Heated mittens for patients with hand osteoarthritis: A randomized controlled trial

Trial acronym: HOT

Trial registration number: www.clinicaltrials.gov NCT04576403

Study protocol version: Version 1.0 - 06/10/2020

SAP version number with date: Version 1.0 25/11/2022

SAP revision history, justifications, and timing: SAP version 1

Author: Marius Henriksen[1], PT, MSc, PhD

Signature and date: 25. nov 2022

Senior biostatistician: Robin Christensen[1,2], BSc, MSc, PhD

Signature and date: 2023, Feb. 01

Statistical analyst: Sabrina Mai Nielsen[1,2], BSc, MSc, PhD

Signature and date: 29. nov 2022

Chief investigator/clinical lead: Karen Ellegaard[1], PT, MSc, PhD

Signature and date: 3.feb 2023

Other contributors:

Cecilie Rødgaard Bartholdy[1], investigator

Anna Døssing[1], investigator

Henning Bliddal[1], investigator

Affiliations:

1. The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark

2. Research Unit of Rheumatology, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, Odense, Denmark

1. STUDY OVERVIEW

Background and rationale Hand osteoarthritis (HOA) is a common disease characterized by reduced function, stiffness, and pain. In the most recent American College of Rheumatology guideline for the management of HOA (2020) heating is recommended as a treatment but without scientifically sound evidence of beneficial effects. Further, our local HOA patient partners have via interviews unanimously stated that heat is their preferred treatment for symptom reduction including finger stiffness. Studies assessing the potential benefits of non-pharmacological treatments of HOA are scarce and both ACR, European League against Rheumatism (EULAR) and Osteoarthritis Research Society International (OARSI) recommend exploration of this area. Thus, we find it relevant to investigate if a daily intervention with electrically heated mittens can help reduce pain, improve function and reduce stiffness of the hands in patients with HOA.

Objectives The aim of this randomized trial is to investigate the effect of electrically heated mittens after 6 weeks (assessed in week 7) on physical function in patients with HOA compared to sham mittens (inactivated electrical heating).

Methods This study is designed as a randomized trial with two parallel groups with change from baseline in physical function of the hand (measured by the AUSCAN questionnaire) as primary endpoint after 6 weeks, with investigators, outcome assessors, and participants being blinded to treatment allocation.

PICOTS

***Population**:* Individuals with HOA.

Intervention: Electrically heated mittens worn at least 15 minutes every day, preferably in the morning, for 6 weeks.

Comparator: Identically appearing mittens with the electrical heating element being deactivated worn at least 15 minutes every day, preferably in the morning, for 6 weeks.

Outcome (primary): Change from baseline to week 7 in the physical function subscale of the Australian/Canadian Osteoarthritis Hand Index (AUSCAN).

Time: 6 full weeks.

Study design: Randomized trial with two parallel groups.

Further statistical details

Randomization: Computer-generated randomization list was developed based upon permuted random blocks of variable size (2 to 6 in each block). The allocation ratio was 1:1 stratified for the presence of OA of the first carpometacarpal joint (CMC-1).

Sample size: A sample size of 200 in total will provide strong statistical power to detect a difference between groups in the primary outcome of 8 AUSCAN-function points (normalized 0-100 scale; equals 81 points on original 0-900 scale). For a two-sample pooled t-test of a normal mean difference with a two-sided significance level of 0.05 ($P<0.05$), assuming a common standard deviation of 19 AUSCAN-function points (normalized 0-100 scale; equals 171 points on original 0-900 scale), a total sample size of 180 HOA patients (i.e. approximately 90 participants per group) has a power of 80.2% to detect a mean difference of 8 AUSCAN-function points (small effect size of 0.42). To account for dropouts, it was decided to include 200 patients in the trial.

Framework: This is a superiority trial assessing if electrically heated mittens are superior to sham mittens for improvements in physical function of the hands.

Statistical interim analyses and stopping guidance: None.

Timing of final analysis: When this statistical analysis plan was finalized and signed, recruitment to the HOT trial had not been completed. We expect recruitment to be completed by the end of March 2023. We will close the database 2 months after the last participant's last visit at the latest. Statistical analyses are expected to be completed after additionally 2 months at the latest.

Timing of outcome assessment: (see next section).

Confidence intervals and P values: All 95% confidence intervals and P-values will be two-sided.

Multiplicity: No explicit adjustments but hierarchical testing of primary and key secondary outcomes.

Statistical software: R version 4.0.3 (or newer).

2. TABULAR PRESENTATION OF TIMING OF OUTCOME MEASUREMENTS

	Baseline	Treatment period						Primary endpoint
Week	0	1	2	3	4	5	6	7
Window (days from first treatment)	-7 to -1		11 to 17		25 to 31			43 to 50
Clinical examination (TJC & SJC)	X							X
AUSCAN	X		X		X			X
VAS pain/global	X							X
Analgesics	X[a]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[a]
Grip strength	X							X
Mitten use		X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	

AUSCAN, Australian/Canadian Osteoarthritis Hand Index; SJC, swollen joint count; TJC, tender joint count; VAS, visual analog scale.
a: Interview-based.
b: Recorded by the participants in a diary.

3. ELABORATIONS ON OUTCOMES AND DATA

Data management:
All outcomes: Change from baseline in all outcomes will be calculated as the baseline values minus follow-up values.

AUSCAN: The AUSCAN consists of 3 subscales, AUSCAN function, AUSCAN pain, and AUSCAN stiffness, each with 9, 5, and 1 items, respectively. Each item is scored on a 0-100 mm visual analogue scale (VAS) scale (0, best; 100, worst). Hence, the total scores of each subscale are 0-900 (AUSCAN function), 0-500 (AUSCAN pain), and 0-100 (AUSCAN stiffness).

Tender and Swollen Joint Counts: For the TJC and SJC, 15 joints in each hand are assessed: CMC-1, MCP1 to 5, IP, PIP2 to 5, and DIP2 to 5 in both hands. The scales span from 0 (no tender or swollen joints) to 30 (all joints on both hands are tender or swollen).

VAS pain: In the protocol (page 10) the description of the VAS pain outcome measurement could be read as if the VAS pain assessment was only targeted one hand. In reality, it was targeted the average of both hands and the outcome reflects the average pain intensity across both hands.

Analgesics: Use of analgesics is recorded as use of *i*) Acetaminophen/paracetamol, ii) NDAIDs, and iii) Other analgesics. The amount is recorded as either '*Almost daily*', '*2-3 times per week*', '*Rarer*', or '*Never*'. The recording at baseline and primary endpoint is based on an interview with an investigator. In the treatment period (week 1-6) the recording is made by the participant in a diary and recorded once per week. We collapse '*Almost daily'* and '*2-3 times per week'* into a category called '*Analgesic user'* and '*Rarer'* and '*Never'* into the category '*Analgesic non-user*'.

Data validation:
All variables used in the analyses, including the derived variables, will be checked for missing values, outliers, and inconsistencies and queried.

Data template:
Based on this SAP, the statistical analyst will develop a tailored data template illustrating the data structure required for the statistical analyses.

4. PROTOCOL DEVIATIONS WITH BEARING ON THIS STATISTICAL ANALYSIS PLAN

The following details in this SAP represent deviations from the trial protocol.

Header in the protocol	Change	Reason
(currently no deviations)		

5. OUTLINE

The anticipated (predefined) outline of the manuscript is illustrated below.

Results that only will be presented in the manuscript text include:

- Adherence, defined as number of days that the participants self-report mitten use of at least 15 minutes. The mean number of days for each group will be reported. In case of skewed data medians with interquartile ranges (IQRs) together with difference in medians with 95% confidence intervals will be presented.
- Mitten usage - time. The total amount of self-reported (diary) mitten usage (in minutes) will be reported as descriptive statistics using means and standard deviations for each group, together with a group difference in means with a 95% confidence interval. In case of skewed data medians with IQRs together with difference in medians with 95% confidence intervals will be presented.
- Mitten usage - frequency. The mittens can be used several times per day if participants wish to do so and will be reported as the median with range together with the most frequently reported frequency.
- Mitten intensities. The mittens have 3 heating intensities red (max), yellow (medium), green (min) and the participants record the used intensity in the diary. These will be summarized as proportion of participants who mainly use red, yellow and green intensities during the entire 6 weeks using the number of days with mitten use as denominator for each participant.
- Analgesics use. We will elaborate changes from baseline in the status of analgesic use that may aid in the interpretation of the patient reported pain and function scores.
- Harm outcomes (including adverse events, serious adverse events, and deaths) will be reported as numbers and % in each group. The safety set will be used as the denominator to calculate the percentages. The safety population includes patients who have reported at least one day with at least 15 minutes of mitten usage.
- At the end-of-trial visit, the participants were asked to guess whether the mittens were heated or sham mittens. We will report the degree of blinding success as the agreement between actual allocation and guesses. We will use Cohen's kappa for the analysis as this accounts for correct guesses by chance. A kappa value of 0 indicates successful blinding, and a kappa value of 1 reveals that all the participants can correctly identify a treatment so that the blinding has been completely broken. A positive value implies failure of blinding, whereby most participants correctly guess the treatment allocation above random guessing. A negative value from 0 to -0.20 indicates that participants have been unable to tell the treatment allocations, while a more extreme negative one implied blinding failure in the other direction. We define a kappa value of -0.20 to 0.20 as successful blinding, 0.21 to 0.40 as slightly broken, 0.41 to 0.60 as moderately broken, 0.61 to 1 as severely broken.

Figure 1. Flow diagram

Anticipated plot design, illustrating potential reasons for exclusion:

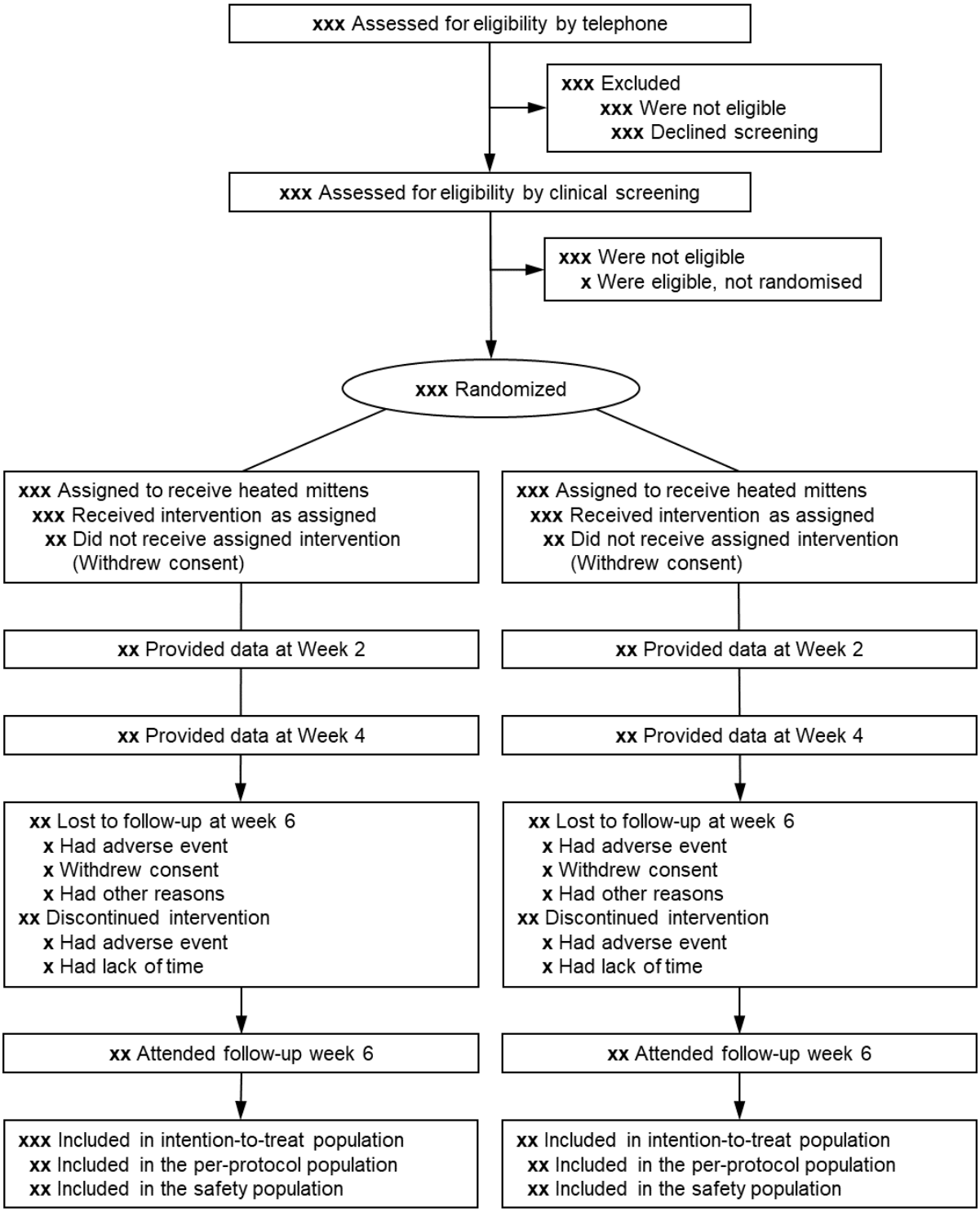


Table 1. Baseline characteristics in the intention-to-treat population

Characteristics	Heated mittens (n =)	Sham mittens (n =)	Total (n=)
Demographics			
Age, years			
Females, n (%)			
Height, m			
Weight, kg			
Body Mass Index, kg/m^2			
Disease duration, years			
Stratification factor			
CMC-1 OA, n (%)			
AUSCAN scores			
Physical function[a], 0-900 score			
Pain, 0-500 score			
Stiffness, 0-100 score			
Clinical assessments			
Tender joint count, 0-30			
Swollen joint count, 0-30			
Number of CMC-1 dislocations, 0-2			
Visual analog scales, 0-100			
Hand pain			
Global rating of hand OA related problems			
Performance measure			
Grip strength right hand, N			
Grip strength left hand, N			
Analgesics			
Analgesic user, n (%)			

Values are mean (SD) unless otherwise stated in the table.
CMC-1 OA, first carpometacarpal osteoarthritis; AUSCAN, Australian/Canadian Osteoarthritis Hand Index.
[a]Primary outcome measure.

Further statistical information related to table 1:

Data will be presented as mean with standard deviation (SD) when normally distributed or as median with interquartile range in case of skewed data. Dichotomous and categorical data will be presented in proportions. Normality of the data will be assessed using Q-Q plots, and histograms.

Table 2. Change from baseline in primary and secondary outcomes at week 7 in the ITT population

	Heated mittens (n =)	Sham mittens (n =)	Difference between groups (95% CI)	*P*-value
Primary outcome				
AUSCAN physical function, 0-900 score				
Key secondary outcomes				
AUSCAN pain, 0-500 score				
Global rating of hand OA related problems, 0-100 VAS score				
Grip strength right hand, N				
Grip strength left hand, N				
Other secondary outcomes				
AUSCAN stiffness, 0-100 score				-
Tender joint count, 0-30				-
Swollen joint count, 0-30				-
Hand pain, 0-100 VAS score				-
Analgesics discontinued, n(%)				-

Values are least squares means (standard error) unless otherwise stated in the table.
95%CI, 95% confidence interval; AUSCAN, Australian/Canadian Osteoarthritis Hand Index; ITT, intention-to-treat; VAS, visual analog scale. N, Newtons.

Further statistical information related to table 2:

The analysis population will include all randomized participants following the intention-to-treat (ITT) principle. In case baseline information is missing, the analysis population will be following a modified intention-to-treat (mITT) principle, including only those with available baseline data for the outcome.

Continuous data with repeated measures (i.e., AUSCAN subscores) will be analyzed using linear mixed models estimating the group mean changes from baseline and the differences between groups at week 7. The analyses will include the participant as random effect, group (2 levels) and week (4 levels; week 0, 2, 4 and 7) as fixed effects, the group*week interaction, as well as the baseline value of the outcome under analysis and the stratification factor (CMC-1 OA) as covariates. Missing data will be handled implicitly by the repeated measures mixed linear models, assuming data missing at random (MAR).

Continuous data measured only at baseline and 7 weeks will be analyzed using ANCOVA models with group as fixed effect (2 levels), as well as the baseline value of the outcome under analysis and the stratification factor (CMC-1 OA) as covariates. Missing data will be handled with multiple imputation using multivariate imputation by chained equations assuming MAR. Imputations will be conducted separately by group, and the imputation model will be conditioned on relevant variables, including baseline variables, and covariates and outcome of the analysis models. Imputation models with many auxiliary variables preserve relationships among variables and provide more precise and accurate imputations (Collins et al. 2001). Estimates will be pooled across 100 imputed datasets (Graham et al. 2007; Rubin 1987). The imputations will be performed using the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011).

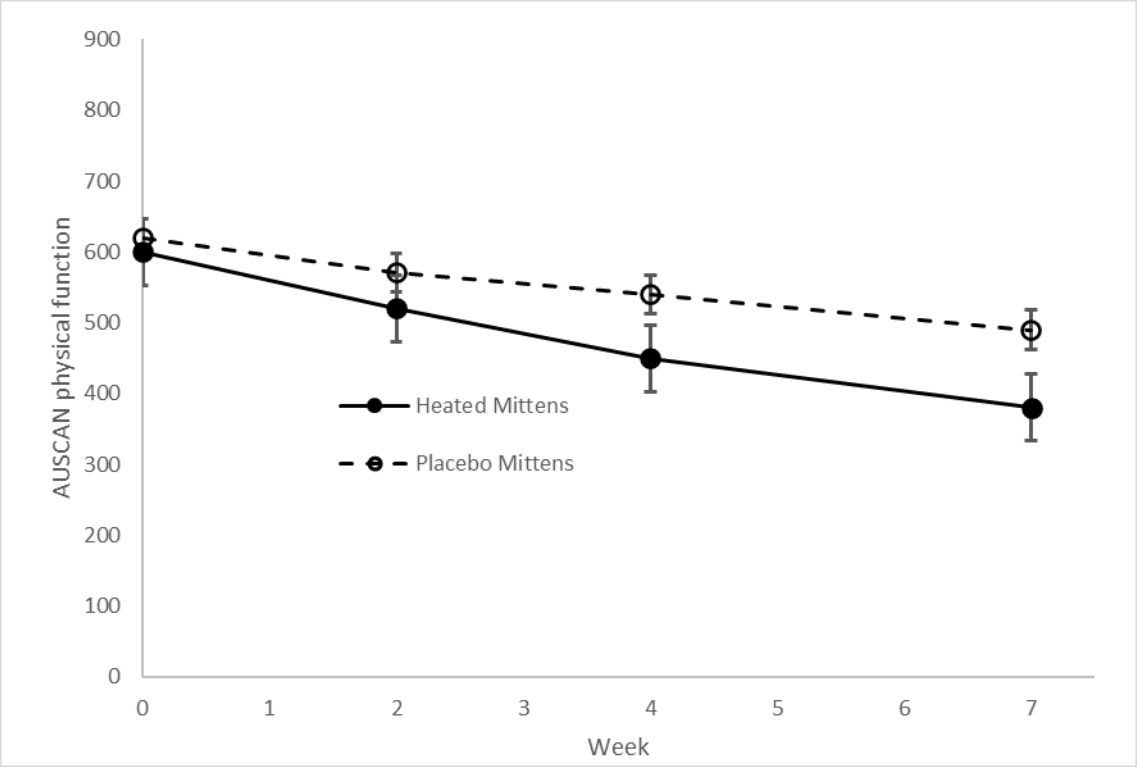
Assumptions will be checked by visual inspection of residual plots assessing the normality of residuals. In case the distributional assumptions do not hold, we will use an appropriate transformation (e.g., log-transformation in the case of right-skewed data and report the results as geometric means and geometric mean ratios), and/or, use non-parametric methods.

We will not apply explicit adjustments for multiplicity, rather we will analyze the key secondary outcomes in prioritized order (i.e. “inverse gatekeeping procedure”): The hypothesis testing of the key secondary outcomes will be performed in sequence until one of the analyses fails to show statistical significance. The hierarchy is illustrated by the order of key secondary outcomes in table 2 (top-down). For the other secondary outcomes no hypothesis testing will be done.

Figure 2. Trajectories for the primary efficacy outcome measure (AUSCAN physical function) from baseline to 7 weeks follow-up in the ITT population

Hypothetical trajectories of the primary outcome measure.

Values are least squares means over time from baseline to 7 weeks follow-up for heated mittens group (solid points) and sham mittens group (hollow points). Error bars indicate standard error of the estimates



Further statistical information related to figure 2

Least-squares mean estimates and standard errors for AUSCAN physical function by group will be estimated based on a model similar to that of the primary analysis.

SUPPLEMENTARY MATERIAL

The anticipated (predefined) supplementary material of the manuscript is illustrated below.

Supplementary file 1. Protocol

Supplemental file 2. Predefined protocol violations with bearing on the interpretation of the trial

	Heated mittens (n =)	Sham mittens (n =)
Major protocol violations		
Primary outcome taken outside visit window [a]		
Prohibited concomitant treatments received [b]		

Values are number (percentage) using the ITT population as denominator.
a: Primary outcome assessment is scheduled to occur between 43 and 50 days from first treatment
b: Major surgery, hand surgery, steroid injections, oral steroids

Explanation:

Protocol deviations will be classified prior to unblinding of the treatment. The number (%) with major and minor protocol deviations will be summarized by treatment group with details of type of deviation provided. The ITT analysis population will be used as the denominator to calculate the percentages.

Supplementary file 3. This SAP

Supplementary table 1. Change from baseline in primary and secondary outcomes at week 7 in the ITT population, imputing missing data with BOCF

[same design as Table 2, but with no p-value column]

Sensitivity analyses using the same model as in the primary analysis, but with missing data imputed using baseline observation carried forward (BOCF) technique. This assumes data missing not at random (MNAR), in contrast to the primary analysis that assumes MAR.

Supplementary table 2. Change from baseline in primary and secondary outcomes at week 7 in the per-protocol population

[same design as Table 2, but with no p-value column]

Subgroup analyses using the same model as in the primary analysis, but only including the per-protocol population. The per-protocol population is defined as participants who were randomly assigned to treatment, have a primary outcome data both a baseline and at the primary endpoint assessment, report mitten use of at least 15 minutes on at least 30 days, and who have no major protocol violations (see table in the supplementary file 2).

6. REFERENCES

Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychol Methods. 2001;6(4):330–351.

Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, Williamson PR, Altman DG, Montgomery A, Lim P, Berlin J, Senn S, Day S, Barbachano Y, Loder E. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017 Dec 19;318(23):2337-2343. Doi: 10.1001/jama.2017.18556. PMID: 29260229.

Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci. 2007;8(3):206-213. Doi:10.1007/s11121-007-0070-9

Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York, NY: Wiley; 1987.

Stef van Buuren, Karin Groothuis-Oudshoorn (2011). Mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67. DOI 10.18637/jss.v045.i03.

7. SAP REPORTING GUIDELINE

This SAP has been reported according to the items recommended by Gamble et al. (2017). Explanation and elaboration of the items are available in their eAppendix 2: https://jamanetwork.com/journals/jama/fullarticle/2666509

Table A. SAP Guidance Document: Recommended Items to Address in a Clinical Trial SAP[a]

Section/Item	Index	Description	Location in this SAP
Section 1: Administrative Information			
Title and trial registration	1a	Descriptive title that matches the protocol, with SAP either as a forerunner or subtitle, and trial acronym (if applicable)	*Front page*
	1b	Trial registration number	*Front page*
SAP version	2	SAP version number with dates	*Front page*
Protocol version	3	Reference to version of protocol being used	*Front page*
SAP revisions	4a	SAP revision history	*Front page*
	4b	Justification for each SAP revision	*Front page*
	4c	Timing of SAP revisions in relation to interim analyses, etc	*Front page*
Roles and responsibility	5	Names, affiliations, and roles of SAP contributors	*Front page*
Signatures of:	6a	Person writing the SAP	*Front page*
	6b	Senior statistician responsible	*Front page*
	6c	Chief investigator/clinical lead	*Front page*
Section 2: Introduction			
Background and rationale	7	Synopsis of trial background and rationale including a brief description of research question and brief justification for undertaking the trial	*1. Study overview*
Objectives	8	Description of specific objectives or hypotheses	*1. Study overview*
Section 3: Study Methods			
Trial design	9	Brief description of trial design including type of trial (eg, parallel group, multiarm, crossover, factorial) and allocation ratio and may include brief description of interventions	*1. Study overview*
Randomization	10	Randomization details, eg, whether any minimization or stratification occurred (including stratifying factors used or the location of that information if it is not held within the SAP)	*1. Study overview*
Sample size	11	Full sample size calculation or reference to sample size calculation in protocol (instead of replication in SAP)	*1. Study overview*
Framework	12	Superiority, equivalence, or noninferiority hypothesis testing framework, including which comparisons will be presented on this basis	*1. Study overview*
Statistical interim analyses and stopping guidance	13a	Information on interim analyses specifying what interim analyses will be carried out and listing of time points	*1. Study overview*
	13b	Any planned adjustment of the significance level due to interim analysis	*1. Study overview*
	13c	Details of guidelines for stopping the trial early	*1. Study overview*
Timing of final analysis	14	Timing of final analysis, eg, all outcomes analyzed collectively or timing stratified by planned length of follow-up	*1. Study overview*
Timing of outcome assessments	15	Time points at which the outcomes are measured including visit "windows"	*2. Tabular presentation of timing of outcome measurements*
Section 4: Statistical Principles			
Confidence intervals and *P* values	16	Level of statistical significance	*1. Study overview*
	17	Description and rationale for any adjustment for multiplicity and, if so, detailing how the type 1 error is to be controlled	*1. Study overview*
	18	Confidence intervals to be reported	*1. Study overview*

Adherence and protocol deviations	19a	Definition of adherence to the intervention and how this is assessed including extent of exposure	*5. Outline (Results to be reported in text)*
	19b	Description of how adherence to the intervention will be presented	*5. Outline (Results to be reported in text)*
	19c	Definition of protocol deviations for the trial	*4. Protocol deviations*
	19d	Description of which protocol deviations will be summarized	*5. Outline (Supplementary file 2)*
Analysis populations	20	Definition of analysis populations, eg, intention to treat, per protocol, complete case, safety	*5. Outline (for each table and figure)*
Section 5: Trial Population			
Screening data	21	Reporting of screening data (if collected) to describe representativeness of trial sample	*5. Outline (Figure 1)*
Eligibility	22	Summary of eligibility criteria	*1. Study overview*
Recruitment	23	Information to be included in the CONSORT flow diagram	*5. Outline (Figure 1)*
Withdrawal/follow-up	24a	Level of withdrawal, eg, from intervention and/or from follow-up	*5. Outline (Figure 1)*
	24b	Timing of withdrawal/lost to follow-up data	*5. Outline (Figure 1)*
	24c	Reasons and details of how withdrawal/lost to follow-up data will be presented	*5. Outline (Figure 1)*
Baseline patient characteristics	25a	List of baseline characteristics to be summarized	*5. Outline (Table 1)*
	25b	Details of how baseline characteristics will be descriptively summarized	*5. Outline (Table 1)*
Section 6: Analysis			
Outcome definitions		List and describe each primary and secondary outcome including details of:	
	26a	specification of outcomes and timings. If applicable include the order of importance of primary or key secondary end points (eg, order in which they will be tested)	*5. Outline (Table 2)*
	26b	specific measurement and units (eg, glucose control, hbA_{1c} [mmol/mol or %])	*5. Outline (Table 2)*
	26c	any calculation or transformation used to derive the outcome (eg, change from baseline, QoL score, time to event, logarithm, etc)	*5. Outline (Table 2)* ***and*** *3. Elaborations on outcomes and data*
Analysis methods	27a	what analysis method will be used and how the treatment effects will be presented	*5. Outline (Table 2)*
	27b	any adjustment for covariates	*5. Outline (Table 2)*
	27c	methods used for assumptions to be checked for statistical methods	*5. Outline (Table 2)*
	27d	details of alternative methods to be used if distributional assumptions do not hold, eg, normality, proportional hazards, etc	*5. Outline (Table 2)*
	27e	any planned sensitivity analyses for each outcome where applicable	*5. Outline (Supplementary table 1)*
	27f	any planned subgroup analyses for each outcome including how subgroups are defined	*5. Outline (Supplementary table 2)*
Missing data	28	Reporting and assumptions/statistical methods to handle missing data (eg, multiple imputation)	*5. Outline (Table 2)*
Additional analyses	29	Details of any additional statistical analyses required, eg, complier-average causal effect analysis	-
Harms	30	Sufficient detail on summarizing safety data, eg, information on severity, expectedness, and causality; details of how adverse events are coded or categorized; how adverse event data will be analyzed, ie, grade 3/4 only, incidence case analysis, intervention emergent analysis	*5. Outline (Results to be reported in text)*
Statistical software	31	Details of statistical packages to be used to carry out analyses	*1. Study overview*
References	32a	References to be provided for nonstandard statistical methods	*5. References*
	32b	Reference to Data Management Plan	-
	32c	Reference to the Trial Master File and Statistical Master File	-
	32d	Reference to other standard operating procedures or documents to be adhered to	-

Abbreviations: CONSORT, Consolidated Standards of Reporting Trials; hbA_{1c}, hemoglobin A_{1c}; QoL, quality of life; SAP, statistical analysis plan.
a Reproduced from Gamble et al. (2017).