UNIVERSITY OF LEICESTER

STATISTICAL ANALYSIS PLAN

A three arm cluster randomised controlled trial to test the effectiveness and cost-effectiveness of the SMART Work & Life intervention for reducing daily sitting time in office workers (SMART Work & Life Study)

SAP Version: 1.2
Date: 08/02/2021

Based on protocol: SMART Work and Life Protocol v1.7 040919 FINAL
Trial registration: 11618007

Sponsor:	University of Leicester
Trial Statistician:	Nishal Bhupendra Jaicim University of Leicester Leicester Clinical Trials Unit Maurice Shock Building University Road Leicester LE1 7RH
Chief Investigator:	Dr Charlotte Edwardson Diabetes Research Centre University of Leicester Leicester General Hospital Leicester, LE5 4PW
Trial Manager:	Jill Clanchy University of Leicester Leicester Clinical Trials Unit Maurice Shock Building University Road Leicester LE1 7RH

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

UNIVERSITY OF LEICESTER

Revision History

Version	Date	Author(s) and Role	Summary of Changes/Comments
0.1	19/05/2020	Nishal Bhupendra Jaicim (NBJ) Statistician	
0.2	20/05/2020	NBJ Cassey Brookes (CB) Principal Statistician	CB comments
0.3	12/06/2020	NBJ CB Dr Charlotte Edwardson (CE) Chief Investigator Prof Laura Gray (LG) Senior Statistician	CE and LG comments
0.4	17/06/2020	NBJ, CB, CE, LG	CE and LG comments
0.5	18/06/2020	NBJ, CB, CE, LG	CB, CE and LG comments
0.6	07/07/2020	Co-Investigators: Alex Clarke-Cornwell (ACC) Dr Stacy Clemes (SC) Prof David Dunstan (DD) Dr Fehmidah Munir (FM) Prof Genevieve Health (GH) Prof Malcolm Granat (MG) Prof Melanie Davies (MD) Prof Stuart Biddle (SB) DMEC Members: Denise Howell (DH) Dr Trish Gorely (TG) TSC Members: Dr Derrick Bennett (DB) Dr Daniel Bailey (DB)	Comments from Co-Investigators, DMEC and TSC Members
0.7	27/07/2020	NBJ, CB, CE, LG, ACC, SC, DD, FM, GH, MG, MD, SB, DH, TG, DB, DB	CE and LG comments
0.7	27/07/2020	NBJ, CB, CE, LG, ACC, SC, DD, FM, GH, MG, MD, SB, DH, TG, DB, DB	CE and LG comments
0.8	04/08/2020	NBJ, CB, CE, LG, ACC, SC, DD, FM, GH, MG, MD, SB, DH, TG, DB, DB	CE and LG comments
1.0	06/08/2020	NBJ, CB, CE, LG, ACC, SC, DD, FM, GH, MG, MD, SB, DH, TG, DB, DB	CE comments
1.1	03/02/2021	NBJ, CB, CE, LG, ACC, SC, DD, FM, GH, MG, MD, SB, DH, TG, DB, DB	Section 5.5 corrections (Changes to planned analysis)

UNIVERSITY OF LEICESTER

1.2	08/02/2021	NBJ, CB, CE, LG, ACC, SC, DD, FM, GH, MG, MD, SB, DH, TG, DB, DB	Section 5.5 and appendix correction – scoring of Fatigue Questionnaire

SAP approval for finalised version:

Trial Statistician	**Nishal Bhupendra Jaicim**	**08/02/2021**
	Signature	**Date**
Chief Investigator	**Dr Charlotte Edwardson**	
	Signature	**08/02/2021** Date
Senior Statistician	**Professor Laura Gray**	**08/02/2021**
	Signature	**Date**

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

UNIVERSITY OF LEICESTER

LIST OF ABBREVIATIONS

AE	Adverse event
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
BMI	Body mass index
BPM	Beats per minute
CI	Chief Investigator
CC	Complete case
CM	Centimetre
CONSORT	Consolidated Standards of Reporting Trials
COVID-19	Coronavirus Disease 2019
CPS2	Copenhagen Psychosocial Questionnaire (Version 2)
CRF	Case report form
CSR	Clinical study report
CTU	Clinical trials unit
DMEC	Data monitoring and ethics committee
EEG	Electroencephalograph
FDA	Food and Drug Administration
HADS	Hospital Anxiety and Depression Scale
HDL	High-Density Lipoproteins
ICC	Intraclass correlation coefficient
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ITT	Intention-to-treat
KG	Kilogram
LCTU	Leicester Clinical Trials Unit
LDL	Low-Density Lipoproteins
MI	Multiple imputation
mmol/mol	Millimoles per mole
mmol/l	Millimoles per litre
NIHR	National Institute for Health Research
NFR	Need for recovery
PANAS	Positive and Negative Affect Schedule
PAST	Past-day Adults' Sedentary Time
PK	Pharmacokinetic
PP	Per-protocol
PSS	Perceived Stress Scale
PSQI	Pittsburgh Sleep Quality Index
QoL	Quality of life
RCT	Randomised controlled trial
SAE	Serious adverse event
SAF	Safety population
SAP	Statistical analysis plan

SAR	Serious adverse reaction
SD	Standard Deviation
SNQ	Standardized Nordic Questionnaire
SUSAR	Suspected unexpected serious adverse reaction
SWAL	Smart Work And Life
TSC	Trial Steering Committee
UWES	Utrecht Work Engagement Scale
WHO-5	Scale 5-item World Health Organization Well-Being Index
WLS	Work Limitations Questionnaire

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

Contents

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

1 Introduction

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for the Edwardson_16_324_SMArt_WorkandLife. All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association and the Royal Statistical Society for statistical practice.

The reader of this SAP is encouraged to also read the trial protocol (v1.7 04/09/19).

The purpose of this SAP is to outline the planned end of trial analyses that are to be performed on the data to support the completion of the Study Report (SR). The SAP will be amended if there are substantial changes to the planned analyses, and in any case will be finalised before the database lock for this study. Exploratory post-hoc or unplanned analyses not necessarily identified in this SAP may be performed on these data as required. These analyses will be clearly identified in the SR.

Throughout the document: Any verbatim text from the protocol is provided inside a box:

Text from the protocol

1.1 Study Objectives

1.1.1 Primary Objectives

The original primary objective of this study was to determine the long-term effectiveness and cost-effectiveness of the multi-component SMART Work & Life intervention (when provided with and without a height adjustable desk) for reducing objectively measured average daily sitting time in office workers, compared with no intervention at 24 months.

However, due to the COVID-19 pandemic, the funder, the National Institute for Health Research (NIHR), requested that the analysis should be carried out using the 12-month follow-up data. 24-month data will no longer be collected.

1.1.2 Secondary Objectives

If the primary objective is achieved and both interventions are shown to be effective, a secondary objective will be to determine if one intervention is more effective than the other.

In addition, other secondary objectives will investigate whether SMART Work & Life, delivered with and without a height adjustable desk, leads to short (assessed at 3 months) and medium (assessed at 12 months) change in:

- Average daily sitting time (3 and 12 months) across all valid days and on workdays and non-workdays
- Average sitting time during work hours
- Average time spent standing overall (i.e. daily) and during work hours and on workdays and non-workdays

• Average light and moderate-to-vigorous physical activity overall (i.e. daily) and during work hours and on workdays and non-workdays
• Average time spent stepping and number of steps overall (i.e. daily) and during work hours and on workdays and non-workdays
• Adiposity (Body Mass Index (BMI), percent body fat, waist circumference)
• Blood pressure
• Blood markers (e.g. blood glucose, cholesterol, triglycerides)
• Psychosocial variables (e.g. vitality, fatigue, stress, anxiety, depression, quality of life)
• Work-related outcomes (e.g., work engagement, job performance and satisfaction, presenteeism and sickness absence)
• Musculoskeletal issues
• Sleep

Due to the impact of the COVID-19 pandemic, reduction in average daily sitting time at 12 months will now be the primary objective (not secondary). Furthermore, secondary objectives will not be assessed over the longer term as 24-month data will no longer be collected.

We will also conduct a full process evaluation and a full economic evaluation.

Note: The process evaluation analysis and the economic evaluation will not be carried out by the LCTU and are not included within this SAP.

1.2 Study Design

1.2.1 Overview

This is a three-arm cluster randomised controlled trial (RCT) involving 78 clusters (~26 per arm) and 756 office workers (~252 per arm). Clusters (different office spaces) will be randomised to receive one of the following conditions: 1) The multi-component SMART Work & Life intervention with a height-adjustable desk or desk platform (intervention 1), or 2) The multi-component SMART Work & Life intervention without a height-adjustable desk or platform (intervention 2) or 3) usual practice (control condition).

Baseline measurements will precede randomisation. Measurements will be repeated, using identical standardised procedures, at 3 months to assess any short-term changes and 12 months to assess any longer-term changes.

1.2.2 Participants

Office-based employees aged ≥18 years of age within local Councils in the Leicester, Manchester and Liverpool areas.

1.2.3 Intervention arms

1.2.3.1 Intervention arm 1 (SWAL+Desk)

Organisational strategies:
1) we will seek buy-in from the management by explaining the importance of reducing and breaking up sitting at work and how this may lead to workplace benefits without negatively affecting performance and productivity;
2) a brief awareness session (online/video) that will reinforce the benefits for the workforce and employers of reducing sitting time in and outside of work, and encourage them to brainstorm organisational strategies that could take place, review any current policies around being active at work and as well as creating new policies around topics such as standing and walking meetings, provision for lunch time walking, internal competitions and displaying signs around the workplace. We will also encourage managers to review the layout of their office space to promote increased movement of staff e.g., location of printers, waste bins, water coolers;
3) Modelling of the positive behaviour from managers will also be emphasised.

Environmental strategies:
1) Small-scale environmental restructuring in the office and at home (e.g., relocation of printers and waste bins);
2) Motivational and reminder signs around the office space and at home to sit less and move more;
3) A height-adjustable desk or desk platform to allow the individual to sit or stand to work. The individual will get a choice of desk/desk platform within a set budget. This allows flexibility for office set up, participant preference and avoids testing the effectiveness of a specific type of desk rather than the concept.

Individual and group strategies:
1) An initial education session that covers health consequences of sitting and benefits of reducing and regularly breaking up sitting. During the session they will brainstorm strategies to reduce sitting at work and outside of work, think about barriers to reducing and breaking up sitting and ways to overcome these. At the end of the session individuals will be encouraged to set a goal around sitting less and an action plan to achieve this. The focus on overall daily sitting will be emphasised rather than just workplace sitting;
2) Self-monitoring of sitting behaviour across the whole waking day will be encouraged through the use of free computer prompts, timers and mobile phone apps. The importance of self-monitoring and the apps will be introduced during the education session and individuals will be encouraged to download the tools;
3) Workplace champions will receive training to deliver brief coaching/refresher sessions. These sessions will be used to review key messages, discuss progress, review goals and action plans, discuss barriers and any benefits experienced. These coaching /refresher sessions will likely take place at 3, 6 and 12 months;
4) Social support, from colleagues and family members, will be encouraged through regular activity competitions inside and outside of work.

1.2.3.2 Intervention arm 2 (SWAL)

This group will receive all of the intervention components listed in the previous sections above minus the height-adjustable desk allowing us to investigate how important providing a simple, but fairly expensive, environmental change is for significant reductions in sitting.

1.2.3.3 Control arm

Office clusters assigned to the usual practice control arm will be asked to continue with their usual occupational health promotion conditions. Participants in the control arm will be asked to complete the same study measurements as those in the intervention arms, at the same time points.

1.2.4 Sample size

The primary outcome is change in objectively measured average daily sitting (total) time across all valid days at 12 months. This was modified from 24 months following NIHR's guidance due to COVID-19. The study has been powered to detect a difference of an average of ≥60 minutes per day between both intervention arms and the control arm, which reflects the goal of the intervention (the study is not powered to assess the difference between the intervention arms).

1.2.4.1 Original sample size

Initial power calculations showed that with a total sample size of 420 participants, 10 clusters per arm, the study would have over 90% power to detect a 60-minute change in overall average sitting time with a 2-tailed significance level of 5%. The calculations assumed an SD of 90 minutes [1], a conservative ICC of 0.05 [2], a coefficient of variation to allow for variation in cluster size of 0.54 (cluster size range of 15-45), and an average cluster size of 20 (based on data from councils interested in taking part). The trial was designed to test two intervention arms independently with the control arm, so to keep an overall significance level of 5% the number of clusters was inflated by a factor of 1.23 [3]. The sample size was also inflated by 30% to allow for potential individual loss to follow-up and non-compliance with the primary outcome; a further inflation was applied to allow for 1 whole cluster drop out per arm. Thus, the total proposed sample size was 660 subjects to be recruited from 11 clusters per arm. The sensitivity of power was assessed against alternative ICC values of 0.021 and 0.10 [1, 2]. Adequate power for RCTs is accepted as 80% and with these ICCs the power was above the required level at 98% and 81%, respectively. Also, the calculations were based on a similar trial that used an ICC=0.021 for overall sitting [1], while we chose a more conservative ICC=0.05.

1.2.4.2 Re-estimated sample size

At the start of recruitment, the observed average cluster size and variability of cluster sizes were different to those assumed in the original sample size calculation. With the DMEC's guidance, the sample size was recalculated to ensure the study was adequately

powered. Changing the average cluster size from 20 to 10, the variability in cluster size from 0.54 to 1.42 (cluster size range of 4-38), the inflation for loss to follow-up and non-compliance with the primary outcome from 30% to 40%, while keeping all other assumptions the same, required 690 participants from 72 clusters.

1.2.5 Randomisation and blinding

A unique ID will be assigned as each participant is consented into the study. Once all participants in a particular office cluster have been measured, the office cluster will be assigned to an arm by a CTU statistician using a pre-generated list. Randomisation was stratified by area (Leicester; Salford; Liverpool) and cluster size (Small <10; Large ≥10).

1.3 Visit schedule

Measure	Baseline	3 months	12 months
Objective sitting, standing and physical activity	✓	✓	✓
Self-report sitting and breaks	✓	✓	✓
Office/desk dwell time	✓	✓	✓
Job performance	✓	✓	✓
Job satisfaction	✓	✓	✓
Work engagement (UWES)	✓	✓	✓
Occupational fatigue (NFR)	✓	✓	✓
Fatigue (physical and mental)	✓	✓	✓
Musculoskeletal symptoms (SNQ)	✓	✓	✓
Presenteeism (WLQ)	✓	✓	✓
Work demands	✓	✓	✓
Social norms and cohesion	✓	✓	✓
Quality of Life	✓	✓	✓
Sleep duration and quality (PSQI)	✓	✓	✓
Self-reported sickness absence	✓		✓
Sickness absence via employee records	✓		✓
Anthropometric and blood pressure	✓	✓	✓
Biochemical	✓	✓	✓
Diet, smoking and alcohol	✓	✓	✓
Mental health	✓	✓	✓
Medical history and medication	✓	✓	✓
Demographics	✓		
Job description	✓	✓	✓
Client Service Receipt Inventory	✓	✓	✓
Strategies for sitting less and moving more often	✓	✓	✓
Workplace audit	✓		
Workplace champion characteristics	✓		
Support for sitting less and moving move	✓	✓	✓

2 Outcomes and other variables

2.1 Primary Outcome

2.1.1 Definition and Derivation of Primary Outcome

The primary outcome is average daily sitting (total) time across all valid days, objectively measured using the activPAL device (worn 24hrs/day for 7 days by waterproofing). activPAL data will be processed by the CI (blinded) and each participant's average daily sitting time will be calculated by summing average daily sitting time across valid days and dividing by number of valid wear days. A valid activPAL wear day is defined as having ≥10 hours wear time per day, ≥ 1,000 steps per day and <95% of the day spent in any one behaviour.

2.1.2 Hypothesis to be investigated

The null hypothesis for the primary analysis is that there is no difference between the intervention arms and control arm in the primary outcome at 12 months follow-up.

2.2 Secondary Outcomes

2.2.1 Definition of Secondary Outcomes

We will investigate whether SMART Work & Life, delivered with and without a height adjustable desk, leads to differences in a range of secondary outcomes over the short (assessed at 3 months) and medium term (assessed at 12 months) compared to the control arm. Please see "Secondary Objectives" in Section 1.1.2 for more details.

Please see below a list of all secondary outcomes collected. Please also see scoring protocols for questionnaire-based measures in Appendix 1.

<u>Anthropometrics and blood pressure measurements</u>

- Waist circumference (cm)
- Weight (kg)
- Body composition:
  - Fat mass (kg)
  - Body fat (%)
- Body mass index (BMI) (kg/m^2)
- Systolic blood pressure (mmHg) (3 measures taken. Average of last two calculated)
- Diastolic blood pressure (mmHg) (3 measures taken. Average of last two calculated)
- Heart rate (bpm) (3 measures taken. Average of last two calculated)

<u>Biochemical assessments</u>

- Glycated haemoglobin (mmol/mol)
- Glycated haemoglobin (%)
- Total cholesterol (mmol/l)
- HDL cholesterol (mmol/l)
- LDL cholesterol (mmol/l)

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

- Triglycerides (mmol/l)
- Fasting glucose (mmol/l)
- Cluster metabolic risk score

<u>Objectively measured sitting and physical activity (activPAL):</u> <u>The variables below will be derived by calculating the average across the number of valid days. The variables will be analysed in 4 different time periods: 1) daily (i.e. across all waking hours on all valid days); 2) during working hours; 3) on workdays; 4) on non-workdays.</u>

- Average sitting time (minutes)– total (3 months) and in prolonged bouts lasting 30+ mins (3 and 12 months)
- Average standing time (minutes) – total
- Average stepping time (minutes) – total as well as at a step cadence threshold of 100 steps/min (in bouts lasting 1+ min)
- Average number of steps
- Average number of transitions from sitting to an upright posture

The variables below will also be summarised descriptively at each time point and time period:

- Average number of valid days
- Average waking wear time (minutes)
- Average percent of the day spent sitting (%)
- Average percentage of the day spent standing (%)
- Average percentage of the day spent stepping (%)
- Average percentage of total sitting time spent in prolonged sitting time (%)

<u>Objectively measured physical activity (wrist accelerometer) – The variables below will be derived by calculating the average across the number of valid days. The variables will be analysed in 4 different time periods: 1) daily (i.e. across all waking hours on all valid days); 2) during working hours; 3) on workdays; 4) on non-workdays.</u>

- Average time spent in light physical activity (minutes)
- Average time spent in moderate-to-vigorous physical activity (minutes)
- Average sleep duration (minutes) *
- Sleep efficiency (%) *

* These variables will be calculated daily, for workdays and non-workdays.

The variables below will also be summarised descriptively at each time point and time period:

- Average number of valid days
- Average wear time (minutes)

<u>Self-reported sitting, standing, walking, breaks in sitting, time at desk and in office</u>

- Percentage of the workday spent sitting, standing and walking, percentage of work time in prolonged sitting (taken from the adapted version of the Occupational Sitting and Physical Activity Questionnaire (Chau et al 2012))

- Estimated hours spent sitting whilst working and number of breaks in sitting per hour whilst working (Clarke et al 2011).
- Percentage of working day spent at desk space, in office space and sitting
- Adapted version of the past day recall of sedentary time (PAST) questionnaire, which asks about sitting time outside of work in certain contexts (transport, TV viewing, computer use, other) on weekdays and weekends (Clark et al 2015).

Dietary behaviours, smoking and alcohol

- Dietary behaviours (snack frequency, soft drink consumption frequency, fruit and vegetable consumption), using questions from the Whitehall II study
- Alcohol intake, using questions from the Whitehall II study

Self-reported sleep

- Self-reported sleep duration and quality - Pittsburgh Sleep Quality Index (PSQI)

Physical and mental fatigue

- Fatigue severity - Fatigue Scale (0=less than usual; 3=much more than usual):

Work-related health

- Job performance: 7-point likert scale (1=dissatisfied; 7=extremely satisfied)
- Job satisfaction: 7-point likert scale (1=very poorly; 7=extremely well)
- Extent to which participants intentionally changed work priorities and objectives to accommodate the change in sitting behaviour (6-point fully anchored scale)
- Work engagement - Utrecht Work Engagement Scale (UWES) (0=never; 6=always)
- Occupational fatigue - The Need for Recovery (NFR) Scale
- Musculoskeletal symptoms - Standardised Nordic Questionnaire (SNQ)
- Presenteeism - Work Limitations Questionnaire
- Workload and relations - Health and Safety Executive Management Standards Indicator Tool (HSE MSIT) (1=never; 5=always)
- Data on sickness absence, collected using both self-report and from employer records and include frequency and duration of self-certified and certified sickness. Reasons for sickness absence will also be recorded. Sickness absence at baseline and follow-up will be compared by collecting these data for 12 months prior to start of the study (baseline) and for 12 months during the study (follow-up).

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

Social norms, cohesion and support for sitting less and moving more often

- Organisational social norms - eight items on a 5-point Likert scale (Dunstan et al 2013)
- Presence and extent of cohesion, cooperation and community in workplace teams, using the 'social community' sub-scale of the Copenhagen Psychosocial Questionnaire-II (CPS2) (Kristensen, 2001). This sub-scale uses three 6-point Likert scale items ("always" to "hardly ever").
- Participants will be asked about the support they have received from the organisation, manager, colleagues and family for sitting less and moving more often (Brackenridge et al 2016).

Mental health, well-being and quality of life

- Health-related quality of life - EQ-5D-5L
- Anxiety and depression - Hospital Anxiety and Depression Scale (HADS). (0 to 3 likert scales)
- Stress - Perceived Stress Scale (PSS). (0=never; 4=very often)
- Positive and Negative Affect Schedule (PANAS). (1=Very slightly or not at all; 5=Extremely)
- Wellbeing - WHO-5 scale. (0=at no time; 5=all of the time)

2.2.2 Hypotheses to be investigated

The null hypothesis is that no difference exists between the intervention arms and control arm in change in the secondary outcomes from baseline, at 3 months and at 12 months. Statistical testing will only be carried out for the following key outcomes: sitting time, prolonged sitting time, standing time and stepping time – daily across any valid days and on work days only and during work hours only calculated from the activPAL data variables. Differences between groups in other outcomes will be evaluated descriptively. Please see Section 5.4.1 for more details.

2.3 Subgroups and/or interactions

Subgroup analyses will be conducted only for the primary outcome and for average sitting during work hours.

We want to investigate the intervention effect for the following subgroups:

- Site: Leicester vs. Liverpool vs. Manchester
- Cluster size: Small (<10) vs. Large (≥10)
- Type of worker: part-time (21-34.9 hours/week) vs. full-time (≥35 hours/week)
- Sex: male vs. female
- Age: < median vs. ≥ median
- BMI: normal (< 25 kg/m^2) vs. overweight/obese (≥ 25 kg/m^2)

2.4 Compliance

We will ensure good compliance with this device by checking each device on return and requesting a re-wear if the participant does not provide enough valid days (e.g., at least four).

activPAL valid wear days and valid wear time across all valid days, across work hours as well as workdays and non-workdays will be summarised for the whole group and between arms. The primary analysis will include participants who provide a minimum of 1 valid wear day in the respective time periods (i.e. daily, across work hours, workdays, non-workdays) both at baseline and follow-up. 1 valid day has been chosen to maximise our sample and is line with previous similar studies [4, 5].

3 Analysis Sets/Populations

3.1 Complete Case Population

The primary analysis will test the effect of the intervention on outcomes using a complete case (CC) population. That is, all clusters randomised and the recruited participants in these clusters, excluding those with missing outcome data (i.e. without at least 1 valid day of activPAL data at baseline and follow-up) and complete stratification variable data. The analysis will follow the intention-to-treat principle, that is, clusters and participants will be analysed in the arm to which they were randomised.

3.2 Intention-to-treat Population / Full analysis set

A full intention-to-treat (ITT) analysis population will consist of all clusters randomised and the recruited participants within these clusters. Clusters and participants will be analysed in the arm to which they were randomly allocated, regardless of if they received the assigned intervention, or any protocol deviations or violations.

We will use multilevel multiple imputation to deal with missing data in the following types of variables: the outcome variable, covariates in the final analysis model and any additional auxiliary variables in the imputation model [6-9]. This will be a sensitivity analysis just for the primary outcome and the key secondary outcome, average sitting time across work hours at 12 months.

3.3 Per-protocol (PP) Population

The effect size will also be estimated using a PP analysis, which will only include those who were compliant with the protocol and follow-up visits.

Participants with the protocol deviations in the Protocol Deviations Section 3.5 will be excluded. No missing data will be imputed in this population. This will be a sensitivity analysis for the primary outcome and the key secondary outcome, average sitting time across work hours at 12 months.

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

3.3.1 Additional exploratory PP analyses

An additional exploratory PP analysis will be carried out comprising participants who were compliant with the intervention implementation at different levels, which will be determined from the process evaluation data. This will be done for the primary outcome, average daily sitting time at 12 months, and for one key secondary outcome, average sitting time during work hours at 12 months.

3.4 Safety Population

There will be no safety population.

3.5 Protocol deviations

This section outlines protocol deviations that will affect inclusion in populations, e.g. exclusion from the per-protocol population.

Participants who did not provide valid activPAL primary outcome data at baseline or at the 12-month follow-up will be excluded from this analysis.

Control arm participants with access to a standing desk at 12 months will be excluded, and so will participants in clusters belonging to the intervention arms who didn't have a workplace champion assigned or whose workplace champion left their role within the first three months.

Also, any ineligible clusters that did not have the minimum number of participants required (i.e. four or more) will be excluded.

In addition, participants with time window deviations for their follow up (>+/- 2 months) in terms of their activPAL data will also be excluded.

Furthermore, one of the inclusion criteria for the study was that participants spent the majority of their day sitting. This information was self-reported and was screening criteria prior to the consent and baseline measurement visit and was subsequently assessed using the objective data collected via the activPAL device. We will therefore exclude any participants who did not spend the majority (>50%) of their day sitting at baseline as measured by activPAL.

Detailed information on adherence to the different intervention components and feedback to these components will be investigated separately as part of the process evaluation. This process evaluation will not be carried out by the LCTU.

UNIVERSITY OF LEICESTER

4 General Issues for Statistical Analysis

4.1 Derived / Computed Variables

4.1.1 activPAL sitting time outcomes

Standardised sitting time

$$\text{Standardised work hours minutes} = \frac{\text{observed minutes x 480}}{\text{observed minutes of wear time}}$$

$$\text{Standardised daily minutes} = \frac{\text{observed minutes x 720}}{\text{observed minutes of wear time}}$$

4.1.2 Other

Ethnicity
The ethnicity variable will be categorised into White vs. Other.

Clustered cardiometabolic score
The cardiometabolic risk variables are waist circumference, triglycerides, HDL cholesterol, systolic blood pressure, diastolic blood pressure, and fasting plasma glucose. A continuous clustered cardiometabolic risk score on the basis of these variables will constructed. Briefly, after normalization (log 10), all cardiometabolic variables (average blood pressure will be used as an index for systolic and diastolic blood pressure) will be standardized, i.e., z-scores will be computed: $z = \left(\frac{\text{value} - \text{mean}}{\text{SD}}\right)$. For HDL-cholesterol (protective for cardiometabolic risk), the z-score will be multiplied by -1. All z-scores will be summed, and the sum will divided by 5 to compile the cardiometabolic risk score with units of SD.

4.2 Interim Analysis and Multiple Testing

The Data Monitoring and Ethics Committee (DMEC) for the study met by teleconference on April 2, 2020 to review progress and interim data. The interim analysis was carried out to investigate the futility of the trial to date in terms of differences in average daily sitting time between the intervention groups and the control group at 12-month follow-up. On the basis of the data reviewed, the DMEC recommended continuation of the trial according to the current version of the protocol (version 1.7 04/09/2019) with no changes. Please refer to Interim SAP for more details on the interim analysis.

With regards to multiple testing, in this study there are two primary comparisons for the primary outcome (each intervention group vs control). The hypothesis tests and p-values will be two-sided, where a p-value of <0.025 will be considered to be statistically significant. Estimates will be presented with 97.5% confidence intervals and p-values.

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

UNIVERSITY OF LEICESTER

There will be no formal correction for multiple significance testing for the secondary analysis of the primary outcome, for the sensitivity and subgroup analyses for the primary outcome and for the secondary outcomes.

4.3 Analysis Software

The clinical data will be extracted from a MACRO database. Accelerometer data will be processed by the CI (blinded to arm) and transferred to the CTU via validated EXCEL sheets. The analysis will be performed with a current version of SAS, Stata or R. Multiple imputation will be implemented using the *jomo* package in R [9]. The version will be recorded in the Statistical Report.

5 Statistical Methodology

The statistical analysis will be based on external guidelines (e.g. ICH E3 and E9). The date of data extraction from the database will be included in each report.

5.1 Disposition

A flow of clusters through the trial will be summarised in a CONSORT diagram [10], as appropriate for cluster trials that will include the eligibility, reasons for exclusion, numbers randomised to the arms, lost to follow-up and numbers analysed.

Participant disposition will be presented with respect to completion status, reason for non-completion, protocol deviations, intervention compliance and length of stay in the trial. Results will be tabulated and summarised over time by arm and in total. Data completeness will also be summarised.

5.2 Demographic and Baseline Characteristics

Cluster and participant level baseline characteristics will be summarised by arm and in total.

Continuous data that are approximately normally distributed will be summarised in terms of the mean and standard deviation. This will be evaluated using histograms. Skewed data will be presented in terms of the medians and interquartile range. Ordinal and categorical data will be summarised in terms of frequency counts and percentages.

We will also carry out a descriptive comparison of baseline data between completers (i.e. participants who provide valid activPAL data at 12 months) vs. non-completers within treatment arms and overall.

5.3 Primary Outcome Analysis

5.3.1 Primary Analysis of Primary Outcome

The primary analysis will be performed using a linear multilevel model. Analysis of Covariance (ANCOVA) will be used with each participant's sitting time (measured using activPAL) at 12-month follow-up as the outcome, adjusting for their sitting time at

baseline and for the average waking wear time across baseline and 12-month follow-up. The model will also include a categorical variable for randomisation group (control as reference) and terms for the stratification factors (area and cluster size).

Office clusters will be included as a random effect to model worker heterogeneity within office sites. The variance-covariance matrix for the random effect will be assumed to be identity and the models will be estimated using restricted maximum likelihood. Also, level 1 and level 2 model residuals are assumed to follow a normal distribution and to have constant variance and these assumptions will be investigated using residual diagnostic plots. Alternative parameterisations and variable transformations will be considered where appropriate and if assumptions are not satisfied.

The models will only include participants who provide at least 1 valid wear day from the activPAL data at both baseline and 12 months.

For both comparisons, the estimate of the difference between intervention arm and control for average daily sitting time at 12 months and the corresponding 97.5% confidence intervals and p-values will be presented, statistical tests will be two-sided. Furthermore, the intra-class (clusters) correlation coefficient (ICC) and 95% confidence interval will be given to assess the strength of the clustering effect.

5.3.2 Secondary Analysis of Primary Outcome

If in the primary analysis of the primary outcome both interventions are shown to be effective, a secondary exploratory analysis will evaluate if one intervention is more effective than the other.

This analysis will use similar methodology to the primary analysis; however, there will be no formal adjustment for multiple significance testing as this is an unpowered analysis. The estimate of the difference between the intervention arms for average daily sitting time at 12 months will be presented with a 95% confidence interval and p-value, and the statistical test will be two-sided.

5.3.3 Sensitivity Analyses

The sensitivity analyses will be conducted using similar methodology as the primary analysis of the primary outcome (Section 5.3.1). However, there will be no formal adjustment for multiple significance testing.

5.3.3.1 Per Protocol Population

The effect size will also be estimated using a per-protocol analysis. The per protocol (PP) population are those who do not have the protocol deviations outlined in the Protocol Deviations Section 3.5.

5.3.3.2 Intention-To-Treat Population

Sensitivity analysis will be performed to assess the impact of missing data on the results found and to account for uncertainty associated with imputing data (full ITT analysis).

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

Originally, the ITT analysis stated that, where applicable, missing baseline categorical values will be replaced using the missing indicator method, missing baseline continuous variables will be replaced using cluster mean imputation and missing outcomes will be imputed using multilevel multiple imputation. However, post database lock the methods to carry out this analysis were changed. Please see more details below.

The ITT analysis will be performed using multilevel multiple imputation (MMI) in R with the jomo package. The steps below will be used as part of the MI process:
- The imputation model will register imputation of the outcome variable at baseline, 3 months and at 12 months. The imputation model will also have auxiliary variables as outcomes if they have missing data or as covariates if they do not have any missing data. The auxiliary variables are BMI at baseline, BMI at 3 months, gender, ethnicity, age, cluster size category (Small <10; Large ≥10) and area (Leicester; Salford; Liverpool). The model will also include average waking wear time across baseline and 12 months (again either as an outcome or a covariate depending on missingness) as it will be adjusted for in the model for the primary analysis.
- The MI will be multilevel with cluster ID as the cluster level variable.
- The multilevel MI in the jomo package in R is carried out using a joint modelling (JM) approach and with JM, for individuals for which one or more (but not all) outcome in the imputation model is missing, the imputation is carried out from the conditional distribution for one element of a multivariate normal model given the others. This means that in addition to the covariates specified, each of the outcome variables in the imputation model will also inform the imputation of the other outcomes if that information is available.
- The multilevel MI will use 20 imputations, 10,000 burn-in iterations and 10,000 between-imputation iterations and the imputations will be carried out separately by intervention arm [11]. A seed will be set in order to make the results reproducible.
- Once the imputations are carried out, the same model as the primary analysis of the primary outcome will be estimated using the lmer command in R. The model will be fitted for each of the 20 imputed datasets and then the estimates will be combined using Rubin's rules. For reproducibility, 95% confidence intervals on the basis of the final combined parameter estimates will subsequently be computed using the confint command. If both comparisons against the standard care arm in the primary analysis are statistically significant, a secondary analysis comparing the two intervention arms will also be carried out as part of these ITT analyses.

5.3.3.3 Effects on the number of valid activPAL days

We will assess the effect of the number of valid activPAL days chosen for the primary analysis and how the results obtained are affected by this change. This analysis will be performed by including participants who wore the activPAL for the following criteria:

- 4 valid days or more at both baseline and 12 months.
- 1 valid day or more of work days at both baseline and 12 months.
- 3 valid days or more of work days at both baseline and 12 months.

5.3.3.4 Standardising occupational/waking hours

To assess the impact of variation in occupational or waking hours between participants, time spent sitting will be normalised to an 8-hour workday for sitting during work hours and a 16-hour waking day for daily sitting. Average wear time across baseline and follow-up will not be included in the statistical models for these outcomes.

5.3.4 Subgroup Analyses

Subgroup analyses will be conducted for the primary outcome.

Methodology similar to that proposed for the primary analysis of the primary outcome in Section 5.3.1 will be used to assess the treatment effect in different subgroups of participants as outlined in Section 2.3. There will be no formal adjustment for multiple significance testing.

For each subgroup being assessed an indicator variable for subgroup assignment will be included in the model. An interaction term between intervention arm and subgroup will be included to assess the level of heterogeneity in treatment effect between the subgroups. An estimate of the treatment effect (i.e. difference between subgroups) and 95% confidence interval will be presented for each subgroup alongside the p-value for the interaction term.

For the Site subgroup analysis, if the model does not converge, the Site variable will be dichotomised into Leicester vs. Other.

5.4 Secondary Outcome Analyses

5.4.1 Primary Analysis of Secondary Outcomes

Secondary outcomes, including those measured at other time-points, will be analysed using similar methodology to the primary outcome. This will only apply to the following key secondary outcomes: sitting time, prolonged sitting time, standing time and stepping time – daily, during work hours and on work days and on non-work days calculated from the activPAL data variables. That is, the models for each of these outcome variables will adjust for their respective variable at baseline and for the respective average wear time period (i.e. daily, work hours, work days or non-work days) across baseline and follow-up (i.e. 3 months or 12 months). The models will also include a categorical variable for randomisation group (control as reference) and terms for the stratification factors (area and cluster size). These models will also only include participants with 1 valid day or more of the respective time period (i.e. daily, work hours, work days or non-workdays) at both baseline and follow-up (i.e. 3 months or 12 months). No corrections for multiple testing will be made. P-values and 95% CIs will be presented for these variables only.

Outcomes that are ordinal (i.e. ≤5 categories, [12]) or binary will be analysed using multilevel logistic regression models. There will be no formal adjustment for multiple significance testing. The estimates of the treatment effect will be presented with the

associated standard error. 95% confidence intervals and p-values will not be presented. The statistical tests will be two-sided.

For the other secondary outcomes defined in Section 2.2, continuous data that are approximately normally distributed will be summarised in terms of the mean and standard deviation. Skewed data will be presented in terms of the medians and interquartile range. Ordinal and categorical data will be summarised in terms of frequency counts and percentages. We will summarise all variables by intervention arm.

5.4.2 Sensitivity Analyses

Sensitivity analyses will be conducted for one secondary outcome: average sitting time during work hours at 12 months. Methodology similar to that proposed for the sensitivity analyses of the primary outcome will be employed (Section 5.3.3).

We will assess the effect of the number of valid activPAL days chosen for the primary analysis of this outcome and how the results obtained are affected by this change. This analysis will be performed by including participants who wore the activPAL for the following criteria:

- 3 valid days or more of work days at both baseline and 12 months.

5.4.3 Subgroup Analyses

Subgroup analyses will be conducted for one secondary outcome: average sitting time during work hours. Methodology similar to that proposed for the subgroup analyses of the primary outcome will be employed (Section 5.3.4).

5.5 Changes to the Planned Analysis

1. Section 2.3 (Subgroups and/or interactions): clarified definition of part-time vs. full-time workers.

2. Sections 2.4 (Compliance) and 5.4.1 (Primary analysis of secondary outcome): clarified definition of minimum number of valid activPAL wear days.

3. Section 5.4.2. (Sensitivity analyses for secondary outcomes): The effect of the number of valid activPAL days on sitting time across work hours at 12 months will be evaluated by carrying out a sensitivity analysis including only participants with 3 valid days or more of work days at both baseline and 12 months. Originally, it was stated that these sensitivity analyses for this outcome would use similar methodology to that proposed for the primary outcome, i.e. including participants with the following data both at baseline and 12 months: a) 4 valid days or more; b) 1 valid day or more of work days; 3 valid days or more of work days.

4. Section 5.3.3.2 (Intention-To-Treat Population): Originally, the ITT analysis stated that, where applicable, missing baseline categorical values will be replaced using the missing indicator method, missing baseline continuous variables will be replaced

using cluster mean imputation and missing outcomes will be imputed using multilevel multiple imputation. However, post database lock the methods to carry out this analysis were changed. A description of the methods is provided in Section 5.3.3.2.

5. Section 3.5 (Protocol Deviations): corrected that clusters whose workplace champions left their role within the first three months will be excluded, as will any ineligible clusters that did not have the minimum number of participants required (i.e. four or more). Also clarified that participants with time window deviations for their follow up (>+/- 2 months) in terms of their activPAL data will also be excluded.

6. Section 5.3.1 (Primary analysis of primary outcome): corrected structure of the variance-covariance matrix for the random effect from unstructured to identity.

7. Section 5.3.3.4 (Standardising occupational/waking hours): Models for standardised sitting time will not adjust for average wear time across baseline and follow-up. This was not stated originally in the SAP.

8. Appendix 1 - Scoring protocols for questionnaire-based secondary outcomes:
   a. Section 1 – time at work and sitting: added derivation of outcomes
   b. Section 11 – Correction as memory question does not need to be reversed
   c. Section 14 – EQ-5D-5L: scores used to derive of TTO Value Set have been corrected. The TTO value set will not be derived for any participants who have any missing items.
   d. Section 16 – Diet: added derivation of outcomes
   e. Section 17 – Additional questions: added derivation of outcome
   f. Provided interpretation of scores

6 Safety and Adverse events (AEs)

6.1 Adverse Events and Tolerability

As this is not a trial of an investigational medicinal product, only AEs that are potentially related to or may impact on the study interventions will be recorded. These are:

- Skin irritation from thigh monitor
- Skin irritation from wrist monitor
- Pain related to the intervention (e.g., desk use) or other intervention components
- Feeling unwell during blood test

AEs will be presented by arm and overall, under the headings seriousness and relatedness to the intervention. A listing providing full details of each event will also be produced. If there are too few events, only the listing will be produced.

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

7 References

1. Dunstan DW, Wiesner G, Eakin EG et al. Reducing office workers' sitting time: rationale and study design for the Stand Up Victoria cluster randomized trial. BMC Public Health. 2013;13(1):1057.

2. Snijders T, Boskers R. Multilevel Analysis: An introduction to basic and advanced multilevel modeling. Thousand Oaks, CA: Sage Publications Inc; 1999.

3. George, S. L. (1984). The required size and length of a phase III clinical trial. In *Cancer Clinical Trials: Methods and Practice*, (Edited by M. E. Buyse, M. J. Staquet, R. J Sylvester), Oxford University Press, New York, NY.

4. Edwardson CL, Yates T, Biddle SJH, Davies MJ, Dunstan DW, Esliger DW, et al. The effectiveness of the stand more AT (SMArT) work intervention: a cluster randomised controlled trial. BMJ. 2018;22.

5. Healy GN et al. A Cluster Randomized Controlled Trial to Reduce Office Workers' Sitting Time: Effect on Activity Outcomes. Med Sci Sports Exerc2016;48:1787-97

6. Groenwold RH, White IR et al (2012) Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. CMAJ 184:1265–1269

7. Sullivan T, White I, Salter A, et al. Should multiple imputation be the method of choice for handling missing data in randomized trials? Stat Methods Med Res. 2016;1–17.

8. Hossain A, Diaz-Ordaz K and Bartlett JW. Missing continuous outcomes under covariate dependent missingness in cluster randomised trials. Stat Methods Med Res 2016; 26: 1543–1562.

9. Quartagno, M., & Carpenter, J. (2018). Package 'jomo'. Retrieved from cran.r-project.org/web/packages/jomo/

10. Campbell et al. "CONSORT statement: extension to cluster randomised trials." BMJ 328.7441 2004; 702-708.

11. White IR, Royston P, Wood AM: Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011, 30:377–399.

12. S. J. Walters, "Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36," Health and Quality of Life Outcomes, vol. 2, article 26, 2004.

8 Appendix 1 – Scoring protocols for questionnaire-based secondary outcomes (Section 2.2.1)

Section 1 – Time at work and sitting

Average daily number of hours worked:
- Add up total number of days worked (Q1a)
- Add up total number of hours worked (Q1b)
- Divide total number of hours worked by total number of days worked

Average daily proportion of workday spent in office at council:
- Add up total daily proportions (Q1c)
- Divide total daily proportions by total number of days worked

Average daily proportion of workday spent sitting down:
- Add up total daily proportions (Q1e)
- Divide total daily proportions by total number of days worked

Average time spent sitting: a) in transport; b) watching TV; c) on computer; d) on other activities – not including time at work (Q6):
- Per day weekly average: add up total weekly time and divide by 7
- Per day weekday average: add up total weekday time and divide by 5
- Per day weekend average: add up total weekend time and divide by 2

Average time spent sitting overall – not including time at work (Q6):
- Add up sitting time over all four domains (transport, tv, computer, other).
- Per day weekly average: divide by 7
- Per day weekday average: divide by 5
- Per day weekend average: divide by 2

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

Section 2: Musculoskeletal Problems

Musculoskeletal symptoms - Standardised Nordic Questionnaire (SNQ)

- Trouble in last 12 months (Y/N)
  - Neck
  - Upper extremity (shoulder, upper back, elbow or hand)
  - Lower back
  - Lower extremity (hip, knee, or feet)
  - Any site

- Prevented from doing normal activities in last 12 months due to this trouble (Y/N)
  - Neck
  - Upper extremity (shoulder, upper back, elbow or hand)
  - Lower back
  - Lower extremity (hip, knee, or feet)
  - Any site

- Trouble in last 7 days (Y/N)
  - Neck
  - Upper extremity (shoulder, upper back, elbow or hand)
  - Lower back
  - Lower extremity (hip, knee, or feet)
  - Any site

- Rating of pain in last 12 months (1=no pain; 10=most pain can imagine)
  Higher scores indicate greater pain.
  - Neck
  - Upper extremity = $\frac{\text{shoulder} + \text{upper back} + \text{elbow} + \text{hand}}{4}$
  - Lower back
  - Lower extremity = $\frac{\text{hip} + \text{knee} + \text{feet}}{3}$
  - Any site = $\frac{\text{neck+shoulder+upperback+elbow+hand+lowerback+hip+knee+feet}}{9}$

UNIVERSITY OF LEICESTER

Section 4: Work Engagement
Utrecht Work Engagement Scale (UWES) (0=never; 6=always)

- Vigour = $\frac{\text{burst energy+vigorous+morning}}{3}$

  Higher scores indicate greater vigour.

- Dedication = $\frac{\text{enthusiastic+inspired+proud}}{3}$

  Higher scores indicate greater dedication.

- Absorption = $\frac{\text{intense+immerse+carried away}}{3}$

  Higher scores indicate greater absorption.

- Overall = $\frac{\text{energy+vigorous+enthusiasm+inspired+morning+intense+proud+immersed+carried away}}{9}$

  Higher scores indicate greater work engagement overall.

Section 5: Work Recovery
The Need for Recovery (NFR) Scale. (0=No; 1=Yes)

- Work Recovery:

$$\frac{\text{hard to relax + worn out + exhausted + fresh (reversed) + second day + no concentration + no interest + hour to recover + leave alone + tired + fatigued}}{11}$$

$$Q(\text{reversed}) = 1 + 0 - Q$$

Higher scores indicate greater need for recovery.

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

UNIVERSITY OF LEICESTER

Section 6: Work Limitations
Work Limitations Questionnaire
Positive answer for each question should be lowest score (1). Negative answer should be highest score (5).
When deriving questionnaire score need to exclude 6=does not apply responses.

- Time management = $\frac{\text{get going easily + start job asap}}{2}$

  Higher scores indicate worse time management.

- Physical demands = $\frac{\text{one position +motion}}{2}$

  Higher scores indicate greater physical demands.

- Mental/interpersonal difficulties = $\frac{\text{difficulty concentrate+difficulty speak}}{2}$

  Higher scores indicate greater mental/interpersonal difficulty.

- Output demands = $\frac{\text{handle workload+finish on time}}{2}$

  Higher scores indicate greater output demands.

- Overall productivity = $\frac{\text{time management+physical demands+mental/interpersonal +output}}{4}$

  Higher scores indicate worse overall productivity.

Section 8: Work Demands

Health and Safety Executive Management Standards Indicator Tool (HSE MSIT) (1=never; 5=always)

- Demands= $\frac{\text{diff.demands+unachievable+Intense+neglect tasks+sufficient breaks+long hours+fast+unrealistic}}{8}$

  Higher scores indicate greater demands.

- Control= $\frac{\text{when break+speed+how work+what work+way work+flexible}}{6}$

  Higher scores indicate greater control.

- Support = $\frac{\text{feedback+rely+talk+encourage+emotional support+help+support+respect+listen}}{9}$

  Higher scores indicate greater support.

Section 9: Social Norms and Cohesion

Social Norms: 1=Strongly Disagree; 5 = Strongly Agree
Social Norms: 1=Always; 5 = Never / Hardly ever

- Norms = $\frac{\text{work health (1A)+ work stand (1B)+colleagues desk (1C)+boss desk (1D)+colleagues meeting (1E)+boss meeting (1F)+ walk colleagues (1G)+walk boss (1H)}}{8}$

  Higher scores indicate better social norms.

- Cohesion = $\frac{\text{atmosphere (2A)+cooperation (2B)+community (2C)}}{3}$

  Higher scores indicate less social cohesion.

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

Section 10: Sleep Quality

Pittsburgh Sleep Quality Index (PSQI):

7 Components:	Question(s):
1. Subjective sleep quality	Q6. Overall sleep quality during past month (0=very good; 3=very bad)
2. Sleep latency	Q2. Time to fall asleep categorised: 0 = ≤15mins (No difficulty) 1 = 16-30 mins 2 = 31-60 mins 3 = >60 mins (Severe difficulty) Q5a. Cannot sleep ≤30 mins (0=not during past month; 3= ≥3 times a week) Sum of Q2 and Q5a = Component 2 score. Greater value = greater latency: 0 = 0 1-2 = 1 3-4 = 2 5-6 = 3
3. Sleep duration	Q4. Sleep hours at night during past month categorised: 0 = >7 hours (No difficulty) 1 = 6-7 hours 2 = 5-6 hours 3 = <5 hours (Severe difficulty)
4. Sleep efficiency	Sleep efficiency = $\left(\frac{\text{\# hours slept}}{\text{\# hours in bed}}\right) \times 100$ # hours slept = Q4 # hours in bed = calculated from responses to Q1 and Q3 Sleep efficiency. Higher values indicate better sleep efficiency: 0 = >85% (No difficulty) 1 = 75-84% 2 = 65-74% 3 = <65% (Severe difficulty)
5. Sleep disturbance	Q5b-Q5j should be scored as follows: 0 = Not during past month 1 = Less than once a week 2 = Once or twice a week 3 = Three or more times a week Sum Q5b to Q5j = Component 5 score. Higher values = greater disturbance: 0 = 0 (No difficulty) 1-9 = 1 10-18 = 2 19-27 = 3 (Severe difficulty)

6. Use of sleep medication	Q7. Use of medication to sleep over past month, should be scored as follows: 0 = Not during past month 1 = Less than once a week 2 = Once or twice a week 3 = Three or more times a week
7. Daytime dysfunction	Q8. Trouble staying awake over past month, should be scored as follows: 0 = Not during past month 1 = Less than once a week 2 = Once or twice a week 3 = Three or more times a week Q9. Problem to keep up enthusiasm to get things done over past month, should be scored as follows: 0 = No problem at all 1 = Only a very slight problem 2 = Somewhat of a problem 3 = A very big problem Sum of Q8 and Q9 = Component 7 score. Greater value = greater dysfunction: 0 = 0 (No difficulty) 1-2 = 1 3-4 = 2 5-6 = 3 (Severe difficulty)

Global PSQI Score = Sum of seven component scores
Higher scores indicate worse sleep quality.

Section 11: Fatigue

Physical and mental fatigue
Fatigue severity - Fatigue Scale (0=less than usual; 3=much more than usual):

- Physical (0-21) = tired + rest + sleepy + problem starting + energy + strength + weak
  Higher scores indicate higher fatigue.
- Mental (0-12) = concentrate + speaking + right word + memory
  Higher scores indicate higher fatigue.
- Global (0-33) = tired + rest + sleepy + problem starting + energy + strength + weak + concentrate + speaking + right word + memory
  Higher scores indicate higher fatigue.

Trial: Edwardson_16_324_SMArt_WorkandLife
SAP Version: 1.2
Version Date: 08/02/2021

Section 12: Anxiety and Depression
Hospital Anxiety and Depression Scale (HADS). (0 to 3 likert scales).
Positive answer should be lowest score (0). Negative answer should be highest score (3).

- Anxiety (0-21) = butterflies + frightened + restless + ease relaxed + sudden panic + tense + worrying thoughts
  Higher scores indicate greater anxiety.
- Depression (0-21) = appearance + cheerful + enjoy book tv + enjoy things + funny side + look forward + slowed down
  Higher scores indicate greater depression.

Section 12: Wellbeing
WHO-5 Scale. (0=at no time; 5=all of the time).

- Well-being index (0-100) = (cheerful + calm + active + fresh + interest) × 4
  Higher scores indicate greater well-being.

Section 12: Stress
Perceived Stress Scale. (0=never; 4=very often):

- PSS score (0-40) = upset + unable to control + nervous + confident (reversed) + your way (reversed) + unable to cope + control irritations (reversed) + on top of things (reversed) + angered + unable to overcome difficulties Q(reversed) = 4 + 0 – Q

  Higher scores indicate greater stress.

Section 13: Mood
Positive and Negative Affect Schedule (PANAS). (1=Very slightly or not at all; 5=Extremely)

- Positive (10-50)=interested + excited + strong + enthusiastic + proud + alert + inspired + determined + attentive + active
  Higher scores indicate higher levels of positive affect.
- Negative (10-50) = distressed + upset + guilty + scared + hostile + irritable + ashamed + nervous + jittery + afraid
  Lower scores indicate lower levels of negative affect.

UNIVERSITY OF LEICESTER

Section 14: General Health

Health-related quality of life - EQ-5D-5L:

The EQ-5D-5L is summarised in a continuous score (called a TTO value set), ranging from -0.285 to 1.000, where a higher score indicates higher health utility. The scoring algorithm, available on the EuroQol website [EuroQolweb], is

$$1 - \sum_{i=1}^{5} score\_i$$

where *score_i* is the score of question *i* according to the following table:

Selected option	1	2	3	4	5
Question					
1 Mobility	0	0.058	0.076	0.207	0.274
2 Self care	0	0.050	0.080	0.164	0.203
3 Usual activities	0	0.050	0.063	0.162	0.184
4 Pain & discomfort	0	0.063	0.084	0.276	0.335
5 Anxiety and depression	0	0.078	0.104	0.285	0.289

If a subject selects option 3 "Moderate problems walking about" in Q1, score_1 = 0.076.

The TTO value set will not be derived for any participants who have any missing items.

Higher scores indicate higher health utility.

Section 16: Diet

Snacks:

- Calculate number of participants who had each snack once a day or more often.
- Calculate number of participants who had any snack once a day or more often.

Alcohol:

- Spirits units: one UK unit for each measure of spirits.
- Wine units: one UK unit for each glass of wine.
- Beer units: two UK units for each pint of beer.
- Total units: sum of spirits, wine and beer units.

Section 17: Additional Questions

Support to sit less and move more often:

- Calculate summary scores for each party (organisation, manager, colleagues, family).