Effectiveness of group-delivered therapy for insomnia

Statistical Analysis Plan (SAP)

Version 2

Valid from 06.03.2024

Administrative information

Sponsor name	Norwegian University of Science and Technology (NTNU)
Sponsor address	Postboks 8905, MTFS, 7491, Trondheim, Norway
EudraCT number / REC no	Regional Committee for Medical Research Ethics - Mid Norway ref. 465241
Trial title	Group-delivered cognitive behavioural therapy versus waiting list in the treatment of insomnia in primary care: study protocol for a pragmatic, multicentre randomized controlled trial
Trial registration number	ISRCTN16185698

SAP and protocol version

SAP version and date	This SAP is version 2.0, dated 06.03.2024
Protocol version	This document has been written based on information contained in the study protocol version 1.0, dated 09.01.2023

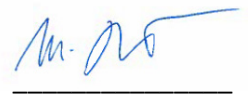
SAP revision history

Protocol version	SAP version	Section number changed	Description and reason for change	Date changed
1.0	2.0	3.4.2	Corrected a typographical error (changed “outcomes” to “objectives”) and specified that secondary objective IV will be analyzed at 24-months follow-up.	06.03.2024
1.0	1.0	NA	NA	25.01.2024

Roles and responsibilities

Main author:	Maria Hrozanova, Department of Public Health and Nursing, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.
Statistician:	Turid Follestad, Department of Clinical and Molecular Medicine, NTNU, Trondheim, Norway.
Project supervisor:	Eivind Schjelderup Skarpsno, Department of Public Health and Nursing, NTNU, Trondheim, Norway; Department of Neurology and Clinical Neurophysiology, St. Olavs Hospital, Trondheim, Norway.
Contributors to SAP:	Ingebrigt Meisingset, Department of Public Health and Nursing, NTNU; Unit for Physiotherapy Services, Trondheim Municipality, Trondheim, Norway.
	Håvard Kallestad, Department of Mental Health, NTNU, Trondheim, Norway; Department of Mental Health Care, St. Olavs Hospital, Trondheim, Norway.
	Ståle Pallesen, Department of Psychosocial Science, University of Bergen, Bergen, Norway; Norwegian Competence Center for Sleep Disorders, Haukeland University Hospital, Bergen, Norway.

Signatures



Maria Hrozanova
Main author

Turid Follestad
Statistician

Eivind Schjelderup Skarpsno
Project supervisor

Abbreviations

CACE	Complier-average causal effect
CBT-I	Cognitive behavioural therapy for insomnia
CPAP	Continuous positive airway pressure
ISI	Insomnia severity index
ITT	Intention-to-treat
NTNU	Norwegian University of Science and Technology
RCT	Randomized controlled trial
SAP	Statistical analysis plan
SD	Standard deviation

Table of contents

1. Introduction

This statistical analysis plan (SAP) should be read in conjunction with the study protocol entitled "Group-delivered cognitive behavioural therapy versus waiting list in the treatment of insomnia in primary care: study protocol for a pragmatic, multicentre randomized controlled trial" published in BMC Primary Care (Hrozanova et al., 2023). The information available here provides a more detailed description of the "Statistical analysis plan" section. The study was approved by the regional Committee for Medical Research Ethics- Mid Norway on the 13th of May 2022 (ID 465241).

The structure of this SAP follows the guidelines by Gamble et al. (2017) and the checklist available from http://lctc.org.uk/SAP-Statement. All analyses will be reported according to CONSORT statement for standardization and reproducibility of randomized controlled trials (RCT) (Cuschieri, 2019).

1.1. Purpose and scope of the statistical plan

This document details the proposed analysis of the main outcomes from the trial entitled "Group-delivered cognitive behavioural therapy versus waiting list in the treatment of insomnia in primary care". Any deviations from the analyses outlined in this SAP will be described and justified in all papers from the project, including the inclusion of any analyses suggested by journal editors and referees. Modifications will be carefully considered and, as much as possible, will follow the broad principles outlined here.

First and foremost, this SAP describes the analysis of the primary and secondary outcomes. Statistical principles for the planned moderating and mediating analyses are also described. Analyses of subsequent sub-studies (e.g., mixed-method process evaluation, see Hrozanova et al. (2023)) are expected to follow the broad principles of this SAP but are not described in detail here.

The details presented in this SAP will not prohibit accepted practices, such as data transformation prior to analysis. When possible, such data management and modelling decisions will be undertaken prior to revealing the treatment allocation. The final analysis strategy will be available on request when the principal papers are submitted.

1.2. Background and rationale

Insomnia is the most common sleep disorder in the general population and in clinical practice (Winkelman, 2015). Cognitive behavioural therapy for insomnia (CBT-I) is considered most effective for long-term alleviation of chronic insomnia (Qaseem et al., 2016; Riemann et al., 2017; Rios et al., 2019), but administration of CBT-I is limited by a lack of trained therapists and long waiting lists (Grandner & Chakravorty, 2017). To increase the availability of CBT-I, a group-delivered treatment based on the core CBT-I principles was developed by the Norwegian Health Directorate. The treatment has been implemented in several municipalities in Norway, but its effectiveness has never been evaluated. Likewise, factors that may influence the therapeutic response of the group-delivered CBT-I treatment (e.g., chronotype, sleep reactivity to stress, unhelpful beliefs and attitudes about sleep) have not been evaluated. Such factors are important to identify, as up to 40% of insomnia patients do not adequately respond to CBT-I (Galbiati et al., 2019). Insomnia patients with evening chronotype, unhelpful beliefs about sleep, short sleep, comorbidity, or high reactivity to stress may have a blunted response to CBT-I (Baron & Hooker, 2017; Blanken et al., 2019; Faaland et al., 2022; Montserrat Sánchez-Ortuño & Edinger, 2010). Adherence to the therapeutic regime may also predict outcome (Mellor et al., 2022).

1.3. Objectives

1.3.1. Primary objective

The primary objective is to investigate the effectiveness of group-delivered CBT-I in primary care on insomnia severity at 3 months post-treatment.

1.3.2. Secondary objectives

The secondary objectives include:

I. Investigate the effectiveness of group-delivered CBT-I on insomnia severity at 1 and 6-months post-treatment.
II. Investigate the effectiveness of group-delivered CBT-I on health-related quality of life, fatigue, mental distress, and sleep diary data at 1, 3 and 6-months post-treatment.
III. Investigate whether chronotype moderates the effectiveness of group-delivered CBT-I at 3 months post-treatment.
IV. Investigate the rates of sick leave, the use of relevant medications, and utilization of healthcare services at 12- and 24-months post-treatment.

1.3.3. Exploratory objectives

I. Investigate whether potential treatment moderators (e.g., sleep reactivity) influence the effectiveness of group-delivered CBT-I.
II. Conduct exploratory mediation analyses to identify mechanisms behind change in the primary and secondary outcomes.

2. Study design

2.1. Type of design

The study is a pragmatic, multicentre RCT.

2.2. Randomization and treatment assignment

Treatment allocation is based on block randomization stratified by centre. Participants are randomized based on a 2:1 ratio to either the intervention group (group-delivered CBT-I) or to a control group (waiting list). Randomization is performed by the project leader using a secure digital platform for multicentre clinical studies, eForsk, and is based on a computerized allocation sequence stratified by each of the 26 participating centres. The allocation sequence is generated by a third party, the Clinical Research Unit at the Faculty of Medicine and Health Sciences at NTNU.

2.3. Determination of sample size

In published RCTs investigating the effectiveness of group-delivered CBT-I on insomnia severity (Alessi et al., 2016; Bothelius et al., 2013; Chan et al., 2022; Sadler et al., 2018; Song et al., 2022), the average observed effect corresponded to a large Cohen's *d* effect size of 0.86. We chose the medium Cohen's *d* effect size of 0.50 for the sample size calculation as the current trial differs from earlier studies in that it employs fewer treatment sessions, allows the use of sleep medication, and as the control group is not prevented from seeking other forms of treatment during the RCT. A power analysis was carried out using a two-tailed t-test with 5% alpha level and 90% power to detect a medium Cohen's *d* effect size of 0.50, with an allocation ratio of 2:1 (G*Power, version 3.1.9.6) (Faul et al., 2007). The needed sample size was 192 participants. Accounting for a 30% attrition rate based on previous studies (Espie

et al., 2001; Morgan et al., 2003; Ong et al., 2008; Vincent & Lionberg, 2001), the needed sample size was 250 participants.

To have sufficient power for moderation analyses of chronotype, we additionally calculated the sample size for a two-tailed t-test between morning and evening chronotype comparing the difference in outcome between the CBT-I and control groups. We applied a 5% alpha level and 80% power to detect a small to medium Cohen's *d* between-group effect size of 0.40, an allocation ratio of 2:1, and accounting for a 30% attrition rate. The effect size of 0.40 was based on previous studies investigating the moderating effects of chronotype on the effectiveness of CBT-I (Faaland et al., 2022; Lien et al., 2019). The final sample size was 292 participants (98 in the control group, 194 in intervention group).

2.4. Framework

The study is designed as a multicentre, pragmatic RCT to test the effectiveness of group-delivered CBT-I compared to a waiting list on insomnia severity in Norwegian primary care.

2.5. Statistical interim analyses

No interim analyses are planned for the trial, as we do not expect any major adverse effects of the treatment, and as all intervention group participants are followed up during the treatment.

2.6. Stopping guidance

No stopping guidance is implemented in the trial.

2.7. Timing of final analysis

The primary and secondary outcomes will be assessed at 1-, 3- and 6-months follow-up. As such, all data from baseline, 1 month, 3 months, and 6 months will be exported, analysed, and published together. Data from 6-months follow-up will be available by the end of January 2024.

Long-term follow-up using national registry data at 12- and 24-months post-treatment will be accessed and analysed separately at the respective timepoints, but published together after 24-months follow-up.

2.7.1. Timing of outcome assessments

There are six measurement timepoints of outcome assessments in the trial:

1. T1: baseline measurements after randomization and prior to intervention start
2. T2: immediately post-intervention (4-5 weeks after T1)
3. T3: 3 months post-intervention (primary follow-up timepoint)
4. T4: 6 months post-intervention
5. T5: 1-year post-intervention (registry data only)
6. T6: 2 years post-intervention (registry data only)

Acceptable time frame for answering questionnaires at the respective timepoints is defined as half the time between each timepoint. For instance, the acceptable time frame of assessments at T1 is the first 2.5 weeks after participants received the questionnaires. For T2, this period amounts to 4 weeks. For T3 and T6, respectively, this period amounts to 6 weeks. The schedule of administering outcome assessments is provided in Table 2.

Table 2. Overview of the utilized instruments, measurement timepoints and objectives of their use in the randomized controlled trial.

Objective	Instrument	Description of recorded data	Timepoint
Demographic information and baseline patient characteristics	Self-rated questions, see 4.3	Various, see 4.3	T1
Primary outcome	Insomnia severity (Insomnia Severity Index [ISI])	7 items scored on a 5-point scale, sum score 0-28, higher values indicate worse insomnia	Screening, T1, T2, T3, T4
Secondary outcomes	Health-related quality of life (EuroQol EQ5D-5L)	5 items scored on a 5-point scale from 1 to 5, index score -0.285 (worst imaginable health)-1 (perfect health)	T1, T2, T3, T4
	Fatigue (Chalder Fatigue Scale)	13 items scored on a 4-point scale, sum score 0-39, higher values indicate worse fatigue	T1, T2, T3, T4
	Mental distress (Hopkins Symptoms Checklist)	5 items scored on a 4-point scale, sum score 0-20, higher values indicate worse mental distress	T1, T2, T3, T4
Moderator	Chronotype (Brief Horne-Östberg Morningness-Eveningness Questionnaire)	5 items, sum score 4-25, higher levels indicate higher levels of morningness.	T1, T2, T3, T4
	Sleep reactivity (Ford Insomnia Response to Stress Scale)	9 items scored on a 4-point scale, sum score 9-36, higher values indicate higher sleep reactivity	
Mediators	Unhelpful beliefs about sleep (Dysfunctional Beliefs and Attitudes about Sleep questionnaire-16)	16 items scored on a 11-point scale, sum score is an average of all scores (0-16), higher scores indicate stronger dysfunctional beliefs and attitudes about sleep	T1, T2, T3, T4
	Sleep-related self-efficacy	1 item scored on a 10-point scale, higher score indicates higher self-efficacy	T1, T2, T3, T4
Subjective assessment of sleep	7-day sleep diary	Various, see 5.2.1	T1, T2, T3, T4
Evaluation of medication use, sick leave, and healthcare utilization	Norwegian Patient Registry	Various	T5, T6
	Norwegian Prescription Database	Various	T5, T6
	National Insurance Administration	Various	T5, T6

3. Statistical principles

3.1. Confidence intervals and p-values

All p-values will be two-sided. For the primary outcome analysis, the significance level will be set to 0.05. We will have the problem of multiple comparisons in mind when we interpret results from the

secondary analyses. Estimates will be presented as mean differences or odds ratios, and their precision will be quantified with 95% confidence intervals.

3.2. Uptake, protocol deviations and protocol violations

3.2.1. Uptake

Adults interested in participating in the study filled out an online screening questionnaire, which was evaluated against inclusion and exclusion criteria by the research team. Adults who met criteria for participation were sent an online consent form. Those who consented were included in the study and randomized. Intervention participants were informed about the next available treatment group, and asked whether they could attend. At this point, some intervention participants withdrew as the time or place for the intervention did not suit them (e.g., especially due to time-scheduling conflicts, if treatment groups took place during the hours of the workday). These intervention participants were asked to answer baseline questions relating to sociodemographic and patient characteristics, but did not receive the full T1 questionnaire, or questionnaires at timepoints T2, T3, and T4. All others who were randomized were sent questionnaires at all timepoints unless they withdrew from the study at any point after baseline.

3.2.2. Adherence and Protocol deviations

Since a 2:1 randomization ratio was implemented, we expected to include twice as many participants in the intervention group than in the control group. However, more intervention participants were excluded due to their withdrawal after randomization and prior to T1 (described in section 3.2.1).

The group-delivered CBT-I intervention consists of four sessions. The first three sessions include new learning outcomes, while the fourth session is built around repetition of the learning outcomes and planning how to continue working with the core components of CBT-I after the course. No new knowledge is introduced in the fourth session. Therefore, participant adherence to the intervention is defined as having attended minimum the first three sessions of the intervention. Non-compliance with the intervention is defined as not having attended any sessions, only having attended 1 or 2 sessions, or 3 sessions that did not include the first three sessions of the intervention.

3.3. Analysis

3.3.1. Intention-to-treat (ITT) analysis (full analysis set)

The ITT principle will be used for analysing the effectiveness of group-delivered CBT-I, the study intervention. The primary analysis population includes all eligible, randomly assigned participants who had at least one outcome measurement. Participants who withdrew from the trial will be included in the analysis until the point they withdrew. This will be the full analysis set.

3.3.2. Complier-average causal effect analysis

A complier-average causal effect (CACE) analysis of the primary and secondary outcomes will be carried out to determine the impact of the treatment effect when accounting for non-compliance of the intervention. CACE is a measure of the causal effect of an intervention on the participants who received it as intended by the original group allocation. Non-compliance is defined in section 3.2.2. CACE analyses will be conducted after the ITT analyses are finalized, as there is a risk that they will not be blinded.

3.3.3. Subgroup analyses

We will conduct subgroup (i.e., interaction) analyses based on chronotype to investigate whether chronotype at baseline moderates treatment effectiveness of group-delivered CBT-I on insomnia severity at 3 months after baseline. Two sets of subgroups based on the Brief Horne-Östberg Morningness-Eveningness Questionnaire (Chelminski et al., 2000) may be defined:

- Morning type (score 18-25), neither type (score 12-17), evening type (score 4-11).
- Definitely morning type (score 22-25), moderately morning type (score 18-21), neither type (score 12-17), moderately evening type (score 8-11), definitely evening type (score 4-7).

The final subgroups will be established based on the final distribution of participants in the different subgroups. In order to ensure adequate power for the analyses, some of the above-defined subgroups may be merged, or percentiles may be utilized to create 3 or 5 equally large groups.

3.3.4. Exploratory subgroup analyses

We will conduct a number of exploratory subgroup (i.e., interaction) analyses. The following variables will be investigated as potential as moderators of treatment effectiveness of group-delivered CBT-I on insomnia severity at 3 months after baseline:

- Sleep reactivity. Subgroups will be based on the Ford Insomnia Response to Test (Drake et al., 2004), and will include high (scores ≥ 16) and low (scores < 16) sleep reactivity.
- Duration of insomnia, duration of prior insomnia treatment, and treatment group size. Subgroups will be based on percentiles of the given variables in the sample.
- Frequency of physical activity. Subgroups will be conceptualized as high physical activity (2-3 times per week and approximately every day) and low physical activity (never, less than once per week and once per week).
- Insomnia phenotypes. We will define two insomnia phenotypes based on sleep diary data at baseline: (1) insomnia with short sleep duration, (i.e., < 6h), and (2) insomnia with normal sleep duration (i.e., ≥ 6h). Moreover, since it is plausible that combinations of different nighttime symptoms provide the basis for insomnia subtypes, we will explore how different subtypes influence the effectiveness of the intervention. These will include: (1) insomnia with early morning awakenings, (2) insomnia with difficulties initiating sleep, and (3) insomnia with nighttime awakenings.

3.3.5. Exploratory mediating analyses

Exploratory analyses will be conducted to identify whether, and to what extent, the change in the following variables at 1 months after baseline mediated the effectiveness of group-delivered CBT-I on insomnia severity and secondary outcomes at 3 months after baseline. We will conduct the mediation analyses according to existing guidelines (Lee et al., 2021), using the following variables:

- Dysfunctional beliefs and attitudes about sleep. Quantified with the Dysfunctional Beliefs and Attitudes about Sleep Questionnaire (Morin et al., 2011), where higher scores indicate greater dysfunctional beliefs and attitudes about sleep.
- Sleep-related self-efficacy. Quantified with a custom self-rater item, where higher scores indicate greater sleep-related self-efficacy.

3.4. Allocation concealment, blinding and order of analysis

3.4.1. Allocation concealment

Treatment allocation to the intervention and control groups is described in section 2.1. The project leader, who performs the treatment allocation, is blinded to the block sizes and the allocation sequence, and will thus not influence the allocation process in any way. Participants and the intervention administrators will not be masked to the participants' group assignment.

3.4.2. Blinding in statistical analyses and interpretation of results

Upon completion of data collection, de-identified patient data will be exported. The participant identification key will be confidential and unavailable to the project group at the time of analysis. The treatment allocation will be concealed by a researcher not involved in the project, who will change the treatment allocation to groups A and B. The researcher will then run the outlined statistical analyses for both primary and secondary objectives (see section 1.3.1 and 1.3.2; note that secondary objective IV will be analyzed at 24-months follow-up), which will be presented in table format without the number of participants belonging to each group. These results will be shared with the project group, who will interpret the results blinded to treatment allocation. Thereafter, the group allocation concealment will be removed.

4. Presentation of study population

4.1. Screening data, eligibility, recruitment, and follow-up

See the published protocol (Hrozanova et al., 2023) for detailed description of screening procedures, eligibility, and recruitment. Once the data collection is finalized, a CONSORT flow diagram will be created to show the number of participants that were screened, excluded (incl. reasons for exclusion), randomized and included, how many dropped out, how many participated in the different measurement timepoints and finally, how many finalized the data collection. Reasons for withdrawal and loss to follow-up will be specified when possible. See the CONSORT flow diagram in Figure 1.

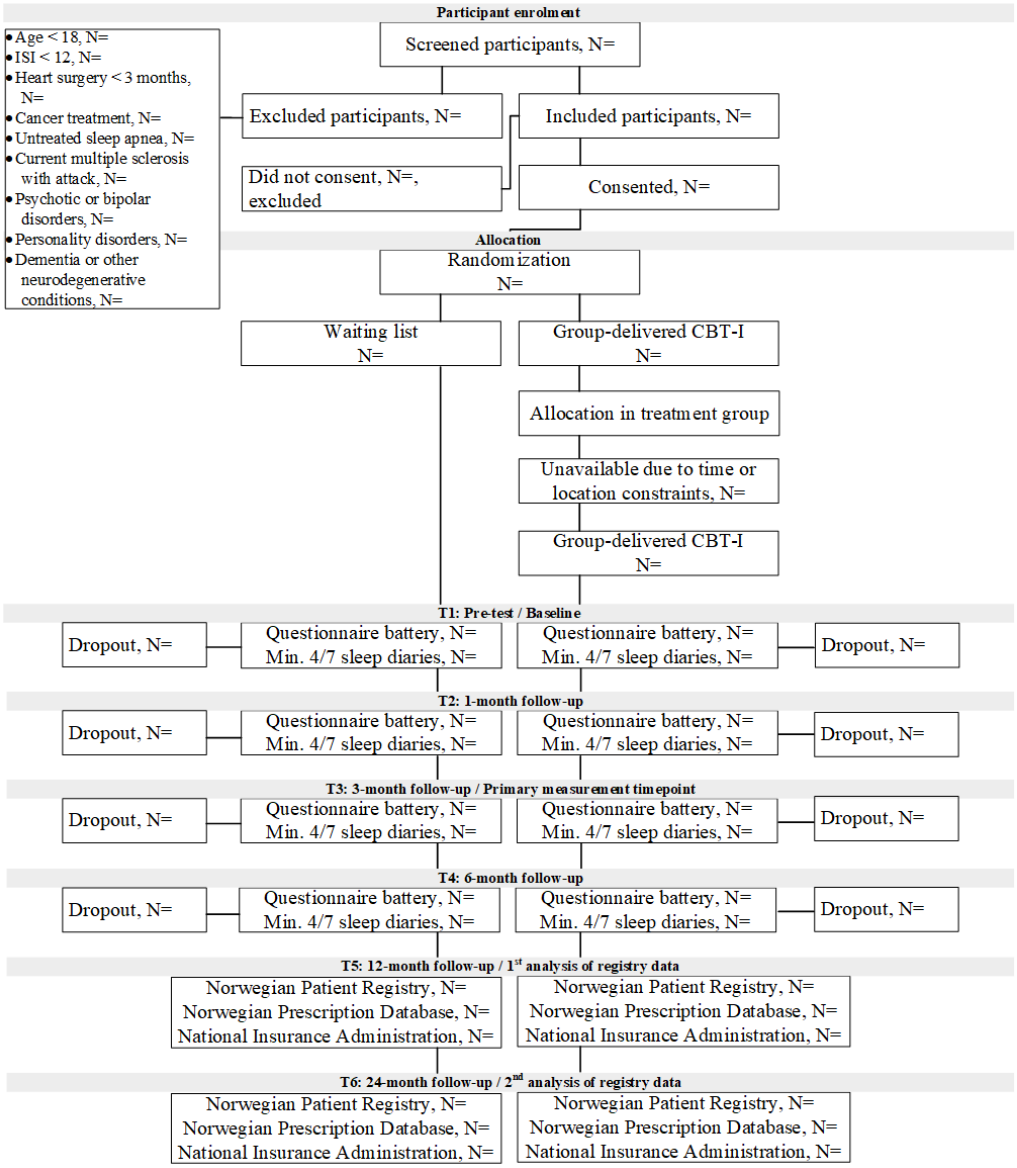


Figure 1. CONSORT flow diagram showing screening, eligibility, recruitment and follow-up data.

4.2. Withdrawal/Follow-up

Participants in the intervention group had an opportunity to withdraw from the study when they were offered a place in one of the upcoming group-delivered CBT-I courses in their municipality. Since these courses were ongoing and planned by the participating centres, it was possible that the time or place of the intervention did not suit the participants' schedules. In that case, the participants informed the project leader or the contact person at their centre, and withdrew from the study. Not all baseline characteristics were collected for these participants, but most of these participants answered questions on education, employment status, and physical health. This level of withdrawal will be presented in a figure or in text. We will compare these participants to those who accepted the offer of treatment based on their screening data (e.g., insomnia severity, sex, age).

Participants who withdrew from the study after the completion of baseline measurements were considered dropouts. Unless they explicitly asked to have their data removed, the collected data of these participants will be used up to the timepoint of drop out. This level of withdrawal will be presented in a figure or in text.

4.3. Baseline patient characteristics

The following baseline patient characteristics will be presented for both the intervention and control groups separately, and combined for the whole study population. In addition, we will present participants' baseline characteristics by intervention compliance status. Continuous data will be summarized by mean (SD) or median (percentiles), as appropriate due to their distributions, and categorical data by n (%).

Sociodemographic factors

- Age (years)
- Sex (female)
- Education level: Not applicable; Vocational education; Highschool; College 2 years; Bachelor; Master; PhD
- Employment status: Yes, part time; Yes, full time; No
    - If employed:
        - Three-shift working arrangement: Yes; No
        - Work ability
    - If unemployed:
        - Main reason for unemployment: Unable to find a job; Retired; Illness or disability; Temporarily laid off; Parental leave; Student; Stays at home; Wanted time off for a while; Waiting to start a new job; Other reasons

Physical health

- Bodily pain in the last week: 0 (no pain) – 10 (worst possible pain)
- Body mass index (kg/m^2)
- Frequency of physical activity: Never; Less than once per week; Once per week; 2-3 times per week; Approximately every day
    - Exercise intensity: Without getting out of breath or sweaty; To get out of breath or sweaty; I go almost all out
    - Exercise duration: < 15 minutes; 15-29 minutes; 30-60 minutes; > 60 minutes

Substance use

- Alcohol intake frequency: Never; Monthly or less; 2-4 times per month; 2-3 times per week; 4+ times per week
    - Typical consumption of alcohol in units: 1-2; 3-4; 5-6; 7-9; 10+
- Cannabis use in the last 12 months: Yes / No
- Use of other narcotic substances than cannabis in the last 12 months: Yes / No

Sleep-related factors

- Duration of current sleep problems (years)
- Earlier treatment for current sleep problems: Yes / No
    - Type of treatment: Sleep medication; Talk therapy; Web-based therapy; Self-help; Meditation, mindfulness, relaxation; Massage; Acupuncture; Homeopathy; Hypnosis; Light therapy, chronotherapy; Continuous positive airway pressure (CPAP), mandibular advancement device; Other
    - Duration of earlier treatment (hours)
- Sleepwalking: Never / rarely; Sometimes; At least 3 times per week
- Sleep talking: Never / rarely; Sometimes; At least 3 times per week

- Discomfort or crawling sensation in the legs: Never / rarely; Sometimes; At least 3 times per week
- Delayed sleep phase: Yes / No

We will report baseline values for primary and secondary outcomes (see 5.1 & 5.2). Descriptive statistics will be presented depending on the type and distribution of variables.

5. Analysis

5.1. Primary outcome

5.1.1. Definition of primary outcome measures

The primary outcome measure is defined as the participants' insomnia severity at 3 months after baseline (i.e., start of treatment or waiting period). Insomnia severity is measured by the Insomnia Severity Index (ISI) (Morin et al., 2011) which consists of 7 items, each measured on a 5-point scale. The composite score ranges from 0 to 28, where higher values indicate higher insomnia severity.

Insomnia severity is also assessed at 1 month and 6 months after baseline. These measurement timepoints will be analysed and presented as secondary analyses, described below.

5.1.2. Analysis of primary outcome

The primary outcome will be analysed using the ITT principle, outlined in section 3.3.1. A three-level linear mixed effect model will be used to estimate the mean differences between the two groups in change in insomnia severity from baseline and 3-month follow-up with 95% confidence intervals. The linear mixed effect model will be based on data from all time points even though 3-months is the primary measurement timepoint. Random effects will include participant and centre. Fixed effects will include time, group, and time–group interaction. Baseline covariates (e.g., sex, age, education level, occupational status, duration of insomnia symptoms, insomnia severity) will be included if inspection of descriptive baseline characteristics indicate that the intervention and control groups differ according to these variables, which might be the case due to the possibility of withdrawal prior to intervention start (see 4.2). Any participants with missing outcomes at the primary measurement timepoint, i.e., 3 months after baseline, will not be excluded; rather, the mixed effects model will implicitly account for missing data. The results will be unbiased as long as the data are missing at random. Model assumptions will be reviewed and if they are violated, bootstrapping, logarithmic transformations, or non-parametric testing will be used as appropriate.

5.1.3. Sensitivity analyses

If the groups show differences in attrition rates at 3 months follow-up, prespecified sensitivity analyses will examine the robustness of primary outcome results to different assumptions regarding missing data. Specifically, if there is a substantial amount of missing data at the primary follow-up time points, pattern-mixture models will be used to investigate the possible influence of missing not at random.

5.2. Secondary outcomes

5.2.1. Definition of secondary outcome measures

Secondary outcome measures are defined as the following:

Insomnia severity:

- Participants' insomnia severity at 1 and 6 months after baseline, measured on a 5-point scale using the ISI (Morin et al., 2011). Higher values indicate higher insomnia severity.

- Participants' clinically relevant change in insomnia severity using the ISI (Morin et al., 2011) at 1, 3 and 6 months after baseline. A clinically relevant change will be defined as a 6-point decrease in insomnia severity from the participants' baseline score.

Health-related factors at 1, 3 and 6 months after baseline:

- Participants' health-related quality of life measured on a 5-point Likert scale using the EuroQol EQ5D-5L questionnaire (Herdman et al., 2011), where higher values indicate lower health-related quality of life. The five items will be transformed into an index value for health status using the UK value set. The index score ranges between − 0.285 (worst imaginable health state) and 1 (perfect health).
- Participants' fatigue measured on a 4-point Likert scale using the Chalder Fatigue Scale (Chalder et al., 1993), where higher values indicate higher fatigue.
- Participants' mental distress measured on a scale from 1.0 to 4.0 using the Hopkins Symptom Check List (Schmalbach et al., 2021), where higher values indicate higher mental distress.

Participants' subjective assessment of sleep patterns as assessed by an adapted consensus 7-day sleep diary (Carney et al., 2012) at 1, 3 and 6 months after baseline. We will calculate an average of each of the following variables which have been reported over at least 4 of the 7 consecutive days, respectively:

- Overall daytime functioning measured on a 5-point Likert scale, where higher values indicate worse daytime functioning.
- Napping frequency measured on a binary scale (Yes / No), and napping length assessed in minutes.
- Intake of prescription and non-prescription sleep medication measured on a binary scale (Yes / No).
- Sleep onset latency measured in minutes.
- Frequency and duration of awakenings measured in minutes. Sleep efficiency will be calculated as a ratio of time asleep / time in bed and expressed in %.
- Overall subjective sleep quality measured on a 5-point Likert scale, where higher scores indicate better subjective sleep quality.
- Sleep duration in minutes will be calculated by determining the number of minutes between sleep onset and final sleep offset, subtracting the total duration of nighttime awakenings.

Data from the Norwegian Patient Registry, Norwegian Prescription Database and National Insurance Administration will be accessed at 12- and 24-months after baseline and will contain data for the whole follow-up period. The following will be accessed and analysed:

- National Insurance Administration, assessment of sick leave. In the National Social Security System Registry managed by the National Insurance Administration, all individuals in Norway receiving any form of benefits are registered with their social security number. The registry includes data on use of medical benefits like sick-leave payments, sick leave certificates, work assessment allowance and disability pensions. During the respective follow-up periods of 12- and 24-months, we will access and analyse the following sick-leave outcomes (all-cause as well as due to musculoskeletal and mental disorders): total number of days on sick leave during the follow-up period and risk of long-term sick leave (≥ 31 consecutive days). Time on graded sick leave will be transformed to whole workdays.
- Norwegian Prescription Database, assessment of relevant prescribed medication. The Norwegian Prescription Database contains data about dispensed drugs in Norway, which is

based on prescriptions to individuals with a valid social security number. A user is defined as a person who has had at least one prescription dispensed in a pharmacy during a defined period. We will obtain data on use of antidepressants, antipsychotics, antiepileptics, benzodiazepines, z-hypnotics, antihistamines, and central nervous system stimulants. For sleep medications, we will obtain data on specific drugs, including those with active ingredients zopiclone, nitrazepam, melatonin, clomethiazole and others. For these medication groups and specific medications, we will assess the proportion of users, i.e., participants that obtain prescriptions of these medications during 12- and 24-months of follow-up, and the defined daily dose of medication use.

- Norwegian Patient Registry, assessment of healthcare resource utilization. The Norwegian Patient Registry is managed by the Norwegian Health Directorate and contains health information about all patients who have received treatment, or who are waiting for treatment in the Norwegian specialist healthcare service. We will obtain data on the referrals to sleep and mental healthcare specialists and rate of sleep diagnoses at 12- and 24-months post-treatment.

5.2.2. Analysis of secondary outcomes

For continuous secondary outcomes assessed at multiple follow-up timepoints, similar three-level linear mixed effect models as for the primary outcome will be utilized. Binary secondary outcomes (e.g., clinically relevant change in insomnia severity) assessed at multiple follow-up timepoints will be analysed using three-level logistic mixed effect models. The mixed effect models will be used to estimate the mean differences or odds ratios between baseline and 3 months follow-up in secondary outcomes with 95% confidence intervals between the two groups. Random effects will include participant and centre. For both continuous and binary secondary outcomes, model assumptions will be reviewed and if they are violated, bootstrapping will be used as appropriate.

CACE analyses for primary and secondary outcomes at 1, 3 and 6 months after baseline:

The association between baseline characteristics and treatment compliance (see definition in section 3.2.2) will investigate factors associated with compliance. Baseline characteristics will be reported for control, complier, and non-complier groups.

The CACE estimates of the difference between the mean ISI score for the compliers (i.e., participants with sufficient adherence to treatment) in the intervention group compared to the would-be-compliers in the control group at the primary measurement timepoint will be obtained. An unbiased estimate of the effect of intervention compliance on the treatment effect will be produced (Peugh et al., 2017). CACE estimates of the primary outcome (i.e., ISI at 3 months), will be obtained using the instrumental variable approach. Further models will be fitted with baseline covariates used in primary and secondary outcome analyses. The impact of re-defining compliance as attendance at a minimum of 1 and 2 sessions, and any 3 sessions other than the first three of the intervention will be assessed. CACE estimates of treatment effect will be reported in addition to the ITT estimates. The following CACE analyses will be carried out:

- The influence of intervention compliance on the treatment effect assessed by insomnia severity at 1, 3 and 6 months after baseline, measured using the ISI (Morin et al., 2011).
- The effect of intervention compliance on the clinically relevant change in insomnia severity (defined above) using the ISI (Morin et al., 2011) at 1, 3 and 6 months after baseline.
- For the primary outcome, an additional CACE analysis will be employed to obtain adjusted estimates for the difference in insomnia severity for participants who completed the first three sessions of the intervention vs. participants who completed all sessions of the intervention.

- CACE analyses for secondary outcomes outlined above will be carried out in line with the same principles.

5.2.3. Sensitivity analyses

For the CACE analyses, sensitivity analyses will be carried out by adjusting the definition of compliance to attending at least two or at least three sessions, and multiple imputation will be done on the primary CACE analysis as a sensitivity analysis to assess the effect of missing data. An additional sensitivity analysis will also include those intervention-group participants who were excluded prior to baseline measurements at T1, and who answered the baseline questions relating to sociodemographic and patient characteristics (described in section 3.2.1). We will compare these participants' sociodemographic and patient characteristics to the intervention-group participants included in the ITT analyses. In addition, we will repeat the primary outcome analysis including these participants, substituting their missing insomnia severity scores at baseline with those from the screening questionnaire.

5.3. Subgroup analyses and mediator analyses

Investigation of potential treatment moderators at baseline influencing the effectiveness of group-delivered CBT-I on insomnia severity at 3 months after baseline will be based on the following statistical principles. Chronotype is used as an example, as the final sample size in the RCT was calculated for this analysis. Other potential treatment moderators, outlined in section 3.3.3 will be analysed according to the same statistical principles.

- Chronotype at baseline was measured by 5 items using the reduced Horne-Östberg-Morningness-Eveningness Questionnaire (Chelminski et al., 2000), and categorized into evening type (scores 4-11), morning type (scores 18-25) and intermediate type (scores 12-17). Analysed with three-level linear mixed effect models, random effects will include participant and centre, while fixed effects will include the predictors time, group, and chronotype variables in addition to insomnia severity, the outcome. We will investigate the main effect of time and chronotype, the two-way interaction between group and time, and time and chronotype, as well as the three-way interaction (group * time * chronotype). As for the analysis of primary outcome, baseline covariates (e.g., sex, age, education level, occupational status, duration of insomnia symptoms, insomnia severity) will be included if inspection of descriptive baseline characteristics indicate that the intervention and control groups differ according to these variables, which might be the case due to the possibility of withdrawal prior to intervention start (see 4.2).

Exploratory mediator analyses, outlined in section 3.3.4, will be based on existing statistical principles (Kraemer et al., 2002; Lee et al., 2021), and conducted in the following way:

- We will assess whether dysfunctional beliefs and attitudes about sleep and sleep-related self-efficacy mediate the effectiveness of the intervention on changes in insomnia severity by using novel and recommended analytical approaches within a causal framework. We will apply four way decomposition analysis, which breaks down the total effect of the exposure on the outcome into components due to mediation alone, to interaction alone, to both mediation and interaction, and to neither mediation nor interaction (VanderWeele, 2014). The results from the analyses will be presented both as beta coefficient, but also as proportion of the changes in the primary outcomes that are explained by the mediator alone after adjusted for the other effects (interaction and both mediation and interaction).

5.4. Missing data

The analysis of the primary outcome will use linear mixed models, incorporating all available data from each participant with at least one outcome measurement. This method should provide unbiased estimates of the effectiveness of group-delivered CBT-I under the assumption that any missing data is missing at random.

A breakdown of participants with missing data for each outcome at each timepoint will be presented, in addition to the baseline characteristics of the entire sample, and those included in the primary analysis.

5.5. Statistical software

Statistical analyses will be carried out using R.

6. References

Alessi, C., Martin, J. L., Fiorentino, L., Fung, C. H., Dzierzewski, J. M., Tapia, J. C. R., Song, Y., Josephson, K., Jouldjian, S., & Mitchell, M. N. (2016). Cognitive Behavioral Therapy for Insomnia in Older Veterans Using Nonclinician Sleep Coaches: Randomized Controlled Trial. *Journal of the American Geriatrics Society*, *64*(9), 1830-1838. https://doi.org/10.1111/jgs.14304

Baron, K. G., & Hooker, S. (2017). Next steps for patients who fail to respond to cognitive behavioral therapy for insomnia (CBT-I): the perspective from behavioral sleep medicine psychologists. *Current Sleep Medicine Reports*, *3*(4), 327-332.

Blanken, T. F., Benjamins, J. S., Borsboom, D., Vermunt, J. K., Paquola, C., Ramautar, J., Dekker, K., Stoffers, D., Wassing, R., Wei, Y., & Van Someren, E. J. W. (2019). Insomnia disorder subtypes derived from life history and traits of affect and personality. *The Lancet Psychiatry*, *6*(2), 151-163. https://doi.org/10.1016/S2215-0366(18)30464-4

Bothelius, K., Kyhle, K., Espie, C. A., & Broman, J. E. (2013). Manual-guided cognitive-behavioural therapy for insomnia delivered by ordinary primary care personnel in general medical practice: a randomized controlled effectiveness trial. *J Sleep Res*, *22*(6), 688-696. https://doi.org/10.1111/jsr.12067

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., & Morin, C. M. (2012). The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep*, *35*(2), 287-302. https://doi.org/10.5665/sleep.1642

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. *J Psychosom Res*, *37*(2), 147-153. https://doi.org/10.1016/0022-3999(93)90081-p

Chan, N. Y., Lam, S. P., Zhang, J., Chan, J. W. Y., Yu, M. M. W., Suh, S., Yang, C. M., Okajima, I., Li, A. M., Wing, Y. K., & Li, S. X. (2022). Efficacy of Email-delivered Versus Face-to-face Group Cognitive Behavioral Therapy for Insomnia in Youths: A Randomized Controlled Trial. *J Adolesc Health*. https://doi.org/10.1016/j.jadohealth.2021.11.005

Chelminski, I., Petros, T. V., Plaud, J. J., & Ferraro, F. R. (2000). Psychometric properties of the reduced Horne and Ostberg questionnaire. *Personality and Individual Differences*, *29*(3), 469-478. https://doi.org/Doi 10.1016/S0191-8869(99)00208-1

Cuschieri, S. (2019). The CONSORT statement. *Saudi J Anaesth*, *13*(Suppl 1), S27-S30. https://doi.org/10.4103/sja.SJA_559_18

Drake, C., Richardson, G., Roehrs, T., Scofield, H., & Roth, T. (2004). Vulnerability to stress-related sleep disturbance and hyperarousal. *Sleep*, *27*(2), 285-291.

Espie, C. A., Inglis, S. J., Tessier, S., & Harvey, L. (2001). The clinical effectiveness of cognitive behaviour therapy for chronic insomnia: implementation and evaluation of a sleep clinic in general medical practice. *Behav Res Ther*, *39*(1), 45-60. https://doi.org/10.1016/s0005-7967(99)00157-6

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Faaland, P., Vedaa, O., Langsrud, K., Sivertsen, B., Lydersen, S., Vestergaard, C. L., Kjorstad, K., Vethe, D., Ritterband, L. M., Harvey, A. G., Stiles, T. C., Scott, J., & Kallestad, H. (2022). Digital cognitive behaviour therapy for insomnia (dCBT-I): Chronotype moderation on intervention outcomes. *J Sleep Res*, e13572. https://doi.org/10.1111/jsr.13572

Galbiati, A., Sforza, M., Fasiello, E., Castronovo, V., & Ferini-Strambi, L. (2019). Impact Of Phenotypic Heterogeneity Of Insomnia On The Patients' Response To Cognitive-Behavioral Therapy For Insomnia: Current Perspectives. *Nat Sci Sleep*, *11*, 367-376. https://doi.org/10.2147/nss.S198812

Gamble, C., Krishan, A., Stocken, D., Lewis, S., Juszczak, E., Dore, C., Williamson, P. R., Altman, D. G., Montgomery, A., Lim, P., Berlin, J., Senn, S., Day, S., Barbachano, Y., & Loder, E. (2017). Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*, *318*(23), 2337-2343. https://doi.org/10.1001/jama.2017.18556

Grandner, M. A., & Chakravorty, S. (2017). Insomnia in Primary Care: Misreported, Mishandled, and Just Plain Missed. *Journal of clinical sleep medicine*, *13*(8), 937-939. https://doi.org/10.5664/jcsm.6688

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*, *20*(10), 1727-1736. https://doi.org/10.1007/s11136-011-9903-x

Hrozanova, M., Meisingset, I., Kallestad, H., Pallesen, S., Nordstoga, A. L., & Skarpsno, E. S. (2023). Group-delivered cognitive behavioural therapy versus waiting list in the treatment of insomnia in primary care: study protocol for a pragmatic, multicentre randomized controlled trial. *BMC Prim Care*, *24*(1), 61. https://doi.org/10.1186/s12875-023-02018-4

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry*, *59*(10), 877-883. https://doi.org/10.1001/archpsyc.59.10.877

Lee, H., Cashin, A. G., Lamb, S. E., Hopewell, S., Vansteelandt, S., VanderWeele, T. J., MacKinnon, D. P., Mansell, G., Collins, G. S., Golub, R. M., McAuley, J. H., group, A. G., Localio, A. R., van Amelsvoort, L., Guallar, E., Rijnhart, J., Goldsmith, K., Fairchild, A. J., Lewis, C. C., . . . Henschke, N. (2021). A Guideline for Reporting Mediation Analyses of Randomized Trials and Observational Studies: The AGReMA Statement. *JAMA*, *326*(11), 1045-1056. https://doi.org/10.1001/jama.2021.14075

Lien, M., Bredeli, E., Sivertsen, B., Kallestad, H., Pallesen, S., Smith, O. R. F., Faaland, P., Ritterband, L. M., Thorndike, F. P., & Vedaa, O. (2019). Short and long-term effects of unguided internet-based cognitive behavioral therapy for chronic insomnia in morning and evening persons: a post-hoc analysis. *Chronobiol Int*, *36*(10), 1384-1398. https://doi.org/10.1080/07420528.2019.1647435

Mellor, A., Kavaliotis, E., Mascaro, L., & Drummond, S. P. A. (2022). Approaches to the assessment of adherence to CBT-I, predictors of adherence, and the association of adherence to outcomes: A systematic review. *Sleep Med Rev*, *63*, 101620. https://doi.org/10.1016/j.smrv.2022.101620

Montserrat Sánchez-Ortuño, M., & Edinger, J. D. (2010). A penny for your thoughts: patterns of sleep-related beliefs, insomnia symptoms and treatment outcome. *Behav Res Ther*, *48*(2), 125-133. https://doi.org/10.1016/j.brat.2009.10.003

Morgan, K., Thompson, J., Dixon, S., Tomeny, M., & Mathers, N. (2003). Predicting longer-term outcomes following psychological treatment for hypnotic-dependent chronic insomnia. *J Psychosom Res*, *54*(1), 21-29. https://doi.org/10.1016/s0022-3999(02)00569-x

Morin, C. M., Belleville, G., Belanger, L., & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*, *34*(5), 601-608. https://doi.org/10.1093/sleep/34.5.601

Ong, J. C., Kuo, T. F., & Manber, R. (2008). Who is at risk for dropout from group cognitive-behavior therapy for insomnia? *Journal of Psychosomatic Research*, *64*(4), 419-425. https://doi.org/10.1016/j.jpsychores.2007.10.009

Peugh, J. L., Strotman, D., McGrady, M., Rausch, J., & Kashikar-Zuck, S. (2017). Beyond intent to treat (ITT): A complier average causal effect (CACE) estimation primer. *J Sch Psychol*, *60*, 7-24. https://doi.org/10.1016/j.jsp.2015.12.006

Qaseem, A., Kansagara, D., Forciea, M. A., Cooke, M., Denberg, T. D., & Clinical Guidelines Committee of the American College of, P. (2016). Management of Chronic Insomnia Disorder in Adults: A Clinical Practice Guideline From the American College of Physicians. *Ann Intern Med*, *165*(2), 125-133. https://doi.org/10.7326/M15-2175

Riemann, D., Baglioni, C., Bassetti, C., Bjorvatn, B., Dolenc Groselj, L., Ellis, J. G., Espie, C. A., Garcia-Borreguero, D., Gjerstad, M., Goncalves, M., Hertenstein, E., Jansson-Frojmark, M., Jennum, P. J., Leger, D., Nissen, C., Parrino, L., Paunio, T., Pevernagie, D., Verbraecken, J., . . . Spiegelhalder, K. (2017). European guideline for the diagnosis and treatment of insomnia. *J Sleep Res*, *26*(6), 675-700. https://doi.org/10.1111/jsr.12594

Rios, P., Cardoso, R., Morra, D., Nincic, V., Goodarzi, Z., Farah, B., Harricharan, S., Morin, C. M., Leech, J., Straus, S. E., & Tricco, A. C. (2019). Comparative effectiveness and safety of pharmacological and non-pharmacological interventions for insomnia: an overview of reviews. *Systematic Reviews*, *8*(1), 281. https://doi.org/10.1186/s13643-019-1163-9

Sadler, P., McLaren, S., Klein, B., Harvey, J., & Jenkins, M. (2018). Cognitive behavior therapy for older adults with insomnia and depression: a randomized controlled trial in community mental health services. *Sleep*, *41*(8). https://doi.org/ARTN zsy10410.1093/sleep/zsy104

Schmalbach, B., Zenger, M., Tibubos, A. N., Kliem, S., Petrowski, K., & Brahler, E. (2021). Psychometric Properties of Two Brief Versions of the Hopkins Symptom Checklist: HSCL-5 and HSCL-10. *Assessment*, *28*(2), 617-631. https://doi.org/10.1177/1073191119860910

Song, Y., Kelly, M. R., Fung, C. H., Dzierzewski, J. M., Grinberg, A. M., Mitchell, M. N., Josephson, K., Martin, J. L., & Alessi, C. A. (2022). Change in Dysfunctional Sleep-Related Beliefs is Associated with Changes in Sleep and Other Health Outcomes Among Older Veterans With Insomnia: Findings From a Randomized Controlled Trial. *Annals of Behavioral Medicine*, *56*(1), 35-49. https://doi.org/10.1093/abm/kaab030

Twisk, J., Bosman, L., Hoekstra, T., Rijnhart, J., Welten, M., & Heymans, M. (2018). Different ways to estimate treatment effects in randomised controlled trials. *Contemporary Clinical Trials Communications*, *10*, 80-85. https://doi.org/10.1016/j.conctc.2018.03.008

VanderWeele, T. J. (2014). A unification of mediation and interaction: a 4-way decomposition. *Epidemiology*, *25*(5), 749-761. https://doi.org/10.1097/EDE.0000000000000121

Vincent, N., & Lionberg, C. (2001). Treatment preference and patient satisfaction in chronic insomnia. *Sleep*, *24*(4), 411-417. https://doi.org/10.1093/sleep/24.4.411

Winkelman, J. W. (2015). Clinical practice. Insomnia Disorder. *The New England Journal of Medicine*, *373*(15), 1437-1444. https://doi.org/10.1056/NEJMcp1412740