point Likert scale ranging from 0 (not at all) to 3 (extreme). Total score range is 0–62, and scores above $\geq 21$ is a recommended cutoff indicating a clinically significant episode of major depression (Beck et al., 1961).

2.4.5  |  Satisfaction with life scale (SWLS)

The SWLS (Cronbach's $\alpha = .89$, based on T1) is a short 5-item scale measuring overall contentment with life compared to personal standards and expectations, using a 5-point Likert scale ranging from 1 (never true) to 7 (always true) (Diener, 1994).

2.5  |  Dropout and lost to follow-up

We defined dropouts as those who did not complete the treatment, or those who by other reasons did not provide posttreatment data. To reduce loss of therapy sessions and avoid early dropout, we contacted all participants who were missing from therapy sessions. Lost to follow-up are those not attending follow-up testing (T3–T5), typically for reasons like work/studies, moving too far away, injuries or illness, pregnancy, not willing to attend physically to the testing facilities, and not responding to email, mail or electronic questionnaires.

2.6  |  Statistical analyses

The power calculation for this trial is provided elsewhere (Mathisen et al., 2017). The SPSS version 24 was used for all analyses. Differences between diagnostic groups at baseline were explored with independent t-tests or Mann–Whitney test. Attrition rates were analyzed separately for each time (posttreatment and follow-ups, T2–T5) with independent t-test or chi-square analyzes as appropriate. A significance level of $p < .05$ was used for all bivariate comparisons.

All inferential tests defined the two therapy arms, CBT and PED-t, as a between-group factor, whereas the repeated posttest scores represented a within-group factor. The low ($n = 23$) number of waitlist participants was excluded from group comparisons due to low statistical power, but are presented as a reference group in the figure and tables. Linear mixed regression models were built to estimate the coefficients for the between (PED-t vs. CBT) and the within factors (baseline vs. any of the three posttest measures). This analysis yields relatively unbiased estimates despite drop out given that data are missing completely at random or missing at random. Moreover, it can be safely used without conducting beforehand multiple imputations (Twisk, de Boer, de Vente, & Heymans, 2013). Standard errors were estimated with the restricted maximum likelihood function. Type III $F$-tests were used. The dependency in the four repeated outcome measures tended to decline across time, and hence an autoregressive model was fitted to the residual covariance matrix. The fixed factors were: Group (0-PED-t, 1-CBT) representing the overall therapy difference, Time (repeated measures) representing change across measurements, and the Group $\times$ Time interaction in order to detect treatment differences at certain time points only. The between-group analyses used the baseline values as a covariate to increase the statistical power (Egbewale, Lewis, & Sim, 2014). Differences between the therapy arms were examined with planned comparisons at each time point (least square difference tests). In addition, linear time trends were estimated by specifying time as a covariate instead of a fixed (and dichotomous) effect. The within-group analyses included all four measurements in the Time factor. Due to the number of tests, differences with $p$-values $<.01$ were considered as statistically significant.

A comparable statistical approach was used for the dichotomous outcome variables, replacing the analysis with a generalized linear model using a binomial distribution, a logit link function (reference coded 0), and the Satterthwaite approximation to settle a more appropriate degrees of freedom. For that reason, we provide 99% confidence bands for the estimated means.

Standardized Hedge's $g$ effect-sizes for continuous data were calculated as a ratio of the estimated mean difference (based on the mixed model) to the observed pooled SDs. Values around .2, .5 and .8 were interpreted as weak, medium and strong, respectively (Hedges & Olkin, 1985). Following the intention to analyze results according to a noninferiority trial, only effect-sizes of at least 0.45 in differences between groups were considered clinical significant (Mathisen et al., 2017).

3  |  RESULTS

Of the 418 women who responded during the recruitment period (March 2014–August 2016), 229 did not fulfill the inclusion criteria,
164 were randomized to therapy groups, and finally 76 and 73 initiated therapy in PED-t and CBT, respectively (Figure 1). Reasons for being excluded were either general reasons (i.e., sex, age, BMI, physical injuries, living distance to treatment facility, \( n = 54 \)), diagnostic reasons (i.e., no ED-diagnosis or having AN, \( n = 54 \)), having severe comorbidity (i.e., suicidality, posttraumatic stress disorder, obsessive–compulsive disorder, bipolarity, personality disorder, referred to inpatient treatment, \( n = 37 \)), other self-chosen reasons (i.e., did not find the time, did not want to attend anyway, \( n = 60 \)), or other competing treatment (i.e., recently received CBT or preferred other concurrently offered therapy, \( n = 24 \)).

The mean percentage (SD) attendance rate to therapy was 80.6% (11.4) in PED-t, and 82.1% (45.7) in CBT. In PED-t the compliance rate to exercise sessions, as supervised in therapy + unsupervised at home (being evaluated from attendance to therapy and evaluated from training diaries) was 69.8% and 56.7% for resistance- and interval exercises, respectively. There were significant differences between diagnoses at baseline, with higher body weight and BMI among those

**FIGURE 1** Recruitment, retention, and attrition during test periods. Overview of recruitment, screening, randomization and attendance to the baseline (T1), posttreatment (T2), and follow-up measures (T3–T5). Note: PED-t, physical exercise and dietary therapy; CBT, cognitive behavior therapy; BN, bulimia nervosa; BED, binge-eating disorder; EDE-Q, eating disorder examination questionnaire; LFU, lost to follow-up; Fairburn and Beglin (2008); Sheehan et al. (1998)
with BED, and higher EDE-Q global score and number of purging episodes among those with BN (Table 1).

The numbers of dropout from treatment were not significantly different across groups or ED diagnoses. Among those who dropped out at posttest (n = 37) a higher mean (CI95) baseline score for depression was found compared to completers (n = 112), (3.36 ± 1.9, 1.43, g = .44, p = .02). Significantly more CBT-participants were lost to follow-up (LFU) at 6-months follow-up (PED-t = 19 (24.4%), CBT = 37 (47.4%), p = .005), 12-months follow-up (PED-t = 20 (25.6%), CBT = 39 (50.0%), p = .003), and 24-months follow-up (PED-t = 20 (25.6%), CBT = 25 (38.5%), p = .03). Overall, there was no difference in LFU across the BN and BED diagnoses (p = .02); however, more with BN in CBT were LFU at all follow-ups compared to BN in PED-t (p = .03). There were no differences (p > .06) between attending participants at 6-months follow-up (n = 100, 64.1%) and LFUs (n = 56, 35.9%) with respect to the scoring at posttest, as was the case at 12-months follow-up with respect to 6-months follow-up scores (attendees; n = 97, 62.2%; LFU; n = 59, 37.8% (p > .05), and at 24-months follow-up with respect to 12-months follow-up scores (attendees; n = 106, 61.6%; LFU: n = 43, 25%) (p > .05).

### 3.1 Effects of PED-t and CBT on ED symptoms (EDE-Q)

#### 3.1.1 Change in EDE-Q symptom scores from baseline to post and follow-up tests

Compared to baseline, both therapy groups improved in EDE-Q global score at posttreatment and at all follow-ups (PED-t: T2, g = 1.36, p < .001; T3, g = .86, p < .001; T4, g = 1.17, p < .001; T5, g = 1.21, p < .001, and CBT: T2, g = .93, p < .001; T3, g = .63, p < .001; T4, g = 1.21, p < .001; T5, g = .99, p < .001). The same relate to EDE-Q subscales weight concern (PED-t: T2, g = 1.02, p < .001; T3, g = .61, p < .001; T4, g = .91, p < .001; T5, g = .87, p < .001, and CBT: T2, g = .60, p < .001; T3, g = .41, p < .004; T4, g = .82, p < .001; T5, g = .64, p < .001), shape concern (PED-t: T2, g = 1.04, p < .001; T3, g = .70, p < .001; T4, g = .99, p < .001; T5, g = 1.03, p < .001, and CBT: T2, g = .50, p < .001; T3, g = .74, p < .001; T4, g = .82, p < .001; T5, g = .84, p < .001), eating restraint (PED-t: T2, g = 1.24, p < .001; T3, g = .80, p < .001; T4, g = .92, p < .001; T5, g = .90, p < .001, and CBT: T2, g = .87, p < .001; T3, g = .72, p < .001; T4, g = 1.24, p < .001; T5, g = .93, p < .001), and eating concern (PED-t: T2, g = 1.28, p < .001; T3, g = .97, p < .001; T4, g = 1.28, p < .001; T5, g = 1.51, p < .001, and CBT: T2, g = .91, p < .001; T3, g = .56, p < .001; T4, g = 1.25, p < .001; T5, g = 1.32, p < .001). Additionally, both therapy groups reported a significant alleviation of binge eating from baseline to posttreatment (p < .001 in both groups), and from baseline to each follow-up (PED-t T3 p < .001; T4 p < .001; T5 p = .003, and CBT T3 p < .001; T4 p = .007; T5 p < .001). Both therapy groups also reported reduction in episodes of self-induced vomiting from baseline to posttreatment (p < .001 in both groups), and from baseline to all follow-ups (PED-t T3 p < .001; T4 p < .005; T5 p < .001, and CBT T4 p = .005; T5 p < .001), except at T4 in PED-t (p = .02) and at T3 in CBT (p = .12). There were no significant changes in the number of episodes with laxative use (p > .02), but both groups reduced the number of episodes with driven exercise from baseline to posttreatment (p < .001 in both groups), and from baseline to all follow-ups (PED-t T3 p < .001; T4 p = .005; T5

### TABLE 1 Demographics of participants

<table>
<thead>
<tr>
<th></th>
<th>Bulimia nervosa (n = 117)</th>
<th>Binge-eating disorder (n = 55)</th>
<th>PED-t (n = 76)</th>
<th>CBT (n = 73)</th>
<th>Reference (n = 23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>27.2 (5.6)</td>
<td>29.0 (6.1)</td>
<td>28.2 (6.2)</td>
<td>27.7 (5.3)</td>
<td>26.5 (5.6)</td>
</tr>
<tr>
<td>BW, kg</td>
<td>66.0 (12.2)</td>
<td>81.9 (15.1)*</td>
<td>71.7 (16.0)</td>
<td>71.4 (14.4)</td>
<td>67.2 (14.1)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>23.5 (3.8)</td>
<td>28.9 (5.0)*</td>
<td>25.3 (5.1)</td>
<td>25.5 (4.7)</td>
<td>24.1 (4.9)</td>
</tr>
<tr>
<td>Illness duration, years</td>
<td>12.0 (6.6)</td>
<td>13.0 (8.2)</td>
<td>13.0 (7.5)</td>
<td>12.1 (6.7)</td>
<td>10.6 (7.4)</td>
</tr>
<tr>
<td>EDE-Q global score</td>
<td>3.79 (0.9)</td>
<td>3.49 (0.9)**</td>
<td>3.70 (0.8)</td>
<td>3.67 (1.0)</td>
<td>3.81 (1.0)</td>
</tr>
<tr>
<td>Binge-eating episodes*</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>9.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Self-induced vomiting*</td>
<td>5.0</td>
<td>0.0*</td>
<td>2.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Laxative use*</td>
<td>0.0</td>
<td>0.0*</td>
<td>5.0</td>
<td>4.8</td>
<td>12.8</td>
</tr>
<tr>
<td>Driven exercise*</td>
<td>10.0</td>
<td>10.0</td>
<td>5.0</td>
<td>6.0</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Note: Baseline demographic and clinical information presented per diagnosis, and per therapy- or reference group. Results are mean (SD), except binge eating, and purging episodes which are median (CI95).

Abbreviations: BED, binge-eating disorder; BMI, body mass index; BN, bulimia nervosa; BW, body weight; CBT, cognitive behavior therapy; EDE-Q global, eating disorder examination questionnaire; PED-t, physical exercise and dietary therapy.

*Episodes last 28 days.

*p < .001; **p = .04; ***p = .03; ****p = .014.
3.1.2 | Therapy effect differences in EDE-Q symptom scores (between analyses with baseline as covariate)

The fixed effect of Time was significant (F_{3,282} = 8.04, p < .001) indicating a general improvement across the follow-ups. However, the more important Group and Group * Time effects were not significant (F_{1,134} = 2.79, p = .097; F_{3,282} = 2.47, p = .062, respectively). Planned comparison tests nevertheless indicated an advantage of PED-t above CBT at the first posttest (T2) exclusively in EDE-Q global score (g = 0.54, p = .004) and the EDE-Q subscale "body weight concern" (g = 0.53, p = .003). Analyzing Time as a linear trend indicated a general decline across the follow-ups (b_{time} = -.177, p = .025), but it was not significantly different by Group (b_{time*group} = .146, p = .115).

Neither the fixed effect of Time, Group, nor Group * Time effects were significant for binge eating and purging episodes, and no planned between-group comparison test were significant either.

3.2 | Effects of PED-t and CBT on full- and partial remission

3.2.1 | Full remission

In both therapy groups, the number of participants in remission was significantly higher posttreatment (PED-t p < .001; CBT p = .001) and at follow-ups (T3-T5: PED-t p < .001; CBT p < .001) compared to baseline, with no differences between the groups (T2 p = .03; T3 p = .6; T4 p = 1.0; T5 p = .7) (Figure 2).

3.2.2 | Partial remission

The number in partial remission in the PED-t group was higher posttreatment (p < .001) and all follow-ups (T3 p = .006; T4 p = .01; T5 p = .005), compared to baseline. The changes in the CBT group approached, but were not significantly different posttreatment (p = .02), with no differences to baseline at follow-ups (T3 p = .7; T4 p = .04; T5 p = .2). There was no difference in numbers in partial remission between the two groups at any time (T2 p = .5; T3 p = .2; T4 p = .6; T5 p = .7) (Figure 2).

3.3 | Effects of PED-t and CBT on clinical impairment (CIA) from ED

3.3.1 | Change in CIA scores from baseline to post and follow-up tests

Compared to baseline, both therapy groups had reduced CIA score posttreatment (p < .001) and at follow-up tests (p < .001, Table 2). The effect sizes were strong for PED-t, and ranged between medium and strong for CBT.

3.3.2 | Therapy effect differences in CIA scores (between analyses with baseline as covariate)

The fixed effect Group (F_{1,118} = 6.86, p = .010) was significant. The Time (F_{2,230} = 3.20, p = .024) and Group * Time (F_{3,230} = 1.83, p = .143) effects were not significant. The planned comparison tests (between-effects column in Table 2) showed an advantage of PED-t above CBT posttreatment and at first follow-up (g = 0.55, p = .005) (T2 and T3 in Table 2), but not at later follow-ups (T4-T5). Analyzing time as a linear trend indicated a decline across follow-up, still not significant.

**FIGURE 2** Remission from diagnosis. Estimated mean percent (SE) in each group in full remission (left) and partial remission (right) at posttreatment (T2) and follow-ups (T3-T5). Note: PED-t, physical exercise and dietary therapy; CBT, cognitive behavior therapy; T2, posttreatment; T3, 6-month posttreatment; T4, 12-month posttreatment; T5, 24 months follow-up; a, significant change from baseline, p < .01; c, significant change from T2, p < .003
### TABLE 2  Satisfaction with Life Scale (SWLS), Beck Depression Inventory (BDI), and Clinical Impairment Assessment (CIA)

<table>
<thead>
<tr>
<th></th>
<th>T1 CI 99</th>
<th>T2-T3</th>
<th>T2-T4</th>
<th>T2-T5</th>
<th>Between effects, p-value, effect size (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWLS PED-t</td>
<td>16.50</td>
<td>4.60</td>
<td>2.92</td>
<td>3.81</td>
<td>5.04</td>
</tr>
<tr>
<td></td>
<td>14.37–18.63</td>
<td>2.73, 6.46</td>
<td>1.00, 4.84</td>
<td>1.92, 5.69</td>
<td>3.16, 6.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g = .69</td>
<td>g = .44</td>
<td>g = .64</td>
<td>g = .76</td>
</tr>
<tr>
<td>CBT PED-t</td>
<td>16.63</td>
<td>2.27</td>
<td>n.s.</td>
<td>2.53</td>
<td>3.98</td>
</tr>
<tr>
<td></td>
<td>14.46–18.80</td>
<td>0.26, 4.28</td>
<td>n.s.</td>
<td>0.32, 4.74</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g = .53</td>
<td></td>
<td>g = .60</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>16.17</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>12.31–20.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI PED-t</td>
<td>16.50</td>
<td>−6.73</td>
<td>−4.95</td>
<td>n.s.</td>
<td>−4.5</td>
</tr>
<tr>
<td></td>
<td>13.67–19.28</td>
<td>−9.93, −3.54</td>
<td>−8.20, −1.71</td>
<td>n.s.</td>
<td>−7.72, −1.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g = 0.93</td>
<td>g = 0.66</td>
<td></td>
<td>g = 0.47</td>
</tr>
<tr>
<td>CBT</td>
<td>14.56</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>11.70–17.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>18.52</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>13.45–23.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIA PED-t</td>
<td>28.22</td>
<td>−13.97</td>
<td>−12.63</td>
<td>−12.46</td>
<td>−16.06</td>
</tr>
<tr>
<td></td>
<td>24.87–31.57</td>
<td>−17.68, −10.26</td>
<td>−16.41, −8.86</td>
<td>−16.30, −8.61</td>
<td>−19.85, −12.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g = 1.18</td>
<td>g = 1.07</td>
<td>g = 1.01</td>
<td>g = 1.36</td>
</tr>
<tr>
<td>CBT</td>
<td>26.90</td>
<td>−7.05</td>
<td>−5.99</td>
<td>−11.71</td>
<td>−10.88</td>
</tr>
<tr>
<td></td>
<td>23.52–30.30</td>
<td>−10.99, −3.11</td>
<td>−10.20, −1.78</td>
<td>−16.17, −7.30</td>
<td>−15.25, −6.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g = 0.78</td>
<td>g = 0.56</td>
<td>g = 1.01</td>
<td>g = 1.08</td>
</tr>
<tr>
<td>Reference</td>
<td>29.99</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>23.91–36.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b_{\text{time}} = −1.36, p = .068), and neither did it significantly co-vary by group (b_{\text{time}\times\text{group}} = .63, p = .524).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 3.4 Effects of PED-t and CBT on depressive symptoms (BDI)

##### 3.4.1 Change in BDI scores from baseline to post and follow-up tests

Compared to baseline, participants in PED-t displayed reduced BDI scores posttreatment \(p < .001\) and at follow-ups \(T3 p < .003\) and \(T5 p < .001\), Table 2). The effect sizes ranged between medium and strong. There were no significant changes in BDI scores among participants in CBT. The estimated mean number of participants (CI 99%) with BDI score above clinical cutoff was reduced in PED-t at T2 with a marginal statistical significance (from 27.8% 16.6, 42.6 to 13.9% 5.9, 13.3% 3.4, \(p = .014\)), but with no further differences to baseline at follow-ups. There were no significant changes in number of participants with scores above cutoff in CBT (T1 = 24.8% 14.1, 39.3).

##### 3.4.2 Therapy effect differences in BDI scores (between analyses with baseline as covariate)

The fixed effect Group was significant \(F_{1,124} = 7.21, p = .008\), whereas Time \(F_{2,256} = 2.99, p < .031\) and Group * Time \(F_{2,256} = .71, p = .544\) was not. The planned comparison tests showed a tentative advantage from PED-t above CBT posttreatment \(p = .016\), but not in the remaining posttests. The linear trend analysis of Time was not significant \(b_{\text{time}} = −.831, p = .164\), and did not significantly co-vary with group \(b_{\text{time}\times\text{group}} = .159, p = .842\). We found no differences between groups by time in percentage of participants scoring above BDI clinical cutoff \(p > .11\).
3.5 | Effects of PED-t and CBT on subjective well-being (SWLS)

3.5.1 | Change in SWLS scores from baseline to post and follow-up tests

Compared to baseline, both therapy groups’ improved SWLS scores posttreatment ($p < .003$) and at follow-ups (T3-T5 $p < .01$, Table 2). The effect sizes were medium in both groups.

3.5.2 | Therapy effect differences in SWLS scores (between analyses with baseline as covariate)

The fixed effect of Group ($F_{1,137} = 2.50, p = .117$) and of Group * Time ($F_{3,276} = 1.03, p = .380$) was not significant, whereas the effect of Time approached significance ($F_{3,276} = 1.03, p < .022$). The trend analysis of Time was not significant ($b_{\text{time}} = .522, p = .138$; $b_{\text{time} \times \text{group}} = -.383, p = .418$).

4 | DISCUSSION

Overall, we found support to our hypothesis, as the PED-t performed equal to CBT in alleviating symptoms of BN, BED, as well as in improving subjective well-being and psychosocial impairment. In addition, PED-t reduced symptoms of depression. We also found support for the hypothesis that the PED-t produced a more rapid treatment response (i.e., the T1 to T2 change in EDE-Q global score, EDE-Q subscale “body weight concern”, CIA and a tentative difference in BDI), however, contradicting our hypothesis, was the lack of difference between groups in drop out from treatment.

The effect size of changes within groups was with medium to strong effects, hence, implying changes with important clinical impact. This also relates to the acute differences identified between groups posttreatment in EDE-Q global score, body weight concern and clinical impairment. Nevertheless, both therapies performed well in an outpatient group setting, with a summarized response rate (i.e., the total of participants that either reached full or partial remission) to therapy of ~50% after PED-t and ~30% after CBT. This corresponds well with previous findings on treatment outcome from group-CBT, that is, ~0–52% remission after group-CBT (Linardon, 2018; Linardon, Messer, & Fuller-Tyszkiwicz, 2018; Linardon & Wade, 2018).

Early symptom changes promote treatment success and predict a favorable long-term outcome (Graves et al., 2017). An important asset of CBT compared to other treatments of EDs, is that CBT produces such early changes (Agras, Walsh, Fairburn, Wilson, & Kraemer, 2000; Fairburn, Jones, Peveler, Hope, & O’Connor, 1993; Hay, 2013; Södersten et al., 2017; Wilson, Fairburn, Agras, Walsh, & Kraemer, 2002), thus leading to a shorter time of suffering from EDs. A similar asset may be transferred to the PED-t compared to CBT in terms of less impairment, and a remission from the BN and BED diagnoses. It is not clear why the previously described rapid effect from CBT was not found in our study, but one explanation could be attributed to the randomization procedure. Considering the study setting and how the study was profiled as a physical activity trial, the rapid effect may reflect the motivation, interest and preference for the new therapy (PED-t) (and not the CBT) among those who chose to participate. If so, the rapid treatment response may not be expected to occur in a non-RCT context.

Although alleviations of core ED symptoms is necessary for a treatment success, improvements in secondary outcomes is important to consolidate symptom improvement and to strengthen positive circuits created by a rapid symptom change. The participants’ subjective well-being was improved in both therapies, and about 60–70% of all participants rated their well-being equal to or above average during follow-up, which is close to the frequency of 83% in national cohort studies (Diener & Diener, 2009). Additionally, the positive effect from PED-t on depression scores is in line with other studies examining effects of physical activity in persons diagnosed with depression (Fox, 1999; Rosenbaum et al., 2014), and summarized evidence holds physical activity at least as effective as psychological or medical interventions (Stubbs et al., 2018). Excluding severe depression may account for low BDI-scores throughout, however findings of symptoms of depression in the current sample were comparable to previous reports in similar populations (Keski-Rahkonen & Mustelin, 2016).

Both the CBT and the PED-t shared a focus on enhancing self-regulative mechanisms and in particular with respect to eating as well as changing negative body and shape self-evaluations and how such evaluations used to be detrimental for self-worth. It is rather intriguing that these endpoints were successfully approached through very dissimilar means, that is, by verbal behaviors (CBT) and practical experiences of body functionality (PED-t). The dissimilarity makes future search for possible common mechanisms of change highly interesting. Neurobiological and cognitive factors may be considered, given that these factors have been associated with effects of physical activity (Lubans et al., 2016; Mullen & Hall, 2015) and that they constitute important core features of BN and BED (Donnelly et al., 2018).

Strengths to this study are the RCT-design with 24 months follow-up, the equal length and duration of the treatments, and the unlikely diffusion of effects from previous or concurrent treatments. Limitations in need to be mentioned are the low compliance rate to PED-t exercise module, the large number of CBT-participants lost to follow-ups as well as the failure to control for therapist factors, and for other kind of treatments that may have been sought in the follow-up period. Concerning the loss of CBT-participants, we argue the statistical model handles this most favorable, with full information maximum likelihood methods reducing any risk of biased data (Witkiewitz et al., 2014). In addition, generalizability is limited to comparable gender, BMI and age, and may also have been hampered by the fact that severe comorbidity was excluded. Still, the level of symptom distress do suggest comparability to the level of anxiety-, mood-, and personality disorders among persons with BN and BED (Friborg et al., 2014; Godart et al., 2007; Keski-Rahkonen & Mustelin, 2016; Martinussen et al., 2017; Swanson, Crow, Le Grange, Swendsen, & Merikangas, 2011). Finally, data on treatment fidelity are presently unavailable, and the recruitment information for this trial (i.e., offering...
therapy with physical activity, and a limitation of BMI range) may have resulted in a selected sample of patients.

4.1 Conclusion and implications

The effects of PED-t are overall on par with the well-established CBT for BN and BED. Its focus on physical exercise and diet may appeal to many sufferers. In addition, PED-t may represent an alternative for those who do not want CBT, or in cases where CBT is unavailable. This can be accommodated given the high availability of professionals with expertise in exercise medicine and nutrition compared to CBT-certified therapists. Following the logic behind the construction of the PED-t it may be worthwhile to test the incremental effects of combining PED-t and CBT. Replications of the present findings are needed with both independent research groups and within new and preferably more clinical contexts. Equally needed is the exploring of mechanisms of change.

ACKNOWLEDGMENTS

The project was funded by the Norwegian Women’s Public Health Association. Operating funds were supported partly by the Norwegian School of Sport Sciences, the UiT—The Arctic University of Norway, and by an assigned fund from the Research Council of Norway. Dairy sports recovery products have been supported for the PED-t intervention by Tine Dairy Producer (Norway) as part of their charity funding activity. We are very thankful to our master students in exercise physiology who provided the lab-, exercise, and test-services; to head of office at the department of Sport Medicine Solveig Sunde for taking care of concealed randomization; and to Guro Pettersen and Ole Johan Friborg for punching all data used in this study.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

There is no ORCID information available for this document.

REFERENCES


Twisk, J., de Boer, M., de Vente, W., & Heymans, M. (2013). Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. Journal of Clinical Epidemiology, 66(9), 1022–1028. https://doi.org/10.1016/j.jclinepi.2013.03.017


