Transdiagnostic group CBT vs. standard group CBT for depression, social anxiety disorder and agoraphobia/panic disorder (TRACT-RCT): A pragmatic, multicenter non-inferiority randomized controlled trial

Statistical Analysis Plan

Primary and secondary analyses

Version 1.8 2019.04.10

Introduction

This statistical analysis plan (SAP) describes the methods to be used to analyze the primary and secondary outcome data from the TRACT-RCT trial. The plan is based on the study protocol (Arnfred et al., 2017) registered at www.clinicaltrials.gov (Identifier: NCT02954731). Analyses follow the ICH-9 statistical guidelines for clinical trials and updated CONSORT reporting guideline extensions for non-pharmacological trials and non-inferiority and equivalence trials.[1-3]

The statistical analysis plan will be open to adjustments after the blind review, however finalized and registered as an amendment to the study protocol before breaking the blind.

Research objectives and hypotheses

The use of a single transdiagnostic cognitive behavior therapy manual (i.e. The Unified Protocol for Transdiagnostic Treatment of Emotional disorders (UP))[4,5] instead of several diagnosis-specific manuals for anxiety and depression disorders has the potential to simplify group logistics and reduce waiting time for patients. Applying the UP to heterogeneous groups might be a cost-effective way of delivering evidence-based psychological treatment for the most common mental disorders. While preliminary results for individual therapy with the UP for anxiety is promising, adaptations of the UP for depression and for group delivery need adequately testing.

The main objective of this study is to investigate the efficacy of group UP compared with standard diagnosis-specific cognitive behavior therapy (dCBT) for psychiatric outpatients with a primary diagnosis of unipolar depression, social anxiety disorder, panic disorder, or agoraphobia. Main outcome is subjective well-being measured with the World Health Organization Well-being Index, five items (WHO-5)[6] post treatment. Other outcomes are symptom levels, personality traits, emotion regulation, perseverative thinking, social functioning and clients' evaluation of therapy.

We hypothesize that subjective well-being and symptom levels will be equally improved following group UP and group dCBT.

Apart from the main objective, we investigate the effects of the UP on emotion regulation, the role of possible mediators and moderators on outcome, as well as qualitative data on patients' and therapists' experience of psychotherapy. These analyses are specified elsewhere.

Study methods

Design

Please see study protocol (Arnfred et al., 2017)[7] or www.clinicaltrials.gov (Identifier: NCT02954731).

Randomization

Please see study protocol (Arnfred et al., 2017)[7].

Sample size

Please see study protocol (Arnfred et al., 2017)[7].

Correction: In the study protocol we wrote that alpha was set to 5% resulting in a 95% one-sided confidence interval (CI). Following Trial guidelines[1,8], the alpha should be 0.025 resulting in a 97.5% one-sided CI. Further, standard deviation (SD) endpoint WHO-5 was reported to be 17.4 WHO-5 points. As reported in Reinholt et al. (2017)[9] the correct WHO-5 endpoint SD is 20.44. Using these corrected values, then a sample size of 208 patients are required to be 90% sure that the lower limit of a one-sided 97.5% confidence interval (or equivalently a 95% two-sided confidence interval) will be above the non-inferiority limit of -9 WHO-points.

Due to an unexpected, high attrition rate within the first year of inclusion, we decided to adjust sample size calculations by increasing the number of included patients from in total 248 patients to in total 320 patients to get 204 participants completing the treatment as needed for the primary analysis based on our initial sample size calculations. The revised sample size calculation is documented in an amendment to the protocol at www.clinicaltrials.gov (Identifier: NCT02954731).

Timing of outcome assessments

Baseline characteristics will be assessed at the time of inclusion. Outcome measures will be assessed pre-treatment (up to two weeks before start of intervention), post treatment (week 19-21 after allocation), and at six months follow-up (45-47 weeks after allocation). The primary outcome (WHO-5) is further assessed at the beginning of each therapy session (14 sessions).

Timing of final analysis

Primary and secondary outcomes are analyzed post treatment and exploratively repeated at six months follow-up.

Statistical principles

Data presentation

We will analyze and report primary and secondary outcomes in accordance with CONSORT reporting guidelines for non-inferiority and equivalence trials.[2] Results from the primary analysis will be reported as between group mean differences with one-sided 97,5% confidence intervals. The focus will be on hypothesis testing rather than point estimation. For secondary analyses, results will be reported as within and between group mean differences with one-sided 97,5% confidence intervals. Effect sizes (Cohens *d*) are calculated for each measure between groups. A one-sided p-value < .025 is considered significant. The focus of secondary analyses will be on point estimation rather than hypothesis testing.

No adjustment for multiplicity will be undertaken. Since outcomes are interrelated, adjustment will be overly conservative.

Adherence and protocol deviations

We will analyze attrition rates (number of patients receiving less than 8 sessions of therapy) in both groups using Chi-square analysis.

The analysis population

In superiority trials, the Intention-to-treat (ITT) analysis is considered the conservative choice to minimize possible type I-errors, that is falsely concluding that the one treatment is superior to another.[10] The ITT population includes all patients randomized and allocated to treatment whether they violated the protocol or dropped out of treatment or not. In superiority trials, the null- hypothesis asserts that there is no difference between the two interventions being compared. The alternative hypothesis being tested is that the two treatments are different by some prespecified value.

A non-inferiority trial, conversely, tries to show that a new treatment is not unacceptably worse than the treatment used for comparison by a prespecified margin.[11,12] The null hypothesis states that the new treatment is worse than the reference treatment by more than a prespecified non-inferiority margin. In non-inferiority trials, the ITT analysis is conventionally considered a liberal choice because it tends to make the two treatments being compared more similar, possibly falsely inflating type I-errors, that is concluding that the new treatment is non-inferior to the reference treatment. Following CONSORT reporting guidelines, the PP population is, therefore, considered the conservative choice of analysis in non-inferiority trials.[2]

Given the potential biases in both PP and ITT analysis, the European Medicines Agency (EMA) in the paper *'Points to consider when switching between superiority and non-inferiority'* states that inference of non-inferiority should rely on both PP and ITT-analyses.[8] Other authors (i.e. Sanchez & Chen, 2006[13]) argue that showing non-inferiority in both sets of analysis does not guarantee a valid conclusion as the conservatism or non-conservatism of the PP population depends on several factors, including the quality of the study, type of protocol violations and missingness. Sanchez & Chen (2006)[13] recommend minimizing the impact of protocol violations and missingness in the statistical analysis by using a hybrid ITT/PP analysis which excludes non-compliant patients as in the PP population and properly addresses the impact of non-trivial missing data in an ITT analysis. We will use this hybrid model for analysis of primary and secondary outcomes.

For this analysis plan, the PP population is defined as patients completing the therapy (i.e. receiving eight therapy sessions or more). The ITT population is defined as all patients randomized and allocated to treatment.

Trial population

Eligibility

Please see study protocol (Arnfred et al., 2017)[7] or www.clinicaltrials.gov (Identifier: NCT02954731).

Recruitment

Please see study protocol (Arnfred et al., 2017)[7] or www.clinicaltrials.gov (Identifier: NCT02954731).

Participant flow

The number of patients screened, randomized, completed treatment, and completed follow-up outcomes, will be summarized in a CONSORT flow diagram.[3]

Baseline characteristics

The following demographic and illness burden variables will be summarized by treatment condition in Table 1: age, sex, immigrant status, marital status, level of employment, level of education, children living at home, comorbidity, previous psychological treatment, current medication, age of onset, previous episodes with primary diagnosis, previous hospitalization, and personality traits (Standardized Assessment of Personality Abbreviated Scale (SAPAS).[14]

Analysis

Primary outcome

The primary outcome is self-reported subjective well-being measured with the WHO-5 assessed post treatment (19 weeks after allocation; single score continuous outcome).

Secondary outcomes

To avoid multiplicity of testing, the following outcomes were specified as secondary outcomes compared to the study protocol (Arnfred et al., 2017)[7] after enrolling patients in the study but prior to data-analysis:

- General anxiety and depression symptoms assessed with the Hopkins Symptoms checklist (SCL-25)[15] 19 weeks after allocation (single score continuous outcome)
- Observer-rated anxiety symptoms assessed with the Hamilton Anxiety Rating Scale (HAM-A6)[16] 19 weeks after allocation (single score continuous outcome)
- Observer-rated depression symptoms assessed with the Hamilton Depression Rating Scale (HAM-D6)[17] 19 weeks after allocation (single score continuous outcome)
- Self-reported impairment of functioning assessed with the Work and Social Adjustment Scale (WSAS)[18] 19 weeks after allocation (single score continuous outcome)

Exploratory outcomes (depending on diagnosis)

The following outcomes were specified as exploratory outcomes compared to the study protocol (Arnfred et al., 2017)[7] after enrolling patients in the study but prior to data-analysis:

- Panic Disorder Severity Scale (PDSS)[19] assessed 19 weeks after allocation (single score continuous outcome).
- Liebowitz Social Anxiety Scale - Self-Report (LSAS-SR)[20] assessed 19 weeks after allocation (single score continuous outcome).
- Mobility inventory for Agoraphobia (MIA)[21] assessed 19 weeks after allocation (single score continuous outcome).
- Beck Depression inventory-II (BDI)[22] assessed 19 weeks after allocation (single score continuous outcome).
- HAM-A6 analyzed by diagnosis (agoraphobia/social anxiety disorder/panic disorder)
- HAM-D6 analyzed by diagnosis (depression)

Analysis methods

Descriptive analyses

Baseline equivalence for the two groups compared will be assessed descriptively. No formal significance testing will be conducted, since we assume that differences between randomized groups could have occurred by chance.

Inferential analyses

Inferential analyses will be based on between group comparisons of group UP versus group dCBT.

Primary analysis

The primary analysis will compare the estimated endpoint of the primary outcome (WHO-5) between UP and dCBT conditions post treatment (19 weeks from allocation) adjusted for baseline WHO-5 values, variables predictive for missing data, and stratification variables (i.e. site (Copenhagen, Risskov, Slagelse) and diagnosis (depression, social anxiety disorder, panic disorder/agoraphobia)), fitting treatment (UP or dCBT) as a fixed effects factor.

We will use a one-sided 97.5% CI interval approach to test non-inferiority between conditions. The one-sided 97,5% CI for the between group difference point estimate post treatment will be calculated. Non-inferiority of UP compared to dCBT will be accepted if the lower bound of the 97,5% CI lies within the prespecified non-inferiority margin of – 9 WHO points (see Figure 1). If non-inferiority is accepted, we will use a closed testing procedure to test for superiority of UP over dCBT (i.e. lower bound of the 97,5% CI lies above 0).

The model will be fitted using a multilevel, mixed linear regression model, which allow us to use all 16 WHO-5 outcomes and handle missing data within the model. We will use stratification variables (i.e. site and diagnosis) and baseline outcome variables and treatment as fixed factors. We will apply an unstructured covariance matrix to the random effects in the model. If the model does not converge, we will apply other covariance matrixes (i.e. autoregressive, heterogeneous, compound symmetry, etc.) to the model and choose the covariance matrix with the best fit according to the Akaike Information Criterion.

Secondary analyses

For secondary analyses we will use ANCOVA to test for non-inferiority between treatment conditions post treatment. Effect sizes (Cohens *d*) for pre-post differences between treatment conditions will be calculated.

Following conventions (Cohen, 1988) statistically significant effect sizes ≤ 0.2 will be considered not clinically relevant and for this study a non-inferior difference between conditions.

All analyses will be run allowing for adjustment for stratification variables (i.e. site and diagnosis), variables predictive for missing data, and baseline values.

The clinically meaningfulness of results of study will be evaluated calculating numbers and odds ratio for responders and remitted patients using binary logistic regression. 'Responders' are defined as patients receiving a clinically meaningful change in primary outcome, that is ≥10 WHO-5 points 'Remission' is defined as patients no longer meeting diagnostic criteria for their primary DSM-5-diagnosis evaluated with the MINI Neuropsychiatric Interview (MINI)[23] 19 weeks after assessment.

Sensitivity analyses

We will run sensitivity analyses for the primary outcome using different definitions of the PP population, depending on how many therapy sessions they attended (i.e. 6, 10 sessions) and using the full ITT-population. Further, we will run a sensitivity analysis without imputed data to assess the potential impact of imputation on the trial findings and assess the impact of missingness by loading the primary analysis using an extreme scenario countering our hypothesis (i.e. imputing the data based on values from the 10% and 90% fractile population). Finally, we will assess the impact of nesting patients in therapy groups by adding therapy group as a random factor. As this analysis might unblind the analyst, this analysis is run after deciphering the blinding.

Additional analyses

Analysis of primary and secondary outcomes will be repeated exploratory at six months follow up.

Differences between groups in clients' evaluation of the therapy post treatment will be evaluated. According to the study protocol (Arnfred et al., 2017)[7], we intended to assess client satisfaction using the Client Satisfaction Questionnaire[24]. However, due to problems with the web-based administration of this instrument, the CSQ was not administered systematically. Before data-analysis, we instead decided to use items from the client evaluation questionnaire, which we also administered, developed for *'New Beginnings for patients with Anxiety and Depression,* a Danish lay person guided self-help program' developed in collaboration between Stanford University, the Expert Patients Program Community Interest Company (EPPCIC), the National Health Service, UK and The Danish Committee for Health Education.[25]

Subgroup analysis

We will run a confirmative analysis by including interaction terms in the primary analysis model to explore the uniformity of the treatment effect found in the primary analysis across potential subgroups (i.e. stratification variables (site, diagnosis). These results should be interpreted exploratively only, due to the low power for subgroup analyses.

If treatment effects differ across diagnosis, we will conduct supplementary subgroup analysis of primary and secondary outcomes separated by diagnosis.

Adverse events

Adverse events will be listed by group and type of event: (a) Minor adverse events will be reported as number of patients encountering any of these: increasing current medication more than one standard dose step; using according-to-need medication for more than two weeks; increasing current medication to entail more than four types of psychotropic medications; new prescription of supplementary medication (accepted at inclusion); changing previously prescribed antidepressant to another type. (b) Major adverse events will be reported as number of events: New prescription of medication (not accepted at inclusion); contacting psychiatric emergency due to self-injury, panic attack or other deterioration; drop-out of therapy due to experienced lack of effect or deterioration. (c) Critical adverse events will be reported as number of events: psychiatric hospitalization due to deterioration or suicide attempt; somatic hospitalization due to self-injury or suicide attempt; referral to electroshock treatment; suicide.

Model checking and validation

We will check the robustness of models by examining residuals.

Handling of missing data and blinding

Data entry is performed by the data manager and research assistants according to the main study protocol. We will undertake a blinded analysis of missing data and outliers to evaluate how to deal with missing data and outliers, and to explore possible predictors of missing data which should be controlled for in the primary and secondary analyses. The analysis of the primary and secondary outcomes will be conducted with the treatment allocation blinded. In the data extraction for the primary analysis, the intervention type is concealed. The intervention and control arm will randomly be coded as '1' and '2'. Deciphering will not be done before all analyses at post-treatment have been performed, and conclusions for the primary outcome has been formulated.

For this analysis plan, we define missing data as those patients with the absence of data at follow-up for one or more outcomes. Data are assumed to be missing at random.

In the primary analysis, the mixed linear regression model handles missing data by full information maximum likelihood, except for the baseline value of WHO-5. Missing baseline WHO-5 values will be handled by multiple imputation with chained equations (m=100) using stratification variables, post-baseline values and variables predicting missing baseline WHO-5 as predictors in the imputation model.

In the secondary analyses, missing data will be handled by multiple imputations with chained equations (m=100) using baseline values of the scales in question, variables predicting missing data, and stratification variables, as auxiliary variables in the imputation models. In cases where other scales are theoretically strongly correlated to the missing data, these scales will also be used as auxiliary variables in the imputation models.

Missing data within a scale less than 25% will be handled by single imputation of mean-scores of the scale. Missingness above 25% within a scale will lead to missing value at scale level and be imputed as described above.

Statistical software

All analyses will be undertaken in SPSS, version 25 and Stata, version 15.[26,27]

Nina Reinholt, MSc., PhD-fellow, Mental Health Center of Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark

10.4.19

Date **Signature**

Carsten Hjorthøj, PhD, MSc., statistical consultant, Copenhagen Research Center for Mental Health – CORE, Mental Health Center Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark

10/4-19

Date **Signature**

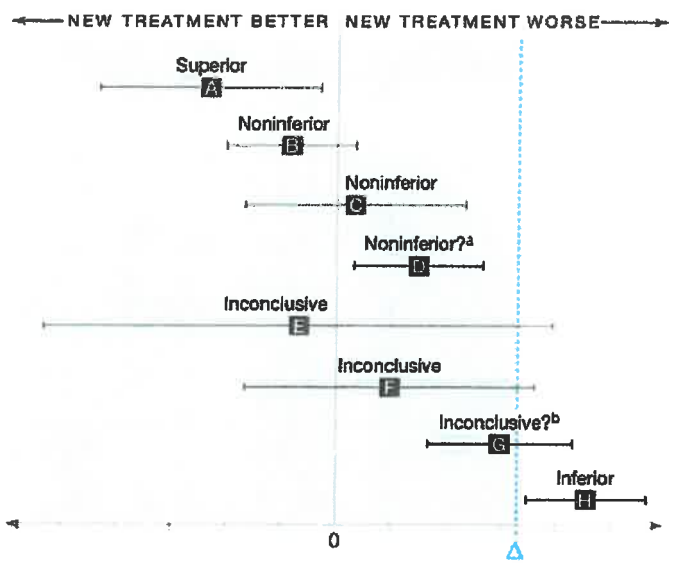
Principal Investigator Sidse Marie Arnfred, MD, PhD, MSc; DMSc, Professor, Psychiatry West, Slagelse, Region Zealand Mental Health Service & Department of Clinical Medicine, Copenhagen University, Denmark

4/4 2019

Date **Signature**

REGION SJÆLLAND
PSYKIATRIEN
– vi er til for dig

Sidse Arnfred
Forskningsoverlæge, dr.med., PhD.
Psykiatrien Slagelse
Forskningslektor i Psykiatri ved Klinisk Institut,
SUND, Københavns Universitet
Fælledvej 6, 4200 Slagelse
Mobil +45 24 22 70 45
Email: sidar@regionsjaelland.dk



Error bars indicate 2-sided 95% CIs. The blue dashed line at x=Δ indicates the noninferiority margin; the blue tinted region to the left of x=Δ indicates the zone of inferiority. A, If the CI lies wholly to the left of zero, the new treatment is superior. B and C, If the CI lies to the left of Δ and includes zero, the new treatment is noninferior but not shown to be superior. D, If the CI lies wholly to the left of Δ and wholly to the right of zero, the new treatment is noninferior in the sense already defined but also inferior in the sense that a null treatment difference is excluded. This puzzling circumstance is rare, because it requires a very large sample size. It also can result from a noninferiority margin that is too wide. E and F, If the CI includes Δ and zero, the difference is nonsignificant but the result regarding noninferiority is inconclusive. G, If the CI includes Δ and is wholly to the right of zero, the difference is statistically significant but the result is inconclusive regarding possible inferiority of magnitude Δ or worse. H, If the CI is wholly above Δ, the new treatment is inferior.
[a]This CI indicates noninferiority in the sense that it does not include Δ, but the new treatment is significantly worse than the standard. Such a result is unlikely because it would require a very large sample size.
[b]This CI is inconclusive in that it is still plausible that the true treatment difference is less than Δ, but the new treatment is significantly worse than the standard. Adapted from Piaggio et al.[6]

Figure 1. From Piaggio, G., Elbourne, D. R., Altman, D. G., Pocock, S. J., Evans, S. J., & CONSORT Group, F. T. (2006). Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *Jama*, *295*(10), 1152-1160.

	UP (n=)	CBT (n=)	Total (n=)[a]
Patient Characteristics			
Age, mean (SD) years			
Female sex			
Immigrant status (born in Denmark)			
Married/partner			
Children living at home			
Education (profession bachelor/university degree)			
Employment			
Fulltime/part time/student			
Sick leave (≥3 months)			
Retired or job seeking			
Previous episodes with primary diagnosis			
Previous hospitalization			
Previous psychotherapy (min. five sessions)			
Psychotropic medication current			
For YES:			
Antidepressants (SSRI)[b]			
Other[c]			
Age of onset, mean (SD) years			
Comorbid diagnoses			
Any			
At least one depressive or anxiety disorder			
At least one comorbid disorder, other[d]			
Personality traits (SAPAS score ≥4)[e]			
Stratification variables			
Recruitment site			
Copenhagen			
Risskov			
Slagelse			
Principal diagnosis			
Major depressive disorder			
Social anxiety disorder			
Panic disorder / Agoraphobia			

Table 1. Baseline demographic, illness burden and stratification variables

a. Data are presented as number (percentage) or mean (SD) unless otherwise indicated.
b. Abbreviation: Selective Serotonin Reuptake Inhibitors (SSRI)
c. Other medication: benzodiazepines (according-to-need), anxiolytics, other allowed medications (maximum three).
d. Other comorbid diagnoses: Generalized anxiety disorder, obsessive-compulsive disorder, personality disorders, posttraumatic stress disorder, somatoform disorders, attention-deficit/hyperactivity disorders (non-interfering) or eating disorders.
e. Abbreviation: Standardized Assessment of Personality – Abbreviated Scale (SAPAS).

References

1. ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. *Stat Med.* 1999;18(15):1905-1942.
2. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA.* 2012;308(24):2594-2604.
3. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, Group C. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med.* 2008;148(4):295-309.
4. Barlow DH. *Unified protocol for transdiagnostic treatment of emotional disorders : therapist guide.* New York: Oxford University Press; 2011.
5. Barlow DH, Farchione, T. J., Sauer-Zavala, S., Latin, H. M., Ellard, K. K., Bullis, J.R., Bentley, K. H., Boettcher, H. T., & Cassiello-Robbins, C. . *Unified Protocol for Transdiagnostic Treatment of Emotional Disorders.* Second edition ed. New York: Oxford University Press; 2018.
6. Newnham EA, Hooke GR, Page AC. Monitoring treatment response and outcomes using the World Health Organization's Wellbeing Index in psychiatric care. *J Affect Disord.* 2010;122(1-2):133-138.
7. Arnfred SM, Aharoni R, Hvenegaard M, et al. Transdiagnostic group CBT vs standard group CBT for depression, social anxiety disorder and agoraphobia/panic disorder: Study protocol for a pragmatic, multicenter non-inferiority randomized controlled trial. *BMC Psychiatry.* 2017;17.
8. Committee for Proprietary Medicinal P. Points to consider on switching between superiority and non-inferiority. *Br J Clin Pharmacol.* 2001;52(3):223-228.
9. Reinholt N, Aharoni R, Winding C, Rosenberg N, Rosenbaum B, Arnfred S. Transdiagnostic group CBT for anxiety disorders: The unified protocol in mental health services. *Cognitive Behaviour Therapy.* 2017;46(1):29-43.
10. International conference on harmonisation; guidance on statistical principles for clinical trials; availability--FDA. Notice. *Fed Regist.* 1998;63(179):49583-49598.
11. Food, Drug Administration HHS. International Conference on Harmonisation; choice of control group and related issues in clinical trials; availability. Notice. *Fed Regist.* 2001;66(93):24390-24391.
12. Committee for Medicinal Products for Human U, Efficacy Working P, Committee for Release for C. Committee for Medicinal Products for Human Use (CHMP) guideline on the choice of the non-inferiority margin. *Stat Med.* 2006;25(10):1628-1638.
13. Matilde Sanchez M, Chen X. Choosing the analysis population in non-inferiority studies: per protocol or intent-to-treat. *Stat Med.* 2006;25(7):1169-1181.
14. Moran P, Leese M, Lee T, Walters P, Thornicroft G, Mann A. Standardised Assessment of Personality - Abbreviated Scale (SAPAS): preliminary validation of a brief screen for personality disorder. *Br J Psychiatry.* 2003;183:228-232.
15. Müller J, Poster C, Beyer T, Furniss T, Actergarde S. Comparison of Eleven Short Versions of the Symptom Checklist 90-Revised (SCL-90-R) in the assessment of general Psychopathology. *J Psychopathol Behav Assess.* 2010;32:246-254.
16. Hamilton M. The assessment of anxiety states by rating. *British Journal of Medical Psychology.* 1959;32:50-55.
17. Hamilton M. A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry.* 1960;23:56-62.
18. Mundt JC, Marks IM, Shear MK, Greist JH. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *Br J Psychiatry.* 2002;180:461-464.
19. Shear MK, Brown TA, Barlow DH, et al. Multicenter collaborative panic disorder severity scale. *Am J Psychiatry.* 1997;154(11):1571-1575.
20. Baker SL, Heinrichs N, Kim HJ, Hofmann SG. The liebowitz social anxiety scale as a self-report instrument: a preliminary psychometric analysis. *Behav Res Ther.* 2002;40(6):701-715.

21. Chambless DL, Caputo GC, Jasin SE, Gracely EJ, Williams C. The Mobility Inventory for Agoraphobia. *Behav Res Ther.* 1985;23(1):35-44.
22. Beck AT, Steer RA, Brown GK. *Manual for the BDI.* San Antonio, Texas: Th Psychological Corporation; 1996.
23. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry.* 1998;59 Suppl 20:22-33;quiz 34-57.
24. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann.* 1982;5(3):233-237.
25. Christensen SM, M. *Evaluering af SATSpulje-projektet: Lær at Tackle Angst og Depression.* København: Sundhedsstyrelsen; 2016.
26. *IBM SPSS Statistics for Windows* [computer program]. Armonk, NY: IBM Corp. ; 2017.
27. *Stata Statistical Software: Release 15* [computer program]. College Station, TX: StataCorp LLC 2017.