SNAP TRIAL

DATA MANAGEMENT PLAN

Version Number and Date: Version 1.0 dated 27 June 2023 Supersedes version: N/A				
	Name	**Position**	**Signature**	**Date**
Authors	Jay Keatley	Data Coordinator		
	Jocelyn Mora	Global Trial Manager		
Reviewed by	Robert Mahar	Statistician		
Approved by	Anna McGlothlin	Senior Statistician		
	Josh Davis	Chief Investigator		
	Steven Tong	Chief Investigator		
Approved by Regional Sponsor		Regional Sponsor Representative		

DOCUMENT HISTORY

VERSION	DATE	REASON FOR REVISION

1 CURRENT PROJECT STATUS

An overview of the current project status is below.

Pre-trial setup	January 2020 to March 2022		
Database go-live	15 February 2022		
First participant recruited to the trial	18 February 2022		
Enrolment target, platform	Up to 7,000 participants		
Enrolment target, registry	Up to 20,000 participants		
Follow-up period, platform and registry	90 days		
Trial enrolment status	☐ In set-up	☒ Enrolment open	☐ Enrolment closed

Collection methods used			
Electronic Case Report Form (eCRF), Spinnaker Database	☐ In set-up	☒ Data entry open	☐ Data locked
Paper Case Report Form (pCRF)	☒ Replaces some eCRFs only during Spinnaker set-up	☐ Tool only - all data must be entered in eCRF	☐ Not in use

2 TABLE OF CONTENTS

3 GLOSSARY OF ABBREVIATIONS

Term	Definition
Acute index hospitalisation	Initial hospital admission to an acute inpatient facility, **does not** include HITH/OPAT/COPAT and stepdown inpatient rehabilitation/post-acute care
COPAT	Complex Outpatient Parenteral Antimicrobial Therapy
CRF	Case Report Form
CSV	Comma Separated Values
DC	Data Coordinator
DSA	Domain Specific Appendix
DSMC	Data and Safety Monitoring Committee
DSWG	Domain-Specific Working Group
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
FOI	Focus of Infection
GTM	Global Trial Manager
GTSC	Global Trial Steering Committee
HITH	Hospital in the home
ITT	Intention-to-treat
MSSA	Methicillin-Susceptible *Staphylococcus aureus*
MRSA	Methicillin-Resistant *Staphylococcus aureus*
OPAT	Outpatient Parenteral Antimicrobial Therapy
pCRF	Paper Case Report Form
Platform	Patients in the platform are those who meet all core eligibility criteria and consent to inclusion in the platform. Occasional patients in the platform will not receive any randomised intervention (if they are not eligible for any available domain).
Platform entry	"Platform entry" is the timepoint when the patient has met core eligibility criteria, given informed consent for the platform, and been randomised
PD	Protocol Deviation
PSSA	Penicillin-Susceptible *Staphylococcus aureus*
PWID	People Who Inject Drugs
RAR	Response Adaptive Randomisation
Registry	Patients in the registry include all those in the platform (as defined above) PLUS the "registry only" patients. Registry-only patients are those who are not in the platform, but who have consented to being in the registry.
RCT	Randomised Control Trial
REMAP	Randomised Embedded Multifactorial Adaptive Platform
RTSC	Regional Trial Steering Committee
SAB	Staphylococcus aureus Bacteraemia
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis Subcommittee
SNAP	Staphylococcus aureus Network Adaptive Platform trial
SOP	Standard Operating Procedure
SSI	Significant Safety Issue
Total index hospitalisation	Initial hospital admission to an acute inpatient facility, including HITH/OPAT/COPAT and stepdown inpatient rehabilitation/post-acute care (if continuous with the initial inpatient admission)
TMF	Trial Master File
TMG	Trial Management Group
UAT	User Acceptance Testing

4 TRIAL INFORMATION

4.1 TRIAL OBJECTIVES

The objective of SNAP is to identify the effect of a range of clinical interventions on all-cause mortality, 90 days after platform entry, in participants with *Staphylococcus aureus* bacteremia (SAB). The SNAP platform aims to collect treatment and outcome data accurately and efficiently to evaluate the comparative effectiveness of alternative treatments. The platform is designed to be adaptive and can accommodate additional pharmacological and non-pharmacological interventions either within existing domains or as part of new domains. This data management plan will be updated when new domains are added to the trial.

4.2 RATIONALE

SAB is a common and severe infection with a 90-day mortality of 15-30% (mortality lower in children but up to 5%) despite the current best available therapies. There are few high-quality data to inform the management of this infection, with less than 3000 patients randomised into any therapeutic trial for SAB before 2020.

We are using an adaptive platform trial to allow us to simultaneously address multiple questions in the management of SAB. We plan to test interventions within 3 initial domains, with the potential to add further domains to the platform. The trial will include 3 silos (susceptibility profiles; penicillin-susceptible (PSSA), methicillin-susceptible (MSSA), and methicillin-resistant *S. aureus* (MRSA)).

For more information, please refer to the core protocol and core protocol publication.

4.3 DOMAINS AND INTERVENTIONS

SNAP is an investigator-initiated, randomised, embedded, multifactorial, adaptive, platform trial, conducted across multiple hospitals globally. Eligible, consenting participants will be randomly allocated to a prescribed regimen of interventions selected from multiple treatment domains. The platform enables the introduction of new treatments and/or domains as the trial progresses.

5 DATA MANAGEMENT PLAN

5.1 PURPOSE

The purpose of this document is to provide detailed data management procedures and related responsibilities for the *Staphylococcus aureus* Network Adaptive Platform (SNAP) trial. This excludes any procedures relating to the handling and processing of Safety Events, which are documented in the Standard Operating Procedures (SOPs) document titled Safety Reporting SOP002 (or regional-specific variation).

This document will be updated throughout the life of the trial to reflect any changes in data management procedures. It should not be used in isolation.

5.2 REVIEW PROCESS

The Data Management Plan (DMP) will be reviewed periodically as and when changes to data management working practices occur, for example following a substantial protocol amendment. When a new version is created, it will be reviewed and approved by those listed on the cover page, and a reason for revision will be listed in the revision history table. When a Data Coordinator leaves the trial, moves between trials, or changes roles, they must ensure the DMP is up to date before they leave.

The review process will:

- address issues identified in the DMP and any other quality documents.
- document any new risks identified during the trial and inform updates to other relevant quality documents (where applicable).
- ensure compliance with the SNAP trial protocols, appendices, and SOPs.

5.3 KEY ELEMENTS

Key elements of data management include database design, data entry, data checking and cleaning, site monitoring, appropriate maintenance of blinding of allocations to interventions and outcomes, data access, operational reports, reporting of interim and final analyses, dataset locking, database locking, and archiving. As an adaptive platform trial, most of these elements will be occurring throughout the lifetime of the trial.

6 DATABASE DESIGN

6.1 OVERVIEW

The Spinnaker database eCRF was developed in conjunction with, and is hosted by, Spiral Software Ltd (Spiral), New Zealand. It is presented in English and is compatible with standard web browsers, tablets, and smartphones. Data protection information for the Spinnaker Database and Spiral is available in Appendix F:

The eCRF is accompanied by a data dictionary spreadsheet that is jointly managed by Spiral and the central TMG. The data dictionary contains a copy of the questions in the eCRF and describes rules governing whether to show/hide certain questions for specific participants, acceptable value ranges, etc.

Raw data exports are stored to the University of Melbourne Research Data Share (Research-NAS), with access regulated by the Research Computing Activity Owner (RCAO; **see Section 11.0**).

Database reports, including operational reports (**See Section 9.3 Operational Reports**), will be available in Spinnaker and as downloadable exports (Excel files in .csv format).

All raw data exports and downloadable reports will have a conventional label, and a timestamp reflecting the date and time in which the export was downloaded.

6.2 TABLE OF CRFS AND FORM DATES

Below is an example of an outline of the eCRFs and form dates for each (i.e., the field on the eCRF that is considered the form date). **See Appendix C. for further examples.** A full version history is maintained in the Spinnaker electronic database.

Module	CRF Name	Form date	Form date field name
Platform Eligibility	Add Patient (before enrolment) Randomisation (after enrolment)	Date of eligibility screening	EL_ScreeningDate
Domain Eligibility	Adjunctive Eligibility	Date of Adjunctive Domain screening result	DomainB_RevealTimeLocal DomainB_RevealTimeUTC
Domain Eligibility	Backbone Eligibility	Date of Backbone Domain screening result	DomainA_RevealTimeLocal DomainA_RevealTimeUTC
Domain Eligibility	EOS Eligibility Day 5-9	Date of EOS Domain screening result (day 7 ± 2 days)	DomainC_D7_RevealTimeLocal DomainC_D7_RevealTimeUTC
Domain Eligibility	EOS Eligibility Day 12-16	Date of EOS Domain screening result (day 14 ± 2 days)	DomainC_D14_RevealTimeLocal DomainC_D14_RevealTImeUTC
Consent	Consent RC	Date and time of consent with regained capacity	CON_RCDateTime

7 DATA CAPTURE

7.1 RECORDS THAT CONTAIN TRIAL DATA

- Spinnaker eCRF: Data is captured in electronic Case Report Forms (eCRFs) on the central database known as Spinnaker.
- Paper CRFs are only used as a tool for this study during the set-up phase and are not considered source data. Sites may also be using paper screening and enrolment logs, or paper questionnaires (for example, the EQ-5D).
- Paper consent documents: Most regions will obtain patient consent on paper consent forms, with a wet-ink signature. These are considered source data and should be stored in accordance with, and for the duration specified by regional regulations.
- Spinnaker extracts: Extracts from the database containing the raw data in comma-separated value (.csv) files are known as Spinnaker extracts. If randomisation (allocation) codes are available in any Spinnaker extracts accessible to a member of the SNAP trial team, that person is considered to have unblinded access.
- Trial Site Metadata refers to the location settings in the database that act as switches to control (*list not exhaustive)*:
    - Patient recruitment (adults/paediatrics/both)
    - Age of adult patients (18+ years/21+ years)
    - Recruitment of pregnant patients (yes/no)
    - Screening variables to be collected (Initials/DOB/age)
    - Participation in sub-studies (PWID)
    - Which domains are actively recruiting (Adjunctive/Backbone/EOS)
    - The interventions available within each domain
    - Limitations on who may provide consent
    - Other nuances not already listed, e.g., region-specific questions
- Interim Reports: Include (but are not limited to) safety reports and interim analyses such as those tabled at the Data and Safety Monitoring Committee (DSMC), working group, and committee meetings.

Investigators, working groups, management teams, and site teams will have varying access to trial data as outlined in the Appendix.

Spreadsheet logs are maintained in the central management team eTMF to log information not recorded within the Spinnaker database. **See Appendix B. Data that supplements the CRF.**

7.2 DATA ENTRY BY SITES

Data is entered by authorised users via https://snap.spinnakersoftware.com.

The eCRFs are designed to minimise missing data and all questions should be answered to achieve a 'green tick' that denotes the form has been completed. Incomplete forms are denoted by an orange question mark and forms not commenced are denoted by a red cross.

There are validity and consistency checks within the database (e.g., range checks for datapoints, and consistency of dates). Warnings appear when a response contradicts what is reasonably expected for that field. Data inconsistencies that remain after forms have been saved (such as conflicts between eCRFs, alerts to protocol deviations, and other matters requiring attention) are denoted by yellow banners at the top of each page, with a brief explanation of any action/s needed.

8 DATA QUALITY CONTROL

8.1 DATA CHECKING AND CLEANING

We will endeavour to check and clean extracted data in close to real time. **See Section 11. Data Extraction for a detailed description of the data extraction process.** Data extracts will be checked with scripts to provide aggregated reports on data completion and inconsistent data points at the whole of trial and regional levels. Line listing of relevant regional and site data completion of specified CRFs and selected data points for inconsistencies will be provided to regional managers and sites.

Aggregated reports on data completion and timeliness of data completion will be provided to the GTSC, and regional trial management groups.

Global and regional trial managers and data coordinators can produce inline queries within the Spinnaker eCRFs to request data checks and corrections at the trial site level. These queries and responses are logged within Spinnaker. It is expected that queries will be addressed within 1 working week.

When data cleaning is complete and quality conditions are met for each participant CRF, a soft lock will be applied. **See Section 12. Data Locking for a detailed description of locking procedures.**

8.2 PRIORITY VARIABLES

Priority variables are defined as those variables that, if found on eCRFs to be missing, out of normal ranges and/or indicative of a protocol violation or unreported adverse event (or any other noteworthy event as defined in the trial protocol), will be cleaned, and chased with more urgency than other data.

Some priority variables are tightly controlled by the database. For example, users will not be able to proceed with patient screening, enrolment, and randomisation if all key variables are not entered. Other variables are associated with data validation checks to ensure that they are complete, within range, or flagged as indicative of a protocol deviation or safety event. These are outlined in the SNAP Trial Data Description.

Priority variables for additional validation checks are:

Priority Variable	**Field Name**
Date of Birth and/or age	EL_Age, EL_DOB
Date and time of collection of index blood culture	EL_DateTimeCollectedIndexBloodCulture
Date of Platform Entry	EL_ScreeningDate
Antimicrobial susceptibility of *Staphylococcus aureus* in the index blood culture isolate	EL_BB_BloodCulture
Lab number of the index blood culture isolate	BAS_IsolateLabNo
Vital status at Platform Day 90	D90_VitalStatus
Date of death	D90_DoD

9 DATA USER ACCESS

Access to the trial database is only granted to authorised trial staff. To request access, staff must complete a Database User Activation Form. The regional trial manager will confirm to the central management team that the database user activation form is complete, a CV and GCP have been filed, and that the staff member is appropriately delegated on the site delegation log. A list of all active and historical SNAP trial database users is available for the data coordinator and regional manager to download from the Spinnaker platform.

9.1 ACCESS TO BLINDED AND UNBLINDED DATA

Trial integrity demands clear demarcation of those with and without access to unblinded data (i.e., where allocation to treatments and trial endpoints are available). Before any domain or cell conclusions, the only groups with access to the relevant unblinded data will be the central data coordinators, the analytic team, and the data and safety monitoring committee. Following the public release of any domain or cell conclusions, any subsequent access to data (including that relevant to the publicly released results) must be carefully considered to ensure avoidance of unblinding of other domain or cell results.

Operational and interim reports for the TSC and other management groups will not include aggregated data on intervention allocations or the trial endpoints. Although response adaptive randomisation is not initially operationalised, it may be used in the future, and knowledge of intervention allocations would then potentially reveal the randomisation probabilities.

9.2 DATA ACCESS PROCEDURES AND ORGANISATIONAL RESPONSIBILITIES

Access to data varies for specific roles within the trial and safeguards are required to appropriately limit the access, amount, and type of data. Organisational responsibilities are detailed in **Appendix D**. and are summarised below.

Site investigators and coordinators will have access to granular data for their site (i.e., all variables for each participant, but not in an aggregated form). This data can be directly accessed through the eCRF front end within Spinnaker. Aggregated data involving intervention allocations and trial endpoints will not be available. Details about site personnel involved in data management (and the trial in general) can be found in the site, central, and regional Delegation Logs.

Monitors are granted temporary access to the database and access is restricted to the site/s they are currently monitoring. Monitors can make in-line data queries on the eCRFs through the Monitoring module in Spinnaker. For additional information, please refer to the regional SNAP Monitoring Plan.

Regional trial managers will have access to the individual participant site-level data, including screening and enrolment, and completeness of data entry for sites in their region. Enrolment and data completion reports will be available as downloadable csv files and reports through Spinnaker. Regional trial managers can make in-line data queries on the eCRFs. Aggregated data which includes both the intervention allocations as well as the trial endpoints will not be available.

Regional data coordinators will have access to individual participant and site-level data, including screening and enrolment and completeness of data entry for sites in their region. Enrolment and data completion reports will be available as downloadable csv files and reports through Spinnaker. Regional

data coordinators can make in-line data queries on the eCRFs. Aggregated data which includes both the intervention allocations as well as the trial endpoints will not be available.

Central trial management will have access to individual participant and site-level data, including screening and enrolment and completeness of data entry for all sites. Enrolment and data completion reports will be available as downloadable csv files and reports through Spinnaker. Central trial managers can make in-line data queries on the eCRFs. Aggregated data which includes both the intervention allocations as well as the trial endpoints will not be available.

The global trial data coordinators have the same access as Central Trial Management and the Analytic Team. Aggregated data which includes both the intervention allocations as well as the trial endpoints will strictly be used for trial management purposes, such as to assess data completeness, facilitate data cleaning and detect inconsistencies using programming code. Reports generated by these activities will not combine aggregated intervention allocations and trial endpoints.

The global trial manager leads the central trial management team and has oversight of the data management processes. Data management activities are delegated to the central data coordinators in this team.

The biobank coordinators will have access to the laboratory isolate report only; this will be available on Spinnaker and as a downloadable csv file. The biobank coordinators will not have access to patient CRFs or aggregated data which includes both intervention allocations as well as the trial endpoints.

The coordinating chief investigators will have access to an enrolment summary and completeness of data entry for all sites. These will be available as downloadable reports through Spinnaker. Aggregated data which includes both the intervention allocations as well as the trial endpoints will not be available.

The analytic team will have access to line listed data for all variables at all sites. These will be available as .csv files downloaded from Spinnaker. Analytic Team members will be provided with unblinded statistician access to the Spinnaker database, to download data exports only.

All other trial staff and investigators will not have access to eCRFs, downloadable csv files, or direct reports from Spinnaker.

9.3 OPERATIONAL REPORTS

Operational reports are useful for the day-to-day management of the trial, and include reports on participant screening, enrolment, data completion, timeliness of data completion, and data inconsistencies. These operational reports will be available to regional managers and data coordinators in Spinnaker and can be downloaded in csv format. High-level aggregated reports may be presented to site investigators, trial committees and management teams, and in public reports where appropriate. The reports may include numbers of participants screened, enrolled in registry or platform, and enrolled in each domain (and cell). The numbers of participants receiving different interventions will not be presented if response adaptive randomisation is introduced.

More granular reports will be available to regional and central trial management teams to assist in data checking, cleaning, and monitoring. **Examples are presented in Section 14. Timescale of Key Activities.**

10 DATA ANALYSIS AND REPORTING

10.1 SCHEDULED ANALYSES: DATA ACCESS, ANALYSIS AND REPORTING

Scheduled analyses will be performed on datasets which have been cleaned and soft locked (**See section 12.2 Data Locking Procedures**), and reports will be prepared by the Analytic Team. These analyses and reports will include unblinded data – i.e., inclusive of intervention allocations and primary, secondary, and safety endpoints. Access to this unblinded data will be strictly limited and the following procedures will be in place to ensure the security of access, transfer and storage of the relevant data, analysis files and reports:

1. The Terms of Reference for the Analytic Team and DSMC will include confidentiality clauses
2. The unblinded datasets will be accessed directly from Spinnaker by named members of the Analytic Team. While other user accounts in Spinnaker are managed by the trial management team, the Unblinded Statistician (UST) account may only be granted by Spiral on the authority of the Global Trial Manager.
3. All email correspondence that includes a discussion of and attachments of interim analyses will be titled 'CONFIDENTIAL SNAP TRIAL INTERIM RESULTS'
4. All files will be password protected

10.2 FINAL ANALYSES: DATA ACCESS, ANALYSIS AND REPORTING

Final analyses will be performed on datasets that have been cleaned and hard locked. Once a domain or cell conclusion has been reached and the GTSC accepts a DSMC recommendation to conclude the relevant domain or cell, data will be prepared for a final analysis. The Statistical Subcommittee and relevant Domain Specific Working Group will complete a Statistical Analysis Plan (SAP) for the relevant domain or cell. Once the SAP is finalised, it will be made publicly available.

The Analytic Team will perform the final analysis as guided by the SAP. The final hard locked dataset will serve as the trial record for the relevant domain or cell. Requests for access to data for this domain or cell will refer to this final hard locked dataset. As this final locked dataset may include unblinded data to other domains or cells for which a domain or cell conclusion has not yet been reached, access to this final locked dataset will remain with the Analytic Team until there is no risk of unblinding of other groups or individuals to these other domains or cells.

Note that the Domain_A_final_hardlocked_dataset captures a single point in time. The overall platform database will continue to accrue data. As participants may be enrolled in more than one domain, the data from each participant may also be updated as subsequent data checking and cleaning for other domains occur.

11 DATA EXTRACTION

11.1 DATA EXTRACTION

Before the unblinding of the analytic team, data will be extracted by a central data coordinator.

When the analytic team becomes unblinded, data will be extracted by a member of that team (or a central data coordinator) to a 1 TB University of Melbourne Research Data Share (Research-NAS) located at the following address: \\research-cifs.unimelb.edu.au\5050-snap. The storage infrastructure is physically located in the University's data centres at Parkville and Noble Park (VIC, Melbourne). Research-NAS is a standard network file share that is only accessible on the University network.

Access to the Research Data Share is provided by the Research Computing Activity Owner (RCAO) who is a University of Melbourne appointed member of the Analytic Team (currently Robert Mahar). Only members of the Analytic Team and the central DCs will have access the Research Data Share.

Statistical analyses will be performed using unblinded data extracted from the trial database to the Research Data Share (Research-NAS) and loaded temporarily into local memory for analysis using the chosen statistical software. At all times the raw data will only be permanently stored on the Research Data Share.

Statisticians on the SNAP Analytic Team will have an 'unblinded statistician' profile on the database which allows access for downloading de-identified data exports only. The Analytic Team will not be able to view any individual participant CRFs on the database.

As part of each analysis, the Analytic Team will run routine consistency checks on the trial data. The Analytic Team may also choose to extract data and perform dummy analyses (with dummy primary outcome data) before the formal interim analyses to develop and check code and ensure the necessary fields are available for the interim analyses. The Central Data Coordinator will work with the Analytic Team to resolve any concerns or issues with the data.

11.2 TIMING OF DATA EXTRACTION

11.2.1 DATA QUALITY CHECKS BEFORE ANALYSIS

Data quality checks and CRF locking will be an ongoing aspect of data management by the central and regional TMG.

Once the 500th patient (and in subsequent multiples of 500) is enrolled, sites will be notified that an interim analysis will take place after the 500th patient has reached the 90-day time point (i.e., approximately 3 months after the 500th patient was enrolled to the Platform). All sites will be asked to ensure data collection and data entry is as complete as possible for all participants up till that time point. Key analysis milestones are planned for every 2 weeks over a 6-week analysis period. Participants with missing outcome data will not be included in the relevant interim analysis.

On the agreed date of data extraction for analysis, the Analytic Team will check with the central TMG that the data is ready. When all parties have agreed that the data are sufficiently accurate and complete for analysis, and that no further querying or chasing of data is required, the Analytic Team will download the data exports from the Spinnaker database.

Please refer to the timescales in section **Error! Reference source not found.** and the Analysis section of the Organisational Responsibilities **Appendix D** for more details.

11.2.2 INTERIM ANALYSES

Interim analyses will be performed by the Analytic Team, who will prepare reports for the meetings of the DSMC. Details of the planned timings of these analyses can be found in the Statistical Appendix.

Before these analyses, the Analytic Team will agree in writing with the central TMG the exact date of data extraction. This will be planned to allow sufficient time for the central and regional trial management teams to be able to query data and request missing data from sites.

Specific timelines for data extraction and associated processes are in section 14 of this document. The *a priori* statistical instructions for each interim analysis will be documented in 'statistical implementation guides' (SIGs).

11.2.3 FINAL ANALYSIS

Final analysis of each domain/silo, and the platform trial, will be conducted on a final data extract after the data has been queried and verified as complete. Please refer to the relevant SAP for details.

11.3 DATA EXTRACTION FOR SUB STUDIES

Various sub-studies will be incorporated into the SNAP Trial throughout the lifetime of the study. Some sub-studies will use data already collected as part of the core trial, and/or collect additional sub-study-specific data in the Spinnaker database.

The global TSC is to be notified before any data extraction for sub-study analysis, and in general, this will not occur before publication of the main study results, with exceptions made by the GTSC on a case-by-case basis only if the publication will not affect trial integrity or constitute prior/duplicate publication. Higher degree by research (HDR) students may report sub-studies in an embargoes thesis chapter without meeting the above criteria, only if they can guarantee that the thesis chapter will not be accessible in the public domain. The central management team will maintain a log of all sub-studies, additional data collection, and what data is required from the Spinnaker database for each sub-study.

Where data extraction has been reviewed and approved by the global TSC, the central data coordinators will liaise with the sub-study lead to organise relevant data exports and access. The analytic team will be notified of any data extraction for sub-studies but will not be actively involved in the extraction or analysis.

Further information will be added when available, regarding the transfer of data to the sub-study lead(s), and logging of all data requests and transfers.

12 DATABASE LOCK AND ARCHIVE

12.1 DATA QUALITY CHECKS BEFORE LOCK

As a locked database can only be updated in extraordinary circumstances, comprehensive efforts need to be made to ensure that data are as clean as possible. The procedure for database lock will only be commenced when all the following data quality conditions have been met:

1. All eligibility screening data is 100% complete and correct (i.e. Spiral has actioned all requested corrections to these locked forms in the Spinnaker database), and
2. The priority variables in section **Error! Reference source not found.** have been validated, and
3. All data queries on priority variables have been satisfactorily resolved, and
4. All CRFs have been completed to the satisfaction of the Global Trial Manager.

12.2 LOCKING PROCEDURES

The trial database will be locked before the final analysis in accordance with the SNAP Trial Database Lock SOP. Data extracts and the Spinnaker interface will indicate if the records have been locked.

In general, data that is 'soft locked' can be unlocked by the regional managers and data coordinators to allow sites to make any changes required. Whereas data that is 'hard locked' is considered final and should not be unlocked or amended (unless required by extraordinary circumstances).

If a domain and silo conclusion is reached, further enrolment into the relevant domain and silo will be blocked in the database. Participants may continue to be enrolled in other domains and silos.

A Statistical Analysis Plan for the final analysis of that domain and silo will be written by the Statistical Analysis team. In parallel, final data cleaning and checking relevant to that domain and silo will occur. After the Statistical Analysis Plan is completed and signed off, a final data extract for that domain and silo will occur. The final data extract will serve as the 'locked' analysis dataset.

When the entire platform concludes, there will be a final lock and archiving of the entire database This final hard-locked dataset will have the following procedures in place to ensure the security of access, transfer, and storage of the relevant data, analysis files, and reports:

1. The Terms of Reference for the Analytic team will Ie confidentiality clauses
2. The unblinded datasets will be accessed from Spinnaker by named members of the Analytic team. While other user accounts in Spinnaker are managed by the trial management team, the Unblinded Statistician (UST) account may only be granted Spiral on the authority of the Global Trial Manager.
3. All email correspondence that includes a discussion of and attachments of final analyses will be titled 'CONFIDENTIAL SNAP TRIAL FINAL RESULTS'
4. All files will be password protected. Files and passwords will be supplied in separate communications.

After the study is completed and the final report has been prepared and approved, all relevant paper documentation including eCRFs and query forms will be archived in accordance with regional and local archiving procedures. Each site will be provided with electronic copies of the CRFs for each participant recruited at that site.

Once the data quality conditions have been met and this has been documented, the Analytic Team will request a database lock by completing a Database Lock Request form and submitting this to the GTM and Spiral, who will then arrange the lock as agreed in that document.

The Analytic Team and central TMG will work with the team from Spiral to arrange the archiving of the trial database and removal of data extracted to the Research Data Store when they have determined that accessing the database is no longer required for any future analyses.

13 DATABASE CHANGES

13.1 MAKING CHANGES TO EXISTING CRFS

Should CRFs need to be updated during the trial, e.g., due to a protocol amendment, this will require approval by the central TMG. The Data Coordinator and Global Trial Manager will be informed of any potential database changes as early as possible. This may involve consultation with the domain-specific working groups, or other relevant working groups.

Version updates will all be logged within the Spinnaker database and in the CRF completion guidelines. Changes may include (*not exhaustive*):

- Adding new questions into existing CRFs; these changes will be version controlled
- Amending a current question to collect additional variables; these changes will be version controlled
- Adding or amending subtext to clarify a question; these changes will not be version controlled
- Adding or editing help text ('I' icon) to guide sites in answering a question; these changes are logged (see **Error! Reference source not found.**) but will not be version controlled
- Changes to alerts or popup messages that appear when specific actions are taken; these changes will not be version controlled.

The change process is like that followed when adding new CRFs and is outlined below:

- TMG notifies Spiral Software Ltd, and the new question/data variable is entered into the Data Description by the DC and/or GTM
- Spiral Software Ltd assigns the change to a 'sprint' and amends the wireframe, which is reviewed by the TMG (DC and TMG with input from the lead CIs)
- Developer from Spiral Software Ltd builds the change into the SNAP trial staging database (https://snap-stage.spinnakersoftware.com/Login/)
- User acceptance testing (UAT) is performed by the DC, GTM, Regional Managers, and TMG using dummy data. Issues/bugs/feedback are raised through the Userback management system
- Updated CRF is released to the live database and is immediately available for data entry.
- Spiral Software Ltd email the GTM and DC confirmation of the sprint release, sprint report was emailed to TMG and filed in the sponsor file

13.2 HANDLING AND DISTRIBUTING NEW VERSIONS OF CRFS

New CRF versions will be released to sites within the database as per the procedure above. Sites will be informed (by email) that a change has been made to an existing CRF and warned of any implications for patients who are still within the 90-day follow-up period.

Every time a change is made to an existing CRF, the central TMG will decide if the change will be released to new participants only, or if this change will affect all or a subset of current trial participants.

The central TMG will confirm to Spiral Software Ltd if the change will be implemented in CRFs for:

- New enrolments only, or
- New enrolments and participants who are currently in the 90-day follow-up period, or
- All participants recruited to date

This will be recorded in the CRF completion guidelines and all active recruiting sites will be informed.

Versioning will be handled within the database as follows:

1. The Spinnaker database will tag each participant with identifiers at the time of randomisation and will be as per the release version.
2. A Log of releases and version 'tags' will be visible in the sub menu Manage > Releases.
3. Each release listed will contain the Date/Sprint name/Version number/Notes or Description columns
4. The date shown will be the date of release (NZ/AU) to the live database. The sprint names will be the same as what is listed in the Spiral management platforms (accessed by the Spiral team, DCs, and GTM: Monday/Userback/Jira).
5. The version number will be formatted as yy.nnn
6. All existing participants from the first release (scheduled Dec 2022) will be backdated and stamped with 22.001
7. Future sprints from that point forward will continue to receive the version numbers sequentially (22.002, 22.003…etc)
8. The Notes/Description column will mention short detail of what was included in the release, such as CRF name or category of tickets.
9. There will also be an "Export Results" button in the top right corner to export the onscreen list to a CSV showing the same data/values.

13.3 MAKING CHANGES TO THE DATABASE

Changes will be made to the trial database as appropriate in the following situations:

1. Following any relevant changes to trial eCRFs
2. In response to faults discovered in the database or occasions where the database is found not to match the data description, it is agreed that any features of the database are not fit for the purpose
3. To clarify the phrasing of a question or provide additional information to assist with interpretation and data collection. Other than amendments to the help text, all such changes must be actioned by Spiral Software Ltd and therefore are extensively documented. An up-to-date Help-Text Log is maintained in the central TMF.
4. In preparation for other database activities such as the development of exports for analysis.

All requests for changes, as well as any testing of those changes and related documentation, will be logged and stored in the Userback software.

Changes, in general, will be implemented without any downtime to the database. Once any changes that affect users have been implemented and 'released', the DCs and/or GTM will email a notification to all active sites and an updated version of the CRF completion guidelines will be distributed.

14 TIMESCALE OF KEY ACTIVITIES

Key data-related activities are summarised below along with target timescales and other requirements.

14.1 CENTRAL MANAGEMENT RESPONSIBILITIES

Access to the University of Melbourne Research Data Share (see section 6.0)	A response will be received by the applicant within two weeks of a written request to the Research Computing Activity Owner.
Review of the Data Management Plan	Annually, or more frequently as required (e.g. to reflect protocol updates).

14.2 SITE RESPONSIBILITIES

Completion of scheduled eCRFs other than Day 90	No more than 7 days after the CRF was due.
Completion of the Day 90 Vital Status (Primary endpoint)	No more than 10 days after the CRF was due.
Completion of the remainder of the Day 90 CRF	No later than platform Day 120
Completion of eCRFs with no specified date (e.g., discharge, protocol deviation)	No more than 7 days after the event.
Response to a query from monitor or TMG	Within 7 days of the query being raised

14.3 REQUESTS FOR DATA SUPPORT

Site user request for database support	Within 2 days of the request
Data corrections that are not able to be made by site teams	Response from central TMG within 7 days of the notification. Actioned by Spiral Software on a monthly cycle.

14.4 REPORTS

Reporting outside of the analysis cycle will be minimised to operational reports to ensure trial integrity. Examples are outlined below.

REPORT TYPE	TIMING
Simplified global CONSORT blinded to the intervention	Will be supplied to the SNAP global community every 2 weeks by the Global Trial Manager (or delegate).
Detailed global CONSORT blinded to the intervention	Supplied to the global Trial Steering Committee every 4 weeks.
Region-specific CONSORT blinded to the intervention	Supplied to the Regional Management Teams every 4 weeks.

Enrolment Report	Available as an on-screen and downloadable report via Spinnaker at any time to the Regional Monitors, Regional Managers and data coordinators (region-specific report).
Data outside expected ranges	Available via Spinnaker at any time to the Regional Monitors, Regional Managers and data coordinators (region-specific report). Available to the central Data Coordinators and Global Trial Manager (global trial report).
Laboratory samples report of specimens for centralised laboratory testing	Available via Spinnaker at any time to the following personnel: Central Data Coordinators and Global Trial Manager have access to a global trial report. Regional Managers, Data Coordinators, and Biobank Coordinators will access a region-specific report.
CRF completion status report of the number of CRFs that have been completed, partially completed, not commenced, and not yet due for each site.	Available via Spinnaker at any time to the following personnel: Chief Investigators, Global Trial Manager, and the Central Data Coordinators have access to a global trial report. Regional Managers and Data Coordinators will access a region-specific report.
List of all current and past account holders with access to the Spinnaker database.	For audit purposes, the Global Trial Manager and the central Data Coordinators have access via Spinnaker at any time.

14.5 ONGOING DATA VALIDATION

Data cleaning and consistency checks by the central and regional TMG	No later than Day 150, or 1 month after all CRFs including Day 90 have been completed (whichever is sooner). Data cleaning should be a continuous/ongoing process by the regional managers and data coordinators.

14.6 ANALYSIS PHASE

Data chase, consistency checks, and cleaning for every 500 participants enrolled in the platform, by the central and regional TMG	Within 2 weeks of the 500th participant reaching Day 90 in addition to other validations listed above

CRF locking, to prevent data edits	Done after data cleaning and consistency checks have been completed and no outstanding queries remain for each participant. Target is no later than Day 150 after platform entry for each participant, or 1 month after all CRFs including Day 90 have been completed (whichever is sooner). Where practicable, CRFs will be locked before data is transferred to the analytic team.
Data transfer to the analytic team	Within 2 weeks of the trigger for analysis being reached
Report completed by the analytic team	Within 4 weeks of the trigger for analysis being reached
Report assessed by the DSMC	Within 6 weeks of the trigger for analysis being reached

APPENDIX A: SUMMARY OF DATA FLOW

An overview of the data flow is given in the following diagram.

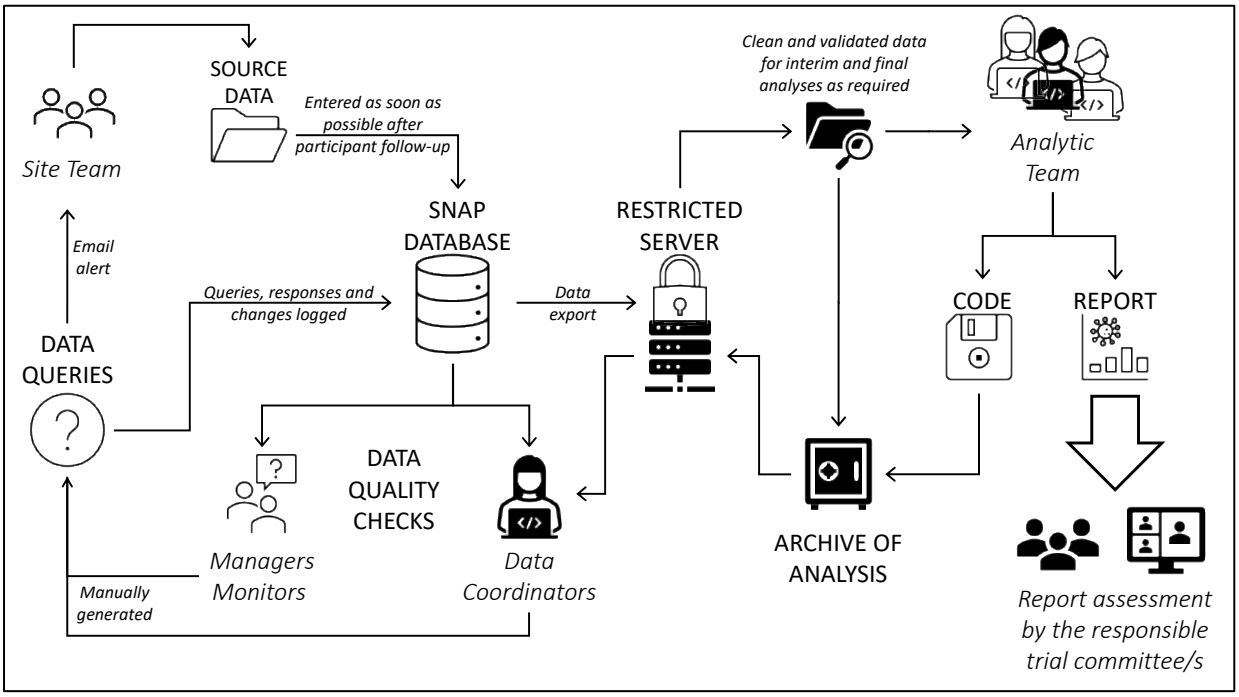

Figure 1: Overview of the processes in the data management for the SNAP Trial.

APPENDIX B: DATA THAT SUPPLEMENTS THE CRF

	Description
Help Text	Help Text is accessible to database users via a pop-up icon beside each question and provides additional guidance on the scope and interpretation. Help Text is created and edited directly within the live database and becomes available immediately on the eCRF. The Help Text log is a dated changelog recording additions and edits to the content.
Data Corrections	Data collected during platform and domain eligibility screening that cannot be edited by sites requires correction by Spiral. A summary of the correction is recorded in this log. File tag numbers link to an Evidence Archive of correspondence containing other details related to each correction.
Data Override	The Data Override is used sparingly to mark an e CRF as 'complete', e.g., when data is unavailable, and the incompleteness of a form prevents data collection from continuing. The log records a summary of the affected eCRFs and reasons for an override. File tag numbers link to an Evidence Archive of correspondence containing other details related to the override request.
Index Blood Culture	The Index Blood Culture Isolate lab number links each isolate with the SNAP Participant ID. Supplementary notes on IBC isolate numbers are recorded in this log to assist with Biobanking e.g., to note that the IBC was collected at another (non-SNAP) hospital to explain the different format of the isolate number.
Serious Adverse Reactions **Protocol Deviations**	These logs summarise the data recorded on paper SAR/PD CRFs and reported to the central team whilst the electronic reports in Spinnaker were under development.

APPENDIX C: TABLE OF CRFS AND FORM DATES

The ranges specified below are inclusive boundaries, i.e., Day 14 means 'up to and including events on day 14'. If no form date field name is present the date will be derived during analysis.

Module	CRF Name	Form date	Form date field name
Platform Eligibility	Add Patient (before enrolment) Randomisation (after enrolment)	Date of eligibility screening	EL_ScreeningDate
Domain Eligibility	Adjunctive Eligibility	Date of Adjunctive Domain screening result	DomainB_RevealTimeLocal DomainB_RevealTimeUTC
Domain Eligibility	Backbone Eligibility	Date of Backbone Domain screening result	DomainA_RevealTimeLocal DomainA_RevealTimeUTC
Domain Eligibility	EOS Eligibility Day 5-9	Date of EOS Domain screening result (day 7 ± 2 days)	DomainC_D7_RevealTimeLocal DomainC_D7_RevealTimeUTC
Domain Eligibility	EOS Eligibility Day 12-16	Date of EOS Domain screening result (day 14 ± 2 days)	DomainC_D14RevealTimeLocal DomainC_D14_RevealTimeUTC
Consent	Consent RC	Date and time of consent with regained capacity	CON_RCDateTime
Platform	Baseline	Date of Platform Entry	PT_RandomisationDateTimeLocal
Platform	Day 1-7	Date of Platform Day 7	
Platform	Day 8-14	Date of Platform Day 14	
Platform	Day 28 Follow up	Date of Platform Day 28	
Platform	Day 42 Follow up	Date of Platform Day 42	
Platform	Acute Discharge	Date of Acute Hospital Discharge	ADIS_DateOfDischargeAcuteBed
Platform	Total Discharge	Date of Total Hospital Discharge	TDIS_DischargeDate
Platform	Day 90 Follow up	Date of Platform Day 90	
Platform	Initial Focus of Infection (FOI)	Date of Platform Day 14 (infections recognised within the first 14 days post platform entry)	
Platform	FOI D15-ADIS	Commences: date of Platform Day 15. Ends: date of Acute Discharge	

Platform	FOI ADIS-D90	Commences: date of Acute Discharge. Ends: date of Day 90	
Platform	FOI D15-D90	Commences: date of Platform Day 15. Ends: date of Day 90 Applicable only to platform participants who were never discharged from acute care	
Registry	Registry CRF	Date of acute hospital admission	RBAS_AdmissionDate
Registry	Registry Discharge	Date of acute hospital discharge	RDIS_DateOfDischargeHospital
Registry	Registry Day 90	Date of day 90 (where index blood culture collection day is day 1)	
Registry	Registry Focus of Infection	Recognised within and up to 14 days from the time of index blood culture collection is day 1)	
Events	Serious Adverse Event (SAR)	Date of SAR	TBC when eCRF is available
Events	Protocol Deviation (PD)	Date of PD	PD_Date

APPENDIX D: ORGANISATIONAL RESPONSIBILITIES

Category	Data Management Responsibilities	Spiral Software Ltd.	Central trial management	Central data coordinator/s	Regional managers/ data coordinators	Global TSC	DSWG	Site clinicians and staff	Regional monitors	Statistical Subcommittee	Analytic Team
Documentation	Data Management Plan creation and updates		R	R						S	S
	Maintain delegation of responsibilities log and send copies to regional sponsors at regular intervals				R			R			
	CRF completion guidelines		R	R							
Access	Data extraction from the trial database			R							R
	User account and site management settings in Spinnaker	S	R	R							
	The University of Melbourne Research Data Share containing raw trial data		M*	S							R*
Database	Database randomisation algorithm	R	R							R	R
	Database: design of CRF, settings, content, user interface	R	R	R	S		S				
	Approve content of CRFs		R	R	R	R	R				
	Build, validate, and test the database before releasing update/s to the database	R	R	R	S						
	Updating the live database and maintaining data integrity through the trial life cycle	R	R	R							
	Manage database change process (e.g. new reports, corrections) during the life of the trial	R	R	R							S
	Data management support to sites and monitors	S	R	R	S						
Records	Timely completion and return of CRFs/data				R			R			

Category	Data Management Responsibilities	Spiral Software Ltd.	Central trial management	Central data coordinator/s	Regional managers/ data coordinators	Global TSC	DSWG	Site clinicians and staff	Regional monitors	Statistical Subcommittee	Analytic Team
	Enrolment of participants and data entry				S			R			
Records cont..	Data query generation				R				R		
	On-site monitoring (both preparing for and conducting site visits)		S		R			R	R		
	Patient record deletion	S	M*,R	R							
	Medical safety data review and MedDRA coding of safety events		R[Ŧ]		R			S			
Validations	Central monitoring		S	R	R			R			
	Timely data query resolution		S		S			R	R		
	Data chase, consistency checks, and cleaning		R	R	S				S		S
	Data correction, editable fields		S	S				R			
	Data correction, fields that are locked	R	S	R				S			
Locking	Locking CRFs (independent of participant lock)		M*, R	R	R				S		
	Locking participants (also locks all CRFs)		M*, R	R	R						
	Locking sites		M*								
Analysis	Data transfer to analysts	S	M*	R							S
	SNAP Trial Platform and Domains: Data analysis, preparation, and circulation of the DSMC report		S			A*					R
	SNAP Trial Sub-studies: Data analysis, preparation, and circulation of the report		S			A*	R				

Key:
R, Primary Responsibility. **R***, Access is granted by the Research Computing Activity Owner who is a University of Melbourne appointed member of the Analytic Team (currently Robert Mahar) after written approval from the Global Trial Manager. **S**, Support or Feedback. **M***, Global Trial Manager approval must be sought before this can occur. Documented in the eTMF. **R[Ŧ]**, Delegated to a rota of Trial Clinicians appointed by the Global Trial Manager. **A***, Approval must be granted before the event can occur. Documented in the eTMF.

APPENDIX E: ACCESS TO BLINDED AND UNBLINDED DATA

Unless otherwise stated, Blinded means unable to view or access aggregated participant allocations and aggregated outcome data.

Central coordinating team

Role	Spinnaker eCRF	Raw Data extracts	Interim analysis reports
Chief investigators	No access apart from to own site if they are also site PIs	No access	Blinded
Global trial manager	Read-only access to one participant unblinded record at a time (all regions); grants unblinded access to aggregated data	Blinded	Blinded
Central trial management team	Read-only access to one participant unblinded record at a time (in the AUS region)	No access	Blinded
Central trial data coordinator(s)	Read-only access to one participant unblinded record at a time (all regions)	Unblinded	Blinded

Trial management committees

Committee	Spinnaker eCRF	Raw data extracts	Interim analysis reports
Global Trial Steering Committee (GTSC)	No access	No access	Blinded
Domain-specific working groups (DSWG)	No access	No access	Blinded
Other committees (Paediatrics, PWID, Clinical Pharmacology, Microbiology, Registry)	No access	No access	Blinded
Statistical Analysis Subcommittee (SAS)	No access	No access	Blinded

Analytic Team Subcommittee	Access to download raw data extracts only	Unblinded	Unblinded

Regional personnel

Role	Spinnaker eCRF	Spinnaker extracts	Interim reports
Members of Regional Committees (RTSC)	No access	No access	Blinded
Regional Data Coordinators	Read-only access to one participant's unblinded record at a time (in their region)	No access	Blinded
Regional Managers	Read-only access to one participant's unblinded record at a time within their region only	No access	Blinded
Site staff (site principal investigators, study coordinators, nurses)	Unblinded to own site records only	No access	Blinded (simplified consort only)
Monitor(s)	Read-only access to one participant's unblinded record at a time, at a site that is monitored	No access	Blinded

APPENDIX F: DATA PROTECTION

IDENTIFIABLE DATA

The following patient identifiable data may be collected in the Spinnaker database during platform screening:

- Date of Birth & Initials
- Age & Initials
- Date of Birth only
- Age Only

However, the type of identifiable data collected from each patient will depend on site/regional regulations and approvals. Please refer to regional-specific appendices for further information for each country.

Identifiable data collected during screening is only visible to the local site team and regional management teams who are assigned to the site. All identifiable information (Initials, DOB) is encrypted in the raw data exports, and will not be accessible or visible to the unblinded study team.

PERSONAL DATA

Personal data is any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data.

According to General Data Protection Regulation (GDPR), personal data means:

"any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."

Any Personal Data collected or used in the course of the trial will be treated as Confidential Information at all times and will be stored securely with all security measures that would be necessary for compliance with Data Protection and GDPR.

SPINNAKER DATA SECURITY

Spinnaker Data are backed up to a server within the Amazon Web Service (AWS) zone and on S3 encrypted server according to the AWS Standard Security Protocol. Backups occur every hour and every 6 hours onsite and offsite respectively. AWS firewalls both internally and externally are in place for intrusion protection.

Spinnaker software platform ensures encryption of personal data using Advanced Encryption Standards (AES) for participant confidentiality and integrity. A data logging system is used to track user access to personal data included any changes, deletions, updates made by the user. For further information on Data Protection for the Spinnaker Platform and Spiral please contact Spiral at audrey@spiral.co.nz.

APPENDIX G: RANDOMISATION/REGISTRATION

Randomisation is performed within the SNAP trial database (Spinnaker).

RANDOMISATION RATIOS

Randomisation ratios will be balanced across all domains when the study opens for enrolments. At some point in the future, the Trial Steering Committee may elect to utilize imbalanced ratios, at which time the Analytic Team will provide weighted ratios which will be used by Spiral to adapt the randomisation ratios in the platform.

SNAP TRIAL ID/STUDY NUMBER/PARTICIPANT ID

During platform screening, each participant is given a 6-letter alphabetic ID. After platform eligibility screening has been completed and the participant is randomised, a new eight-digit SNAP Trial ID is assigned. The screening ID remains visible alongside the SNAP Trial ID when viewing platform screening data in Spinnaker or when viewing the eligibility data extracts.

The SNAP Trial ID comprises three letters designating the site code of the enrolling site, and the next five are allocated sequentially per site, in the order of enrolment into SNAP. Registry-only participants are intercalated within this sequence and are designated with the prefix "R-" before the site code. Thus, a sequential list of participants at a site may be as follows:

SIT00001, SIT00002, R-SIT00003, SIT00004, R-SIT00005, …

whereby the first two participants were randomised to the platform, the next was enrolled as a registry-only participant, the fourth was a platform participant, the fifth was enrolled as a registry-only participant, and so on.

The site codes used are listed in the international site tracker, available to all regional trial managers, and are assigned by the DC. If a site not on this list becomes accredited to recruit participants to the trial, the site must first be added to the tracker. All site codes are unique and reflect the name of the site in some manner.

RANDOMISATION AND REVEAL PROCEDURES

The participant's randomised allocations are only revealed for each domain if they meet all of the domain-specific inclusion criteria and none of the domain-specific exclusion criteria.

The timing of the reveal for each domain can be found in the patient timeline (Core Protocol).