

Supplement: Methodological details

Assessment of eligibility

Assessment of eligibility was performed by research staff, all clinically trained mental health professionals, supervised by a psychiatric specialist, and specifically trained for the purpose. The assessment was conducted through clinical interview, partly guided by 1) MINI International Neuropsychiatric Interview,¹ 2) Standardized Assessment of Personality - Abbreviated Scale (SAPAS),² 3) Attention deficit hyperactivity disorder symptom checklist for adults (Adult Self-Report Scale, ASRS v1.1)³ and 4) Mini-Mental State Examination (MMSE)⁴ (when dementia was clinically suspected).

Eligible were ≥ 18 year old sick leave benefit recipients for \geq four weeks with an ICD10 diagnosis of either depression (F32-F33), generalized anxiety disorder (F41.1), social phobia (F40.1) or panic disorder (F41.0). Participants were required to have Danish proficiency and give written consent. Excluded were those who were pregnant, in at least moderate risk of suicide according to clinical assessment and stratification according to the MINI International Neuropsychiatric Interview,¹ had clinically significant substance abuse disorder according to clinical assessment by assessor, other unstable medical condition judged by assessor to yield significant obstacles for mental health care treatment in the research project, showed signs of dementia by assessors' clinical judgement, or could not indicate willingness to abstain from seeking mental health care outside the study, whilst receiving such through the study.

Statistical analyses

A detailed statistical analysis plan, following the consensus template for such, recommended by a wide range of relevant bodies,⁵ was submitted for publication on www.clinicaltrials.org on July 5th, 2019, before any analyses of primary outcome data was scrutinized. It was followed notoriously (with few deviations, see supplements) regarding, for each outcome, its calculation, analysis method, results presentation, covariate adjustment, statistical method assumption control (and assumption fail-alternatives), sensitivity analyses and handling of missing data. All analyses were performed using base functions in R,⁶ and packages *survival*, *mice*, *lme4*, and *sandwich*. All analyses adhered to the intention-to-treat-principle (except the outcome *employee productivity*):

- For all time to return to work outcomes we performed cox-regression to calculate hazard rate ratios.
- For proportion in work outcomes we calculated odds ratios using logistic regression.

- For weeks in work outcomes we calculated rate ratios using Poisson regression and a Sandwich-Hubert estimator to correct standard errors due to expected severely skewed data.
- For self-reported (all numerical) outcomes we calculated pairwise group differences of estimated marginal means, using linear mixed-effects model, with unstructured covariance. Exceptions were outcomes without baseline scores—*client satisfaction* and *employee productivity*—where we used generalized linear models.
- Adjustments were, in all analyses, made for only randomization stratification variables, as recommended for RCTs by European Medicines Agency.⁷
- Missing data regarding self-report questionnaires was handled by generation of 100 multiple imputations by Chained Equations (MICE), using stratification variables, primary outcome data, and the four self-reported secondary outcome measures as predictor variables. No missing register data was expected, and only complete case analyses were planned.
- Subgroup analyses were performed for all outcomes in following strata, all pre-planned: per diagnosis (anxiety; depression); per employment at baseline (employed; unemployed); per two IBBS teams (Team City (Copenhagen municipality) and Team North (remaining municipalities)); per first and last temporal half. Furthermore, all analyses were performed adjusted for the interaction of diagnoses and intervention.
- Sensitivity analyses were, for register based data-outcomes, performed by including the missing data (due to consent withdrawal or register failure), with all their outcomes handled as either the worst possible (no return to work at follow-up) vs. best possible (return to work at first observation). For self-reported outcomes all missing data was single imputed as the mean of the outcome variable ± 2 standard deviations in the best/worse case scenario-manner.

Predefined exploratory outcomes are found in supplements.

➤ Register-based:

→ *Total numbers of weeks at work* (12-month follow-up), with one week at work being without sick leave benefit and any salary from an ordinary job during that week

➤ Self-reported at 6- and 12-month follow-up:

→ *Levels of depression, anxiety, somatization and distress*, by 4 Dimensional Symptoms Questionnaire, 4DSQ,⁸

→ *Levels of exhaustion*, measured with Karolinska exhaustion disorder scale, KEDS,⁹

→ *Illness perception*, by Illness Perception Questionnaire, IPQ,¹⁰

- *Health related quality of life*, by EQ5DL,¹¹
- *Quality of life*, measured by *Quality of Life Scale*, QoL,¹²
- *Return to work-self efficacy*, by RTW-SE,¹³
- *Generalised self-efficacy* measured with *Generalized Self-Efficacy Scale*, GSE,¹⁴
- *Employee productivity*, measured with the Stepford Presenteeism scale, SPS,¹⁵ (NB: this outcome was measured only among participants in any vocational activity (supported or competitive), and is hence not an *intention-to-treat* analysis)
- Self-reported at 6-month follow-up:
 - *Client satisfaction*, by *The Client Satisfaction Questionnaire*,¹⁶

All outcomes at 24-month follow-up will be reported later.

Changes to predefined outcomes

Before any analysis of any vocational outcomes, and before 24-month follow-up, we discarded the recurrent sick leave-outcome due to failure to find reasonable statistical measures.

Intervention delivery description

Pre-planned:

- Self-reported at 6-month follow-up: Any use of psychotherapy-like interventions, regardless of funding source, at general practitioners, psychiatrists, psychologists, coaches, psychotherapists or group therapy.

Post-hoc we decided to describe the following parameters:

- Delivered outside the study interventions:
 - Number of publicly funded consultations at general practitioners, psychiatrists and psychologist
 - Use of psychiatric admissions, out-patient consultations and emergency room
 - Use of vocational rehabilitation services: number courses and their duration.
- Delivered within the study interventions:
 - Number of contacts and duration of treatment course with health care staff
 - Number of roundtable-meetings, and their relative placement in the treatment
- Delivery across intervention placement:
 - Employment consultant consultations, meetings and virtual contacts

Interventions

The three intervention groups contained the following elements:

- INT – all interventions delivered as a part of the study, by IBBIS-teams:
 1. Best practice vocational rehabilitation
 2. Best practice mental health care
 3. Systematic integration of vocational rehabilitation and mental health care
- MHC – some interventions delivered by IBBIS-teams, some outside IBBIS:
 1. Standard vocational rehabilitation
 2. Best practice mental health care in IBBIS-teams
 3. No systematic integration of vocational rehabilitation and health care
- SAU – all interventions delivered outside IBBIS-teams:
 1. Standard vocational rehabilitation (vocational rehabilitation)
 2. Usual primary sector health care
 3. No systematic integration of vocational rehabilitation and health care

Description of content elements:

- Standard vocational rehabilitation (in SAU and MHC) was delivered by municipal job centres and private vocational rehabilitation providers, payed by the municipalities.
- Usual primary sector health care (in SAU) was provided by the in Denmark freely available general practitioners, and whatever treatment they might prescribe, deliver or refer to, granting either *just partly*- or non-subsidised psychotherapy, or fully subsidised treatment by a psychiatrist.
- Best practice mental health care (in MHC and INT), was delivered systematically and free of charge, adhering to Danish national guidelines for depression and/or anxiety, complying with NICE guidelines.¹¹ A stepped-care principle was applied, with initial step according to baseline symptom severity, and intensifying treatment in case of response absence or deterioration, identified through frequent symptom monitoring. Across diagnoses, principal steps were, from mildest to most severe cases:
 1. Psychoeducation
 2. Cognitive behavioural therapy (CBT) or medication
 3. CBT and medication
 4. Referral to hospital-based treatment

Psychoeducation (three sessions) and CBT (10-12 sessions) was delivered by *care managers* who were nurses, occupational-/physiotherapists or social workers, with ≥ 1 year of clinical psychiatric experience and CBT certified. They were on continuous supervision by a psychologist with CBT

specialty, and a psychiatrist for general case supervision. When medication was needed, this was prescribed by a general practitioner, but in close collaboration with an IBBIS psychiatrist.

- Best practice vocational rehabilitation (in INT) was informed by a literature search for existing evidence regarding most efficient vocational rehabilitation, suggested these key elements:
 - 1) Assessment of individual's work capacity and barriers for the individual's return to work, utilizing pre-existing questionnaires regarding work: return to work-readiness,¹² functioning,¹³ and self-efficacy;¹⁴
 - 2) Vocational rehabilitation plan tailored with, by and uniquely for the individual;
 - 3) Focus on early, graded return to work, at current workplace, inspired by the guidelines in the activation model¹⁵ and the intervention SHARP-at work;¹⁶
 - 4) Support in job finding focusing on best possible match (inspired by IPS, Individual Placement and Support);¹⁷
 - 5) Coordination with any relevant public authority;
 - 6) Involvement of relatives, if desired.

Vocational rehabilitation was delivered by employment consultants, concurrently working part-time in job centres, and specially trained to deliver in-study best practice vocational rehabilitation, while also being the participants' sickness benefit case manager.

- Systematic integration of vocational rehabilitation and mental health care (in INT) was performed using the principles of *relational coordination* stipulated by Gittels in 2006,¹⁸ entailing these integrational activities:
 - 1) Early in each course, readily after completion of both the employment consultant's vocational assessment and the care manager's first session, at least one physical meeting, coined *the roundtable-meeting*, with both the participant, the care manager and the employment consultant, forming a *joint plan*, encompassing both intervention types;
 - 2) co-location of all team-members;
 - 3) common training and
 - 4) joint supervision of care managers and employment consultants.

References

- 1 Sheehan D V., 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. In: *Journal of Clinical Psychiatry*. 1998; 22–33.
- 2 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**: 228–32.
- 3 Kessler RC, Adler L, Ames M, *et al.* The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. *Psychol Med* 2005; **35**: 245–56.
- 4 Cole MG, McCusker J, Elie M, Dendukuri N, Latimer E, Belzile E. Systematic detection and multidisciplinary care of depression in older medical inpatients: a randomized trial. *Can Med Assoc J* 2006; **174**: 38–44.
- 5 Gamble C, Krishan A, Stocken D, *et al.* Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA* 2017; **318**: 2337.
- 6 R: A Language and Environment for Statistical Computing. Vienna, Austria. 2019; **0**. <https://www.r-project.org/>.
- 7 European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP). European Medicines Agency - Adjustment for baseline covariates in clinical trials. Eur. Med. Agency. 2015.
- 8 Terluin B, van Marwijk HWJ, Adèr HJ, *et al.* The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006; **6**: 34.
- 9 Saboonchi F, Perski A, Grossi G. Validation of Karolinska Exhaustion Scale: Psychometric properties of a measure of exhaustion syndrome. *Scand J Caring Sci* 2013; **27**: 1010–7.
- 10 Moss-Morris R, Weinman J, Petrie KJ, Horne R, Cameron LD, Buick D. The revised Illness Perception Questionnaire (IPQ-R). *Psychol Health* 2002; **17**: 1–16.
- 11 Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004; **13**: 873–84.
- 12 Burckhardt CS, Anderson KL. The Quality of Life Scale (QOLS): reliability, validity, and utilization. *Health Qual Life Outcomes* 2003; **1**: 60.
- 13 Lagerveld SE, Blonk RWB, Brenninkmeijer V, Schaufeli WB. Return to work among employees with mental health problems: Development and validation of a self-efficacy questionnaire. *Work Stress* 2010; **24**: 359–75.
- 14 Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. In: Wright S, Johnston M, eds. *Measures in health psychology: A user's portfolio. Causal and control beliefs*. Windsor, England, 1995: 35–7.
- 15 Koopman C, Pelletier KR, Murray JF, *et al.* Stanford presenteeism scale: health status and employee productivity. *J Occup Environ Med* 2002; **44**: 14–20.
- 16 Attkisson C, Zwick R. The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann* 1982; **5**: 233–7.