SUPPLEMENT 3

Title: *Protocol for a multi-center randomized controlled trial of digital cognitive behavior therapy for insomnia compared with digital patient education about insomnia in individuals referred to secondary mental health services in Norway.*

Håvard Kallestad[1,2,3], Simen B. Saksvik[2], Øystein Vedaa[2,4], Knut Langsrud[2,3], Gunnar Morken[1,2,3], Stian Lydersen[2], Melanie R. Simpson[5], Signe K. Dørheim[6], Bjørn Holmøy[7], Sara G. Selvik[8], Kristen Hagen[9], Tore C. Stiles[10], Allison G. Harvey[11], Frances Thorndike[12], Lee Ritterband[13], Børge Sivertsen[2,4,14], Jan Scott[2,15].

[1]Department of Research and Development, St. Olavs University Hospital, Trondheim, Norway
[2]Department of Mental Health, Norwegian University of Science and Technology, Trondheim, Norway
[3]Department of Østmarka, St. Olavs University Hospital, Trondheim, Norway
[4]Department of Health Promotion, Norwegian Institute of Public Health, Bergen, Norway
[5]Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway
[6]Department of Mental Health Care, Stavanger University Hospital, Stavanger, Norway
[7]Deparment of Mental Health Care, Akershus University Hospital, Follo, Norway
[8]Department of Mental Health Care, Namsos Hospital, Namsos, Norway
[9]Department of Mental Health Care, Molde Hospital, Molde, Norway
[10]Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway
[11]Department of Psychology, University of California, Berkeley, USA
[12]Pear Therapeutics, LLC, USA
[13]Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, USA
[14]Department of Research and Innovation, Helse-Fonna HF, Haugesund, Norway
[15]Institute of Neuroscience, Newcastle University, Newcastle, England

Corresponding author:
Author 1 (Håvard Kallestad, havard.kallestad@ntnu.no) as shown in Scholar One

Håvard Kallestad
St. Olavs University Hospital
Department of Research and Development
PO Box 3250 Sluppen
7006 Trondheim
Norway

Outcome assessments

Sleep measures:

Primary outcome measure:

The Insomnia Severity Index (ISI)[31] is a 7-item questionnaire assessing the severity of insomnia symptoms the last 14 days. Each item is rated on a 0 to 4 rating scale with higher scores indicating more severe symptoms. The ISI has good psychometric properties and is validated for online use.[32] Range is 0-28 with higher values represent higher levels of insomnia symptom severity. A cut-off of 7 or below is used to indicate normal sleep and used to indicate remission after the interventions, and a reduction of 8 points during the interventions is used to indicate response to the intervention.[31]

Secondary outcome measures:

Other measures of sleep and chronotype:

Consensus Sleep Dairy: Prospective daily sleep-wake patterns will be assessed with the consensus sleep diary[33] which will be completed by the participant at baseline and each follow-up point for at least 10 out of 14 consecutive days. Individuals will be asked to record their bed-time, sleep onset latency, wake after sleep onset, number of nocturnal awakenings, time of final awakening in the morning, and rise-time, in addition sleep quality will be rated on a 5-point scale (with a higher score indicating better sleep quality).

Reduced (also known as Brief) Morningness – Eveningness Questionnaire (rMEQ): The rMEQ is a widely used measure of chronotype i.e. time preference for daily activities, including bedtimes. The rMEQ has five items yielding scores from 4 to 25, with lower scores indicating a preference for "eveningness" and higher scores indicating a preference for "morningness". Scores can be categorized as: definitely evening type (score <8), moderately evening type (score 8-11), neither type (score 12-17), moderately evening type (score 18-21), and definitely morning type (score >21).[34]

Bergen Insomnia Scale[35]*:* comprises six items that assesses symptoms of insomnia based on the insomnia criteria found in the Diagnostic and Statistical Manual of Mental Disorders-IV-TR (American Psychiatric Association).[36]

Level of psychopathology & functional impairment:

The Outcome Questionnaire – 45.2 (OQ-45.2) is a 45 item self-report scale assessing mental health status. The scale is specifically designed for assessing patient progress throughout therapy and is widely used clinically, and as an outcome measure in clinical trials. It has excellent internal consistency and is shown to be highly correlated with well-known outcomes such as the Symptom Checklist 90R, Beck Depression Inventory, The State Trait Anxiety Inventory, The inventory of interpersonal problems, The Social Adjustment Scale, and the SF-36.[37] The scale is scored on a scale of 0 (=never) to 5 (=almost always) giving a range of 0 to 180, with higher scores indicating higher levels of psychopathology. The OQ-45.2 has three validated subscales: symptom distress, interpersonal relations, and social role functioning (perceived level of difficulties in the workplace, school or home duties). Results will be reported for the sum score, and for the three subscales. The OQ-45.2 has an established clinical cut-off value and reliable change index.[37]

The Hospital Anxiety and Depression Scale (HADS)[38] is a 14-item questionnaire assessing non-vegetative symptoms of anxiety and depression on a 0 to 3 likert scale. The sum score can be used as a measure of general psychological distress and is widely used in the community, general practice and psychiatric settings. It has a range of 0 to 42 points with higher scores indicating higher levels of psychological distress.

Nightmares and nocturnal mentation:

The Nightmare Frequency Questionnaire (NFQ)[39]: One item from the NFQ assessing of frequency of nights with nightmares (in days per week, month or year).

Nightmare intensity[40]*:* One item from the clinical administered PTSD-scale (CAPS)[40] assessing the intensity of nightmares on a four point Likert-type scale ranging from 1 (= minimal) to 4 (=extreme).

Nocturnal mentation[41]*:* Three items from the Dream Recall Frequency Scale (DRFS) are used to assess nocturnal mentation and can be used as a proxy for REM sleep fragmentation.[41] The items are scored on a 0 (=never) to 8 (=almost every day/night) and assess how often the participant is aware of having thoughts during sleep, how often night-time thoughts affect mood next day, and how often the participant tells others about their night-time thoughts, giving a total score of 0 to 24.

Measures of somatic symptoms and health

The Chalder Fatigue Scale (CFS)[42] is a 11-item questionnaire assessing levels of daytime physical and psychological fatigue on a 0 to 3 likert scale (0=less than usual, 3=much more than usual). The scale has a range of 0 to 33 with higher scores indicating higher levels of fatigue. Two additional items assess duration of fatigue (0=less than a week, 4=six months or more) and how much of the time the individual experience fatigue (0=25% of the time, 3=all the time).

Euroqol-5D[43] is a 5-item self-report questionnaire assessing general health state on a 0 to 5 likert scale. It measures levels of problems with walking, performing self-care, doing usual activities, pain/discomfort, and anxiety/depression. It is widely used across Europe in assessments of health resources utilization as it allows measurement of Quality Adjusted Life Years (QALYs) in individuals presenting with a wide range of physical and mental disorders.

Alcohol Use Disorders Identification Test – Consumption (AUDIT-C)[44] is a three-item self-report questionnaire assessing frequency of alcohol use (0=never, 4=four times each week or more), number of units typical for a drinking day (0=1-2 units, 4=10 or more units), and frequency of binge-drinking (0=never, 4=daily or almost daily). It is a short-version of the 10-item AUDIT developed by the World Health Organization.

Headache Impact Test – 6 (HIT-6)[45] is a six-item self-report questionnaire assessing intensity and consequences of headaches the last month, rated from never to always. The scale has a range from 36 to 78 with higher values indicating higher severity of headaches. The four headache impact severity categories are little or no impact (49 or less), some impact (50–55), substantial impact (56–59), and severe impact (60–78).

Work performance:

Work Productivity and Impairment Questionnaire General Health (WPAI:GH)[46] will be used to assess work performance and impairment in daily living. It is a six-item questionnaire that measures: sickness absenteeism; sickness presenteeism; overall work impairment and; activity impairment.

Cognitive function

The Cognitive Complaints in Bipolar Disorder Rating Assessment (COBRA) will be used to assess cognitive function. It is a 16-item self-reported instrument of subjective cognitive dysfunctions including executive function, processing speed, working memory, verbal learning and memory, attention/concentration and mental tracking. Items are rated using a 4-point scale. The higher the score, the more subjective complaints. Although the assessment was initially introduced for use with individuals with bipolar disorders, the rating can be used on other clinical populations.

Intervention-Related Assessments

The assessment package includes several brief self-report instruments that give insights into the expectation of and experiences associated with the intervention provided. These include:

Expectations: At baseline, one item scored on a 1 to 9 scale indicating to what extent the patient thinks a digital sleep intervention can work for them (1= not at all, 9 = perfectly).

Self-efficacy: At baseline, self-efficacy related to making behavioral changes will be assessed with a 13-item questionnaire scored on a 1 to 10 scale (1=not confident at all, 10=extremely confident). Seven items assess self-efficacy of behaviors related to CBT-I (e.g. following a plan for bed time and rise time), and six items assess self-efficacy in situations where it may be challenging to keep the behavioral changes (e.g. if you are depressed or feel hopeless, in weekends, etc).

The Negative Effects Questionnaire (NEQ)[47]: will be used to assess negative effects of digital interventions. The NEQ is a self-report measure that contains 20 items that are scored on a five-point Likert-scale (rated 0-4) where higher scores indicate higher levels of negative effects. After each item, the individual is asked whether they consider the effect to be caused by the intervention received or caused by other circumstances (yes/no), as well as one open-ended question.

Impact on future treatment: will be assessed post intervention with two items. The first will assess if this intervention has impacted motivation for face-to-face treatment for mental disorders, the second will assess if this intervention has impacted motivation for specifically working with sleep interventions at a later time (both scored on a 1 to 9 scale with 1=very demotivated, 9=very motivated).

The Use of Sleep Strategies (USS): is a six item self-report questionnaire have been developed to assess how often individuals use six different therapeutic techniques (keep a stable rise time, refrain from sleeping during daytime, use the bed and bedroom only for sleeping, practiced sleep restriction, practiced stimulus control) and their perception of its utility. The techniques are integral to CBT-I but are also described in sleep psychoeducation or hygiene programmes.

Resource use and national registries

Using objective data available from national registries we will collate information on participants to allow to explore group differences before and/or after randomization-
a) reasons for referral;
b) diagnoses (International Statistical Classification of Diseases and Health Related Problems, ICD-10);
c) substance use, person injury (similar to Core minimum data set in World Health Organization (WHO) guidelines), incident leading to any hospital admission;
d) number of appointments at mental health care clinics, type and timing of treatment and admissions, and date of the first appointment for each patient during the study period (data on a-d from the *Norwegian Patient Registry (NPR));*
e) dose, timing and type of prescribed medications (According to the WHO Anatomical Therapeutic Chemical Classification System) and changes recorded during the RCT (data from *the Norwegian Prescription Database (NorPD));*
e) costs of treatment offered by the public services (data from the database named 'Kontroll og Utbetaling av Helserefusjon');
f) sick leave or in receipt of disability benefits (data from the administrative database called Forløpsdatabasen (FD Trygd);
g) cause of death (Dødsårsaksregisteret);

Subgroup data collection of objective sleep and circadian assessments

A subgroup of approximately 40 patients recruited at St Olavs Hospital will have concurrent assessments of sleep-wake patterns with actigraphy for 9 weeks during the intervention period.[49]

Actigraphy: Actigraphy is a wrist-worn accelerometer that can be used for assessment of

movement and indirectly as an objective assessment of sleep and circadian measures. [48] We will use GENEActive actiwatches (Activeinsights, Kimbolton, UK), which has a triaxial sensor for acceleration, a light sensor, and a temperature sensor.