Statistical Analysis Plan (SAP) Statement: Explanation and elaboration of essential items

Section 1: Administrative information

1. Title and Trial registration

Item 1a: Descriptive title that matches the protocol, with 'Statistical analysis plan' either as a fore runner or sub title, and trial acronym (if applicable)

Explanation

The title provides vital information required for trial identification. The title should unambiguously state which trial the SAP relates to and should therefore be identical to the trial protocol with 'Statistical analysis plan' either as a fore runner or sub title. Ideally the title should identify the study design, population, interventions, and, if applicable, trial acronym.

> **Example**
> "Statistical analysis plan for the Stroke Oxygen Study (SO2S): a multi-center randomized controlled trial to assess whether routine oxygen supplementation in the first 72 hours after a stroke improves long-term outcome." [1]

Item 1b: Trial registration number

Explanation

A trial registration number should be provided which uniquely identifies a clinical trial and its existence on a publicly-accessible registry. The International Committee of Medical Journal Editors (ICMJE) mandates the registration of clinical trials in a primary register of the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) or in ClinicalTrials.gov before recruitment of the first patient as a condition of consideration for publication[2]. This identifier should be clearly listed in all relevant documentation including the protocol and the SAP.

> **Example**
> "Trial registration: ISRCTN50133740." [3]

2. SAP Version

Item 2: SAP version number with dates

Explanation

Sequentially numbering and dating each SAP version avoids any confusion over which document is the most recent. Transparent tracking of version numbers and amendments facilitates trial conduct, review and oversight. The first final version of a document will be Version 1.0. It is recommended that subsequent final documents will have an increase of "1.0" in the version number (1.0, 2.0, etc.).[4] While the document is under review, subsequent draft versions will increase by "0.1", e.g., 1.1, 1.2, 1.3, etc. When the revised document is deemed final, the version will increase by "1.0" over the version being revised, e.g. the draft 1.3 will become a final 2.0.

Example
Version: 1.0 Date: July 3, 2014[5]

3. Protocol Version

Item 3: Reference to version of Protocol being used

Explanation

Referencing the version of the protocol being used is helpful as it links the SAP to the protocol and serves as a reminder that the SAP is not a standalone document and needs to be read in conjunction with the corresponding version of the protocol. This avoids the need for the author to duplicate information from the protocol in the SAP. If there have been protocol amendments after the SAP has been written then the SAP needs to be reviewed against the amendments, and updated where necessary. The information in Table 2 may be extended to record that the SAP has been reviewed in light of protocol amendments but no changes were required.

Example
This document has been written based on information contained in the study protocol version 5, dated 11 December 2012.[6]

4. SAP Revisions – revision history, with justification and timing

Item 4a/4b/4c: SAP Revision History

> *Justification for each SAP revision*
> *Timing of SAP revisions in relation to interim analyses etc.*

Explanation

A clear explanation of the changes made between each version of the SAP is essential, along with a justification for the revision and the date. This is important to maintain transparency. After the first version of the SAP is agreed and signed off, the SAP revision history should include the following information: the previous version number, the SAP section changed, details of the change made along with justification for the revision, and date of revision. A justification for each SAP revision is necessary to document the reasons for changes. This ensures the external validity of the trial as it demonstrates that changes are not being made based on unblinded trial data. From a regulatory perspective when SAP revisions occur after unblinded interim analyses have been conducted the people involved in deciding, writing, or approving the SAP should ideally have no knowledge of unblinded data particularly if the trial will be used for a licence application. In other situations it may be sufficient for the justification to document the reason for the change is not based upon comparative data and for the approver to have no knowledge of unblinded data.

Table 2: Example of SAP revision history[7]:

Protocol version	Updated SAP version no.	Section number changed	Description of and reason for change	Date changed
1.0	2.0	Appendix D	Organisms added to the appendix	21/02/2014
2.0	2.0	No changes required	SAP reviewed against protocol amendments	31/07/2014

5. Roles and Responsibility – non-signatory names and contribution

Item 5: Names, affiliations, and roles of SAP contributors

Explanation

Individuals who contribute significantly to SAP development should have their contributions described. Listing the SAP contributors, their affiliations and their roles in the SAP development process provides due recognition, accountability, and transparency. Naming of authors and statements of author's contributions is standard for SAPs published in journals such as *Trials*, but rare

in unpublished SAPs. Contributors may be non-signatory members if only the statistician writing the SAP, supervising senior statistician and the chief investigator/clinical lead will sign and approve the SAP.

6. Roles and Responsibility – signatures

Item 6a: Signature of person writing the SAP

Explanation

The signature of the person writing the SAP is crucial as it identifies who is responsible for the SAP and that they have approved the SAP. In all circumstances this should be signed and dated. If an update has been made then the author of the update should sign the updated version.

Item 6b: Signature of senior statistician responsible

Explanation

The signature of the senior statistician responsible for overseeing the trial is important as it highlights that the SAP has been reviewed and approved by an experienced statistician. In some circumstances the senior statistician may be the person writing the SAP and such a dual role should be reflected in the signatories. The signature should always be dated.

Item 6c: Signature of chief investigator/clinical lead

Explanation

The signature of the chief investigator/clinical lead demonstrates that they have reviewed and approved the SAP. Once the final version has been approved and signed off it avoids any post-hoc changes being made without the justification and approval of all signatory members to maintain internal and external trial validity. The signature should always be dated.

Section 2: Introduction

7. Background and rationale (optional)

Item 7: Synopsis of trial background and rationale including brief description of research question and brief justification for undertaking the trial

Explanation

The full rationale for undertaking the trial and trial background are explained in detail in the protocol so only a brief synopsis is necessary within a SAP to avoid duplication of information. The synopsis should include justification for undertaking the trial, why the trial is needed and description of the research question. This item would be regarded as essential if the SAP is to be accessible externally (e.g. published in a journal or on a website) but is optional if the SAP is an internal document only.

> **Example**
>
> "To be brief, chronic fatigue syndrome is characterised by chronic disabling fatigue in the absence of an alternative diagnosis, present in 0.2 to 2.6% of the population. The National Institute for Health and Clinical Excellence (NICE, UK) recommends two treatments: cognitive behaviour therapy (CBT) and graded exercise therapy (GET), but patient organisations recommend a third treatment: adaptive pacing therapy (APT). A definitive randomised trial was therefore needed to compare all three treatments with specialist medical care (SSMC) and to compare the established treatments (CBT, GET) against the new treatment (APT)."[4]

8. Objectives

Item 8: Description of specific objectives or hypotheses

Explanation

The trial objectives reflect the scientific questions to be answered by the trial, defining its rationale and scope. This information may be provided in sufficient detail within the protocol, in which case a reference would be sufficient. If the protocol contains insufficient detail as protocols usually target clinical rather than statistical readers, then additional detail may be required within the SAP. The trial hypotheses should be stated as these provide information on the framework (e.g. superiority, non inferiority) and regions of statistical testing (one or two-sided tests).

Example

Research hypothesis

The null hypothesis is that there is no difference in time to first blood stream infection between the standard and impregnated (antibiotic and heparin combined) groups. The alternative hypothesis is that there is a difference between the two groups.

Study objectives

The primary objective of this trial is to determine the effectiveness of heparin bonded or antibiotic impregnated CVCs (combined) compared with standard CVCs for preventing hospital acquired blood stream infection

Secondary objectives are:

a. To determine the cost effectiveness of heparin bonded or antibiotic impregnated CVCs compared with standard CVCs, based on the primary outcome and costs of acute care from the perspective of the NHS.

b. To determine the effectiveness of type of CVC in 3-way comparisons of heparin bonded versus antibiotic impregnated versus standard CVCs for preventing hospital acquired blood stream infection, based on culture, quantitative bacterial DNA, and clinical measures of infection."[7]

Section 3: Trial Methods

9. Trial design – description of trial design

Item 9: Brief description of trial design including type of trial (e.g. parallel group, multiarm, crossover, factorial), allocation ratio and brief description of interventions

Explanation

Specify the type of trial design. This can influence many aspects such as methods used, risk of bias, trial conduct, costs, results and interpretation. For example, factorial or adaptive designs can involve more complex methods, analyses, and interpretations than parallel group superiority trials. Although most trials use equal randomisation (i.e. 1:1 for two groups), it is still important to provide the allocation ratio. For drug trials, specifying the phase of the trial (I-IV) may also be relevant.

> **Example**
> "The trial is a two centre, randomised, parallel-group, placebo-controlled trial. Treatment allocation is a 1:1 ratio. Patients are randomised to either gabapentin or matched placebo control."[6]

10. Randomisation

Item 10: Randomisation details e.g. whether any dynamic allocation (e.g. minimisation) or stratification occurred (including stratifying factors used or the location of that information if not held within the SAP)

Explanation

Details regarding the randomisation process should be provided within the protocol. Additional detail such as the method of randomisation, e.g. minimisation or stratification, specific information relating to block sizes or specific factor levels used within minimisation or stratification should be stored in a restricted access area. Reference to where this information is stored should be provided in the SAP. This is to protect against predictability of the randomisation sequence by those providing clinical input to the SAP. This allows the statistician executing the SAP to identify stratification factors for use according to ICH E9.

11. Sample size

Item 11: Full details of the sample size calculation or alternatively reference to sample size calculation in protocol (instead of replication in SAP)

Explanation

The sample size calculation may be included in full in the SAP or a reference to the sample size calculation in the protocol or other document may be provided. The sample size calculation is an important piece of information for every trial as it determines how many patients are required in the primary analysis to ensure the trial is adequately powered to detect a clinically important difference. The size of that minimum clinically important difference may be used to interpret results. Justification of the sample size should be given including, if appropriate, the expected rate of attrition. All relevant information on which the calculation is based e.g., effect size, power, significance level etc., should be provided with any references to support parameter specifications together with details of any software used. Sufficient detail must be provided to enable another statistician to reproduce the calculation.

12. Framework

Item 12: Superiority, equivalence or non-inferiority trial hypothesis testing framework, and which comparisons will be presented on this basis

Explanation

Specifying the framework of a trial refers to its overall objective to test the superiority, equivalence or non-inferiority of one intervention from another. However, if for example the main objective is to determine equivalence of the primary outcome, secondary outcomes may be intended to demonstrate superiority. The SAP should clearly specify the framework for each outcome or provide a global statement.

Example

The SLEEPS trial protocol states the secondary objective is "to determine whether clonidine reduces side-effects and improves clinical outcomes due to its effects on reduction of sympathetic outflow, improved organ perfusion and protection in ischaemic reperfusion injury". Therefore, the secondary outcomes are testing for superiority rather than equivalence like for the primary outcome. [9]

13. Statistical Interim analyses and stopping guidance

Item 13a: Information on Interim analyses specifying what interim analyses will be carried out and listing of time points

Explanation

Information needed to conduct interim analyses should be detailed including statistical methods to be used, who will perform the analyses, what interim analyses will be carried out and when they will be performed e.g. timing and frequency. If interim analyses are not planned then this should be stated for clarity. If details of interim analyses are included in the protocol, or another document e.g. DMC Charter, then depending on the level of detail given the appropriate document could be referenced to avoid duplication. If separate SAPs have been written for interim analyses then these should be referenced.

> **Example**
> "One formal statistical interim analysis is planned on the primary endpoint for the Ir vs IrCs comparison. This interim analysis was planned to take place when the study was at least 18 months into recruitment and at least half the number of patients required for the final analysis (as per the sample size calculation) were recruited (i.e. 375 patients)."[8]

Item 13b: Any planned adjustment of the significance level due to interim analysis

Explanation

If analyses are to be performed on the accruing data at multiple time points then methods must be used to control the type 1 error in order to avoid increasing the risk of a false positive result. Various statistical methods have been developed to control this inflated risk such as Heybittle-Peto or O'Brien-Fleming techniques and the chosen approach should be clearly specified, justified and referenced. The DMC Charter could also be referenced, if applicable.

Item 13c: Details of guidelines for stopping a trial early

Explanation

Details should be provided on the guidelines to be used for stopping a trial early. It should be clear whether a statistical method will be considered within the early stopping guideline.

14. Timing of final analysis

Item 14: Timing of final analysis e.g. all outcomes analysed collectively or timing stratified by planned length of follow-up

Explanation

Information on the timing of final analyses should be included, if relevant. Information on timing of final analysis should explain whether all outcomes are analysed collectively or whether timing is stratified by length of follow-up required. Details should be provided on whether there are short-term and long-term outcomes and how they will be reported i.e. will all outcomes be analysed collectively or will the short-term outcomes be published earlier and the long-term outcomes reported at a later date.

15. Timing of outcome assessments

Item 15: Time points at which the outcomes are measured

Explanation

The time points at which outcomes are measured is helpful information that can be found in the protocol often in table format. The SAP should either refer to the relevant section of the protocol for details or include this information. If outcomes are required to be measured within a particular time window in relation to each planned visit in order to contribute to the analysis then this should also be specified.

Example

"The schedule of study procedures is given in the Table 8-1. The expected visit dates and visit windows are defined in Table 11.3-1. The start time for each calculation is the participants date of birth corrected for gestational age. Then additionally, 26 weeks, 52 weeks, 156 weeks and 260 weeks are added to determine the expected date for 6 months, 12 months, 3 years and 5 years follow up visits."

Section 4: Statistical Principles

16. Confidence intervals and p-values

Item 16: Level of statistical significance

Explanation

Where cut-off values are to be used to declare statistical significance then it is important for authors to document the significance level to be used including whether tests will be one- or two-sided. The significance level used for the primary outcome should be consistent with that used in the sample size calculation but secondary outcomes may use different levels.

> **Example**
> "All applicable statistical tests will be 2-sided and will be performed using a 5% significance level." [6]

Item 17: Description of any planned adjustment for multiplicity, and if so, including how the type 1 error is to be controlled

Explanation

Authors should pre-define what methods will be used to conduct any adjustment for multiplicity as different methods can lead to different conclusions[12]. The rationale for adjustment and method(s) chosen should be justified. If no adjustment for multiplicity is planned then an explicit statement should be included. Justification for the absence of multiplicity adjustments for secondary outcomes is probably unnecessary unless a claim is to be made on them, however for multiple testing on the primary outcome (e.g. different doses) justification should be given. If gatekeeping methods are to be used then authors should state the order of testing.

Item 18: Confidence intervals (CI) to be reported

Explanation

The CI is essential to the interpretation of statistical analyses reported for any of the primary or secondary outcomes. The level of CI to be reported should be decided at the design stage to avoid bias being introduced by modification based on trial data. The confidence levels used may be consistent across outcomes or vary by primary, secondary, exploratory and safety outcomes and this should be clearly specified.

19. Adherence and Protocol Deviations

Item 19a: Definition of adherence to the intervention and how this is assessed including extent of exposure

Explanation

Authors should pre-specify their definition of adherence to the intervention. Non-adherence to the intervention can include not completing the intervention, (e.g. not consuming all prescribed drugs or consuming a lower dose than is prescribed). This may be reported to aid generalizability of results or may be linked to an analysis population specification.

Example

"Compliance is assessed based on the percent of subjects who have taken the scheduled number of pills. It is defined as:

% compliance = (number of pills taken / number of pills supposed to have been taken)*100%.

The number of pills supposed to have been taken will be calculated as the duration of treatment (end of study medication – start of study medication + 1) multiplied by 4. In this study 2 pills are taken in the morning and 2 pills in the evening.

Item 19b: Description of how adherence to the intervention will be presented

Explanation

Along with defining adherence to the intervention it is also crucial to describe how adherence to the intervention will be presented. This process avoids any bias being caused by adherence being defined after unblinding of data.

> **Example**
>
> "The number and % of participants taking more than 75% of the prescribed treatment will be presented in a table for i) randomisation to visit 3 and ii) visit 3 to visit 4. Results will be provided by treatment group".[summary of example 1 in 16a]
>
> *Or*
>
> "Descriptive statistics on the percent compliance (N, mean, SD, median, minimum, maximum) will be summarized by randomisation group."

Item 19c: Definition of protocol deviation for the trial

Explanation

A protocol deviation is defined as a failure to adhere to the protocol such as the wrong intervention being administered, incorrect data being collected and documented, errors in applying inclusion/exclusion criteria or missed follow-up visits. A protocol deviation should be defined as major or minor. A deviation may be considered a serious breach if it affects efficacy, the safety, physical or mental integrity of the participants in the trial, or the scientific value of the trial. Protocol deviations should be defined prior to unblinding of data to avoid any bias being caused and due consideration given to inclusion of participants within analysis populations[14]. Protocol deviations may be defined in another document and referenced within the SAP.

> **Example**
>
> "The following are pre-defined major protocol violations with a direct bearing on the primary outcome:
> 1) Taking of rescue medication (loperamide) during the primary outcome assessment period i.e. weeks 11-12 of the treatment period.
> 2) Taking of antibiotics during the primary outcome assessment period i.e. weeks 11-12 of the treatment period."[15]

Item 19d: Description of which protocol deviations will be summarised (may include details of whether deviation is major or minor and impact on analysis populations and approach to summarising protocol deviations e.g. number and type of protocol deviation, per group)

Explanation

A description should be provided on how protocol deviations will be summarised. Providing details of whether the deviation is major or minor is helpful if sensitivity analyses are to be conducted by removing patients with major deviations to assess impact on overall conclusions or to align with

analysis populations. The approach to summarising the protocol deviation should also be made clear e.g. number and type of protocol deviations by intervention group or listing of all deviations.

> **Example**
>
> "Protocol deviations are classified prior to unblinding of treatment. The number (and percentage) of patients with major and minor protocol deviations will be summarised by treatment group with details of type of deviation provided. The patients that are included in the ITT analysis data set will be used as the denominator to calculate the percentages. No formal statistical testing will be undertaken."[9]

20. Analysis populations

Item 20: Definition of Analysis populations e.g. intention-to-treat (ITT), per-protocol, complete case, safety.

Explanation

The analysis populations should be specified in advance. This includes how the analysis populations will be defined and which outcomes will be analysed according to each analysis population. It is important to clearly define populations, even if terms are considered standard. For example, there is no consistent definition of ITT and the phrase has different meanings for different authors. [16]

> **Example**
>
> The intention-to-treat population will include all randomised patients, regardless of their eligibility, according to the treatment they were randomised to receive.
>
> *Or*
>
> A per-protocol population will be considered if >5% of the total number of patients in this comparison are major protocol violators. The per-protocol analysis set consists of subjects who were randomly assigned to treatment, have both a baseline and at least 1 post-baseline measurement on the primary efficacy variable, have a minimum exposure of 36 days to the double-blind treatment regimen, and have no major protocol violations such as violations of entry criteria, errors in treatment assignment and use of excluded medications.[8]
>
> *Or*
>
> The safety population will consist of all randomised patients in this comparison who have received at least one dose of study treatment. Patients will be analysed according to the treatment they actually received.

Section 5: Trial Population

21. Screening Data

Item 21: Reporting of screening data (if collected) to describe representativeness of trial sample

Explanation

If a trial collects screening data then it is important that the data are appropriately presented to describe the representativeness of the trial sample. This information is not only important for the trial but also important for future trials in the area. The process for screening patients e.g. how patients will be screened and what data will be collected, should be fully described within the trial protocol. According to the CONSORT guidelines[17] as a minimum the number of patients who are assessed for eligibility should be provided with this information presented in a flow diagram, however, more detailed tabulations may be provided. The SAP should describe how this data will be summarised and presented.

> **Example**
> "The following summaries will be presented for all screened patients:
>
> Enrolment: the number of days recruiting, the number of patients screened, the number of patients recruited, the number of patients recruited per day, the number of screened patients not recruited, and the reason for non-recruitment. This summary will be provided overall and by study centre." [18]
>
> *Or*
>
> "The total number of eligible babies was not collected during the conduct of this study as it was considered heavy on resources and would not be sufficiently reliable." [19]

22. Eligibility

Item 22: Summary of eligibility criteria

Explanation

The trial inclusion and exclusion criteria should be specified in the protocol. Details of how eligibility data will be summarised should be provided. Some CONSORT diagrams provide details of the number of patients screened followed by a breakdown of how many patients were eligible and how many were excluded due to violating each inclusion/exclusion criteria.

> **Example**
>
> "The number of ineligible patients randomised, if any, will be reported, with reasons for ineligibility."[19]

23. Recruitment

Item 23: Information to be included in the CONSORT flow diagram

Explanation

Information included within a CONSORT flow diagram displays the progress of all participants through the trial. The CONSORT guidelines say that "you must complete a flow diagram in order to be compliant with the CONSORT 2010 standard."[17] They provide a CONSORT flow diagram template that can be used and adapted to create a trial specific flow diagram. All necessary information that is displayed in a CONSORT flow diagram should be listed in the SAP so it is clear where the patient throughput will begin to be summarised and how, specific follow-up time points that will be presented along with information on withdrawals and loss to follow up. Alternatively, a study specific CONSORT flow diagram template can be included in the SAP highlighting the information that will be collected.

> **Example**
>
> "The "CONSORT" diagram comprising the number of people screened, eligible, consented, randomised, receiving their allocated treatment, withdrawing/lost to follow-up" [20]
>
> *Or*
>
> "A CONSORT flow diagram (appendix A) will be used to summarise the number of patients who were:
>
> - assessed for eligibility at screening
>   - eligible at screening
>   - ineligible at screening*
> - eligible and randomised
> - eligible but not randomised*
> - received the randomised allocation
> - did not receive the randomised allocation*
> - lost to follow-up*
> - discontinued the intervention*
> - randomised and included in the primary analysis
> - randomised and excluded from the primary analysis*
>
> *reasons will be provided."[9]

24. Withdrawal/Follow-up – level of withdrawal

Item 24a: Level of withdrawal e.g. from intervention and/or from follow up

Explanation

In this section, all the possible levels of withdrawal should be listed, which may differ from trial to trial. Participants may withdraw from the intervention but continue with follow-up; withdraw from follow-up but allow data collected to date to be used; withdraw from follow-up and withdraw consent for data collected to date to be used; or be lost to contact/follow-up. Some clarification within the SAP about how each level of withdrawal will be categorised and presented is important.

> **Example**
> "The level of consent withdrawal will be tabulated (classified as "consent to continue follow-up and data collection" "consent to continue data collection only", "complete – no further follow-up or data collection")." [21]

Item 24b: Timing of withdrawal/lost to follow up data

Explanation

Timing of withdrawals and lost to follow up is important information. This information allows you to see if there are any patterns in lost to follow up or withdrawals between the different time points and intervention groups. Timing of withdrawal from follow-up or lost to follow up data can be presented in a Kaplan-Meier graph, a table or incorporated into a CONSORT flow diagram. For each follow-up time point information on the number of withdrawals and reasons for withdrawal, number included in the analysis and the number died (if applicable) should be provided.

> **Example**
> "This will be presented in CONSORT diagram format rather than as a table, with numbers and reasons for withdrawal and/or exclusion from analysis given at each stage (delivery, 6 months, 1 year, 2 years."[22]

Item 24c: Reasons and details of how withdrawal/lost to follow up data will be presented

Explanation

Patients can withdraw and be lost to follow up for many different reasons e.g. moved home, unable to participate any longer, withdrawn by clinician reasons etc. It is useful for the trial team to attempt to ascertain reasons for all withdrawals and loss to follow up. According to ICH E6 "Although a subject

is not obliged to give his/her reason(s) for withdrawing prematurely from a trial, the investigator should make a reasonable effort to ascertain the reason(s), while fully respecting the subject's rights".[23] Details of how this data will be presented should be included in the SAP. This information may be presented by intervention arm within a CONSORT flow diagram or in a table.

> **Example**
> "The numbers (with reasons) of losses to follow-up (drop-outs and withdrawals) over the course of the trial will be summarised by treatment arm."[15]

25. Baseline patient characteristics

Item 25a: List of baseline characteristics to be summarised

Explanation

Presentation of baseline characteristics by trial arm is crucial for every trial as it allows the reader to see whether the characteristics are balanced across intervention groups. Details of which baseline characteristics will be summarised in the final report should be specified. Any factors on which the randomisation has been stratified or minimised should be included so that balance across the randomised groups can be demonstrated.

> **Example**
> "Patients will be described with respect to age, gender, time since diagnosis, cancer type, performance status, the number of previous chemotherapies and presence of brain metastases at baseline, both overall and separately for the two randomised groups." [24]

Item 25b: Details of how baseline characteristics will be descriptively summarised

Explanation

It is important to describe how baseline characteristics will be summarised and presented in the final analysis report. Formal statistical comparisons of baseline data by randomised groups are not normally advocated[25] [26] but if such comparisons are planned these should be justified. It is recommended that prognostic baseline characteristics are presented for the analysis population included in the primary analysis of the primary outcome as well as for all randomised participants in order to assess whether attrition has introduced selection bias and/or upset the balance achieved at randomisation.

Example

"Categorical data will be summarised by numbers and percentages. Continuous data will be summarised by mean, SD and range if data are normal and median, IQR and range if data are skewed. Minimum and maximum values will also be presented for continuous data. Tests of statistical significance will not be undertaken for baseline characteristics; rather the clinical importance of any imbalance will be noted."[9]

Section 6: Analysis

26. Outcome definitions

List and describe each primary and secondary outcome including details of:

Item 26a: Specification of outcomes and timings.

Item 26b: Specific measurement and units (e.g. glucose control hbA1c (mmol/mol or %))

Item 26c: Any calculation or transformation used to derive the outcome (e.g. change from baseline, quality of life (QoL) score, time to event, logarithm etc).

Explanation

The SAP should define each outcome explicitly clearly identifying primary and secondary variables. If multiple primary variables are used then the considerations outlined in ICH E9[27] (2.2.5 Multiple Primary Variables) should be explored with direction provided on interpretation. Consistency should be ensured with the chosen approach to adjust for multiplicity in item 17. If an outcome is recorded at multiple timepoints which of these timepoints are required for the specific outcome. Detailed explanations should be provided, for example for survival outcomes making it clear what the length of survival is (e.g. calculated from the time of randomisation or time of administration of intervention) and censoring information.

The SAP should identify the specific measurement variable and its units if applicable (e.g. overall survival in days) and provide descriptions and details of any data manipulations or derivations to be performed by the statistician. Detail needs to be provided on what data manipulations or derivations will be performed and how they will be carried out. This may be relevant when data collection units may vary, for example HbA1c measurement in % or mmol or quality of life scores. If the calculation of a score is more complex, but a validated algorithm is available, then providing a reference and a link to the algorithm is sufficient. Scoring, including handling of missing data, should follow that proposed by the instrument developers, unless there is good reason to use an alternative technique, which should be described and justified. Sufficient detail needs to be provided in order for the reader to understand how the scores or results are to be calculated for each outcome.

27. Analysis methods

List and describe each primary and secondary outcome including details of:

Item 27a: - What analysis method will be used, and how the treatment effects will be presented

Explanation

Conclusions can be affected substantially by the analysis method(s) used, therefore it is extremely important to pre-specify the analysis method(s) so there is no possibility of the method being chosen because it gives the most positive results. If transformations are to be applied, then these should be specified along with the rationale for the transformation and the resulting interpretation.

For each outcome, the SAP should specify what analysis method(s) will be used for statistical comparisons and which trial participants will be included in this analysis if applicable. The SAP should also define what summary measures will be reported such as any descriptive statistics to be displayed, what the unit of each effect estimate will be and whether confidence intervals and p-values will be reported. If more than one method is to be used to analyse the primary outcome, e.g. adjusted and unadjusted for covariates, then the primary analysis method should be identified.

> **Example**
> "The number and percentage of deaths by 90 days after randomisation will be reported for each treatment group. The primary-effect estimate will be the relative risk of 90-day mortality, reported with a 95% CI. The absolute risk reduction and 95% CI will also be reported. Deaths by 90 days after randomisation will be compared between the treatment groups, unadjusted, using the Fisher exact test."[29]

Item 27b: List and describe each primary and secondary outcome including details of: any adjustment for covariates

Explanation

For each analysis, the SAP should specify whether adjustment will be used, and if so, the covariates to be used (including the categories if applicable), and how these will be included in the model (e.g. as fixed effects, random effects etc.). For the primary outcome, it must be clear whether the adjusted or unadjusted analysis is the primary analysis as failing to pre-specify can lead to bias.

> **Example**
> "…adjusted for baseline variables, will also be conducted, using multilevel logistic regression. Baseline variables adjusted for in the multilevel logistic regression model will be the components of the MEDS score (age, metastatic cancer, nursing home residence, altered mental status, septic shock, respiratory difficulty, low platelet count and low neutrophil count) and a site-level random effect."[29]

Item 27c: List and describe each primary and secondary outcome including details of: methods used for assumptions to be checked for statistical methods

Explanation

All statistical tests require that a number of assumptions hold in order for the test to be valid and the conclusions drawn from the analysis to be correct. However, a two-stage analysis may lead to bias.[30-33] If checks on the underlying assumprions are to be performed then it is important for these to be pre-specified.

> **Example**
> "The PH assumption will be checked by adding time-dependent covariates and graphing scaled Schoenfeld residuals against time (Therneau & Grambsch, 2000). If PH is found not to fit the data adequately, an AFT alternative will be fitted and the adequacy of its fit assessed using Q-Q plots (Bradburn et al. 2003)." [20]

Item 27d: List and describe each primary and secondary outcome including details of: alternative methods to be used if distributional assumptions (e.g. normality, proportional hazards etc) do not hold

Explanation

Many distributional assumptions can be checked during blind data review[27] and the SAP updated accordingly. However, if assumptions can only be checked once the treatment allocations are known then the SAP should pre-specify the alternative methods to be used if the underlying assumptions do not hold. The approach taken should be considered carefully as bias may be introduced either by choosing the method of analysis based on the results of tests of assumptions[31 33] or from performing hypothesis tests in which the underlying assumptions are not upheld. Three possible approaches may be considered: i) pre-specify alternative analyses and how the statistician will choose between them in the SAP so that the process is transparent; ii) select a method of analysis that is robust to assumptions, e.g. survival analysis will be carried out using the restricted mean survival time (RMST) method as this does not assume proportional hazards (PH); or iii) state the method of analysis to be used in the SAP and specify that a sensitivity analysis will be performed using an alternative set of assumptions and the results compared.

> **Example**
> "If PH is found not to fit the data adequately, an AFT alternative will be fitted and the adequacy of its fit assessed using Q-Q plots (Bradburn et al. 2003). If this too does not fit, a residual life analysis (Royston & Parmar 2011) will be used as the basis for summarising the treatment effect"[20]

Item 27e: List and describe each primary and secondary outcome including details of: any planned sensitivity analyses for each outcome

Explanation

For each outcome, where applicable, the SAP should specify whether any sensitivity analyses will be conducted, and what these analyses will be with the same level of detail as for the primary and secondary analyses. The SAP should also include details of the analysis population to be used for each sensitivity analysis. The SAP should state whether there is a minimum percentage of missing data required to trigger the need for sensitivity analyses.

> **Example**
>
> "A sensitivity analysis will be performed to include the patients that were not included in the primary analysis because they did not fully complete the loading dose and two hour maintenance period. They will be assumed to be not adequately sedated i.e. AS=0. The per-protocol analysis and sensitivity analyses will test the robustness of the primary complete-case analysis."[9]

Item 27f: List and describe each primary and secondary outcome including details of: any planned subgroup analyses for each outcome

Explanation

All pre-planned subgroup analyses should clearly specify the baseline characteristics to be considered, the cut-offs for the subgroup categories, the statistical method that will be used and how the results will be presented (e.g. in a forest plot). However while a large number of subgroup analyses should be avoided due to issues with multiplicity this may be appropriate for example when the aim is to demonstrate consistency across subgroups.

> **Example**
>
> "The following pre-specified subgroup analyses will be performed on the primary outcomes stratified by:
> • whether randomised in the 1st or 2nd 24 hours after birth
> • gestational age at birth as per minimisation: 23w, 24w, 25w, 26/27w, 28/29/30w.
> • male versus female
> • colonised versus not colonised at 2 weeks
> • gestational age <28+0 versus ≥28+0
>
> Results will be presented on forest plots with the interaction results alongside." [19]

28. Missing data

Item 28: Missing data- reporting and assumptions/statistical methods to handle missing data (e.g. multiple imputation)

Explanation

The majority of trials will have some missing data[34], which can introduce bias depending on the type of "missingness" (e.g. missing completely at random, missing at random, missing not at random[35]). Therefore, it is important that the SAP states how missing data will be handled and reported including details of any statistical methods and their assumptions, to be used to handle missing data. It should explain if there are any plans to impute missing outcome data, including a list of variables that will be

used in the imputation process if multiple imputation (MI) is to be used. Using different statistical methods to handle missing data can lead to differing conclusions so it is crucial to pre-specify exactly what methods will be used under what circumstances, and which will be considered the primary analysis. It is highly recommended that sensitivity analyses are conducted to assess the robustness of trial results when using different methods to handle missing data[36] and these should be clearly described in the SAP.

Example

"Multiple imputation (MI) will be used to account for participants who have an observed outcome at 6 months, but are missing the outcome at 12 months, as well as participants who completed some, but not all, of the questions on the CPG disability score at 12 months. 20 imputations will be performed, and results will be combined using Rubin's Rules. Only participants who will be included in the analysis will be included in the imputation model. Imputation will be performed separately within each treatment arm. The imputation model will include the three questions which form the CPG disability score at baseline, 6 months, and 12 months, as well as site of recruitment, age, gender, the HADS depression score at baseline, and employment status (employed or in full time education vs not employed or in full time education) (14 variables in total). In the intervention arm, multilevel imputation will be performed, with 'course' included in the imputation model as a random effect. Missing data in any of the covariates to be adjusted for in the analysis (site of recruitment, age, gender, HADS depression score, CPG disability and baseline) will be accounted for using the same multiple imputation model as above. We will perform three sensitivity analyses for the primary outcome to assess the robustness of the results to other methods of account for missing data. The first sensitivity analysis involves specifying a different imputation model than that used in the primary analysis, and the last two sensitivity analyses involve re-analyse the primary outcome using two approaches which are not based on MI." [37]

---

29. Additional Analyses

Item 29: Details of any additional statistical analyses required e.g. complier-average causal effect (CACE) analysis38

Explanation

Any additional analyses to be conducted should be specified with reasons these are required, a description of the additional analysis and how it will be conducted. This may include pre-specified exploratory analyses that are hypothesis generating or confirmatory of issues identified in other trials.

> **Example**
>
> "The delivery of a complex intervention may improve with time as those delivering the intervention gain experience and familiarity. Typically, such improvements will be more rapid at first and then tail off over time to reach a steady state; termed a "learning curve". Modelling the learning curve enables estimation of the treatment effect for an experienced team. A site-level learning curve for patients randomly allocated to early, goal-directed, protocolised resuscitation (EGDPR) will be modelled by repeating the multilevel logistic regression on the primary outcome and including a power curve (aX-b) for the sequential observation number (X) for each EGDPR patient within each site." [29]
>
> Or
>
> "Complier Average Causal Effect (CACE) analysis: Instrumental variable regression will be used to investigate the effect of compliance to treatment dose, assuming linear dose-response relationship. The estimate of increased or decreased treatment effect with every 1% increase of compliance will be presented."[15]

30. Harms

Item 30: Sufficient detail provided on summarising harms e.g. information on severity, expectedness and causality; details of how AE's are coded or categorised; how adverse events (AE's) data will be analysed, i.e. grade 3/4 only, incidence case analysis, intervention emergent analysis

Explanation

Consideration of safety data is key for every clinical trial. It is important that safety data is reviewed and details are provided in the SAP on how safety data will be summarised in the final analysis report including the analysis population to be used. Information may be provided on the severity, causality and expectedness of the adverse event, information on how the adverse events will be coded or categorised and by whom. The method of summarising the adverse event data should be described ensuring it is clear whether the descriptive summary will use number of events or number of patients and any analyses to be conducted (e.g. will the adverse events be compared descriptively or will formal statistical testing be undertaken.)

> **Example**
> "The number of treatment related serious adverse events (SAE), including treatment related deaths, are reported divided by their relationship as 'definitely', 'probably' and 'possibly' related to treatment. The proportions of patients with grade 3/4 toxicity or SAE will be compared descriptively across treatments and differences assessed for clinical significance." [39]
>
>
> "The number (and percentage) of patients experiencing each AE/SAE will be presented for each treatment arm categorised by severity. For each patient, only the maximum severity experienced of each type of AE will be displayed. The number (and percentage) of occurrences of each AE/SAE will also be presented for each treatment arm. No formal statistical testing will be undertaken." [7]

31. Statistical Software

Item 31: Details of statistical packages to be used to carry out analyses (optional)

Explanation

Details of the statistical packages to be used to conduct the statistical analyses may be provided in the SAP. While version numbers of software may change during the lifetime of the trial and so should not be specified in the SAP they should be included within final reports.

> **Example**
> "The analysis will be carried out using Stata version 12. Other packages such as R, SAS, or REALCOM may be used if necessary." [24]

32. References

Item 32a: References to be provided for non-standard statistical methods

Explanation

References should be provided in a SAP for any non-standard statistical methods that will be used. If there is any doubt on whether a method is non-standard then it is better to include a reference.

Item 32b: Reference to Data Management Plan

Explanation

Reference should be made to the Data Management Plan (DMP) with the version number that was used when writing the SAP. This is important as both documents should be linked with information in

the DMP that is also important for the final analysis report. If there is no DMP, then the location of this information (e.g. data handling and cleaning) should be provided.

Item 32c: Reference to the Trial Master File and Statistical Master File

Explanation

The Statistical Master File is part of the Trial Master File but is often held separately with restricted access . The Statistical Master File may hold details of the randomisation process or specific protocol deviations that the statistician needs to refer to when executing the statistical analysis plan. If a Statistical Master File is held separately to the Trial Master File then both should be referenced.

Item 32d: Reference to other Standard Operating Procedures or documents

Explanation

Reference should be made to any other Standard Operating Procedures (SOPs) or documents that are adhered to and followed when writing the SAP.

Statement References

1. Sim J, Gray R, Nevatte T, et al. Statistical analysis plan for the Stroke Oxygen Study (SO2S): a multi-center randomized controlled trial to assess whether routine oxygen supplementation in the first 72 hours after a stroke improves long-term outcome. Trials 2014;**15**(1):229.
2. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. New England Journal of Medicine 2004;**351**(12):1250-51.
3. Lawlor DA, Peters TJ, Howe LD, et al. The Active for Life Year 5 (AFLY5) school-based cluster randomised controlled trial protocol detailed statistical analysis plan. Trials 2013;**14**(234):21.
4. Walwyn R, Potts L, McCrone P, et al. A randomised trial of adaptive pacing therapy, cognitive behaviour therapy, graded exercise, and specialist medical care for chronic fatigue syndrome (PACE): statistical analysis plan. Trials 2013;**14**(1):386.
5. Shephard N BM, Biggs K et al. The HubBLe Trial : Statistical Analysis Plan. [Unpublished SAP].
6. Lewis S HA. Statistical Analysis Plan: A pilot randomised controlled trial of the efficacy and mechanism of action of gabapentin for the management of chronic pelvic pain in women. (GAPP) [Unpublished SAP].
7. Harron K, Mok Q, Dwan K, Ridyard CH, Moitt T, Millar M, Ramnarayan P, Tibby SM, Hughes DA, Gamble C, Gilbert RE. The CATCH trial -CATheter Infections in CHildren: a randomised controlled trial and economic evaluation comparing impregnated and standard central venous catheters in children. HTA monograph in press.
8. Brown S TH, Olivier C, Napp V, Seymour M. Panitumumab, Irinotecan and Ciclosporin in colorectal cancer therapy – PICCOLO: Ir vs. IrCs comparison. Statistical Analysis Plan. [Unpublished SAP].
9. Wolf A, McKay A, Spowart C, et al. Prospective multicentre randomised, double-blind, equivalence study comparing clonidine and midazolam as intravenous sedative agents in critically ill children: the SLEEPS (Safety profiLe, Efficacy and Equivalence in Paediatric intensive care Sedation) study. Health Technology Assessment 2014;**18 (71)**.
10. Young PJ, Weatherall M, Saxena MK, et al. Statistical analysis plan for the HEAT trial: a multicentre randomised placebo-controlled trial of intravenous paracetamol in intensive care unit patients with fever and infection. Critical Care and Resuscitation 2013;**15**(4):279-86.
11. Marson AG, Appleton R, Baker G, et al. A Randomised Controlled Trial Examining the Longer-term Outcomes of Standard Versus New Antiepileptic Drugs: The SANAD Trial. Health Technology Assessment 2007;**11**(37).
12. Committee for Proprietary Medicinal Products. Points to consider on multiplicity issues in clinical trials. London, September 19, 2002: CPMP/EWP/908/99.
13. Dore C CS. VIDEO: Vitamin D Evaluation in Osteoarthritis. Statistical Analysis Plan. [Unpublished SAP].
14. Medicines for Healthcare products Regulatory Agency. *Good Clinical Practice Guide*. United Kingdom: The Stationary Office, 2012.
15. Tan W MA, Spiller R. Efficacy and mode of action of Mesalazine in the treatment of Diarrhoea-predominant irritable bowel syndrome (IBS-D). Mesalazine for IBS-D (MIBS) Statistical Analysis Plan [Unpublished SAP].
16. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;**319**(7211):670-74.
17. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Medicine 2010;**8**(1):18.
18. Bradburn M CJ, Goodacre S, Campbell M. The 3Mg Trial : Randomised controlled trial of intravenous or nebulised magnesium sulphate or standard therapy for acute severe asthma. Statistical Analysis Plan. [Unpublished SAP].

19. Hardy P CK, Juszczak E, Heal P, King A, Nelis C, Millar M, Wilks M, . PiPS: Trial of probiotic administered early to prevent infection and necrotising enterocolitis. Statistical Analysis Plan. [Unpublished SAP].

20. Bradburn M MC, Taylor L, Young C. A randomised controlled trial in patients with Respiratory Muscle Weakness due to Motor Neurone Disease of the NeuRx/4 Diaphragm Pacing Trial. (DiPALS) Statistical analysis plan. . [Unpublished SAP].

21. Nan H PJ. Phase I study of Continuous Hyperfractionated Accelerated RadioTherapy – Escalated Dose Study (CHART-ED): CHART-ED Statistical Analysis Plan. [Unpublished SAP].

22. Cooper S, Lewis S, Thornton JG, et al. The SNAP trial: a randomised placebo-controlled trial of nicotine replacement therapy in pregnancy–clinical effectiveness and safety until 2 years after delivery, with economic evaluation. Health Technology Assessment 2014;**18**(54).

23. International Conference on Harmonisation. Topic E6 (R1) Guideline for Good Clinical Practice (CPMP/ICH/135/95) European Medicines Agency, 1996.

24. Dutton S FD, Sydes M. Phase III randomised, double-blind, placebo-controlled trial of gefitinib (Iressa®) versus placebo in oesophageal cancer progressing after chemotherapy: COG Statistical Analysis Plan. [Unpublished SAP].

25. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Journal of clinical epidemiology 2010;**63**(8):e1-e37.

26. Senn S. Testing for baseline balance in clinical trials. Statistics in medicine 1994;**13**(17):1715-26.

27. International Conference on Harmonisation. Topic E9 Statistical Principles for Clinical Trials (CPMP/ICH/363/96). European Medicines Agency, 1998.

28. Appleton R, Jones AP, Gamble C, et al. The use of MElatonin in children with neurodevelopmental disorders and impaired sleep: a randomised, double-blind, placebo-controlled, parallel study (MENDS). Health Technology Assessment 2012;**16**(40).

29. Power G, Harrison DA, Mouncey PR, et al. The Protocolised Management in Sepsis (ProMISe) trial statistical analysis plan. Critical Care and Resuscitation 2013;**15**(4):311-17.

30. Kahan BC. Bias in randomised factorial trials. Statistics in medicine 2013;**32**(26):4540-49.

31. Shuster JJ. Diagnostics for assumptions in moderate to large simple clinical trials: do they really help? Statistics in Medicine 2005;**24**(16):2431-38.

32. Raab GM, Day S, Sales J. How to select covariates to include in the analysis of a clinical trial. Controlled clinical trials 2000;**21**(4):330-42.

33. Campbell H, Dean C. The consequences of proportional hazards based model selection. Statistics in medicine 2014;**33**(6):1042-56.33

34. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clinical Trials 2004;**1**(4):368-76.

35. Rubin DB. Inference and missing data. Biometrika 1976;**63**(3):581-92.

36. National Research Council Panel on Handling Missing Data in Clinical Trials. *The Prevention and Treatment of Missing Data in Clinical Trials* Washington, DC: The National Academies Press, 2010.

37. Kahan B CD, Homer K, Diaz-Ordaz K, Joshi M, Bremner S, Eldridge S, Taylor S, Underwood M. Improving the self-management of chronic pain: COPERS Statistical Analysis Plan. [Unpublished SAP].

38. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods. Statistical Methods in Medical Research 2005;**14**(4):369-95.

39. Stocken D BL. EUROPEAN STUDY GROUP FOR PANCREATIC CANCER - TRIAL 3(v2) (ESPAC-3). Adjuvant chemotherapies in resectable pancreatic cancer. Statistical Analysis Plan for the ESPAC-3(v2) Trial. [Unpublished SAP].