Statistical Analysis Plan

We would like to acknowledge the Cambridge University Hospitals Clinical Trials Unit for the development of the template (version CCTU/TPLV2), which was modified by the Michigan Institute for Clinical & Health Research (MICHR).

1 How to Use This Template

Refer to institutional SOP on writing SAP if applicable for the key requirements of a statistical analysis plan (SAP). Fill in the details as dictated by the protocol and case report form for a study and consult with members of the trial to meet the needs of the final report. The key document for regulatory requirements is the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guidelines E9-Statistical Principles for Clinical Trials, which is referred to frequently throughout this template.

Example text is provided in *italics.* The purpose of example text is to provide a guide to the level of detail, but it is **not** intended to suggest any standard or default. Multiple examples exist, including some bad examples, to prompt the user to consider the issues at stake and make an informed choice appropriate to the study.

There is a choice of what pieces of data to explicitly list. In some contexts, it may not be necessary to list every AE observed, say, in the final study report, but simply retain the data and record where the data is stored.

Delete this section before circulating any draft or final version of a SAP.

This SAP template is meant to serve a wide array of uses including NIH and industry standards, as well as a variety of study types such as clinical trials (phases 1, 2, & 3), case-control, observational, etc. As a result, some sections may not be applicable.

As outcomes are being detailed, keep in mind the reporting requirements for clinicaltrials.gov.

TRIAL FULL TITLE	Does the use of nerve stimulator improve the outcome of ultrasound guided supraclavicular block (anesthesia) for upper extremity surgery?
SAP VERSION	Version 11
SAP VERSION DATE	September 27, 2022
TRIAL STATISTICIAN	William R. Grubb, DDS, MD
Protocol Version (SAP associated with)	Version 11
TRIAL PRINCIPAL INVESTIGATOR	William R. Grubb, DDS, MD
SAP AUTHOR(s)	Dennis Hall, MD, Scott Mellender. MD, Geza Kiss, MD

2 SAP Signatures

I give my approval for the attached SAP entitled <STUDY TITLE> dated <SAP VERSION DATE>

Statistician (Author)
Name:

Signature: 

Date: September 27, 2022

Statistician Reviewer (As applicable)
Name:

Signature: 

Date: September 27, 2022

Principal Investigator
Name:

Signature: 

Date: September 27, 2022

3 Table of Contents

4 Abbreviations and Definitions

Provide a list of the abbreviations and acronyms used in the Statistical Analysis Plan (SAP) with definitions. All terms will appear in alphabetical order.

This section should be completed on an ongoing basis during the preparation of the document and checked carefully after preparing the rest of the SAP to ensure that all the abbreviations are captured. MS Word has a Find tool that can search for wild cards; <[A-Z]{2,}> will pick up acronyms.

Although the abbreviations are listed, it is standard practice to spell out abbreviated terms and to indicate the abbreviation in parentheses at their first appearance in the text.

AE	Adverse Event
CRF	Case Report Form
IMP	Investigational Medical Product
SAP	Statistical Analysis Plan

5 Introduction

5.1 Preface

Include a very brief summary (approximately one paragraph) of the important background information from the protocol. This not intended to be used for the design of the study, but rather the significance.

5.2 Scope of the analyses

Include a brief statement of the purpose of the analyses. For example:
These analyses will assess the efficacy and safety of [IMP] in comparison with the [standard] and will be included in the clinical study report.

Provide an example of the Scope.

6 Study Objectives and Endpoints

6.1 Study Objectives

(ICH E3; 8.)
This section describes the overall purpose of the study and is taken directly from the protocol. However, additional elaboration may be helpful.

6.2 Endpoints

(ICH E9; 2.2.2)
List separately the primary, secondary, exploratory endpoints for the study as taken directly from the protocol.

7 Study Methods

7.1 General Study Design and Plan

(ICH E3;9)

Identify the study design, including the following

- Study configuration and experimental design: x-period cross-over, longitudinal, 2x2 factorial, observational, cohort. However, not every design can be abbreviated to a label of few words, and enough detail should always be given to eliminate any ambiguities.
- Type of Comparison: non-inferiority, superiority, equivalence. (ICH E3; 9.2, 9.7.1, 11.4.2.7. ICH E9; 3.3.2)
- Type of control(s): placebo, no treatment, active drug, different dose or administration, historical.
- Level and method of blinding (e.g. double-blind). However, not every method can be abbreviated to a label of few words, and enough detail should always be given to eliminate any ambiguities.
- Method of treatment assignment: Randomization with stratification, minimization.
- At what point in time subjects are randomized relative to treatments, events and study periods.
- Sequence and duration of all study periods: screening, baseline, active treatment, follow-up.

The last two points should routinely be represented by a study flow-chart that can be copied directly from the protocol.

7.2 Inclusion-Exclusion Criteria and General Study Population

(ICH E3;9.3. ICH E9;2.2.1)

This section is intended to describe particulars about all of the subjects in the study. It is distinct from the Analysis Population (section 8.2). This section is intended to describe the intended characteristics of ***all*** the subjects in the study, whereas the Analysis Population section is designed to identify the characteristics that are needed to defined sub-populations used for the analyses.

The SAP may include:

- a list of all inclusion and exclusion criteria directly copied from the protocol.
- or a simple description of the relevant diagnostic or disease related criteria (e.g. a history of chronic back pain for over 10 years.

7.3 Randomization and Blinding

(ICH E3; 9.4.3, 9.4.6. ICH E9; 2.3.1, 2.3.2)

Describe essential components of the randomization and blinding methodology and process in enough detail to enable its reproduction. Include any minimization, stratification or blocking procedures used to avoid or minimize bias. This section may be copied from the protocol but it may be necessary to include additional information details, particularly regarding block size. However, in a double-blind study it is not appropriate to include such information in the SAP but document it within the final study report, in which case document that these details will be provided in the final study report. Document any software packages used to perform the randomization and the method used for retrieving the treatment assignments.

7.4 Study Assessments

(ICH E3; 9.5.1. ICH E9; 2.2.2)

Describe the frequency and timing of all the relevant assessments. A table or flow chart may be appropriate for example

Visit	Baseline	Month 1	Month 3	Month 6	Month 12
Target day of visit		30	90	180	360
protocol assessment time windows (days)		± 3	± 7	± 14	± 14
History and examination	x	x	x	x	x
Weight	x	x	x	x	x
Vital signs	x			x	x
Hematology	x	x	x	x	x
Biochemistry	x	x	x	x	x
Urinary pregnancy test	x				
Tumor response	x	x	x	x	x
Blood samples for predictive markers	x		x	x	
Concomitant medication	x	x		x	x
Administer treatment		x			
QOL questionnaire	x		x	x	x (12 weeks only)
Adverse event monitoring	x	x	x	x	x

Define the time-windows to be used for converting dates into visit numbers for scheduled assessments (e.g. assessments collected from 27 to 33 days post-Randomization are identified as the Month 1 visit). Describe the decision rules that will be used to classify measurements obtained outside of scheduled protocol specified assessment time-windows. Clearly state how data collected outside the time-windows will be handled for analysis (i.e. how far before and after time-window will data be accepted for use in analysis). Describe the methods for handling multiple measurements

that occur within the same assessment time window. For example, the table below defines the analysis time windows that correspond to the targeted visit time frame.

Analysis Time Windows

Visit (target day)	Lower bound (days)	Upper bound (days)
Baseline (0)	N/A	N/A
Month 1 (30)	2	60
Month 3 (90)	61	135
Month 6 (180)	136	270
Month 12 (360)	271	374

This section will go beyond the description of variables provided in the protocol in that it will list and describe all important study variables from a statistical perspective. The description of each variable should include:

- Identification of any number ranges for numeric endpoints along with their corresponding text descriptors.
    - Items are measured on a 0-100 visual analog scale (VAS) for which 0=no pain and 100=worst pain imaginable
    - Items are measured on a 1-4 ordered categorical scale for which 1=no pain, 2= slight pain, 3=moderate pain, 4=extreme pain

If there are numerous variables it may be useful to create subsections corresponding to each variable which are grouped together as in the protocol (e.g. efficacy, safety) and sections 9-13 of this document.

8 Sample Size

(ICH E3; 9.7.2. ICH E9; 3.5)
This section should reproduce the relevant section from the protocol. If any amendments to the sample size have been made during the study, these should be documented and explained here. If any techniques are used to adjust the primary analysis for sample size adjustment they should be described in the relevant section (10.1).

9 General Analysis Considerations

9.1 Timing of Analyses

Give details here of when, or under what criteria, the final analyses will be performed. Give details of what data cleaning and locking processes must take place to comply with standard operating procedure (SOP) specifications. For example:

- *The final analysis will be performed after XXX progressions have been observed*
- *The final analysis will be performed when XXX subjects have completed visit Y or dropped out prior to visit Y.*
- *The final analysis will be performed on data transferred to the file XXX, having been documented as meeting the cleaning and approval requirements of SOPZZZ and after the finalization and approval of this SAP document.*

9.2 Analysis Populations

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 5.2)

This section is designed to identify the characteristics needed for inclusion in particular populations used in the analyses. Clearly define all the populations with a formal title (e.g. Full Analysis, Per Protocol, Safety) and give criteria to determine if a subject or observational unit belongs to that population. The criteria typically relate to adherence to protocol and the taking of observations, which relate to missing data (section 9.4).

A note on terminology: Pharmaceutical companies generally adhere to the International Conference on Harmonization (ICH) guidelines which use the term Full Analysis Set instead of Intention to Treat (ITT) population. In academia, the term ITT is more generally used. In ICH E9 "Statistical Principles for Clinical Trials," the Full Analysis Set is identical to the Intention-to-Treat (ITT) Population: "The intention-to-treat (see Glossary) principle implies that the primary analysis should include all randomized subjects. … In practice this ideal may be difficult to achieve, … the term 'full analysis set' is used to describe the analysis set which is as complete as possible and as close as possible to the intention-to-treat ideal of including all randomized subjects." Also, the distinction between (pure) ITT and modified ITT should be made when defining the study population. Most will define ITT as all randomized subjects, while modifications of the ITT principle exclude subjects who are never treated or who do not have post-treatment values.

It is not enough just to use a standard label for population. Such labels are vague and need further precise definitions within each trial; ***examples*** are given below.

9.2.1 Full Analysis Population (or Intention to Treat or Modified Intention to Treat)

- *All subjects who received any study drug*
- *All subjects who received any study drug and who participated in at least one post-baseline assessment*
- *All subjects who were randomized*

9.2.2 Per Protocol Population

- *All subjects who adhere to the major criteria in the protocol (e.g. all subjects who completed at least two efficacy analyses, whose study drug compliance was between 75% and 125% and who did not take any rescue medication)*
- *All subjects who did not substantially deviate from the protocol as to be determined on a per-subject basis at the trial steering committee immediately before database lock.*

9.2.3 Safety Population

- *All subjects who received any study treatment (including control) but excluding subjects who*

drop out prior to receiving any treatment.

- *All subjects who received any study treatment (including control) and are confirmed as providing complete follow-up regarding adverse event information.*

Discuss each of the following

- Specification of the primary efficacy population
- Specification of the population to be used for each type of data (e.g. background, safety, efficacy, health-economic).

If the primary analysis is based on a reduced subset of the subjects with data (e.g. subjects who complete the active phase of the study) and if the trial is intended to establish efficacy, there should be additional analyses that use all the randomized subjects with any on-treatment data.
It is crucial to assign each subject's inclusion or exclusion status with regard to each analysis population prior to breaking the blind. Such a statement should be included in this section. The exact process for assigning the statuses will be defined and documented prior to breaking the blind along with any predefined reasons for eliminating a subject from a particular population.

9.3 Covariates and Subgroups

(ICH E3; 9.7.1, 11.4.2.1. ICH E9; 5.7)
Provide a general comment identifying the covariates (continuous or categorical, including subgroups) that are expected to have an important influence on specific endpoints (e.g. demographic or baseline measurements, concomitant therapy). Document any model selection procedures (e.g. forward stepwise selection).

Any variables used to stratify or minimize over in treatment allocation should be adjusted for in the primary analysis; otherwise specific reasons should be included (for example, a categorical variable used in a minimization treatment allocation process could be omitted if it introduced too many categories).

State which important demographic or baseline-value-defined subgroups are to be analyzed for different treatment effects (for example comparison of effects by age, gender, ethnic group, prognosis, prior treatment). If there exists an *a priori* hypothesis of subgroup differences, it should be noted in this section. Likewise, it should be noted if subgroup analyses are exploratory.

Subgroup analyses should focus on the evidence for a difference in treatment effects: the interaction effect. It is flawed to present an analysis that provides two p-values, one for each of the two subgroups, and then report that only one subgroup showed a statistically significant difference. Only if the interaction effect is judged to be statistically and clinically significant should subgroup-specific treatment effect estimates be presented. It is acceptable to present exploratory subgroup-specific ***summary*** statistics. The use of forest plot figures is a highly effective way of communicating the relevant information about possible subgroup effects and interactions.

Where applicable, discuss the impact of the sample size on the power of subgroup analyses or reference section 7 if discussed there.

Additionally, The NIH is increasingly focused on the consideration of relevant biological variables such as sex, age, weight, and underlying health conditions. In particular, NIH expects that sex as a biological variable will be factored into research designs, analyses, and reporting in studies unless scientific literature strongly supports inclusion of only one sex. For further guidance, refer to the NIH website.

9.3.1 Multi-center Studies

(ICH E3;9.7.1, 11.4.2.4. ICH E9; 3.2)

This section may be copied directly from the protocol, if appropriate.
Where a multi-center study is intended to be analyzed as a whole, describe the following:

- Procedures to combine individual center results into more usable pseudo-centers with greater numbers of subjects
- Rationale for the combining of centers and the decision rule for whether or not the grouping will be necessary
- Methods to test for qualitative or quantitative treatment-by-center interactions
- Analyses of treatment comparisons that will allow for center differences with respect to response
- Center effects should be considered exploratory in analyses of studies that have not been explicitly designed with enough power to detect center effects.

The general discussion about the analysis of subgroups in section 8.3 is applicable to center effects.

9.4 Missing Data

(ICH E3; 9.7.1, 11.4.2.2. ICH E9;5.3. EMA Guideline on Missing Data in Confirmatory Clinical Trials)

Describe procedures to be used for dealing with premature discontinuation from the study or treatment and the handling of spurious or missing data (e.g. use of multiple imputation, random effects models or complete case analyses). Describe any possible biases these techniques may introduce. Describe the underlying assumptions (e.g. Missing At Random) in both statistical and non-statistical terms. Describe procedures to be used for describing the pattern of permanent (i.e. dropout) or transient missing data.

This section is intended to be a *general* discussion of the approach to missing data. Variable-specific information for imputing missing data, where appropriate, will be documented in section 6.5; analytical methods may be further detailed in section 9.

A minimal requirement would be to quantify the extent of missing data. Explicitly for key primary and secondary endpoints, and possibly an overall graphical representation of the extent of missing data, a “bar-code” style figure for example.

Having to exclude participants with observed endpoints from a regression analysis because of missing baseline covariates is definitely worth creating a simple alternative to avoid. A simple strategy is to impute a value of 0 for missing covariates, and then include a missing-indicator variable for the covariate as an additional covariate in the regression model [White and Thompson (2005)].

9.5 Interim Analyses and Data Monitoring (as applicable)

(ICH E3; 9.7.1, 11.4.2.3. ICH E9; 4.1, FDA Feb 2010 “Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics”)

For complex studies, consider pulling this section out and using as a standalone interim analysis Statistical plan document.

If this template is used to write the SAP for an interim analysis then this section should be limited to documenting the motivation and objectives for the interim analysis, typically by quoting from the

protocol or DMC charter; precise details of the methods used to carry out the interim analysis should be given in section 11.

If however this is the final SAP, then details of any preceding interim analyses, including analysis methods should be recorded in this section, or references given to previous interim SAPs.

Subsections that are not applicable may be deleted entirely.

9.5.1 Purpose of Interim Analyses

Give a description of why the interim analyses are to be performed. Typically, the reason is due to uncertainty about some aspect or aspects of the treatment or treatments and the interim will allow learning to influence the subsequent design of the study at Data Monitoring Committees. This can range from simple uncertainty about safety aspects, the primary endpoint treatment effect that leads to early termination for futility of efficacy, to decisions regarding the choice of dose, endpoint, treatment arm, randomization weighting, subgroup enrichment. The data to be analyzed in the interim analyses should be explicitly specified (for example, baseline data, treatment received, safety).

9.5.2 Planned Schedule of Interim Analyses

It must be detailed when the first interim analysis will occur, and what scope of decisions will be taken at future interim analyses. Technically, details of the interim analysis beyond the next interim can be left open to be decided sequentially at each interim, under the proviso that rules for the analysis to combine the future data at each stage are defined and the scope for adaptations is not enlarged. However, it is recommended to plan as much as possible in advance and give a full predicted schedule of all interim analyses.

9.5.3 Scope of Adaptations

Give an explicit list of which aspects of the trial may be revised at an interim analysis. Document any formal rules governing these adaptations. If an interim SAP will not be produced, or it is appropriate to document the interim analysis in the main SAP, then specify what analyses, summaries or figures will be used to inform the choice of adaptations.

9.5.4 Stopping Rules

Document any formal stopping rules for futility, efficacy or lack of power. Document the probability of each possible eventuality under the null and alternative hypothesis e.g. the probability of stopping for futility or efficacy, or continuing to the next stage.

9.5.5 Analysis Methods to Minimize Bias

It is generally advised to perform a naïve analysis that pools all data at the final analysis as if it were collected in a fixed design. However, this may induce biases in estimation. For example, in group sequential designs the estimate of treatment effects will be biased away from the stopping region; for sample sizes that are revised to reflect the estimated treatment effect at the first interim, the naïve pooled estimate will be biased away from the null.
Any known biases must be discussed, and any methods proposed to correct the biases must be documented. It must be stated in advance which analysis will be the primary analysis used in the case of conflicting interpretations and for “headline” reporting of the trial.

9.5.6 Adjustment of Confidence Intervals and p-values

In a design where formal hypothesis testing is used and the interim analyses provide multiple opportunities to stop for efficacy, the overall trial significance level will be greater than nominal

significance levels used at each stage. Any stopping rules should adjust for this, and correspondingly any confidence intervals or p-values presented must be calculated to adjust for the possibility of stopping earlier and for having reached the observed stage in the trial.
Conversely, a trial that only has the option to stop early for futility will conservatively preserve the overall significance level. Here the nominal confidence interval and p-value at the end of the trial can be used. Investigations should be made into the effect on the power of the trial, and only if the power is substantially reduced should adjustments be used.

9.5.7 Interim Analysis for Sample Size Adjustment

If the sample size is to be adjusted at an interim, specify any rules: for example, conditional power calculations.

The weighting of data from different stages of the trial needs to either, be set in advance independently of (random) sample sizes, or rules given for how the weighting will be determined. The final analyses must specify how these weightings will be used, see section 8.5.5.

9.5.8 Practical Measures to Minimize Bias

It is important to establish and control who will have access to what information at each stage of the trial. Uncontrolled reporting of interim analyses to study centers could lead to investigators responsible for recruiting subjects to change their desire to recruit to a trial, which would induce uncontrollable biases into the subject population. The final analyses could be biased by knowledge of interim results by the analyst. Any level of unblinding, either of individual subjects or of treatment estimates, could induce biases.

It should be explicitly documented:

- who will perform any interim analysis
- who will see any data or analyses at the interim and make decisions
- what information will be publically available following an interim analysis
- what information will be provide to the sponsor and investigators
- who will be unblinded at any point in the trial
- who will perform any final analyses and remain blinded
- if any safety monitoring decision making will remain isolated from efficacy information

9.5.9 Documentation of Interim Analyses

Snapshots of the data available at each interim analysis should be preserved, as should all documentation of analysis plans, programming code and reporting provided at each interim. It should be possible to recreate the decision process from the trial archive in the fullness of time when any limitations of access to information by blinded statisticians become redundant.

If multiple interim reports are prepared, then it must be clear which versions of the SAP, data set, and report are linked together even if there are no changes in the SAP. A copy of a previous version can used to make this clear if needed.

Record what documents will be created and stored thus.

9.6 Multiple Testing

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 2.2.5)

This section can be copied from the protocol, if available.

In a confirmatory trial, the choice of sample size will be justified in terms of the power, which focuses on a single analysis, which in turn focuses only on one primary endpoint. This means there should only be one primary endpoint.

However, exceptions to this maxim do occur. In such circumstances the most acceptable statistical methodology is to either combine the co-primary endpoints through a deterministic function into a single endpoint or adopt a formal closed-testing procedure that examines a variety of hypotheses in such a way that preserves the overall significance level of the analyses; for example, Bonferroni adjustments of the nominal significance level, or gate-keeping approaches to a pre-specified order of hypothesis tests.

In trials that are focused on learning and hypothesis generation, the preservation of the overall significance level is of lesser importance, however there should be a statistical discussion that reflects awareness of the issues, and the analyses should present confidence intervals rather than p-values.

Identical issues arise if there are:

- more than two treatment groups,
- subset analyses,
- multiple time points,
- multiple methods of analysis,
- sensitivity analyses for missing data.

10 Summary of Study Data

Specify the method of describing the study data that will be presented in the final study report. This section provides a general overview of the methods. If any of the items require a unique approach that differs from the general overview, then it should be noted in the appropriate section.
Specify

- How summary tables will be structured (e.g. columns for each treatment and overall in the order: Placebo, Experimental Low Dose, Experimental High Dose, All Subjects.)
- Descriptive or summary statistics that will be displayed for continuous data and for categorical data.
- The analysis populations upon which the tables and figures will be based.

Example:

> *All continuous variables will be summarized using the following descriptive statistics: n (non-missing sample size), mean, standard deviation, median, maximum and minimum. The frequency and percentages (based on the non-missing sample size) of observed levels will be reported for all categorical measures. In general, all data will be listed, sorted by site, treatment and subject, and when appropriate by visit number within subject. All summary tables will be structured with a column for each treatment in the order (Control, Experimental) and will be annotated with the total population size relevant to that table/treatment, including any missing observations.*

It may be worthwhile to document that the sample size of non-missing values for univariable summary statistics may be larger than the sample size of non-missing values in a complete-case analysis used in a typical primary regression analysis, as different patients may have missing values in different variables, such as baseline covariates and the endpoints. Generally, we would not want to exclude observations by repeating the summary statistics in a "complete case population." Exclusions may imply the excluded observations are systematically different, thus contradicting a typical Missing At Random assumption, and could also be considered as a post-hoc subgroup analysis.

Only deviations from the general overview will be noted in the subsequent sub-sections within section 9.

10.1 Subject Disposition

Specify which variables from the CRF will be used to establish how many subjects reached the various stages of the trial, how many dropped out and for what reasons (death, toxicity, treatment failure, withdrew consent). For example, the number screened, randomized, reached visit 1- X, study close, follow-up visits 1- X…

An overview of the time-dependent rates of recruitment should be provided.

If there is any ambiguity arising from multiple sources of visit date, document how this will be resolved.

This section should determine how the population membership and population size for each treatment (to be used in most table headers) will be determined.

A skeleton CONSORT flow diagram should be provided in this section that provides an explicit statement of what statistics are to be provided. http://www.consort-statement.org/.
The reporting requirements of the regulatory agencies need to be met.

If appropriate use standard text: "The summary statistics will be produced in accordance with section 9."

10.2 Derived variables

If any endpoints are derived from a variable or variables calculated from source data recorded in the CRF, then their definition should be provided. Ensure that a primary endpoint that is such a derived variable is clearly identified with a consistent name.

The description of each derived variable should include the method for computing the variable including any special techniques used in the computation (e.g. carrying forward values into missing observations, transformation of values) and specific methods for combining multiple variables into a single value (e.g. EQ-5D Quality of Life questionnaire).

An example of a derived variable is a binary variable indicating if an ordered categorical variable has increased from baseline.

10.3 Protocol Deviations

Define the specific protocol deviations that could impact the analysis (e.g. major deviations and a definition of a major deviation) and specify the methods used to describe and analyze them. Clearly

define which deviations will exclude a subject from each of the analysis populations defined in section 8.2.

If appropriate use standard text: “The summary statistics will be produced in accordance with section 9.”

10.4 Demographic and Baseline Variables

Identify all variables, or groups of variables, that will be considered as demographic or baseline variables, recorded at, or shortly, before randomization or first treatment administration. The table of tables document may be given as reference to provide the finest level of detail needed to produce the report and to avoid duplication. If transformation of data will occur (e.g. age coarsening into 18-40, 41-65, over 65) then define this exactly in this section. It may be appropriate to summarize these data by center.

If appropriate use standard text: “The summary statistics will be produced in accordance with section 9.”

10.5 Concurrent Illnesses and Medical Conditions

Include a description of which, if any, coding system was used and version number (e.g. MedDRA v19.1, WHO drug dictionary). If appropriate use standard text: “The summary statistics will be produced in accordance with section 9.”

10.6 Treatment Compliance

Examples of the assessment of treatment compliance include: remaining pill count, diary records of medication. Any method for calculating a measure of treatment compliance should be defined clearly here. The variables used to assess treatment compliance should be identified. The table of tables document may be given as reference to provide the finest level of detail needed to produce the report and to avoid duplication.

If appropriate use standard text: “The summary statistics will be produced in accordance with section 9.”

11 Efficacy Analyses

Specify the method of summarizing and formally analyzing the efficacy data that will be used. This section includes a description of the general methods that will be used repeatedly for different variables and analysis populations. If any items need to be handled in a different manner, then this should be clarified in the appropriate section below.

The following should be specified in this section

- Grouping of summary table information (e.g. by treatment group and in what order, possibly adding a combined “all subjects” column)
- Summary statistics that will be produced for continuous and categorical data
- Analysis populations that will be used and identification of the primary population.

The following details of the statistical analyses should be considered and included as needed. The details should be placed either in section 10, if broadly applicable across most analyses, or in the relevant subsections below.

- The statistical model underlying the analysis including, strata, covariates, baseline values and interaction terms.
- A statement of the clinical objective rephrased in precise statistical terms (null and alternative hypotheses).
- The nature of the hypothesis: descriptive, exploratory or confirmatory.
- The methods used to obtain parameter estimates, confidence intervals and, if required, p-values.
- Methods used to check any assumptions behind the analyses (histograms, box plots) and approaches to be taken if the data do not meet the assumptions.
- The test statistics, the sampling distribution of the test statistic under the null hypothesis, significance level, alternative hypothesis, whether the test is 1- or 2-sided. If a 1-sided test is to be used, provide justification.
- If Bayesian techniques are to be used, specify which prior distributions will be considered, or how the prior(s) will be obtained, with justification.
- Any procedures for removing non-significant covariates from the model or model selection procedures in general.
- Methods and assumptions for handling longitudinal data or missing data.
- Consider how the results may be compared to summary statistics. A regression analysis may benefit from centering covariates around a mean value, so that the intercept can be given a clear meaning that corresponds to a summary statistic reported elsewhere.
- Consider the possibility of storing the full results, including estimates of all nuisance parameters, and covariance matrices of parameter estimates in an appendix or csv file, for possible use in planning future studies, meta-analysis and to enable replication of analysis.

Examples:

All efficacy variables will be listed by subject within study center. Data will be summarized by treatment group. N, Mean, Standard Deviation, Minimum and Maximum will summarize continuous efficacy variables, whereas number and percent will summarize categorical efficacy variables.

All analyses of the continuous efficacy variables (e.g. VAS pain score) will be performed as analysis of variance with treatment group adjusting for study center and surgical category. Treatment groups will be tested at the 2-sided 5% significance level.

All assumptions for regression models will be assessed by viewing plots of the residual values.

All analyses of categorical efficacy measures will be performed using logistic regression with treatment group and adjustments for study center.

In some cases, it may be more appropriate to clarify items that do not conform to the general methods in individual sections below.
All the variables being considered should be mentioned explicitly in the subsections below even if they are mentioned in this section.

11.1 Primary Efficacy Analysis

Define the primary analysis that will provide the main result of the trial in this section. Note the use of "analysis" singular. This section of the document should be structured in parallel with 5.2 in terms of the ordering of endpoints considered. If appropriate use standard text: "The summary statistics will be produced in accordance with section 9."

11.2 Secondary Efficacy Analyses

Include all secondary efficacy analyses. There may be secondary analyses of the primary endpoint including subgroup analyses and sensitivity analyses, which will be included in this section. If appropriate use standard text: "The summary statistics will be produced in accordance with section 9."

11.2.1 Secondary Analyses of Primary Efficacy Endpoint

e.g., Using the per protocol, instead of the ITT population, for the analysis of the primary efficacy endpoint.

11.2.2 Analyses of Secondary Endpoints

11.3 Exploratory Efficacy Analyses

Further analysis of exploratory endpoints used for hypothesis generation and exploration should be included here. If appropriate use standard text: "The summary statistics will be produced in accordance with section 9."

12 Safety Analyses

Specify the methods of describing the safety data that will be used in the final report. This section includes general descriptions of the methods. If any of the items require a unique approach, then this should be noted in the appropriate subsection below.

Specify:

- Sort order of any listings.
- Grouping of summary information (e.g. by preferred terms and treatment group, including an "All Subjects" column).
- Descriptive statistics that will be displayed for continuous data and for categorical data.
- Analysis populations on which the descriptions will be based.
- How repeat events will be handled when producing summary statistics. For example: *"When calculating the incidence of adverse events, or any sub-classification thereof by treatment, time period, severity, etc., each subject will only be counted once and any repetitions will be ignored; the denominator will be the total population size."*

Only deviations from the aforementioned analytical and summary approaches will be noted in the subsequent subsections of section 11.

Variables being summarized should be listed in the subsections below. The table of tables document may be given as reference to provide the finest level of detail needed to produce the report and to avoid duplication.

Per ClinicalTrial.gov reporting requirements, three tables summarizing adverse events are required:

- All-Cause Mortality (only required if primary outcome completion date is on or after

01/18/2017): A table of all anticipated and unanticipated deaths due to any cause, with number and frequency of such events in each arm/group of the clinical study.

- Serious Adverse Events: A table of all anticipated and unanticipated serious adverse events, grouped by organ system, with number and frequency of such events in each arm/group of the clinical study. (See Adverse Events definition below).

- Other (Not Including Serious) Adverse Events: A table of anticipated and unanticipated events (not included in the serious adverse event table) that exceed a frequency threshold (for example, 5 %) within any arm of the clinical study, grouped by organ system, with number and frequency of such events in each arm/group of the clinical study.

12.1 Extent of Exposure

If appropriate use standard text: "The summary statistics will be produced in accordance with section 9."

12.2 Adverse Events

If appropriate use standard text: "The summary statistics will be produced in accordance with section 9."

It is useful to include an identification of the components of the numerator and denominator that will be used to calculate incidence rates and percentages. For example: *"When calculating the incidence of adverse events, or any sub-classification thereof by treatment, time period, severity, etc., each subject will only be counted once and any repetitions of adverse events will be ignored; the denominator will be the total population size."*

Be certain to specify those adverse events that will be included in the summary and analysis. For example, treatment emergent adverse events are those events that occur after the baseline assessment, and some definitions also include those adverse events that worsen post-treatment.

It may be appropriate only to report the incidence of specific AEs of interest, in which case document these specific AEs. Or it may be appropriate only to report the incidence of AEs that are judged to be related to the treatment.

12.3 Deaths, Serious Adverse Events and other Significant Adverse Events

If appropriate use standard text: "The summary statistics will be produced in accordance with section 9." State whether these are a subset of the Adverse Events, or if these are unique events not previously described. Separation of events into serious and non-serious AEs may be useful, depending on the overall number of AEs.

12.4 Pregnancies (As applicable)

If appropriate use standard text: "The summary statistics will be produced in accordance with section 9." If the study did not perform any pregnancy tests or pregnancies are impossible, for example if it was limited to male or post-menopausal subjects only, then explain this succinctly.

12.5 Clinical Laboratory Evaluations

If appropriate use standard text: "The summary statistics will be produced in accordance with section 9."

Address the issues of

- Normal ranges that differ between study centers. Explicitly tabulating the normal ranges when producing the SAP may be useful and timely to ensure the normal ranges have been provided by all centers.
- How to handle duplicate laboratory test within study periods. Normally summaries are only provided over scheduled laboratory tests. Any unscheduled follow-up tests performed for medical or safety concerns, are normally only listed.

Laboratory tests are often summarized using shift tables. Shift tables may show the change in laboratory values from baseline to either each subsequent visit, the final visit, or the most extreme post-baseline value. An alternative may be a figure showing a scatter plot of the baseline value on the horizontal axis versus the subsequent values, as considered above, with different plotting symbols used to distinguish different treatment groups.

12.6 Prior and Concurrent Medications (As applicable)

The definitions used to distinguish prior and concurrent should be provided. Include a description of which, if any, coding system was used (e.g. MedDRA, WHO drug dictionary). If appropriate use standard text: “The summary statistics will be produced in accordance with section 9.”

12.7 Other Safety Measures

If appropriate use standard text: “The summary statistics will be produced in accordance with section 9.” Vital signs might be appropriately included in this subsection. Many of the points made regarding laboratory tests in section 11.5 are relevant to vital signs.

13 Pharmacokinetics (As Applicable)

Describe pharmacokinetic and pharmacodynamic parameters to be analyzed and the approach to the data summaries and analyses. Include pharmacodynamic data in section 12 only if it is not considered as efficacy data included in section 10. If there are no such data collected, then this section may be deleted. If appropriate use standard text: “The summary statistics will be produced in accordance with section 9.”

If there are a number of variables observed, then it may be necessary to generate subsections below. All variables being summarized must be explicitly mentioned.

14 Other Analyses

Variables that cannot be easily included in the preceding sections should have their own section here. Replace the heading “Other Analyses” with more appropriate text. Some examples of such data are: health economic data, quality of life data, patient satisfaction data.
All the comments from section 9 onwards may be relevant.

15 Reporting Conventions

Describe reporting conventions, for example the precision used for reporting p-values and other numeric values.
Example:

P-values ≥0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001". The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

16 Quality Assurance of Statistical Programming (As Applicable)

For units without SOPs covering statistical programming quality assurance, the information below may provide guidance. For units with SOPs in place, this section can be removed.

Include a brief statement of: study-specific documents used, including version numbers; which software package or packages will be used if known; the operating system of the computer (e.g. Mac, Windows etc.) and the directory/file paths planned to store data, code and output documents. It may be helpful to precede such details with the text, "At the time of writing,".

Describe what quality assurance measures are in place to monitor the quality of any coding. Document who will review which pieces of code, and to what level of detail. For example:

A second review statistician will independently reproduce the primary analyses and summary statistics table X, Y, Z. The reviewing statistician will have an overview of the entire analyses and will explicitly check the code producing tables A, B & C (selected at random) as well as any other pieces of code as desired.

To provide high quality code that is understandable, and allows reproduction of the analysis the following points will be followed.

The population to be used in a table or figure will be explicitly set at the start of a block of code that computes the output, ideally by looking up the population from the table of tables.

Any outputs will have the

- date and time included
- the name of the code file that produced the analysis
- the author

At the start of any code file there will be a set of comments that give

- the author
- the date and time of writing
- references to inputs and outputs
- reference to any parent code file that runs the child code file

17 Summary of Changes to the Protocol and/or SAP

If the statistical analysis plan proposes changes to the statistical approach described in the protocol,

then summarize those changes in this section. Analyses are usually faithful to those specified in the protocol, but occasionally different, or supplemental, analyses are needed. Explain the reason for such changes. This section identifies those analyses that are not from the protocol and greatly aids clarity. Other important, non-statistical changes to the protocol should also be noted in this section, for example the introduction of an additional treatment group.

An example is provided below:

Rationale for Adjustments of Statistical Analysis Plan from Protocol (Version 7, Amendment 5, May 4, 2010)

The changes from the protocol-specified definitions of aims, outcomes and statistical analytic approaches are outlined below. These changes reflect advances in our knowledge of biliary atresia since the design of the study in 2004 that were not incorporated as protocol amendments, but were discussed during the formation of the Statistical Analysis Plan. These changes are documented herein and represent changes made prior to the database lock and unblinding of the study.

1. Clarification of Threshold for Total Bilirubin to Define Good Bile Drainage

PROTOCOL:
In various places in the protocol, good bile drainage is defined as < 1.5 mg/dL or $\leq$ 1.5 mg/dL and poor bile drainage is defined as $\geq$ 1.5 mg/dL or > 1.5 mg/dL, respectively.

SAP:
For consistency, good bile drainage is defined as < 1.5 mg/dL and poor bile drainage is defined as $\geq$ 1.5 mg/dL.

2. Hypothesis, Endpoint, and Analysis of Duration of Improved Bile Drainage

There is some ambiguity in the protocol regarding the definition of duration of improved bile drainage with respect to the endpoint and analysis method. The intent of the protocol is to assess the duration of improved bile drainage, and accordingly the SAP definition is based on time-to-event principles and the SAP analysis is based on time-to-event methods.

PROTOCOL:
Section 4.E Secondary outcome measure: "All measurements will be made at 12 and 24 months of age (unless noted otherwise):

1. Serum total bilirubin concentration (and also at 3 months after portoenterostomy)"

Section 5.A2 Sample Size for Secondary Endpoints: "A serum total bilirubin value >1.5 mg/dL, transplant or death, is indicative of failure of the portoenterostomy. The time to the first value >1.5 mg/dL (when the value is initially >1.5 and then declines to ≤1.5, the time to the first value >1.5 mg/dL after the decline will be used), to transplant or to death will be considered time to failure of the procedure. The two groups will then be compared by a Cox regression model fitted to the time to failure."

Section 5.B Data Analysis:
"Aim 1: To determine whether corticosteroid therapy decreases serum bilirubin concentration after portoenterostomy.

Hypothesis 1b: Improved bile drainage (as defined by a serum total bilirubin level <1.5 mg/dL) will remain for a longer period in infants treated with corticosteroids than in those treated with placebo. Using the same definition of good bile drainage as in the primary hypothesis, the duration of time with good bile drainage (time to onset of poor bile drainage, transplant or death, whichever comes first) will be modeled by Cox proportional hazards model using the same model as above. If there is a statistically significant difference, the profiles of the bilirubin for the two treatment groups will be compared and 95% confidence limits will be computed."

SAP:
Secondary Endpoint. The duration of improved bile drainage will be defined as the time from a subject's first total bilirubin is <1.5 mg/dL to the earlier of his/her first total bilirubin value ≥1.5 mg/dL, liver transplantation or death. The date of the first bilirubin value ≥ 1.5 mg/dL, liver transplantation or death is the failure date. Lost-to-follow-up, withdrawal, and study completion without experiencing the event are censoring events. If a subject is missing total bilirubin values during the study, then duration of good bile drainage is censored at the date of the last known value <1.5 mg/dL for subjects who do not have subsequent liver transplantation or death. If a subject never achieves good bile drainage (i.e., total bilirubin ≥1.5 mg/dL for the duration of the study), then duration is set to 0. Time will be defined as duration of good bile drainage in months; it is not defined as the time from HPE.

18 References

Provide references for any citations in the main body of the SAP.
White IR and Thompson SG (2005). Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine,* **24,** 993-1007

19 Listing of Tables, Listings and Figures

This section is to give precise details for each table, listing or figure to be produced.

Mock tables, listings and figures should be provided. Including the title, number, analysis population, etc.

Tables: # of decimal digits, formatting (e.g., confidence in brackets, in parentheses, in separate columns).

Figures: labels for all axes, legends, plotting symbols.

These detailed specifications can have *minor* revisions during the production phase without needing to revise the SAP, this may include changing the table numbers if a reordering or deletion is appropriate, providing that the specifications in the main body of the SAP sections 4 to 17 are met.

Suggestions for ordering or organizing TLFs:

e.g., Tables E for efficacy and S for safety and nothing for patient disposition.

Disposition is Table of Figure 1.xxx
Efficacy is Table or Figure 2.xxx
Safety is Table or Figure 3.xxxx
Other is Table or Figure 4.xxxx

Statistical Analysis Plan: Does the use of nerve stimulator improve the outcome of ultrasound guided supraclavicular block (anesthesia) for upper extremity surgery?

Having reached the end, you should edit the headers and footers to add in the correct study title, change the version number, the date you finalized the current version (do not use the automatic "today's date" as this will change each time you open the document). Also check the accuracy of the table at the top of the document. Update the table of contents. Leave in this reminder paragraph until the final version is confirmed.