Statistical Analysis Plan

ResMed Germany Inc, Fraunhoferstr. 16, 82152 Martinsried, Germany,

The effects of a personalized sleep improvement App from SleepScore Labs: a single-blinded randomized waitlist controlled trial

Christoph Schöbel (Principal Investigator), Universitätsmedizin Essen, Zentrum für Schlaf- und Telemedizin, Ruhrlandklinik, Tüschener Weg 40, 45239 Essen, Germany

Date and signature

Susanne Fischer (Sponsor), ResMed Germany Inc, Fraunhoferstr. 16, 82152 Martinsried, Germany

Date and signature

Elie Gottlieb (Sleep Scientist), Sleepscore Labs Limited, 2175 Salk Avenue, Carlsbad, CA 92008

Date and signature

Luke Gahan (Statistician), Sleepscore Labs Limited, 6th Floor, 2 Grand Canal Square, Dublin, Ireland

Date and signature

Michael Schredl (Medical Advisor, Statistics), Sleep laboratory, Central Institute of Mental Health, J 5, 68159 Mannheim, Germany

Date and signature

Version 1.0 date 05/23/2023

Content

1 Aim of the statistical analysis plan (SAP) 4
2 Responsibility of the statistical analysis of study data 4
3 Background 4
3.1 Objectives and aims 4
3.2 Design 4
4 Population for analyses 5
4.1 Definitions 5
4.2 Changes to the protocol 5
4.3 Violations of the protocol 5
5 Variables 5
5.1 Demographics and baseline characteristics 5
5.2 Primary variable 5
5.3 Secondary variables 5
5.4 Deduced, grouped and summarized variables 6
6 Treatment of missing values and outliers 6
6.1 Missing values 6
6.2 Outliers 6
7 Statistical methods 6
8 Statistical analyses 7
8.1 Primary analysis 7
8.2 Secondary analyses 7
8.3 Concomitant medication and comorbidities 7
8.4 Compliance and adherence in the app condition 7
8.5 Subgroup analyses and analyses of further secondary parameters 7
9 Software 7
10 Appendices – Tables, Lists, Figures 7
10.1 Tables: 7
10.2 Figures 7
10.3 Appendix 8
10.3.1 Annex I: Flowchart diagram 8
10.3.2 Annex II: Data analyses and reports: Responsibilities and timelines 9
10.3.3 Annex III: Questionnaires Perceived Stress Scale 9
10.3.4 Annex IV: Pittsburgh Sleep Quality Index 10
10.4 References 12

1 Aim of the statistical analysis plan (SAP)

The SAP describes the major statistical hypotheses on primary outcomes, methods and procedures for the analysis of study data. The study protocol versions May 27 2022 and Jan 26 2023 (April 28 2023 – decision pending at this point in time) already describe some statistical methods and study design.

2 Responsibility of the statistical analysis of study data

The sponsor ResMed Germany Inc will be responsible for performing analyses on all self-reported sleep and quality of life (questionnaire) data, and Sleepscore Labs Limited will be responsible for the analyses of objectively measured sleep data (application). Both parties will be responsible for the review of each other`s data. A detailed description of tasks, responsibilities and timelines is outlined in appendix II.

3 Background

Prevention programs can help to avoid sleep problems and train people who are not satisfied with their sleep how they can sleep better can help to improve quality of life and sleep. The Dein Schlaf. Dein Tag. validation study follows 600 healthy participants, who are either randomized to the app condition or to the waitlist condition, with self-reported sleep problems over a period of 12 weeks to investigate whether the use of a smartphone application can improve subjective sleep behavior and objective sleep quality.

3.1 Objectives and aims

Primary outcome

- Changes in score of the sleep quality (SQ) component of the SF-B comparing app and waitlist condition at baseline with data at 6 and 12 weeks of follow-up.

Secondary outcomes

Changes in subjective sleep parameters (questionnaires) comparing app and waitlist condition at baseline with data at 6 and 12 weeks of follow-up.

- Changes in scores of the “feeling refreshed after sleep” (GES) component of the SF-B
- Changes in physiological and psychological component scores of the SF-12
- Changes in quality of sleep (QoS), measured by the Pittsburgh Sleep Quality Index (PSQI) total score
- Changes in scores of the perceived stress scale (PSS)

Changes in objective sleep parameters in the app condition at baseline with data at 6 and 12 weeks of follow-up.

- Changes in sleep onset latency
- Changes in wake after sleep onset time
- Changes in total sleep time
- Changes in sleep efficiency
- Changes in sleep staging
- Changes in nighttime awakenings

3.2 Design

Decentral prospective interventional randomized study. Sleepy, otherwise healthy participants will be randomized to download and use the application (app condition) or to a waitlist (control condition). Both groups will fill out questionnaires online on their sleep behavior and quality of life status. In the app condition, objective sleep parameters will be documented.

4 Population for analyses

4.1 Definitions

“App condition”: Participants randomized to tracking their sleep with the Dein Schlaf. Dein Tag. application.
“Waitlist condition” (control condition): Participants randomized to a waitlist group without using the application.

4.2 Changes to the protocol

Changes from version May 27 2022 to Jan 26 2023 were that informed consent was given electronically (no personal signature). The flowchart was adapted to show that after participants consented, the baseline assessment was performed and then the randomization. The Munich Chronotype questionnaire and the LISST questionnaire were not provided. Changes from version Jan 26 2023 to Apr 28 2023 were that inclusion criteria were changed to allow recruitment of non-iOS users. Decision on the version Apr 28 2023 was pending until this point in time.

4.3 Violations of the protocol

Participants in the app condition who do not track their sleep for 3 nights or more per week will be excluded from parts of the secondary analyses.

5 Variables

5.1 Demographics and baseline characteristics

The following table lists the variables to characterize demographics and baseline characteristics:

Variable name and unit		Unit
Demographics and inclusion		
Age [yrs]		years
Sex [male/female]		
Weight [kg]		kg
Height [cm]		cm
Objective sleep data and sleep documentation		
sleep onset latency		min
wake after sleep onset time		min
total sleep time		min
sleep staging		min
nighttime awakenings		number

5.2 Primary variable

Score of subjective sleep quality (SQ) component of the SF-B in the app group and waitlist condition after 6 weeks and 12 weeks. SF-B SQ component contains scores of ESS (2 questions), DSS (2 questions), VZA (1 question), ASC (6 questions) and 1 question (28), which are summarized to an SQ score. Results are ranging from 1 (poor sleep quality) to 5 (best sleep quality). Scoring and interpretation will be done following the SF-B manual (revised version) © Hogrefe Verlag 2011.

5.3 Secondary variables

- Score of SF-B GES component: Feeling of being refreshed after sleep. Scoring and interpretation will be done following the SF-B manual (revised version) © Hogrefe Verlag 2011.
- Scores of physiological and psychological component of the sleep questionnaire SF-12. Scoring and interpretation will be done following the SF-12 manual (2nd version) © Hogrefe Verlag 2011.
- Score of the perceived stress scale (PSS). PSS consists of 10 questions, divided into a perceived helplessness subscale PH and a =perceived self-efficacy PSE subscale. The range of answers are: 1 =

never, 2 = almost never, 3 = sometimes, 4 = fairly often, 5 = very often. The subscale PH is calculated by summing up Items 1, 2, 3, 6, 9 and 10. The subscale PSE is calculated by summing up items 4, 5, 7 and 8. A total score can be derived by taking the sum of all PH and reversed PSE items. The items 4, 5, 7 and 8 are reverse scored for the total score. High scores indicate a high level of stress. (Annex III).

- Scores of the Pittsburgh Sleep Quality Index (PSQI). PSQI contains 19 questions (+5 that are not included in the scoring – question 10 does not count for quantitative evaluation) – divided into 7 components - that are rated from 0=very good to 3=very bad. The 7 component scores will be added up to a global score ranging from 0 to 21. PSQI scores will be evaluated according to the scoring manual (Annex IV; in case a version has been updated, the update will be used).
- sleep onset latency in minutes
- wake after sleep onset in minutes
- total sleep time in minutes
- sleep efficiency in %
- sleep staging in minutes/stage
- numbers of nighttime awakenings
- time spent on guide (average time and/or number of interactions over the duration of the study)

5.4 Deduced, grouped and summarized variables

Sleep efficiency (%) = Total sleep time/time in bed
Body mass index BMI = weight (kg)/height^2 (m)

6 Treatment of missing values and outliers

6.1 Missing values

Missing values will not be replaced. All available data will be used for analyses. For further details see sections 7 and 8 on statistical methods and analyses.

6.2 Outliers

The analyses comprise all measured and documented values. Excluded are only those values that have been identified or confirmed by an investigator to be a measurement error or a confirmed false entry, e.g. typos in the date of birth.

7 Statistical methods

Demographic and other key study variables will be compared between intervention group and control group at the baseline using t tests for continuous variables and chi squared for categorical variables to test whether the randomization has worked adequately.

The questionnaire data will be analyzed by an ANOVA, with one factor (app condition vs. waitlist control group) and time (Baseline, 6-weeks, 12-weeks). To handle missing values, a mixed-model approach with the same factors will be applied. The critical statistical test would be the interaction term between app condition and the waitlist control group.

For objectively recorded sleep data, repeated measured data will be analyzed using multi-level modeling (MLM) in a nested way (nights within participants), will be performed allowing for analyzing changes with the intervention (app) group. To evaluate changes in primary and secondary outcome measures assessed by validated scales at baseline, 6-weeks, and 12-weeks, we will compare within-group differences for each time-point using a t-test if normally distributed, or the Mann Whitney-U test for non-normally distributed continuous data. In addition, effect sizes for the changes within both groups (questionnaire data) and app group (objective sleep parameters) will be computed.

8 Statistical analyses

8.1 Primary analysis

The primary analysis will be to examine the sleep quality (SQ) component score of the SF-B total for a significant change between the app and the control condition at baseline (week 0), at 6 weeks, and at 12 weeks.

8.2 Secondary analyses

The sleep quality SQ component of the SF-B will be calculated similar to the primary analysis, but using a per-protocol approach (participants who are adherent to treatment), excluding participants who did not sufficiently engage with app throughout the study (<3 tracked nights of sleep per week, as averaged over the entire follow-up period, respectively used the app modules regularly).

The GES component of the SF-B (feeling of being refreshed after sleep), the physiological and psychological component of the SF-12, the total score of the Pittsburgh Sleep Quality Index questionnaire (PSQI) and the score of the perceived stress scale PSS will be tested for significant changes between app condition and control condition at baseline, at 6 weeks and at 12 weeks. This will be done with the paired t-test, if there are no significant deviations from a normal distribution, otherwise the Wilcoxon matched pairs test will be used.

Objective sleep parameters (sleep onset latency, wake after sleep onset, total sleep time, sleep efficiency, sleep stages, nighttime awakenings) will be tested for significant change in the app condition comparing baseline with data at 6 and 12 weeks of follow-up.

8.3 Concomitant medication and comorbidities

Participants with severe comorbidities and intake of sleep medication were excluded.

8.4 Compliance and adherence in the app condition

Usage times will be tracked on a daily basis in a monitoring portal, i.e. the nights where the application was activated and the sleep tracked. The parameter “time spent on the guide” is a measure to show the frequency of interaction of users with the application.

8.5 Subgroup analyses and analyses of further secondary parameters

Further secondary parameters might be tested for differences between app group and control group at week 0, week 6 and week 12: Components of the SF-B: mental comfort before sleep (PSYA), mental weariness before sleep (PSYE), psychosomatic symptoms during sleep (PSS), dream memory (TRME), difficulties falling asleep (ESS), difficulties staying asleep (DSS), early awakening (VZA) and general sleep characterisation (ASC).

9 Software

SAS (SAS Institute) will be the main statistical programming software.

10 Appendices – Tables, Lists, Figures

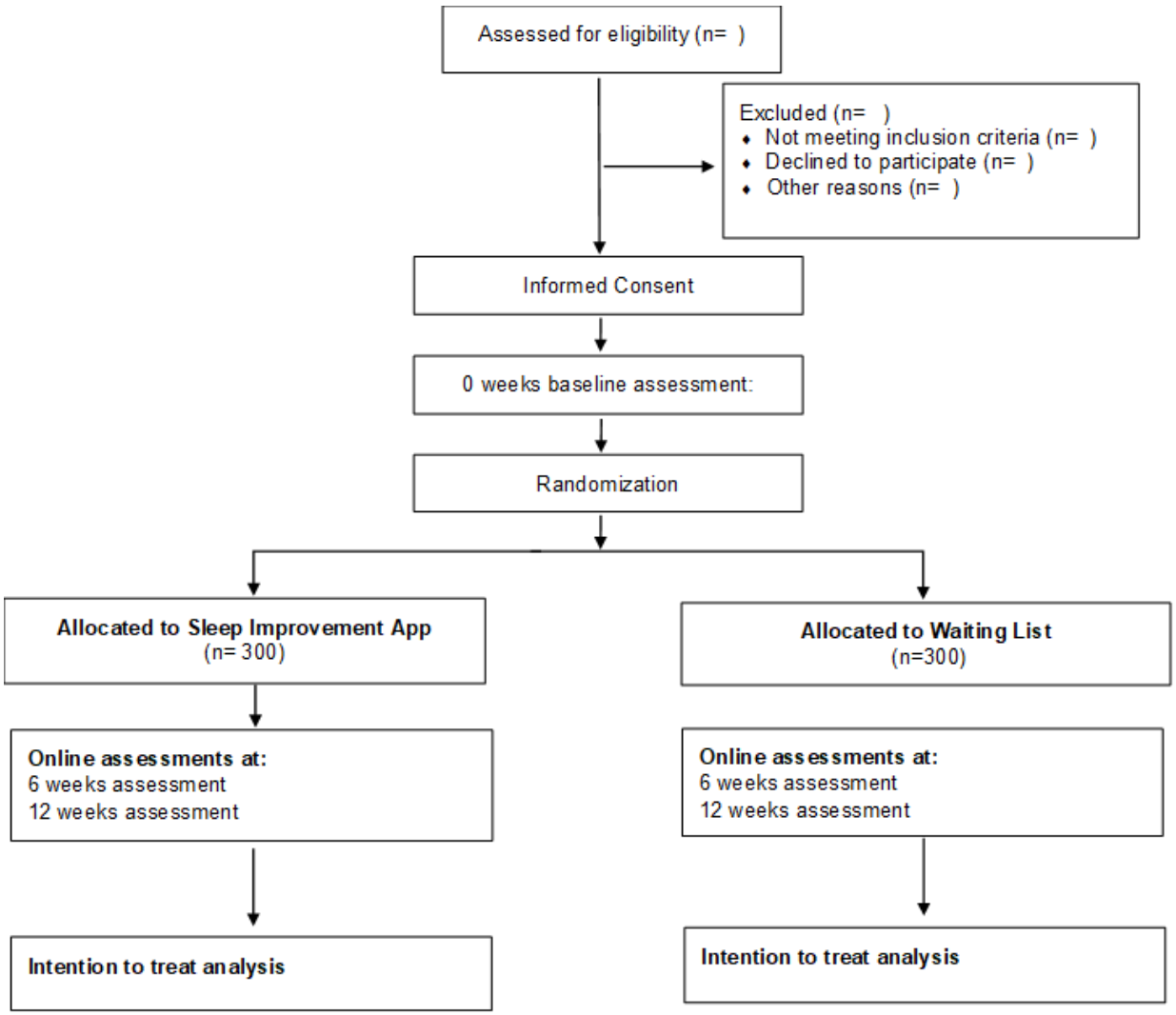
10.1 Tables:

10.2 Figures

The graphical presentation of qualitative parameters (absolute and percentage frequencies) will be done using pie charts or bar charts. Quantitative variables are to be visualized using boxplots or also using bar charts.

10.3 Appendix

10.3.1 Annex I: Flowchart diagram



10.3.2 Annex II: Data analyses and reports: Responsibilities and timelines

Task (What?)	Responsibility (Who?)	Timeline (When?)
Analyses		
Analyses of the questionnaires (subjective sleep data), including calculation of score differences and p-values	RMD	With data available
Analyses of objective sleep data from the application, including score differences and p-values	SSL	With data available
Study report first draft		
Introduction and study design	RMD	Until CW26
Methods, inclusion/exclusion criteria	RMD	Until CW26
Description of the questionnaires	RMD	Until CW26
Description of the application	SSL	Until CW26
Results subjective sleep measurements (questionnaires)	RMD	Until CW29
Results objective sleep measurement (application)	SSL	Until CW29
Outcomes and discussion: Fill the outcomes of the questionnaires with context, e.g. comparison with standard data and interpretation of the relevance of changes in scores.	RMD	Until CW29
Outcomes and discussion: Validate the outcomes of sleep parameters against existing data.	SSL	Until CW29
Review of the draft report	RMD/SSL	Until CW28
Study report revised version		
Review of the report	RMD/SSL	CW30
Finalizing the report	RMD	CW30

RMD: ResMed; SSL: Sleepscore Labs

10.3.3 Annex III: Questionnaires Perceived Stress Scale

Instructions, items and subscales of a German adaption of the 10-item Perceived Stress Scale (PSS-10)

From: Schneider, E. E., Schönfelder, S., Domke-Wolf, M., & Wessa, M. (in press). Measuring stress in clinical and nonclinical subjects using a German adaptation of the Perceived Stress Scale. International Journal of Clinical and Health Psychology.

PERCEIVED STRESS SCALE - 10

Die folgenden Fragen beschäftigen sich mit Ihren Gedanken und Gefühlen während des letzten Monats. Bitte geben Sie für jede Frage an, wie oft sie in entsprechender Art und Weise gedacht oder gefühlt haben.

		nie	Fast nie	Manchmal	Ziemlich oft	sehr oft
1	Wie oft waren Sie im letzten Monat aufgewühlt, weil etwas unerwartet passiert ist?	1	2	3	4	5

2	Wie oft hatten Sie im letzten Monat das Gefühl, nicht in der Lage zu sein, die wichtigen Dinge in Ihrem Leben kontrollieren zu können?	1	2	3	4	5
3	Wie oft haben sie sich im letzten Monat nervös und gestresst gefühlt?	1	2	3	4	5
4	Wie oft waren Sie im letzten Monat zuversichtlich, dass Sie fähig sind, ihre persönlichen Probleme zu bewältigen?	1	2	3	4	5
5	Wie oft hatten Sie im letzten Monat das Gefühl, dass sich die Dinge zu Ihren Gunsten entwickeln?	1	2	3	4	5
6	Wie oft hatten Sie im letzten Monat den Eindruck, nicht all Ihren anstehenden Aufgaben gewachsen zu sein?	1	2	3	4	5
7	Wie oft waren Sie im letzten Monat in der Lage, ärgerliche Situationen in Ihrem Leben zu beeinflussen?	1	2	3	4	5
8	Wie oft hatten Sie im letzten Monat das Gefühl, alles im Griff zu haben?	1	2	3	4	5
9	Wie oft haben Sie sich im letzten Monat über Dinge geärgert, über die Sie keine Kontrolle hatten?	1	2	3	4	5
10	Wie oft hatten Sie im letzten Monat das Gefühl, dass sich so viele Schwierigkeiten angehäuft haben, dass Sie diese nicht überwinden konnten?	1	2	3	4	5

Skala Hilflosigkeit (H): Summe der Items 1, 2, 3, 6, 9, 10; Skala Selbstwirksamkeit (S): Summe der Items 4, 5, 7, 8. Für die Berechnung des Gesamtscores müssen die Items 4, 5, 7 und 8 der Selbstwirksamkeitsskala invertiert werden. Der Gesamtscore berechnet sich aus der Summe der Items der Hilflosigkeitsskala und der Summe der invertierten Items der Selbstwirksamkeitsskala. Höhere Werte deuten auf ein erhöhtes Stresslevel hin.

Original Items in English (Cohen et al., 1983)

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way.

1	PH	In the last month, how often have you been upset because of something that happened unexpectedly?
2	PH	In the last month, how often have you felt that you were unable to control the important things in your life?
3	PH	In the last month, how often have you felt nervous and “stressed”?
4	PSE	In the last month, how often have you felt confident about your ability to handle your personal problems?
5	PSE	In the last month, how often have you felt that things were going your way?
6	PH	In the last month, how often have you found that you could not cope with all the things that you had to do?
7	PSE	In the last month, how often have you been able to control irritations in your life?
8	PSE	In the last month, how often have you felt that you were on top of things?
9	PH	In the last month, how often have you been angered because of things that were outside of your control?
10	PH	In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

Answer range: 1 = never, 2 = almost never, 3 = sometimes, 4 = fairly often, 5 = very often; PH=perceived helplessness subscale; PSE=perceived self-efficacy; Items 4, 5, 7 and 8 are reverse scored for the total score. The PH subscale is computed by summing up Items 1, 2, 3, 6, 9 and 10; the PSE subscale is computed by summing up items 4, 5, 7 and 8; the total score is the sum of all PH and reversed PSE items. Higher scores reflect greater levels of stress

10.3.4 Annex IV: Pittsburgh Sleep Quality Index

Pittsburgh Sleep Quality Index (PSQI)

Form Administration Instructions, References, and Scoring
Form Administration Instructions

The range of values for questions 5 through 10 are all 0 to 3.

Questions 1 through 9 are not allowed to be missing except as noted below. If these questions are missing then any scores calculated using missing questions are also missing. Thus it is important to make sure that all questions 1 through 9 have been answered.

In the event that a range is given for an answer (for example, ‘30 to 60’ is written as the answer to Q2, minutes to fall asleep), split the difference and enter 45.

Reference

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ: The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research* 28:193-213, 1989.

Scores – reportable in publications

On May 20, 2005, on the instruction of Dr. Daniel J. Buysse, the scoring of the PSQI was changed to set the score for Q5J to 0 if either the comment or the value was missing. This may reduce the DISTB score by 1 point and the PSQI Total Score by 1 point.

PSQIDURAT **DURATION OF SLEEP**

IF Q4 ≥ 7, THEN set value to 0
IF Q4 < 7 and ≥ 6, THEN set value to 1
IF Q4 < 6 and ≥ 5, THEN set value to 2
IF Q4 < 5, THEN set value to 3
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIDISTB **SLEEP DISTURBANCE**

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) = 0, THEN set value to 0

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) ≥ 1 and ≤ 9, THEN set value to 1

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) > 9 and ≤ 18, THEN set value to 2

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) > 18, THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQILATEN **SLEEP LATENCY**

First, recode Q2 into Q2new thusly:
IF Q2 ≥ 0 and ≤ 15, THEN set value of Q2new to 0
IF Q2 > 15 and ≤ 30, THEN set value of Q2new to 1
IF Q2 > 30 and ≤ 60, THEN set value of Q2new to 2
IF Q2 > 60, THEN set value of Q2new to 3
Next
IF Q5a + Q2new = 0, THEN set value to 0
IF Q5a + Q2new ≥ 1 and ≤ 2, THEN set value to 1
IF Q5a + Q2new ≥ 3 and ≤ 4, THEN set value to 2
IF Q5a + Q2new ≥ 5 and ≤ 6, THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIDAYDYS **DAY DYSFUNCTION DUE TO SLEEPINESS**

IF Q8 + Q9 = 0, THEN set value to 0
IF Q8 + Q9 ≥ 1 and ≤ 2, THEN set value to 1
IF Q8 + Q9 ≥ 3 and ≤ 4, THEN set value to 2
IF Q8 + Q9 ≥ 5 and ≤ 6, THEN set value to 3
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIHSE **SLEEP EFFICIENCY**

Diffsec = Difference in seconds between day and time of day Q1 and day Q3
Diffhour = Absolute value of diffsec / 3600
newtib =IF diffhour > 24, then newtib = diffhour – 24
IF diffhour ≤ 24, THEN newtib = diffhour
(NOTE, THE ABOVE JUST CALCULATES THE HOURS BETWEEN GNT (Q1) AND GMT (Q3))

tmphse = (Q4 / newtib) * 100

IF tmphse ≥ 85, THEN set value to 0
IF tmphse < 85 and ≥ 75, THEN set value to 1
IF tmphse < 75 and ≥ 65, THEN set value to 2
IF tmphse < 65, THEN set value to 3
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQISLPQUAL **OVERALL SLEEP QUALITY**

Q6
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIMEDS **NEED MEDS TO SLEEP**

Q7
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQI **TOTAL**

DURAT + DISTB + LATEN + DAYDYS + HSE + SLPQUAL + MEDS
Minimum Score = 0 (better); Maximum Score = 21 (worse)
Interpretation: TOTAL ≤ 5 associated with good sleep quality
TOTAL > 5 associated with poor sleep quality

10.4 References

Sheldon Cohen, Tom Kamarck and Robin Mermelstein, A global measure of perceived stress. Journal of Health and Social Behavior Vol. 24, No. 4 (Dec., 1983), pp. 385-396

Schneider, E. E., Schönfelder, S., Domke-Wolf, M., & Wessa, M., Measuring stress in clinical and nonclinical subjects using a German adaptation of the Perceived Stress Scale. International Journal of Clinical and Health Psychology Volume 20, Issue 2, May–August 2020, Pages 173-181

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ:, The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. Psychiatry Research 28:193-213, 1989.