TITLE: Taking the LEAP: a randomized, multicentre, naturalistic, efficacy trial of the compuLsive Exercise Activity theraPy (LEAP) for patients with eating disorders

STATISTICAL ANALYSIS PLAN

Study title	**Taking the LEAP: a randomized, multicenter, naturalistic, efficacy trial of the compuLsive Exercise Activity theraPy (LEAP) for patients with eating disorders**
Short title	LEAP
Principal Investigator	Emma Forsén Mantilla (EFM; PhD, Clin. Psych) Emma.forsen@ki.se Emma.forsen.mantilla@gih.se Tfn: 073-654 32 54
Trial registration and study protocol	The trial is registered with the ISRCTN registry (registration date 2020-03-25), trial ID: ISRCTN80711391. Study protocol: Monell E, Meyer C, Szwajda A, Meyer & Forsén Mantilla E. (2021). Taking the LEAP: study protocol for a randomized, multicentre, naturalistic, efficacy trial of the compuLsive Exercise Activity theraPy- a cognitive behavioral program specifically targeting compulsive exercise in patients with eating disorders. *BMC Psychiatry*, 2021; 21 (369); https://doi.org/10.1186/s12888-021-03356-2
SAP version	1
Date	260130
Written by	EFM, Elin Monell (EM: project leader/coordinator, elin.monell@ki.se) & Agnieszka Szwajda (AS: main analyst, agnieszka.szwajda@ki.se)

CONTENT

1 BACKGROUND 4
2 OBJECTIVES 4

2.1 Primary objective 4
2.2 Secondary objectives 5
2.3 Exploratory objectives 5
3 STUDY METHODS 5

3.1 Trial design 5
3.2 Randomization 6
3.3 Sample size 6
3.4 Framework 6
3.5 Statistical interim analysis and stopping guidance 6
3.6 Timing of final analysis 6
4 STATISTICAL PRINCIPLES 6

4.1 Confidence intervals and significance levels 6
4.2 Intervention Adherence and Protocol deviations 7
4.3 Analysis population 7
4.3.1 Modified Intention-to-Treat (mITT) Analysis Set 7
4.3.2 Per-protocol (PP) analysis set 7
5 TRIAL POPULATION 7

5.1 Screening data 7
5.2 Eligibility 7
5.3 Withdrawal / Follow-up 8
5.4 Baseline patient characteristics 8
6 ANALYSIS 8

6.1 Outcome definitions 8
6.1.1 Primary outcome 8
6.1.2 Secondary outcomes 8
6.1.3 Exploratory outcomes 9

6.2 Analysis methods 9
6.2.1 Primary analysis 9
6.2.2 Per-protocol sensitivity analysis 10
6.2.3 Clustered intervention delivery 10
6.2.4 Diagnosis-adjusted analyses 10
6.2.5 Treatment-by-diagnosis interaction 10
6.2.6 SEDI remission status 10
6.2.7 Additional sensitivity analysis 10
6.2.8 Effect estimates 11

6.3 Exploratory analyses 11
6.3.1 General psychopathology and physical activity 11
6.3.2 Baseline predictors of outcome 11
6.3.3 Treatment exposure and implementation 11
6.3.4 Patient satisfaction 11

6.4 Missing data 11
6.5 Harms 12
6.6 Statistical software 12
7 APPENDIX: SHELL TABLES 13
8 REFERENCES 15

1 Background

Compulsive exercise (CE) is an important factor in ED etiology, development, and maintenance (1-3). In adult patients, increased levels of CE are related to more severe psychopathology, increased suicidality, longer hospitalizations, increased relapse risk, and a more chronic course in anorexia nervosa (AN) (2, 4-6). Patients who continue or initiate CE during treatment have 1-year remission rates at roughly half of that among patients without CE or patients who discontinue CE during treatment (3). Further, patients experience CE as a particularly complicated symptom to overcome and feel uncertain about what constitutes healthy, adaptive physical activity (7, 8). Further, patients experience CE as a particularly complicated symptom to overcome and feel uncertain about what constitutes healthy, adaptive physical activity (7, 8). In our nationwide study, CE was reported by ca 50% of all ED patients (3). Previous studies have found CE prevalence to be highest in AN (2), but in our sample, patients with bulimia nervosa (BN) and patients with other specified feeding and eating disorders (OSFEDs) reported the highest prevalence of CE (3). More importantly, CE was prevalent in all ED diagnoses and almost as common in both genders. Of note, a 2011 nationwide study on Swedish ED patients (9) found lower overall CE rates, but higher rates of binge-eating and purging, than we did. This suggests that CE is becoming more prevalent as a symptom, and/or that there might be gradual symptom-shifting among ED patients over the years. To date, there is no "gold standard" treatment to tackle CE. However, ours and others' results indicate a clear need to systematically address CE in ED treatment. The empirically based, targeted treatment program LEAP (10), aims to educate patients about CE and how it is maintained, challenge maladaptive beliefs and behaviors, and equip patients with strategies promoting healthy exercise. LEAP is founded on the same theoretical principals as the evidence-based ED treatment: Cognitive Behavioral Therapy Enhanced (CBT-E) (11) and may be offered as a complement to CBT-E.

LEAP is delivered in 1 initial individual session and 8 group sessions. LEAP has been evaluated in two pilot studies and one RCT. In these studies focusing adult patients with AN, LEAP reduced pathological exercise, ED psychopathology, general psychopathology, and length of hospitalization; it also improved attitudes and beliefs toward exercise (which in turn has proven important for outcome) (12), BMI, and quality of life (13-15). A systematic review/meta-analysis also suggests treatments combining different interventions (e.g. psychoeducation, psychotherapy and CBT-techniques) like LEAP, are most effective in treating CE (16). However, the positive preliminary results for LEAP need confirmation at other sites and with other ED diagnostic groups. Of note, the patients in the RCT had suffered from AN (known as an illness notoriously difficult to treat) for an average of 6 years, a considerably long time, yet made important progress. The effects of LEAP with other ED diagnostic groups and patients with shorter illness duration may therefore be even greater. To our knowledge, there are no existing systematic efforts targeting CE in Swedish specialist care.

The aim of LEAP is to equip patients with knowledge and skills to challenge maladaptive beliefs and behaviors, regain control of exercise behavior, and be able to participate in exercise appropriate to age, goals, and health-status. In doing so, the anticipated result is a reduction in ED psychopathology in general, and in CE-related behaviors and cognitions, in particular. Given the high prevalence of CE among patients with EDs, an effective targeted intervention could have relevance for a substantial proportion of patients receiving specialist ED care. As several studies have demonstrated a worse prognosis for ED patients with CE, treating this symptom will possibly improve prognosis. Improving recovery rates is of course beneficial for patients, but also for society (in terms of reducing costs of long-term treatments and hospitalizations). Although the need for treating CE is evident, the present effort of systematically attempting to do so, is unique.

2 Objectives

2.1 Primary objective

The primary objective of this study is to evaluate the effectiveness of the LEAP intervention, delivered in addition to treatment as usual (TAU), in reducing compulsive exercise and ED psychopathology among adult patients with anorexia nervosa (AN), bulimia nervosa (BN), or other specified feeding or eating disorder (OSFED) who report compulsive exercise.
The two primary outcomes are:

1. Total score on the Compulsive Exercise Test (CET); and
2. Global score on the Eating Disorder Examination Questionnaire (EDE-Q).

The primary endpoint is the assessment at T2, approximately three months after the initial assessment. T2 was selected as the primary endpoint because it was expected to occur closest to completion of the LEAP intervention for the majority of participants.
The primary effectiveness analysis will follow the modified intention-to-treat (mITT) principle and will compare participants with complete data randomized to TAU+LEAP with those randomized to TAU. Per-protocol analyses, defined according to attendance at a minimum of six of the eight LEAP group sessions (approx. 75%), will also be conducted for the primary and relevant secondary objectives.

2.2 Secondary objectives

Secondary objectives are to:

- examine whether treatment effects on compulsive exercise and ED psychopathology are maintained at T3, approximately six months after the initial assessment;
- evaluate the effects of LEAP on specific dimensions of compulsive exercise and ED psychopathology, including CET and EDE-Q subscales;
- evaluate the effects of LEAP on other clinically relevant outcomes, including compulsive exercise frequency, emotion regulation and ED remission status where available;
- assess the adherence of LEAP among participants allocated to TAU+LEAP.
- examine whether adjustment for baseline ED diagnosis alters the estimated treatment effects; and
- investigate whether treatment effects differ according to baseline ED diagnosis by testing treatment-by-diagnosis interactions.

2.3 Exploratory objectives

Exploratory objectives are to:

- evaluate the effects of LEAP on general psychopathology and physical activity patterns;
- investigate whether baseline patient characteristics, such as BMI, emotion regulation, compulsivity, and eating-disorder diagnosis, are associated with treatment outcomes among participants allocated to TAU+LEAP;
- investigate patient satisfaction and treatment perception
- additional exploratory analyses will examine the relationship between treatment exposure, including LEAP session attendance, LEAP therapist experience and treatment outcomes.

The main trial publication will focus on the primary and secondary objectives only. Exploratory objectives will be analyzed in subsequent publications and hence the exploratory analyses are not outlined with the same level of details as the analysis of the primary and secondary objectives.

3 Study Methods

3.1 Trial design

The trial is a naturalistic two-armed parallel open-label superiority randomized effectiveness trial, with recruitment from five Swedish specialist ED clinics. Eligible adult patients were recruited via the participating clinics and randomized to treatment as usual (TAU) or TAU+LEAP, where LEAP ran in parallel with TAU. Key outcome variables in both arms were measured at initial assessment (T1), after 3 months (T2), and 6 months after initial assessment (T3). Although most TAU+LEAP

participants were expected to complete the LEAP program prior to T2 assessment, there was a possibility that some would terminate LEAP prior to T3 instead.

Please see also the SPIRIT figure in the study protocol (Monell, et al., 2021).

3.2 Randomization

Each eligible patient consenting to participate was assigned experimental condition using an online random number generator, stratified by site. An allocating investigator (AI) performed the randomization procedure and informed LEAP-therapists and participants. The AI has exclusive access to randomization group data. Randomization ratio was initially 1:1, but in December 2022 it was changed to 1:2 (in favor of the intervention group). The reason being that patients randomized to receive LEAP were kept waiting until the group at their unit had a minimum of 4 included patients. In some cases, this wait resulted in them no longer being eligible (e.g., as their TAU had ended) or able to participate due to changed personal circumstances. Thus, the weighted randomization was needed to speed up recruitment to the LEAP groups, with clear benefits for participants.

Blinding: Participants and clinicians could not be blind to group, since LEAP deviates too much from standard ED treatment. However, the PI (EFM) will not have access to randomization group data until after the main treatment effects have been analyzed. The main analyst (AS) will receive pseudonymized data after locking the raw data material. Most assessments were self-reported, but clinicians conducting diagnostic pre-assessments were blinded. The LEAP therapists did not conduct assessments, neither did the PI nor the project coordinator (PC).

3.3 Sample size

To achieve 80% power 98 participants (49 per group under the original 1:1 allocation) were sufficient to detect the hypothesized treatment effect, under the assumptions of a two-sided alpha = .05, effect size f = .25 and an estimated correlation between pre- and post-measures of .50. Assuming an approximate attrition rate of 30%, we aimed at recruiting at least a total of 128 participants. Because the allocation ratio was changed from 1:1 to 1:2 in favour of TAU+LEAP during recruitment, the final number allocated to each treatment group differs from the originally planned equal allocation. No formal adjustment of the target sample size was made following the change in allocation ratio.

3.4 Framework

This is a superiority trial, with the aim of showing that the addition of LEAP to TAU will improve patient symptoms, compared to the control group.

3.5 Statistical interim analysis and stopping guidance

No interim analysis was planned nor performed. The trial stopped recruitment early spring of 2025. At that point 144 participants had been randomized, which was slightly higher than our original target and this was due to study technical circumstances. T3 follow up measurements and interviews with the final group of participants took place in the fall (August/September) of 2025.

3.6 Timing of final analysis

After all participants have finished their 6 months follow-up assessment and data has been curated (i.e. during fall/winter of 2026).

4 STATISTICAL PRINCIPLES

4.1 Confidence intervals and significance levels

All inferential tests will be two-sided. A significance threshold of $p < 0.05$ will denote statistical significance. Precision of estimates will be communicated via 95% confidence intervals (CIs). Standardized effect sizes (e.g., Cohen's d) will be calculated and reported alongside point estimates. See shell table in appendix.

4.2 Intervention adherence and protocol deviations

Intervention adherence will be presented descriptively as the mean percentage of scheduled LEAP sessions attended (where 9 sessions = 100%), aggregated overall and stratified by clinical site. A CONSORT Flow Diagram will map participant tracking from initial screening, group allocation, post-intervention completion (T2), through to final follow-up (T3), detailing specific reasons for attrition.

4.3 Analysis population

4.3.1 Modified Intention-to-Treat (mITT) Analysis Set

The primary analyses will utilize the modified intention-to-treat (mITT) analysis set. It will include all randomized participants, analysed within the group to which they were randomized, regardless of their actual compliance, protocol deviations, or intervention attendance. Since no outcome imputation will be performed, for each outcome and assessment timepoint, only participants with an observed value for that outcome at the relevant assessment will contribute to the corresponding regression analysis. The exact N's relevant for each analysis will be reported.

4.3.2 Per-protocol (PP) analysis set

The Per-Protocol (PP) analysis set will serve as a secondary sensitivity analysis to evaluate treatment effectiveness under optimal protocol adherence. It will include participants randomized to TAU+LEAP who attended at least six of the eight LEAP group sessions before the relevant outcome assessment timepoint. The TAU participants will be retained in the per-protocol comparison if they completed the assessments. All participants in this set will have outcome data available at both baseline and post intervention.

5 TRIAL POPULATION

5.1 Screening data

Baseline screening characteristics (sex, age, diagnosis, BMI) will be reported on group level (see shell table in appendix). Exclusion: Since patients were only ever informed about the study if their clinicians judged them to fulfil the inclusion criteria, no patients reaching the research team were excluded. However, patients could of course opt out after having received the more thorough study information. Available characteristics of individuals opting out after having been forwarded to the research team will also be presented. Please see also SPIRIT figure in study protocol (Monell, et al., 2021).

5.2 Eligibility

Inclusion criteria: adult (≥18 years) male and female patients with AN, BN or OSFED (1. Atypical AN, 2. BN with low frequency and/or limited duration of symptoms, 4. Purging disorder/compensatory behaviors after small food amounts), reporting CE at initial assessment, T1 (e.g., measured by EDE-Q version 6, item 18, see Instruments), and receiving outpatient care or daycare at one of the specialized ED treatment units. As more females than males seek treatment, there will be more females than males in the final sample.

Exclusion criteria: inability to communicate in Swedish, being hospitalized, psychotic disorder, and/or BMI < 14.

5.3 Withdrawal / Follow-up

We will complete a CONSORT flow diagram to show the number of participants recruited, the number of (and reasons for) withdrawals, and the number of participants completing the follow-ups.

5.4 Baseline patient characteristics

Shell table 1 (see appendix) is showing the data that will be reported in the primary RCT article. A similar one (covering the main variables at all three time-points) will be presented in the manuscript.

Presence and severity of potential general psychopathology as measured with the DSM-5 cross-cutting-symptom measure will be described in subsequent papers, as well as basic information regarding everyday movement pattern as measured by the Active Q.

6 ANALYSIS

6.1 Outcome definitions

6.1.1 Primary outcome

The trial has two primary outcomes:

1. **Compulsive Exercise Test (CET) total score**
2. **Eating Disorder Examination Questionnaire (EDE-Q) Global score**

Both primary outcomes will be assessed at T2, approximately three months after initial assessment. T2 was selected as the primary endpoint because it was expected to represent the assessment closest to completion of LEAP for the majority of participants. The two primary outcomes will be analysed separately. The primary effectiveness criterion will be met if at least one of the two primary outcomes demonstrates a statistically significant treatment effect at T2.

The CET is a 24-item questionnaire measuring behavioral, cognitive, and emotional aspects of compulsive exercise (17). Items are scored from 0 to 5, with higher scores indicating greater symptom severity. The total CET score (sum of all subscale mean scores) will be used as a primary outcome and the study objective is to estimate the adjusted group difference in CET total at T2. The five CET subscales (Avoidance and Rule-driven Behavior, Weight Control Exercise, Mood Improvement, Lack of Exercise Enjoyment, and Exercise Rigidity; mean of constituent items) will be considered secondary outcomes.

The Eating Disorder Examination Questionnaire (EDE-Q) is a 28-item questionnaire scored 0–6 (no days-all days) measuring severity of self-reported ED pathology over the past 28 days (18). It generates an overall Global Scale and the subscales Restraint, Eating Concerns, Weight Concerns and Shape Concern. The EDE-Q Global score (the mean of the subscale means) will be used as a primary outcome and the study objective is to estimate the adjusted group difference in EDE-Q Global score at T2. The four EDE-Q subscales will be considered secondary outcomes.

EDE-Q also measures prevalence of ED behaviors such as restriction, binge-eating, purging and CE. This study will use information of presence and frequency of CE (count variable) and the ED psychopathology scales mentioned above as outcomes (continuous variables). Potential outliers in CE frequency defined as values above 4 standard deviations away from the pooled baseline mean, will be replaced with the highest value within the threshold.

6.1.2 Secondary outcomes

Secondary outcomes include:

- T3 measurements of CET total and EDE-Q Global score will be considered secondary outcomes and will be used to examine whether treatment effects observed at T2 are maintained at approximately six months after initial assessment.

- CET subscale scores;
- EDE-Q subscale scores;
- frequency/presence of compulsive exercise as measured with EDE-Q;
- DERS-16 total score;
- SEDI (ED diagnostic/remission status);
- Adherence to LEAP.

The Structured Eating Disorder Interview (SEDI) is a semi-structured interview assessing ED diagnostic criteria according to DSM-5 (19). It is performed at T3; remission status (i.e., no remission/remission) will be used as a secondary outcome (binary data). Initial ED diagnosis is assessed by clinicians at the units (diagnostic interviews used at the units are the SEDI and the EDE). BMI (weight in kg / [height in m]2, central for an AN diagnosis) is obtained at T1 and T3. At initial registration, some patients are weighed at units (mainly patients with AN), for others height and weight are self-reported. At T3, BMI is self-reported through the SEDI interview.

The Difficulties in Emotion Regulation Scale (DERS) short-form (DERS-16) is a 16-item questionnaire rated 1–5 (never-always) measuring emotion dysregulation (20). The Total score (sum of all individual items) is used as a continuous outcome in analyses.

6.1.3 Exploratory outcomes

Exploratory outcomes include:

- DSM-5 Cross-Cutting Symptom Measure domains;
- Physical activity variables from the Active-Q;
- LEAP patient satisfaction questionnaire
- Identifying predictors of treatment response (e.g. BMI, DERS, diagnosis)

The DSM-5 Self-Rated Level 1 Cross-Cutting Symptom Measure (DSM-5 symptoms) measures presence and potential severity of general psychopathology in 13 symptom domains. Descriptive data for all domains will be presented and relevant domains will be used in future exploratory analyses (binary and/or categorical data).

Active Q measures different types of physical activity, frequency and duration (MET hours; continuous data). Descriptive data may be presented and relevant variables (for instance hours of total physical activity) will be explored in future analyses.

A patient satisfaction and treatment perception questionnaire was specifically designed for this study, measuring participants satisfaction with structure and content of LEAP (intervention group only; continuous data).

6.2 Analysis methods

6.2.1 Primary analysis

The primary analysis will follow the modified intention-to-treat principle and will evaluate the two primary outcomes, CET total score and EDE-Q Global score, separately at T2. For each primary outcome, an individual-level linear regression model will be fitted with the outcome at T2 as the dependent variable using a complete case approach. Treatment allocation will be the main predictor, and the corresponding baseline outcome score at T1 and clinic will be included as covariates. Clinic will be included as a categorical fixed effect because randomization was stratified by clinic.
The primary models will therefore be:

- CET T2 (continuos) = treatment allocation (factor) + CET T1 (continuos) + clinic (factor)
- EDE-Q T2 (continuous) = treatment allocation (factor) + EDE-Q T1 (continuous) + clinic (factor)

The coefficient for treatment allocation will represent the adjusted between-group difference in the outcome at T2. T3 will be analysed as a secondary outcome and as a separate assessment point in the same way to evaluate the maintenance of treatment effects. The CET and EDE-Q subscales, the DERS and CE frequency will be analysed using analogous models at both T2 and T3.

6.2.2 Per-protocol sensitivity analysis

The primary mITT analyses will be supplemented by per-protocol sensitivity analyses. The same regression models will be fitted using the per-protocol population, defined as participants allocated to TAU+LEAP who attended at least six of the nine LEAP sessions, together with the corresponding TAU participants. The mITT and per-protocol analyses will be conducted separately at T2 and T3.

6.2.3 Clustered intervention delivery

LEAP was delivered in small treatment groups across 5 sites, whereas TAU was delivered individually. Participants receiving LEAP within the same treatment group may therefore have correlated outcomes because they received the intervention together. The regression models will adjust for site, the randomization factor, and will account for potential within-LEAP-group dependence using a cluster-robust variance estimator with a small-sample correction (CR2/Bell–McCaffrey). LEAP treatment groups will be used as the clustering variable with TAU participants treated as singleton clusters. Satterthwaite-adjusted degrees of freedom will be used for hypothesis testing and confidence intervals.

6.2.4 Diagnosis-adjusted analyses

The primary outcome models will be extended to include baseline ED diagnosis as an additional covariate. These analyses will examine whether adjustment for diagnostic group changes the estimated treatment effect. The diagnosis-adjusted models will include treatment allocation, the corresponding baseline outcome, clinic, and baseline ED diagnosis.

6.2.5 Treatment-by-diagnosis interaction

Treatment-by-diagnosis interaction analyses will examine whether the treatment effect differs according to baseline ED diagnosis. The models will include treatment allocation, baseline outcome, clinic, diagnosis, and the treatment-by-diagnosis interaction term. Where appropriate, treatment effects within individual diagnostic groups will be estimated and presented with 95% confidence intervals.

6.2.6 SEDI remission status

Patients' diagnostic outcomes will be classified as either 'improved' or 'not improved' in relation to baseline based on the SEDI assessment post-intervention. The proportion of patients classified as 'improved' will be analysed using a binary logistic regression model with a logit link function. The model will include treatment group as the primary predictor, baseline diagnosis and site as categorical covariates.

6.2.7 Additional sensitivity analysis

Additional sensitivity analyses will be performed to address potential sources of bias such as missing values and temporal effects.

- Missing post-randomization (T2/T3) outcome data and/or baseline T1 data resulting in exclusion from the primary complete-case analysis;

To examine if patients excluded from the primary analysis due to missing post-treatment outcomes were missing at random and did not bias the effect estimates, we will descriptively compare their

baseline characteristics and baseline scores against the analyzed group (using t-tests/Wilcoxon tests for continuous variables, and Chi-squared/Fisher's exact tests for categorical variables).

- Temporal effects due to the change in allocation ratio and varying waiting time between initial assessment and treatment start (first LEAP group session).

To examine if the change in the allocation ratio affected the observed treatment effect, we will run an additional sensitivity model adjusting the primary analysis with a categorical variable representing the allocation phase (pre- vs. post-change in allocation ratio). To examine if the waiting time between the initial assessment and the first LEAP group session affected treatment outcomes, an additional analysis will be conducted using the TAU+LEAP group data only. The number of days elapsed between these two events will be correlated (using Pearson or Spearman coefficients) directly with individual raw change scores (T2-T1) for the primary outcomes.

6.2.8 Effect estimates

For continuous outcomes, adjusted between-group mean differences will be reported together with 95% confidence intervals and p-values. Standardized effect sizes will additionally be reported where appropriate. For binary outcomes, adjusted odds ratios with 95% confidence intervals will be reported using logistic regression.

6.3 Exploratory analyses

Several exploratory analyses will be conducted to investigate factors that may be associated with variation in treatment outcome and patient satisfaction.

6.3.1 General psychopathology and physical activity

Descriptive data for all domains from the DSM-5 cross-cutting scale will be presented and relevant domains will be used in future exploratory analyses (binary and/or categorical data). Descriptive data on Active-Q will be presented and relevant variables (for instance hours of total physical activity) will be explored in future analyses.

6.3.2 Baseline predictors of outcome

Exploratory analyses will investigate whether baseline patient characteristics are associated with outcome among participants allocated to TAU+LEAP. Potential predictors include baseline BMI, emotion regulation difficulties, compulsivity, ED diagnosis, and other theoretically relevant baseline characteristics. Continuous outcomes will be analysed using linear regression and binary outcomes using logistic regression. Predictor variables will be measured at baseline and will not include post-randomization treatment characteristics.

6.3.3 Treatment exposure and implementation

The relationship between treatment exposure and outcome may be explored within the TAU+LEAP group. Potential variables include therapist experience (early vs later groups) and the number of LEAP sessions attended.

6.3.4 Patient satisfaction

The patient satisfaction and treatment perception scale is only completed by the TAU+LEAP group. Descriptive statistics will be presented and patterns will be explored in future analyses.

6.4 Missing data

Item-level missing data are not expected for the CET or EDE-Q because the electronic questionnaire system requires all items to be completed before submission. Consequently, no item-level imputation will be performed. Missing outcome data may nevertheless occur at the assessment level

if a participant does not complete a given follow-up assessment. The extent and patterns of missing outcome data at T2 and T3 will be reported separately by randomized treatment group, including the number and proportion of participants with available outcome data at each assessment. The primary mITT analyses will include all participants with available outcome data at the relevant assessment. No imputation of missing T2 or T3 outcome values will be performed, but they will be subjected to sensitivity analysis (see above). Reasons for missing assessments, withdrawal, and loss to follow-up will be summarized where available.

6.5 Harms

No harm was expected through participation. However, LEAP participants sacrifice time and energy to take part in a treatment with uncertain benefits while individuals randomized to control group may be disappointed that they did not receive LEAP. Further, although self-rating symptoms is unlikely to be harmful, it may raise concerns. Therefore, in all stages of data collection (in visible materials, in BASS, etc.), contact details to the PI were provided, encouraging participants to contact her in case of questions or concerns. Both PI and PC are clinical psychologists and clinicians and/or participants were encouraged to contact any of them for guidance if needed. As LEAP is delivered in groups, sensitive information was potentially shared, why the therapists made sure that in the first session of every new LEAP group all participants agreed upon a code of conduct. All LEAP therapists were experienced clinicians at the units where participants had active treatment. Thus, if adverse events arose, participants had access to care and could be monitored. LEAP therapists were also instructed to register any issues related to the LEAP intervention groups, including potential adverse events and to report such to the PI (to take appropriate actions). However, no such issues were reported throughout the trial.

6.6 Statistical software

Analysis will be performed using the programming language R or IBM Statistics (SPSS).

7 Appendix: Shell tables

Shell Table 1

Shell Table 1. Baseline characteristics

Baseline characteristics	Control-group	LEAP-group (Intention to treat)	LEAP-group (Per protocol)
Sex (F/M)			
ED diagnosis (%/type)			
	M (SD)	*M (SD)*	*M (SD)*
Age			
BMI			
EDE-Q Total scale			
EDE-Q Restraint			
EDE-Q Eating concerns			
EDE-Q Shape concerns			
EDE-Q Weight concerns			
Frequency of CE			
CET Total scale			
Avoidance and rule-driven behavior			
Weight control exercise			
Mood improvement			
Lack of exercise enjoyment			
Exercise rigidity			
DERS Total Scale			

Shell Table 2

Shell table 2. Results of primary analysis

Linear regression analysis at T2 presenting values for the allocation variable related to each outcome	Outcome	B	SE; CL (low-high)	P	N	Adjusted SE and; P	Cohen's D
Intention to treat (ITT)	**CET total**						
	Avoidance and rule driven behavior						
	Weight control						
	Mood improvement						
	Lack of enjoyment						
	Exercise rigidity						
	EDE-Q Global score						
	Restraint						
	Weight concerns						
	Eating concerns						
	Shape concerns						
	EDE-Q CE frequency						
	DERS Total						
Per protocol (PP)	**CET total**						
	Avoidance and rule driven behavior						
	Weight control						
	Mood improvement						
	Lack of enjoyment						
	Exercise rigidity						
	EDE-Q Global score						
	Restraint						
	Weight concerns						
	Eating concerns						
	Shape concerns						
	EDE-Q CE frequency						
	DERS Total						

8 References

1. Meyer C, Taranis L, Goodwin H, Haycraft E. Compulsive exercise and eating disorders. Eur Eat Disord Rev. 2011;19(3):174–89.
2. Rizk M, Mattar L, Kern L, Berthoz S, Duclos J, Viltart O, et al. Physical Activity in Eating Disorders: A Systematic Review. Nutrients. 2020;12(1).
3. Monell E, Levallius J, Forsen Mantilla E, Birgegard A. Running on empty - a nationwide large-scale examination of compulsive exercise in eating disorders. J Eat Disord. 2018;6(1):11.
4. Rogers ML, Duffy ME, Buchman-Schmitt JM, Datoc AE, Joiner TE. Exercise dependence: Associations with capability for suicide and past suicidal behavior. J Clin Psychol. 2019;75(1):165–77.
5. Solenberger SE. Exercise and eating disorders. Eating Behaviors. 2001;2(2):151–68.
6. Dalle Grave R, Calugi S, Marchesini G. Compulsive exercise to control shape or weight in eating disorders: prevalence, associated features, and treatment outcome. Compr Psychiatry. 2008;49(4):346–52.
7. Mathisen TF, Hay P, Bratland-Sanda S. How to address physical activity and exercise during treatment from eating disorders: a scoping review. Curr Opin Psychiatry. 2023;36(6):427–37.
8. Simon ET, Monell E, Lindstedt K, Wiberg AC, Mantilla EF. "To exercise sustainably" - Patients' experiences of compulsive exercise in eating disorders and the Compulsive Exercise Activity Therapy (LEAP) as a treatment: a qualitative interview study. J Eat Disord. 2024;12(1):151.
9. Welch E, Birgegard A, Parling T, Ghaderi A. Eating disorder examination questionnaire and clinical impairment assessment questionnaire: general population and clinical norms for young adult women in Sweden. Behav Res Ther. 2011;49(2):85–91.
10. Taranis L, Touyz S, La Puma M, Meyer C. Loughborough Eating-disorders Activity Programme “LEAP” Group Cognitive-Behavioural Therapy for Compulsive Exercise in the Eating Disorders: Therapist Manual. Loughborough University: Centre for Research into Eating Disorders; 2011.
11. Fairburn CG, Cooper Z, Shafran R, Bohn K, Hawker D, Murphy R, et al. Enhanced cognitive behavior therapy for eating disorders: the core protocol. In: Fairburn CG, editor. Cognitive Behavior Therapy and Eating Disorders. New York: Guilford Press; 2008.
12. Hay P, Touyz S, Arcelus J, Pike K, Attia E, Crosby RD, et al. A randomized controlled trial of the compuLsive Exercise Activity TheraPy (LEAP): A new approach to compulsive exercise in anorexia nervosa. Int J Eat Disord. 2018.
13. Harris A, Mannan H, Hay P, Aouad P, Arcelus J, Attia E, et al. Assessment and treatment of compulsive exercise in anorexia nervosa - A combined investigation of Compulsive Exercise Activity Therapy (LEAP) and Compulsive Exercise Test subscales. Eat Behav. 2024;52:101825.
14. Meyer C, Touyz S, editors. Managing compulsive exercise among eating disordered patients. . Proceedings of the ninth annual Australian and New Zealand Academy for Eating Disorders (ANZAED) Conference; 2011; Coogee, Sydney, Australia.
15. Touyz S, Meyer C, Taranis L, editors. The Loughborough Eating Disorders Activity Therapy (LEAP): a new approach to driven exercise in Anorexia Nervosa. Eating Disorders Research Society; 2010; Cambridge, UK.
16. Hallward L, Di Marino A, Duncan LR. A systematic review of treatment approaches for compulsive exercise among individuals with eating disorders. Eat Disord. 2022;30(4):411–36.

17. Meyer C, Plateau CR, Taranis L, Brewin N, Wales J, Arcelus J. The Compulsive Exercise Test: confirmatory factor analysis and links with eating psychopathology among women with clinical eating disorders. J Eat Disord. 2016;4:22.
18. Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? Int J Eat Disord. 1994;16(4):363–70.
19. de Man Lapidoth J, Birgegard A. Validation of the structured eating disorder interview (SEDI) against the eating disorder examination (EDE). Stockholm: Karolinska Institutet; 2010.
20. Bjureberg J, Ljotsson B, Tull MT, Hedman E, Sahlin H, Lundh LG, et al. Development and Validation of a Brief Version of the Difficulties in Emotion Regulation Scale: The DERS-16. J Psychopathol Behav Assess. 2016;38(2):284–96.