CARMEN
Centre for Applied Research in Mental Health Care

Statistical analysis plan for The Vega Trial

Value of gym-based exercise training for young adults with severe mental illness: A pragmatic, single-blinded, multicenter randomized controlled trial

SAP v1_23.04.25

Clinical trials number: NCT05461885

WHO Universal Trial Number (UTN): U1111-1271-9928

Regional Ethics Committee number: H-21079211.

Publication date: April 2025

Document version: SAP uploaded to ClinicalTrials.gov

Citation for published version:
Rafn BS, Sørensen V, Andersen MF, Bjerre ED, Baandrup L, Vernal DL, Mors O, Knop FK, Tolver A, Firth J, Skou ST, Ebdrup BH, Midtgaard J (2025). *Statistical analysis plan for the Vega trial: Value of gym-based exercise training for young adults with severe mental illness: a pragmatic, single-blinded, multicenter randomized controlled trial*

Research group

Bolette Skjødt Rafn, *Associate Professor*
Victor Sørensen, *PhD student*
Martin F. Andersen, *PhD student*
Bjørn H. Ebdrup, *Professor*
Lone Baandrup, *Associate Professor*
Ditte Lammers Vernal, *PhD*
Ole Mors, *Professor*
Filip Krag Knop, *Professor*
Joseph Firth, *PhD*
Søren T. Skou, *Professor*
Eik Dybboe Bjerre, *PhD*
Julie Midtgaard, *Professor*

Statistician
Anders Tolver, *PhD*

Corresponding author: bolette.skjoedt.rafn@regionh.dk

Table of content

1 Study synopsis

Exercise is recommended to protect physical health among people with severe mental illness and has the potential to facilitate long-term recovery [1–5]. An inclusive exercise community provides an opportunity for life skill training and social connectedness and may reduce the experience of loneliness and internalized stigmatization which together may improve personal recovery [6–8]. Using a pragmatic randomized design, we aim to examine the effectiveness of a gym-based exercise intervention tailored to young adults with severe mental illness (i.e., Vega Exercise Community) compared to usual care.

This statistical analysis plan (SAP) describes the planned analyses for the primary and secondary objectives and current planned tertiary objectives to be reported in the main article of the study. Analysis plans for subsequent articles using tertiary outcomes and additional data collected in the study will be written separately. Additional outcomes for subsequent articles are listed in the published protocol [9]. This SAP adheres to appropriate guidelines [10] and was made publicly available before any outcome analyses commenced and before unblinding the data.

2 Study objectives

A study protocol elaborating the methods used in this study has been published [9].

2.1 Primary objective

The primary objective is to determine the effectiveness of participation in a supervised, gym-based exercise program in addition to usual care at four-month compared to usual care alone on patient-rated personal recovery among people with severe mental illness (primary outcome).

2.2 Secondary objectives

Secondary objectives are to examine if the gym-based exercise program in addition to usual care compared to usual care alone improves the following:

1. Mental health
2. Health-related quality of life
3. Behavioral and functional symptoms
4. Metabolic health

2.3 Tertiary objectives

As tertiary objectives to be reported in the main article, we will investigate if:

1. Prolongation of subsidized gym membership in addition to motivational text messages (extended support) will be superior to subsidized gym membership alone (minimal support), and to treatment as usual in relation to post-intervention adoption of physical activity.
2. The exercise program is cost-effective.

2.4 Statistical hypotheses

1. Gym-based exercise training in addition to usual care will be superior to treatment as usual in relation to change in personal recovery at four months (primary outcome), mental health, health-related quality of life, behavioral and functional symptoms, and metabolic health (secondary outcomes) in young adults with severe mental illness.

3 Design and Outcomes

The trial design is outlined in Figure 1 and Table 1 outlines the outcomes. All outcomes will be obtained from all participants at baseline and all follow-ups (4months, 6months and 12months). The 12month follow-up is expected to be completed by January 2026.

Figure 1

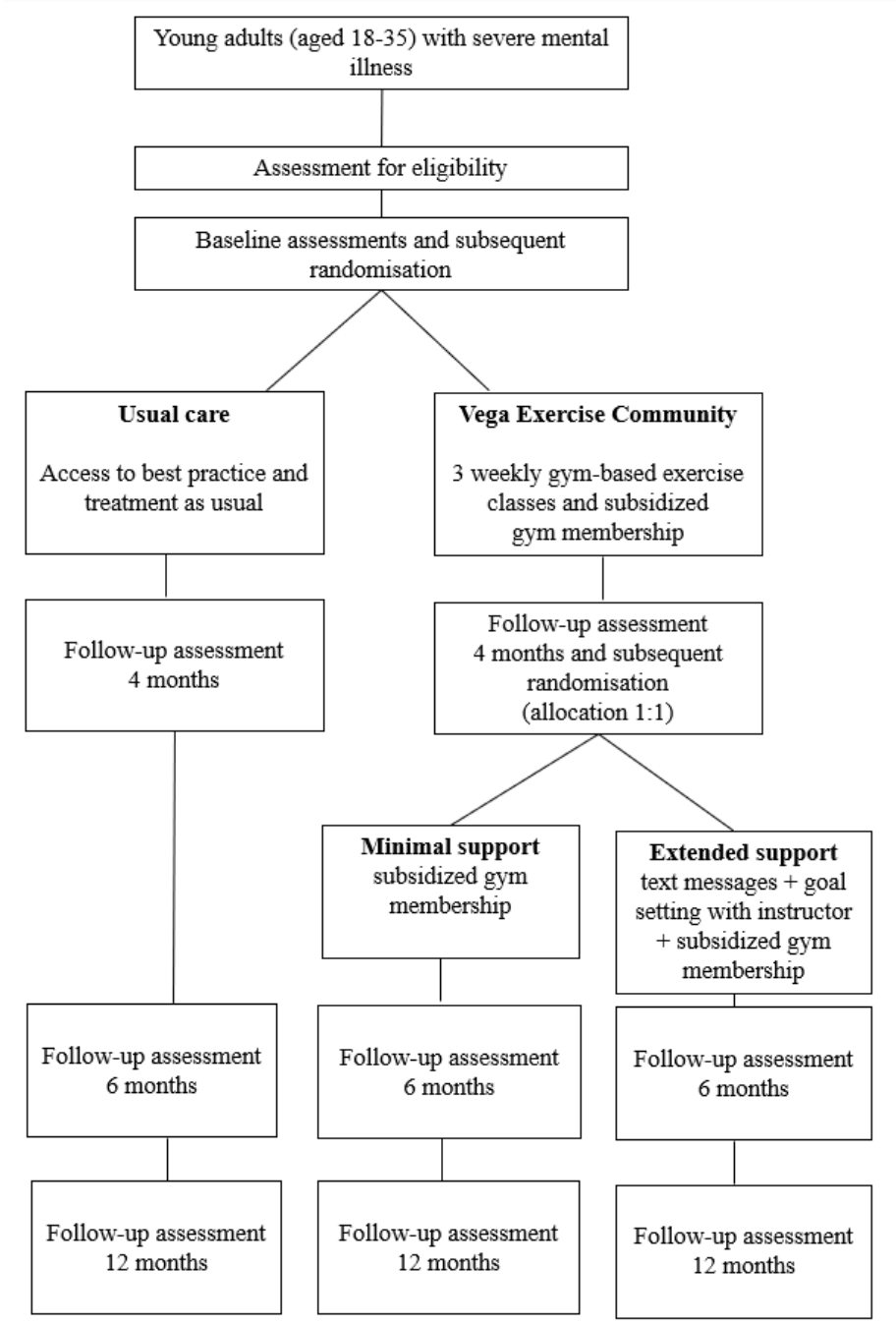

Table 1. Primary and secondary outcomes taken at baseline, 4-month, 6-month and 12-month

Endpoint	Domain	Measure	Instrument / Method	Data supplied by	Score range and cut-off values
Primary	Recovery	Personal recovery	Questionnaire about the Process of Recovery (QPR)	Participant (rating self) *and* participant's primary relative (rating participant)	0-60 (higher is better) Hypothesized difference 5
Main Secondary	Mental health	Mental Component Summary (MCS)	Short-Form-12 (SF-12) Health Survey	Participant (rating self) *and* participant's primary relative (rating self)	0-100 (higher is better) Hypothesized difference 4
Secondary	Health-related quality of life	Physical Component Summary (PCS)	Short-Form-12 (SF-12) Health Survey	Participant (rating self) *and* participant's primary relative (rating self)	0-100 (higher is better)
		Physical role, bodily pain, general health, vitality, social functioning, emotional role			0-100 (higher is better)
	Behavioral and functional symptoms	Affective symptoms	The Patient-Reported Outcomes Measurement Information System (PROMIS) Emotional distress (depression)	Participant (rating self) *and* participant's primary relative (rating self)	Mild Moderate Severe
		Physical activity	International Physical Activity Questionnaire short form (IPAQ-SF)		No range Low, moderate, and vigorous activity and sedentary behavior
		Sleep	Pittsburgh Sleep Quality Index (B-PSQI)		Average (SD) ≤5 good sleep quality >5 poor sleep quality
		Internalized Stigma of Mental Illness	Internalized Stigma of Mental Illness Inventory (ISMI-9)	Participant (rating self)	Total score Add scores and divide by number of items answered
		Substance abuse	Self-developed items		Binary outcome (y/n): Regular/ harmful use for each substance
		Positive and negative symptoms	Modified Colorado Symptom Index (MCSI)		Total score 0-56 (higher is worse)

Endpoint	Domain	Measure	Instrument / Method	Data supplied by	Score range and cut-off values
		Loneliness	Single-item measure (4-point scale)		Binary outcome (y/n) Loneliness defined as: Cut-off ≥3
	Metabolic health	Abdominal circumference, weight	Anthropometry	Blinded assessor (rating participant)	Mean (SD)
		Total and visceral fat mass and muscle mass	Non-invasive bioimpedance analysis		% fat mass Kg muscle mass
		Cardiorespiratory fitness	Cycle ergometer		Mean (SD) mlO2/min/kg mlO2/min
		Blood pressure and resting heart rate	Digital blood pressure monitor		Mean (SD) Systolic Diastolic Rest HR Mean (SD)
		Glycosylated haemoglobin (HbA1c)	Blood samples and biochemical analysis	Routine blood samples	Mean (SD) Mmol/mol
		Blood lipids (total cholesterol, high density lipoproteins, triglycerides)			Mean (SD) Mmol/L

3.1 Descriptive endpoints

Recruitment rate will be reported using data on eligibility (y/n) and reasons for lack of interest in participating for those who are eligible. Additionally, the proportion of participants recruited through each recruitment channel (website vs. clinician referred) will be calculated. Furthermore, retention rate and reasons for withdrawal will be reported.

The following intervention-related variables will be presented descriptively:

1) Attendance to the exercise will be recorded by the instructors at every exercise class. Attendance is reported as the total number of exercise sessions attended out of the total available sessions. Good attendance is defined as participation in 16 exercise sessions out of 48 available sessions from baseline to four months. The following will be reported:

median (IQR), min-max, and % attendance from baseline to four months and from four months to 12 months.

2) Exercise intensity is collected by research staff using heart rate monitors at selected exercise classes. Average and maximum heart rate, as well as number of classes and number of participants monitored will be reported. This will be reported for all sites together.
3) The duration of the exercise classes will be recorded by research staff at selected exercise sessions. This will be reported as mean (SD) minutes.
4) Structure (i.e., warm up, resistance training, high intensity functional training (HIFT) and cooldown) will be reported by instructors at all exercise sessions using a four-point Likert scale (the higher value the better compliance with exercise protocol). Good compliance is defined as ≥3.
5) Organization is defined as having two instructors present at the exercise sessions. This will be collected from the work schedule of the instructors and reported as proportion of sessions with two instructors out of the total sessions during the study period. This will be reported as total for all sites.

Assessors who are blinded to group allocation perform the assessments at 4-, 6- and 12-months follow-up. Following each assessment, the assessors report whether they accidently have been unblinded (i.e., if the participant has revealed the allocation). We will calculate and report the number of blinded assessors who become unblinded.

Because anchor-based estimations of minimal clinical important difference (MCID) on the primary outcome QPR are currently not available, we will supplement the data collection with one question on perceived improvement as an external criterion (“anchor”), i.e. asking participants (on a Likert

scale of 7 levels) how they perceive their change (improvement or worsening) from last visit (1, worse than ever; 2, much worsened; 3, slightly worsened, 4, unchanged, 5, slightly improved; 6, much improved; 7, completely improved) and calculate an empirical derived anchor-based MCID [11]. This anchor question is collected at all follow-up visits (4-, 6-, and 12-months).

Finally, an analysis of Number Needed to Treat (NNT) will be performed. NNT estimates the number pf people who would need to participate in the Vega Exercise Community for one person to have a MCID improvement on QPR from baseline to four months compared to the usual care group.

3.2 Safety endpoints

We will draw data from the medical record and report number of adverse events (AE) and serious adverse events (SAE) (i.e., hospitalization and contact to the emergency department) in both groups. An AE is defined as any undesirable experience during follow-up leading to contact with the health-care system. If an AE results in death, hospitalization, prolonged inpatient hospital care, permanent disability or damage, or if an AE is life-threatening, it will be categorized as an SAE [12]. If available, date of healthcare system contact, and duration will be registered and reported. The responsible psychiatrist (BE) or physical therapist (BSR) will evaluate the causality (whether the event was related to participation in the trial or not) and expectedness (whether the event was expected or not).

In addition, participants were asked to self-report any adverse events during follow-up using specific questions in the follow-up questionnaire. This data was used for prospective monitoring of safety within the trial and will not be reported in the main manuscript, as these events are expected to overlap with the events drawn from the medical records.

4 Statistical considerations

4.2 Sample size and power considerations

The published study protocol outlines the sample size calculations [9]. No interim analysis will be performed. Recruitment was ended in December 2024.

4.3 Statistical analyzes

Descriptive statistical methods will be used to summarize all outcomes for both groups at each follow-up. The main analyses will evaluate the effectiveness of having access to the intervention (gym-based group exercise) applying the intention-to-treat principle with patients analyzed in the treatment group to which they were randomly allocated, adjusted for stratification variables (diagnosis and site). Secondary per-protocol analyses will target the effect of receiving the intervention. The per-protocol population will be defined as intervention group participants who attended the exercise program at least 16 of 48 times during the 4-month intervention period (baseline to four months follow-up).

The primary analysis will compare average change in QPR from baseline to four months follow-up between groups using appropriate methods for the data distribution.

For the analyses of numerical secondary and tertiary outcomes, we will use a suitable variant of (generalized) linear mixed models to account for within-subject correlations across different assessment times (baseline, 4, 6, and 12 months). Treatment arm (intervention vs usual care), time, and their interaction as well as the stratification variables site (Copenhagen, Aarhus, Aalborg) and diagnosis (F20 vs F30) will be included as fixed factors. This approach will use all available data, does not require imputation of missing data, and results are valid under the missing at random assumption. P-values and 95% CIs for relevant group differences will be presented and used to assess superiority of the intervention.

The per-protocol analyses will be adjusted with baseline covariates on diagnosis (F20, F30) and lifestyle characteristics to adjust for confounding that cannot be stratified for and that may arise when a sub-group (the per-protocol population) is analyzed [13]. Supplementary analyses assessing the robustness of results to the definition of the per-protocol population, or analyses replacing binary group allocations with a continuous proxy of compliance (e.g. percentage of exercise sessions attended) may be included. Significance tests will be two-sided and P-values below 5% will be reported as statistically significant.

No adjustment of P-values will be performed, but due to the increased risk of false positive findings, the exploratory nature of secondary and tertiary analyses will be emphasized, and results interpreted with care. As such, the conclusion will be based on the primary outcome, and the secondary and tertiary outcomes will only be used to support the primary outcome conclusion. All analyses will be conducted using R [14].

5 Implementation of Analysis Plan

This SAP will serve as a detailed guide for the statistician performing the analyses. All analyses will be performed by the same statistician (AT) and none of the investigators involved in this trial will perform any of the statistical analyses.

The implementation of the SAP will be as follows:

1. With an allocation ratio of 2:1, blinding of the statistician is not possible. To mitigate the impact of this, all program code used to transform raw data into results will be made available upon request. This approach will allow the results to be tested for re-producibility by others.

2. The results will be presented to the steering committee of the trial without revealing group allocation. Any uncertainties will be clarified, and blinded interpretations of the primary endpoint results will be conducted prior to unblinding of data to the steering committee.

6 References

1. Vancampfort D, Firth J, Correll CU, et al. The impact of pharmacological and non-pharmacological interventions to improve physical health outcomes in people with schizophrenia: a meta-review of meta-analyses of randomized controlled trials. *World Psychiatry*. 2019;18(1):53-66. doi:10.1002/wps.20614

2. Dauwan M, Begemann MJH, Heringa SM, Sommer IE. Exercise Improves Clinical Symptoms, Quality of Life, Global Functioning, and Depression in Schizophrenia: A Systematic Review and Meta-analysis. *Schizophr Bull*. 2016;42(3):588-599. doi:10.1093/schbul/sbv164

3. Sabe M, Kaiser S, Sentissi O. Physical exercise for negative symptoms of schizophrenia: Systematic review of randomized controlled trials and meta-analysis. *Gen Hosp Psychiatry*. 2020;62:13-20. doi:10.1016/j.genhosppsych.2019.11.002

4. Bueno-Antequera J, Munguía-Izquierdo D. Exercise and Schizophrenia. *Adv Exp Med Biol*. 2020;1228:317-332. doi:10.1007/978-981-15-1792-1_21

5. Ashdown-Franks G, Firth J, Carney R, et al. Exercise as Medicine for Mental and Substance Use Disorders: A Meta-review of the Benefits for Neuropsychiatric and Cognitive Outcomes. *Sports Med*. 2020;50(1):151-170. doi:10.1007/s40279-019-01187-6

6. Firth J, Siddiqi N, Koyanagi A, et al. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. *Lancet Psychiatry*. 2019;6(8):675-712. doi:10.1016/S2215-0366(19)30132-4

7. Firth J, Rosenbaum S, Stubbs B, Gorczynski P, Yung AR, Vancampfort D. Motivating factors and barriers towards exercise in severe mental illness: a systematic review and meta-analysis. *Psychol Med*. 2016;46(14):2869-2881. doi:10.1017/S0033291716001732

8. Yanos PT, DeLuca JS, Roe D, Lysaker PH. The impact of illness identity on recovery from severe mental illness: A review of the evidence. *Psychiatry Res*. 2020;288:112950. doi:10.1016/j.psychres.2020.112950

9. Rafn BS, Andersen MF, Sørensen V, et al. Value of gym-based group exercise versus usual care for young adults receiving antipsychotic medication: study protocol for the multicenter

randomized controlled Vega trial. *BMC Psychiatry*. 2023;23(1):634. doi:10.1186/s12888-023-05086-z

10. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. Jama 2017;318(23):2337-2343. DOI:10.1001/jama.2017.18556.

11. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56(5):395-407. doi:10.1016/s0895-4356(03)00044-1

12. U. S. Food Drug Administration. What is a Serious Adverse Event? (http://www.fda.gov/Safety/MedWatch/HowToReport/ucm053087.htm).

13. Hernán MA, Hernández-Díaz S, Robins JM. Randomized trials analyzed as observational studies. *Ann Intern Med*. 2013;159(8):560-562. doi:10.7326/0003-4819-159-8-201310150-00709

14. R Core Team (2025). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/.