Statistical analysis plan (SAP) for Project Meaningful Activities and Recovery (MA&R)

NCT 03963245

JULY 26TH, 2021

Content

1. General information

1.1 Title of the project

Project Meaningful Activities and Recovery (MA&R)

1.2 Trial registration number

ClinicalTrials.gov NCT03963245.

1.3 SAP version number

This is a detailed version of the SAP description in the protocol (Danish latest version 5 sent to Regional Ethics Committee e) and protocol paper (Bjørkedal *et al.*, 2020). This detailed SAP is version 1, July 2021. The headings in this SAP are in many instances aligned with the structure recommended by Gamble et al. (Gamble *et al.*, 2017).

1.4 Names, affiliations, and roles of SAP contributors

This SAP has been drafted by Siv Therese Bogevik Bjørkedal (MSc and PhD student) guided by senior researcher and statistical supervisor Carsten Hjorthøj (MSc, PhD) and principal investigator for Project Meaningful Activities and Recovery Lene Falgaard Eplov (MD, PhD). Siv Therese Bogevik Bjørkedal, Carsten Hjorthøj and Lene Falgaard Eplov are all employed at Copenhagen Research Center for Mental Health (CORE). CH is also an associate professor at the University of Copenhagen.

2. Introduction

2.1 Trial background and rationale

The background and rationale are described detailed in the protocol paper (Bjørkedal et al, 2020).

2.2 Objectives or hypotheses

From the protocol paper (Bjørkedal et al, 2020):

"The primary objective of this RCT is to determine whether MA&R in addition to standard mental healthcare is more effective than standard mental healthcare alone regarding activity engagement (POES-S) [38, 39] in a sample of individuals living with psychiatric disabilities. The intervention is being tested in addition to standard mental healthcare as it is an add-on to usual care. The second objective of this study is to evaluate whether MA&R in addition to standard mental healthcare is more effective than standard mental healthcare alone regarding outcomes related to activity engagement, e.g. functioning, personal recovery and quality of life"

3. Study Methods

3.1 Trial design

A multicenter two arm, parallel randomized controlled trial. Eligible participants where individuals with psychiatric disabilities, who were treated in one of three community mental health centers in Mental Health Center Copenhagen, or were citizens in the municipalizes of Copenhagen, Svendborg or Odense. Included participants were allocated with an 1-1 ration to either Meaningful Activities and Recovery (MA&R) a novel rehabilitation intervention, in addition to usual care, or usual care alone. The enrollment started at September 2018 and lasted to August 2020. MA&R lasts for about eight months and consists of 11 group sessions and 11 one-to-one sessions.

From the protocol (Bjørkedal et al, 2020)

"MA&R is delivered by a peer worker and an Occupational therapist. In addition to the planned sessions, the participants are offered individualized support to engage in meaningful activities. The

MA&R is a manualized intervention, with a workbook developed for the participants to support learning and personal exploration. MA&R utilized various methods such as didactic presentations, peer exchange, direct experiences with activities and personal exploration. The MA&R contains two courses: MA&R I and MA&R II which complement each other. In MA&R I, the participants can explore and identify meaningful activities, while MA&R II supports anchoring new activities and habits in everyday life."

Data was collected twice, baseline before randomization og at follow up, at the end of the intervention.

3.2 Randomization

From the protocol paper (Bjørkedal, et al. 2020):

After baseline data are obtained, the participants are randomly assigned to either the control (standard mental healthcare) or experimental group (MA&R + standard mental healthcare) with a 1:1 allocation using the randomization module in REDCap (Research Electronic Data Capture) [44, 45]. The randomization sequence was generated by a researcher employed outside the research team and entered to RedCap. The randomization is stratified by sex. Varying block sizes, unknown to the research team, are used. To ensure concealment, the randomization schedule is stored away from the research team and the block sizes are not disclosed. The allocation is performed by the researcher, who informs the participants and the counsellors and who assigns participants to the interventions."

Participants gave informed consent (verbal and written) and filled out the baseline questionnaire before randomization.

3.3 Sample size calculation

Below numbers are retrieved from the protocol paper (Bjørkedal et.al 2020).

- Type 1 error set at: 0.05 (α)
- Power: 0.80 (β)
- Difference in mean: 3 points
- within group standard deviation (SD): 6

From the protocol paper (Bjørkedal *et al.*, 2020): "The sample size is calculated based on the primary hypothesis to detect a minimal but clinically significant difference between the intervention group and the control group on POES-S (increased engagement in meaningful activities). In a recent RCT testing, the intervention Balancing Everyday Life (BEL) in a similar population as that in this study, Eklund et al. found a small effect of 1.4 points measured on POES-S (d = 0.27) [21]. As MA&R has a more comprehensive format (in terms of duration and costs) than BEL, the clinically significant difference between the study groups is set to 3 points, corresponding to a moderate effect size (Cohens d = 0.5). Based on the BEL trial, we assume the standard deviation in the study population to be 6. To achieve a statistical power of 80% and a significance level of 5%, a total of 128 participants must be included in this study; 64 in each group to detect this difference."

No interim analyses were planned.
There were no guidelines for stopping the trials early.

3.5 Timing of outcome measurements

Post interventions, about 8 months: Questionnaires were sent out 1-7 days after the last one-on-one session. There is no definite upper limit for when the follow-up data is not relevant anymore; however, the contact procedure included 6 contact attempts for 3-8 weeks. Generally, data was considered not relevant after three months after the questionnaires were sent out. Exceptions from this rule was made, if participants were hospitalized at the time for follow up (and wanted to answer the questionnaires after discharge), or participants needed help for filling out questionnaires and had to wait until the end of the first COVID -19 lockdown before being offered assistance.

4. Trial population

4.1 Eligibility criteria

This is described detailed in the protocol paper (Bjørkedal et al, 2020).

4.2 Information to be included in the CONSORT flow diagram

Numbers of persons referred, screened for eligibility, fulfilled inclusion/exclusion criteria (excluded), followed up (answered questionnaires), loss to follow-up (did not want to answer questionnaire, never answered questionnaire), retracted consent, analyzed according to intention to treat (ITT) analyses.

4.3 Withdrawal

4.3.1 Withdrawal from treatment

It was always possible to drop out of treatment/intervention. Hereafter one could choose to participate in follow-up (answer questionnaire etc.) or not. If a person in the intervention group dropped out of treatment this was registered (+ preferably reason if possible). If the participant was in the control group drop out of treatment as usual was not registered as it was not possible to get this information.

4.3.2 Withdrawal from research

It was always possible to drop out from research and thus not answer questionnaires at follow up. This was registered as e.g. *did not want to answer questionnaire/loss to follow-up* if the participant explicitly communicated this to the project, otherwise unspecified reason to dropout (e.g. *not possible to get in touch with participant/loss to follow-up*) was registered.

According to Danish law it is possible for participants to withdraw their consent and have all person sensitive data deleted. In these cases, data cannot be used in analyses. In these instances, we will keep the randomization result and projects specific ID (not identifiable CPR number) in the database. Dropout from research (loss to follow-up and withdrawn consent) is summarized at 8 months follow-up.

4.4 Baseline characteristics

Sex, age, housing/marital status, educational level, diagnosis, work status, functional level (assessed by researcher). Register-based and self-reported baseline outcome measures will also be summarized.

5. Statistical principles and analysis

5.1. Confidence intervals and p-values

The two-sided significance level for statistical tests will be 5 %. Differences in means and proportions will be presented with a 95% confidence interval (CI) and a p-value. All primary, secondary and exploratory measures will be presented.

5.2 Adherence

Participant adherence to the MA&R is not defined as a minimum number of sessions that the participants have participated in. MA&R combines group sessions, individual sessions and individualized support. Participation

in group and one-on one session will be reported separately for the type of session and totally for the sessions together. Amount of individual support received measured in hours will be reported. MA&R is delivered in an individualized manner, which means that each MA&R course is tailored to each participant. Hence, the amount of support and session participated in, may vary extensively from participant to participant.

Fidelity to the MA&R intervention and its principles is evaluated through fidelity reports. Fidelity to MA&R will be described as a result of one or more fidelity reports (i.e. poor, good or excellent fidelity to the model).

Major deviations (i.e. outcome changes etc.) from this SAP will be presented/summarized in the reporting of results.

5.3 Specification of outcomes and timing

See appendix 1 for an overview over outcomes and timing.

- Activity engagement (measured by POES-S) at eight months follow up is the primary outcome (end point comparison)
- Personal recovery (measured by QPR), quality of life (measured by MANSA) and levels of functioning (measured by WHODAS-12) at eight months follow up are secondary outcomes (end point comparisons)
- WHODAS-36 subscales, WHODAS-36 total score and health related quality of life (EQ-5D-3L) at eight months follow up are exploratory outcomes (end point comparison).

5.4 Analysis methods

The trial is analyzed according to the statistical principle “intention-to-treat”. This means that analyses are based on all included participants as opposed to “per protocol” analyses.
For the primary outcomes in both trials, the null hypothesis tested is that there is no difference in activity engagement between the two groups (intervention group and control group) at 8 months’ follow-up. The null hypotheses are similar in the other analyses assuming no differences between groups.

5.4.1 Baseline variables

The chi^2 test is used to test for differences in nominal baseline variables (sex, education, housing status, marital status, diagnosis, employment status) and the t-test is used to test for differences in ordinal/continuous baseline variables (age, activity engagement, functioning, personal recovery, quality of

life). If the test assumptions in the latter are not met, a non-parametric test will be performed instead (e.g. Kruskal Wallis or Mann Whitney U test).

5.4.2 Self-reported outcomes

Primary, secondary and exploratory outcomes are calculated and presented in means of scores at 8 months follow-up. In the protocol we wrote:

"The main outcome measure is activity engagement, measured by POES-S. To test the research hypothesis, the differences between the intervention group and the control group will be analyzed using ANOVA to determine the statistical significance. Effect sizes to judge clinical relevance will be calculated by Cohen's d. In accordance with the intention-to-treat principles,

"multiple multivariate imputations" will be used and all co-variates of supposed prognostic significance will be used to impute a distribution of missing data.

The continuous power measurements will be analyzed with a repeated measurement model in mixed model analyses with unstructured variance. The prerequisite for using this analysis and for the use of multiple imputations is that data are missing at random or data missing completely at random as opposed to non-ignorable nonresponse. This distinction is important as repeated measurements and multiple imputations are both models based on a statistical estimation of non-existent responses, and the prerequisites for this estimation must be met for the analyses to be valid." (Bjørkedal et al, 2020)

5.5 Sensitivity analyses

No adjustments will be made in the primary analyses other than for stratification variables.

Sensitivity analyses are made with:

- Adjustment for baseline differences for those baseline characteristics with unequal/skewed baseline means and that can thus be associated with the outcome. Baseline characteristics in this case include (in addition to stratification variables) diagnosis, age, site and functioning) and self-reported outcome measures at baseline. Baseline tests will not be used exclusively in judging whether to include in the sensitivity analysis. Inclusion will be based on whether the groups are different to the extent that it could affect the analyses. The selection will be guided by CH.
- Imputations of missing values representing "worst" and "best" case of imputations. We will make two analyses: one where we in both groups replace imputations with "worst" case defined as the 90 percentiles from the imputed value and one where we in both groups replace imputations with "best" case defined as the 90 percentiles from the imputed value.
- The observed data.

Sensitivity analyses will be made for the primary outcome.

5.6 Subgroup analyses

No subgroup analyses will be planned, as the sample size does not allow this.

5.9 Missing data

Multiple imputations will be used to address the issue of missing data. If possible, multiple multivariate normal regression imputations (Markov Chain Monte Carlo (MCMC)) will be used. This is possible if all imputed data follow the same distribution (e.g. are scale variables). A table of auxiliary variables are made based on the list of variables used in the MA&R project. These are chosen for being theoretically associated with the specific outcome. Also, variables that are predictive for missing data are included. This selection will be based on whether there is skewed dropout for the specific outcome variable. In the cases of extreme distributions predictive mean matching will be used instead of multivariate normal imputations. All variables in the model will also be included in the imputation model. We will perform at least 100 imputations for each analysis.

5.10 Additional statistical analyses

Additional to the planned analyses in the protocol we will:

- Calculate the effect size (based on Cohen's d) of the primary outcome POES- S at 8 months follow up.

5.11 Safety data

The following variables will be summarized for each randomization group at 8 months follow-up.

- Number of bed days and admissions (somatic indication)
- Number of bed days and admissions (psychiatric indication)
- Number of somatic outpatient visits
- Number of deaths from suicide and other causes
- Scores of The Clinical Global Impression – Severity of Illness Scale (CGI-S)

5.12 Statistical packages to be used to carry out analyses

We will primarily use the statistics program SPSS. Other statistical packages might be used for some analyses.

6. Appendix 1

Table 1. Overview of outcomes

Outcome measure no	Outcome	Source	Outcome measure	Definition of outcome and calculation	Variable properties	Higher/lower score indicates better outcome	Collected at 8-months follow-up
#1	Primary	Self-reported	Activity engagement (POES-S)	Calculated and presented as mean sum of points/scores	Numerical variable. 9 items. Items rated: 1-4. Scale from 9-36	Higher	x
#2	Secondary	Self-reported	Functioning (WHODAS – 12 items)	Simple scoring or Complex scoring (ITEM RESPONSE THEORY) Computer program available from WHO website Calculated and presented as mean sum of points/scores	Numerical variable. 12 items. Items rated: 1-5. Scale from 12-60 Metric range from 0 -100	Lower	x
#3	Secondary	Self-reported	Recovery (Questionnaire about the processes about recovery (QPR) 15 item)	Calculated and presented as mean sum of points/scores	Numerical variable. 15 items. Items rated: 0-4. Scale from 0-60	Higher	x
#4	Secondary	Self-reported	Quality of life (Manchester short assessment of quality of life (MANSA))	Calculated and presented as mean sum of points/scores.	Numerical variable. 12 items. Items rated: 1-7. Scale from 12-84	Higher	x
#5	Secondary	Self-reported	Health-related quality of life (EQ-5D-3L)	Calculated and presented as a weighted mean on a scale	Numerical variable. 5 items. Each combination of answers on items is provided with a weight/number between 0-1.	Higher	x
#6	Explorative	Self-reported	Functioning (WHODAS 2.0 36 items	Simple scoring or Complex scoring	Each item rated 1-5 Total scale score: 36-180	Lower	x

			(domain score and total score))	Calculated and presented as a mean sum of points/scores	Domain scale score: Each domain contains between 5 to 8 eight items. (complex scoring) metric range 0-100		

Table 2. Safety measures

Safety measure no.	Outcome measure	Source	Definition of outcome and calculations	Variable properties
1	Number of outpatient somatic visits from baseline to follow-up	Register	Number of visits are calculated as an incidence rate ratio based on the follow-up period. Total number is also presented.	Numerical variable. Minimum: 0 visits, no defined upper limit
2	Number of inpatient somatic bed days and admissions from baseline to follow-up	Register	Number of bed days and admissions are calculated as incidence rates and an incidence rate ratio based on the follow-up period. Total number is also presented.	Numerical variables. Minimum: 0 bed days, no defined upper limit
3	Number of inpatient psychiatric bed days and admissions from baseline to follow-up	Register	Number of bed days and admissions are calculated as an incidence rate ratio based on the follow-up period. Total number is also presented.	Numerical variables. Minimum: 0 bed days, no defined upper limit
4	Deaths from suicide and other causes from baseline to follow-up	Register	Total number of deaths is reported + stratified by cause of death: suicide and other	Numerical variable. Minimum: 0 deaths, upper limit is the total number of included participants
5	Symptoms severity (Clinical Global Impression – self report)	Self-reported	Calculated and presented as mean sum of points/scores	Numerical variable. 1 item, item rated 1-7 (Scale from 1-7)

7. References

Bjørkedal STB *et al*. (2020) Meaningful Activities and Recovery (MA&R): the effect of a novel rehabilitation intervention among persons with psychiatric disabilities on activity engagement—study protocol for a randomized controlled trial, *Trials* (2020) 21:789

Gamble, C. *et al.* (2017) 'Guidelines for the content of statistical analysis plans in clinical trials', *JAMA - Journal of the American Medical Association*, 318(23), pp. 2337–2343. doi: 10.1001/jama.2017.18556.