SNAP TRIAL INTEGRITY SUITE: DATA MANAGEMENT PLAN

DOCUMENT APPROVAL

VERSION NUMBER	DATE	CREATED BY	APPROVED BY	SIGNATURE OF APPROVAL
2.1	05 Jun 2026	GTMG	Steven Tong Joshua Davis Zoe McQuilten	*N/A – administrative changes only as documented in Appendix 1.*

Members of the SNAP Trial Global Trial Steering Committee (GTSC), Statistical Subcommittee (Stats-WG), Analytic Team (AT) and Regional Trial Management Groups (RTMGs) have had the opportunity to review this document, and it has been signed by the Chief Investigators and the Chair of the GTSC on behalf of the SNAP Trial.

IF ADAPTING THIS DOCUMENT, PLEASE ACKNOWLEDGE US VIA THE FOLLOWING STATEMENT:

"This document has been adapted from the *Staphylococcus Aureus* Network Adaptive Platform (SNAP) Trial Integrity Suite: Data Management Plan (Version 2.1), available from the SNAP Trial website: https://www.snaptrial.com.au/for-investigators#integrity.

The SNAP trial, globally coordinated by the University of Melbourne, is an international adaptive platform trial aimed at identifying the most effective treatments for *Staphylococcus aureus* bloodstream infections."

1 TABLE OF CONTENTS

1.1 TABLE OF FIGURES

1.2 TABLE OF TABLES

2 SNAP TRIAL INTEGRITY SUITE

The SNAP Trial Integrity Suite is a comprehensive collection of essential documents that provide evidence and guidance for the conduct and oversight of the SNAP Trial globally and sets the minimum standards of compliance for the study.

The SNAP Trial Integrity Suite is comprised of the following documents:

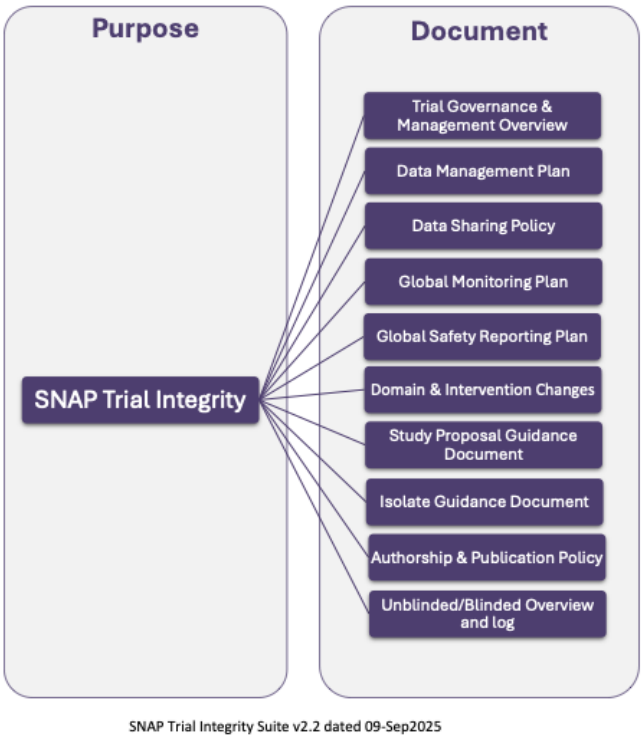

Figure 1 Trial Integrity Suite

3 INTRODUCTION

The SNAP Trial Data Management Plan (DMP) outlines the overarching processes and procedures that will be followed to ensure the accuracy, integrity, and security of data collected during the lifetime of the SNAP Trial. Adherence to this plan is essential for maintaining compliance with applicable regulations and guidelines, including Good Clinical Practice (GCP) and regional regulatory guidelines.

By following the protocols outlined in this document, we aim to minimise the risk of data errors, omissions, or breaches, and to facilitate efficient and reliable data analysis and reporting.

This document specifically outlines the processes and procedures for the following aspects throughout the life of this trial:

- Handling, storing, and analysing data
- Managing access, sharing and unblinding of the data
- Designing and implementing the data collection tools
- Modifying and maintaining the Electronic Data Capture (EDC) systems
- Data lock, archiving and data retention

It will cover all aspects of data handling, including initial EDC development, data entry and validation to data storage, transfer, and archiving.

3.1 PURPOSE & SCOPE

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 document covers the data management and data sharing of the trial as a whole, including the various Electronic Data Capture (EDC) systems that are utilised within the trial. For more specific information on the EDCs used within the SNAP Trial, please refer to the EDC-specific documentation.

This document does not cover in detail any procedures relating to the handling and processing of Safety Events and On-Site Monitoring Activities which are documented in the Safety Reporting Plan and Monitoring Plan respectively.

This document should be read in conjunction with the **Global Safety Reporting Plan**, **Global Monitoring Plan**, **Data Sharing Policy** and **EDC-specific documentation.**

3.2 UPDATE & REVIEW

This document will be updated throughout the life of the trial to reflect any changes in data management procedures.

This document will be reviewed annually at a minimum, or in the event of the following occurring:

- A new domain or intervention is added to the trial
- A substantial amendment to the core protocol documents is implemented that impacts this document
- Changes to data management working practices occur

With each update, all members of the SNAP Trial Global Trial Steering Committee (GTSC), Regional Trial Management Group (RTMG), Statistical Subcommittee (Stats-WG) and Analytic Team (AT) will have the opportunity to review and provide feedback on the changes.

Each version of the Data Management Plan will be ratified by the Global Trial Steering Committee (GTSC) and signed off by the chair on behalf of the committee.

3.3 TRIAL INFORMATION

An overview of the current project status is below:

Table 1 Trial Status

Trial Status	
Pre-Trial Setup Period	January 2020 - February 2022
Current Recruitment Status	Open to Recruitment
First Participant Recruited to the Trial	17 February 2022
Last Participant Recruited to the Trial	*Recruitment ongoing*
Enrolment Target: Platform	8,000 participants
Enrolment Target: Registry	20,000 participants
Follow-Up Period: Platform & Registry	90 days

Table 2 Trial Domains and Interventions

Domain	Interventions		Domain Status			
	Silo	Treatment	Population	Date of Go-Live in EDC[a]	Status (active / not active / closed)	Date of Closure in EDC[b]
Antibiotic Backbone Domain (A)	PSSA	(Flu)cloxacillin	Adults	16 Feb 22	CLOSED	21-Jun-24
			Paediatrics		ACTIVE	
		Penicillin	Adults	16 Feb 22	CLOSED	21-Jun-24
			Paediatrics			
		Cefazolin	Adults	N/A	-	
			Paediatrics	11 Mar 25	ACTIVE	
	MSSA	(Flu)cloxacillin	Adults	16 Feb 22	CLOSED	21-Jun-24
			Paediatrics		ACTIVE	
		Cefazolin	Adults	16 Feb 22	CLOSED	21-Jun-24
			Paediatrics		ACTIVE	
	MRSA	Vancomycin / Daptomycin (clinician choice)	Adults	16 Feb 22	ACTIVE	
			Paediatrics			
		Vancomycin / Daptomycin (clinician choice) plus cefazolin	Adults	16 Feb 22	ACTIVE	
			Paediatrics			
Adjunctive Domain (B)	-	No Clindamycin	Adults	16 Feb 22	CLOSED	17-Jul-25
			Paediatrics			
	-	Adjunctive Clindamycin	Adults	16 Feb 22	CLOSED	17-Jul-25
			Paediatrics			
Early Oral Switch Domain (C)	-	Continued IV	Adults	16 Feb 22	ACTIVE	
			Paediatrics		CLOSED	11-Mar-25
	-	Early Oral Switch	Adults	16 Feb 22	ACTIVE	
			Paediatrics		CLOSED	11-Mar-25
PET/CT Domain (D)	-	No PET/CT scan	Adults	15 Aug 24	ACTIVE	
			Paediatrics	N/A	-	
	-	PET/CT scan	Adults	15 Aug 24	ACTIVE	
			Paediatrics	N/A	-	

[a] Date of Go-Live = date the domain became available for recruitment in the EDC

[b] Date of Closure = the date that recruitment to the domain was no longer available in the EDC

4 GLOSSARY OF TERMS & ABBREVIATIONS

Table 3 Glossary of Terms & Abbreviations

Term	Definition
AT	Analytic Team
Confidential Information	All unpatented inventions, ideas, know-how, concepts, trade secrets, processes, techniques, software, products and any and all other unregistered or unpatented intellectual property, financial and business information and all other commercially valuable information of a party which that party regards as confidential to it or which is evident by its nature or the manner of its disclosure to be confidential.
CRF (eCRF, pCRF)	Case Report Form (electronic Case Report Form, paper Case Report Form) A printed, optical, or electronic document designed to record all of the protocol-required information to be reported to the sponsor on each trial subject. Case Report Form
CSV file	Comma Separated Values file A plain text file that stores tabular data (numbers and text) in a simple, structured format. Each line of the file typically represents a single record or row, with individual fields or columns separated by commas.
Data Extraction (Data Cut)	A data extraction (data cut) refers to a specific extraction event of the entire participant data set from the relevant EDCs that is not modified/updated. It will serve as a snapshot in time and, therefore, it is expected that some participant data may be incomplete or inaccurate, due to the ongoing nature of the adaptive platform trial.
DD	Data Dictionary A document that outlines the structure, content, and variable definitions for a dataset or collection of data. A data dictionary is a critical tool for reproducibility because it allows others to understand your data; it defines and describes the elements of a dataset so that it can be understood and used at a later date.
DSA	Domain-Specific Appendix
DSMC	Data and Safety Monitoring Committee
DSWG	Domain-Specific Working Group
DTA	Data Transfer Agreement
EDC / EDC System	Electronic Data Capture System / Database
EDC Provider	A 3rd-party provider who is contracted to design, develop, maintain and operate an EDC system.
EDC-RP	EDC-Responsible Party The party who contracts the EDC Provider and oversees the design, development, maintenance and operation of the EDC system, and not necessarily the data custodian.
GCC	Global Coordinating Centre
GDPR	General Data Protection Regulation

Term	Definition
GTMG	Global Trial Management Group
GTSC	Global Trial Steering Committee
ICH-GCP Guidelines	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use Guideline for Good Clinical Practice Good Clinical Practice (GCP) is an international, ethical, scientific and quality standard for the conduct of trials that involve human participants. Clinical trials conducted in accordance with this standard will help to assure that the rights, safety and well-being of trial participants are protected; that the conduct is consistent with the principles that have their origin in the Declaration of Helsinki; and that the clinical trial results are reliable.
ISO	International Sponsor Organisation
PD	Protocol Deviation
Personal Data	Under ICH-GCP, personal data encompasses any information that can be directly or indirectly linked to an individual participant in a clinical trial. This includes, but is not limited to, names, addresses, medical records, and other identifiers. Under GDPR, personal data is 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.
PII	Participant-Identifying Information A subset of personal data that refers to any data or combination of data elements that can be used to directly identify, contact, or locate a specific individual.
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
Nested Trial	Includes new domains, new interventions within a domain, and other clinical trials that sit within the SNAP Platform or Registry.
Pseudonymised Data	Pseudonymised data refers to personal data that has been processed in such a way that it can no longer be attributed to a specific individual without the use of additional information. This additional information is kept separately and is subject to technical and organisational measures to ensure non-attribution to an identified or identifiable person. This is sometimes referred to as "de-identified data", but pseudonymised data will be used throughout this document for clarity and consistency.
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.

Term	Definition
RTMG	Regional Trial Management Group
RTSC	Regional Trial Steering Committee
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan A detailed, pre-specified document that outlines the statistical methods, procedures, and rules for analysing clinical trial data to ensure unbiased, transparent, and reproducible results that support the trial's objectives.
SAR	Serious Adverse Reaction
Screening ID	A randomly generated 6-letter alphabetic ID that is assigned to each participant during SNAP Trial screening. This identification number can be used to identify each unique participant, regardless of their inclusion or exclusion into the trial.
SIG	Statistical Implementation Guide The current state of the analytic plan based on any previous platform adaptations.
SNAP	*Staphylococcus aureus* Network Adaptive Platform trial
SNAP	*Staphylococcus aureus* Network Adaptive Platform trial
SOP	Standard Operating Procedure
Stats-WG	Statistical Subcommittee
TMF	Trial Master File A comprehensive collection of essential documents that allows for the evaluation of the conduct of a clinical trial and the quality of the data produced. These documents serve to demonstrate compliance with the standards of Good Clinical Practice and all applicable regulatory requirements.
ToR	Terms of Reference
Trial ID	An 8- or 9-digit number assigned to a participant that is assigned to each participant that is eligible for either the SNAP platform or SNAP registry.
UAT	User Acceptance Testing A critical phase in the software development lifecycle where the end-users or clients test the software to verify that it can handle required tasks in real-world scenarios, according to specifications.

5 DATA MANAGEMENT RESPONSIBILITIES & ACTIVITIES

The specific responsibilities of each party are outlined in each corresponding agreement, Terms of Reference (ToR) or Statement of Work (SoW) documents.

Data management activities that relate to on-site monitoring or safety monitoring will be detailed in the respective global and regional monitoring and safety reporting plans.

- **Appendix 2** provides a summary of all parties' responsibilities with regards to data management.
- **Appendix 3** provides a summary of specific data management activities, the primary and personnel responsible for each and, where applicable, a timeline for each.

5.1 TRIAL MANAGEMENT GROUPS (TMGS)

The specific responsibilities of the Global and Regional Trial Management Groups (GTMGs & RTMGs) are outlined in detail in the TMG Terms of Reference and in each corresponding sponsor agreement.

- Only designated data coordinator(s) from the GTMG will have access to the unblinded data for the purposes of conducting cleaning and validation checks, and for liaising with the Analytic Team (AT) as required. This will be documented clearly by the GTMG.
- Each RTMG performing data linkage may have a designated data coordinator(s) who has access to participant personal data for the specific purpose of performing data linkage. This will be documented clearly by the RTMG

Please see **SNAP Governance Overview Document** for more information on the overarching trial governance and management structures.

5.2 COMMITTEE & WORKING GROUP RESPONSIBILITIES

The specific responsibilities of the Global Trial Steering Committee (GTSC), Data and Safety Monitoring Committee (DSMC), Statistical Subcommittee (Stats-WG), Analytic Team (AT), and other working groups (WGs) are outlined in detail in the committee or working group's Terms of Reference (ToR).

5.3 SITE RESPONSIBILITIES

The specific responsibilities of each site are outlined in the site agreements and investigator agreements.

5.4 TRIAL-INTEGRATED STUDY PROPOSAL LEAD(S) RESPONSIBILITIES

The specific responsibilities of each nested trial / sub-study are outlined in the agreements that are signed prior to commencement of the proposed study.

Please see **Study Proposal Guidance Document** for more information on the different kinds of integrated studies within the SNAP Trial.

5.5 EDC PROVIDER(S)

The EDC-Provider is a 3rd-party provider who is contracted to design, develop, maintain and operate an EDC system.

The specific responsibilities of the EDC provider(s) are outlined in the service agreement, license terms, Statement of Work (SoW) or any other similar contractual agreement between the EDC provider and the EDC-RP.

See **Table 4** to view EDCs used within the SNAP Trial and their respective EDC-Providers and EDC-RPs.

5.6 EDC RESPONSIBLE PARTY(S) (EDC-RP(S))

The EDC Responsible Party (EDC-RP) is the party who contracts the EDC Provider and oversees the design, development, maintenance and operation of the EDC system, and not necessarily the data custodian.

See **Table 4** to view EDCs used within the SNAP Trial and their respective EDC-Providers and EDC-RPs.

5.7 DATA REQUESTER(S) RESPONSIBILITIES

The specific responsibilities of data requesters are outlined in the Data Transfer Agreements (DTAs) that are signed prior to data transfer and summarised in the **Data Sharing Policy.**

6 DATA MANAGEMENT STRATEGY

The data management strategy for this clinical trial is designed to ensure the accurate, efficient, and secure collection, monitoring, processing, analysis, and storage of data throughout the data lifecycle. The strategy will focus on maintaining data quality, integrity, and compliance with regulatory requirements while supporting the overall objectives of the trial.

The Data Management Plan covers the following sections:

- EDC System Design & Implementation – **Section 7**
- Randomisation Design & Validation – **Section 8**
- Data Entry, Handling & Extraction Procedures – **Section 9**
- Data Quality Control & Monitoring Procedures – **Section 10**
- Data Privacy & Confidentiality – **Section 11**
- Managing Blinding to Data & Analysis Results in the Ongoing Trial – **Section 12**
- Statistical Analyses & Reporting – **Section 13**
- Data Sharing – **Section 14 (and Data Sharing Policy)**
- Data Lock, Archiving & Retention – **Section 15**
- Data Storage & Security – **Section 16**

Please review each relevant section below for more detail.

7 ELECTRONIC DATA CAPTURE (EDC) SYSTEM DESIGN & IMPLEMENTATION

Data collection will be carried out using Electronic Data Capture (EDC) systems that are compliant with **Section 7** of this document to ensure consistency and accuracy.

7.1 EDC SYSTEMS UTILISED WITHIN THE SNAP TRIAL

The following Electronic Data Capture (EDC) systems are implemented within the trial. For more specific information on the EDCs used within the SNAP Trial, please refer to the **EDC-specific documentation**.

Table 4 EDC Systems Used in the SNAP Trial

EDC System	EDC-RP	EDC-Provider	Location of EDC-Specific Documentation
Central Spinnaker EDC	GTMG	Spiral Software (https://spiral.co.nz/)	GTMG TMF
Central Biobank EDC	GTMG	TBC	GTMG TMF
Data Linkage EDCs	RTMGs	Varies depending on region	RTMG TMF
Adjunctive EDCs *Only where the EDC will be used in more than 1 participating SNAP region and/or the data analyses will utilise the core statistical model*			
DAPTO-SNAP EDC	Lead – Todd Lee	REDCap (https://project-redcap.org/)	GTMG TMF
EDCs used in only 1 region where data analysis will not utilise the core statistical model	RTMGs	Varies	RTMG TMF

7.1.1 CENTRAL SPINNAKER EDC

The Central Spinnaker EDC will act as the main data collection tool for the SNAP Trial, and will include, at a minimum:

- Core protocol screening, eligibility, randomisation and core data collection
- Domain-specific screening, eligibility, and randomisation

This EDC may include domain-specific data collection, and sub-study / addendum screening, eligibility, randomisation, and data collection, if selected for use by the lead.

7.1.2 CENTRAL ISOLATE BIOBANK EDC

For regions that are participating in the central (global) biobank, the Central Isolate Biobank EDC will act as the main data collection tool for the SNAP Trial isolate-associated data, including genomic and phenotypic data for the purposes of the biobank.

This EDC will not contain any participant data other than the SNAP Screening ID and Laboratory ID (see **Section 11.3** for more information).

7.1.3 DATA LINKAGE EDC (IF RELEVANT IN THE REGION)

The data linkage EDC will act as the region-specific data collection tool for collecting Participant Identifying Information (PII; see **Section 11.2.1**) for the purpose of data collection, in regions approved to conduct this activity.

This EDC may include collection of participant's names, date of birth, health care number and other identifiers that are important in conducting data linkage activities (See **Section 11.2.3** for more information).

7.1.4 ADJUNCTIVE EDC(S)

Other adjunctive EDCs may be implemented for data collection within the SNAP Trial. Data collected in an adjunctive EDC may include:

- Nested Trial / Sub-Study screening, eligibility, randomisation (if applicable) and/or data collection

- Other region-specific data collection requirements

The GTMG will be made aware of and supplied with the EDC-specific documentation for any Adjunctive EDCs used across more than 1 region and/or where the analyses of the data within the adjunctive EDCs will utilise the core statistical model (See **Statistics Appendix** for more information).

The SNAP Participant ID or Screening ID assigned to the participant in the Central Spinnaker EDC (**see Section 11.3**), will be used to link participant records across EDCs, and as such, should be collected in each EDC used within the SNAP Trial.

7.2 GENERAL PRINCIPLES FOR EDC DESIGN

The EDCs used in the SNAP Trial will be designed for ease of use, incorporating built-in validation checks to minimise data entry errors and ensure completeness.

7.2.1 PRAGMATIC DATA COLLECTION

As a pragmatic trial, data collection will be simplified. The required data points will be designed to be easily entered by the sites, using information available from participants' medical records or will be routine data collected as part of their standard care, wherever possible.

7.2.2 DISCOURAGEMENT OF DATA DUPLICATION

In principle, duplicative data collection is discouraged; if a datapoint is already collected, it should not be re-collected, unless explicitly required*.

In principle, if a datapoint is already collected in the Central Spinnaker EDC, it will take precedence in duplicative data collection.

If duplicative data collection is required, justification will be provided to the GTMG for consideration. If data duplication is approved, clear rules will be implemented by the EDC-RP(s) and outlined in the Statistical Analysis Plan (SAP) or EDC-Specific Documentation to describe how the duplicated data will be handled during data cleaning and quality control procedures.

The SNAP Participant ID or Screening ID must be collected in each EDC for linking participant records; this is an exception to the rule (See **Section 11.3 for more information).*

7.2.3 STANDARDISATION OF DATA COLLECTION

As data may be collected in more than one EDC, it is vital that timepoints and terminology is consistent across the trial EDCs.

Timepoints for data collection should be harmonised with the core protocol and/or existing domains and EDCs wherever possible to reduce burden on site staff and for comparability when analysing the data.

Terminology, abbreviations, and definitions used in the EDCs should be consistent with the core protocol and/or existing domains. If this is not possible and a new term is introduced, it should be clearly defined in the relevant protocol, EDC, and EDC-specific documentation.

7.3 REQUIRED FEATURES OF EDCs

Each EDC used in the SNAP Trial will include the following features that will be documented by the EDC-RP in a validation statement and/or in the EDC-specific documentation that is stored in the TMF as per local and regional guidelines.

7.3.1 UNIQUE USER LOGIN

Unique user logins will be required for each individual who accesses the EDC system so there is a clear audit trail of changes. It should be clear when a user is granted access, and when their access has been revoked.

7.3.2 USER APPROVAL/REMOVAL PROCESS

A formal request and approval process will be implemented for users to gain access to the EDC in the SNAP Trial. All trial staff involved in data entry should be appropriately trained, delegated, qualified and approved by the EDC-RP prior to undertaking data entry.

- Trial staff will receive training on the proper data entry procedures, ensuring uniformity in data collection, handling and minimising potential for human error.
- Trial staff should be appropriately delegated the duty of data entry by the site PI
- Trial staff should be appropriately qualified to access the required participant data and enter it into the EDC

A formal removal process will be implemented for users to have their access to the EDC removed when the user is no longer involved in the SNAP Trial and/or require access to the EDC.

7.3.3 ROBUST PASSWORD POLICY

EDCs will implement a robust password policy that will be documented in the validation statement (See **Section 7.5.1**) and made available upon request. Where available, Multi-Factor Authentication (MFA) methods should be implemented.

7.3.4 ACCESSIBLE AUDIT LOGS

To provide transparency and ensure all data modifications are traceable, EDCs will include a human-readable audit log to track all changes to the EDC, including but not limited to changes made to participant data as well as EDC settings, user settings, etc.

The audit log will include the identity of the person making the change, what was modified, and when the change occurred. If a change is unable to be tracked within the EDC, the change should be tracked by the EDC-RP with the evidence supporting the change stored for complete documentation.

If the EDC cannot track all changes (for example, changes to user settings, site settings, question text, etc.), this will be documented, and the EDC-Responsible Party (EDC-RP) will be responsible for ensuring that detailed logs are maintained for these changes.

7.3.5 DATA ENTRY GUIDANCE DOCUMENTS

The EDC-RP will be responsible for ensuring that a detailed user guide is created for accessing the EDC, and for correctly entering data within the EDC system. This document will be updated as required, to ensure that users are provided with the most current information for using the EDC.

7.3.6 DATA DICTIONARY (DD)

A Data Dictionary (DD) is a document that outlines the structure, content, and variable definitions for a dataset or collection of data. A DD is a critical tool for reproducibility because it allows others to understand your data; it defines and describes the elements of a dataset so that it can be understood and used at a later date.

Each EDC utilised within the SNAP Trial should have a maintained DD that is reflective of the EDC at the current moment in time.

7.4 MAKING CHANGES TO THE EDCS

Changes to the EDCs may occur in various situations, including but not limited to:

- Introduction of new aspects of the trial (e.g., new domains or protocol updates)
- Identification of bugs or faults in the EDC functionality
- Determination that a feature of the EDC is not meeting its intended purpose

For each change, a decision will be made by the EDC-RP, in consultation with the EDC Provider, regarding its scope:

- **New participants only:** Changes will apply only to participants enrolled after the update is implemented.
- **Subset of participants:** Changes may apply to specific groups (e.g., certain domains, paediatric participants, participants from a particular region).
- **Protocol version-specific:** Changes may apply only to participants enrolled under a specific protocol version (e.g., protocol version 2.0).
- **All participants:** Changes will apply to the entire EDC system, affecting all completed CRFs and follow-up data.

7.5 REQUIRED PROCESSES FOR INITIAL RELEASE AND CHANGES TO THE EDCS

Each EDC used within the SNAP Trial will undertake and implement the following processes when preparing for initial release, or when making changes. This process should be documented and signed off by the EDC-RP and stored in the TMF according to local and regional guidelines.

Where the EDC is used in more than 1 region and/or will make use of the core statistical model, the signed documentation will be provided by the EDC-RP to the GTMG for confirmation, and a copy of the documentation filed in the GTMG Trial Master File (TMF) prior to implementation and made available to participating regions.

7.5.1 EDC TESTING & COMPLETION OF VALIDATION STATEMENT

The EDC should be well-tested prior to release or update. A validation statement should be completed, documenting the testing process that occurred.

This statement should include at a minimum:

- Description of and validation of the randomisation algorithm (See **Section 8** below)
- Description of the password policy (or reference to the document that details this)
- Confirmation that User Acceptance Testing (UAT) was completed and by whom
- Confirmation that user profiles are specified and the permissions are described
- Confirmation that data fields adequately collect the protocol or appendix outcomes
- Confirmation that show/hide logic implemented is working as expected
- Confirmation that the EDC aligns with the DD specifications including:
    - Data fields and field names
    - Built-in verifications for data entry
    - on-screen alerts (as relevant)
    - email alerts (as relevant)
    - Show/Hide logic

- Data exports match on-screen fields
- Confirmation that the EDC complies with ICH-GCP

A **Template Validation Statement** is included in **Appendix 5** and can be used by the EDC-RP to document the validation process.

7.5.2 DESCRIPTION & VALIDATION OF RANDOMISATION ALGORITHM

The method by which randomisation algorithm was generated and the justification for the method selection should be included in the initial EDC Validation Statement.

The randomisation algorithm should be validated prior to initial go-live, and upon each change to the EDC that affects the availability of domains and/or interventions (i.e. affects the randomisation schedule).

See **Section 8.3** for more information on validating the randomisation algorithm.

7.5.3 CREATING AND UPDATING DATA ENTRY GUIDANCE DOCUMENTS

The EDC guidance documents should accurately reflect how to use the EDC at the current point in time. If changes are made to the EDC that affect the information provided in these documents, they should be updated accordingly and users notified of any changes.

The data entry guidance documents should be version controlled for data auditability and integrity.

7.5.4 DOCUMENTATION OF EDC VERSIONING

Each update/change to the relevant EDC should have a unique version number and/or name and/or date that can be used to identify it, and a timestamp when the changes came into effect.

The versioning of an EDC should be easily comparable to the DD and data entry guidance documents for data auditability and integrity.

7.5.5 UPDATING THE DATA DICTIONARY (DD)

The DD should be reflective of the EDC at the current point in time. When a change is made to the EDC, the DD should be updated accordingly to reflect the changes.

The DD should be version controlled for data auditability and integrity.

7.5.6 PROVIDING NOTIFICATION OF CHANGES

Once the changes are released to the live EDC environment, users should be notified as soon as possible and within 1 month in writing, and updated guidance documents provided as required. Where multiple regions are utilising the EDC, 1 week notice should be provided prior to the implementation of any changes.

8 RANDOMISATION DESIGN & VALIDATION

Randomisation for the trial will primarily be handled within the Central Spinnaker EDC (the core randomisation system), which will be preferred for domain-specific and other randomisations that directly feed into the primary statistical model. For aspects of the trial that are statistically separate from the primary model (e.g. nested trials), a separate randomisation system in an adjunctive EDC may be used; these randomisations may occur simultaneously with the core randomisations or at a later time point, depending on the specific protocol for each nested trial.

The randomisation systems used within the SNAP Trial should be formal, secure, reproducible and unpredictable. Design and validation of the randomisation systems is completed by a qualified statistician.

8.1 CORE RANDOMISATION SYSTEM

The core randomisation system is managed currently by the Central Spinnaker EDC through an integrated randomisation engine linked to the eligibility module for trial participant screening and eligibility:

- It is preferred that all domain randomisations are be integrated into the core randomisation system, unless otherwise justified, approved and documented in the study protocols and EDC-specific documentation.
- If appropriate, nested trial randomisations may be integrated into the core randomisation system.

8.1.1 TIMING OF CORE RANDOMISATION

The participant is randomised to all aspects that are part of the core randomisation system at the time of trial enrolment, only into aspects available at the trial site and to domains/studies they have consented to. However, randomised allocations are only revealed once the participant meets all of the inclusion criteria and none of the exclusion criteria for that aspect.

The participant is only considered 'enrolled' in a trial aspect once reveal has occurred. If the participant is deemed 'not eligible', their randomisation allocation will never be revealed.

- **Site NOT participating in specific trial aspect or participant NO consent** = allocation not generated
- **Site IS participating in specific trial aspect and participant IS eligible** ***(due to meeting all inclusion criteria and being assessed within the required timeframe)*** = allocation IS revealed
- **Site IS participating in specific trial aspect and participant NOT eligible** ***(due to meeting exclusion criteria and/or not being assessed within the required timeframe)*** = allocation NOT revealed

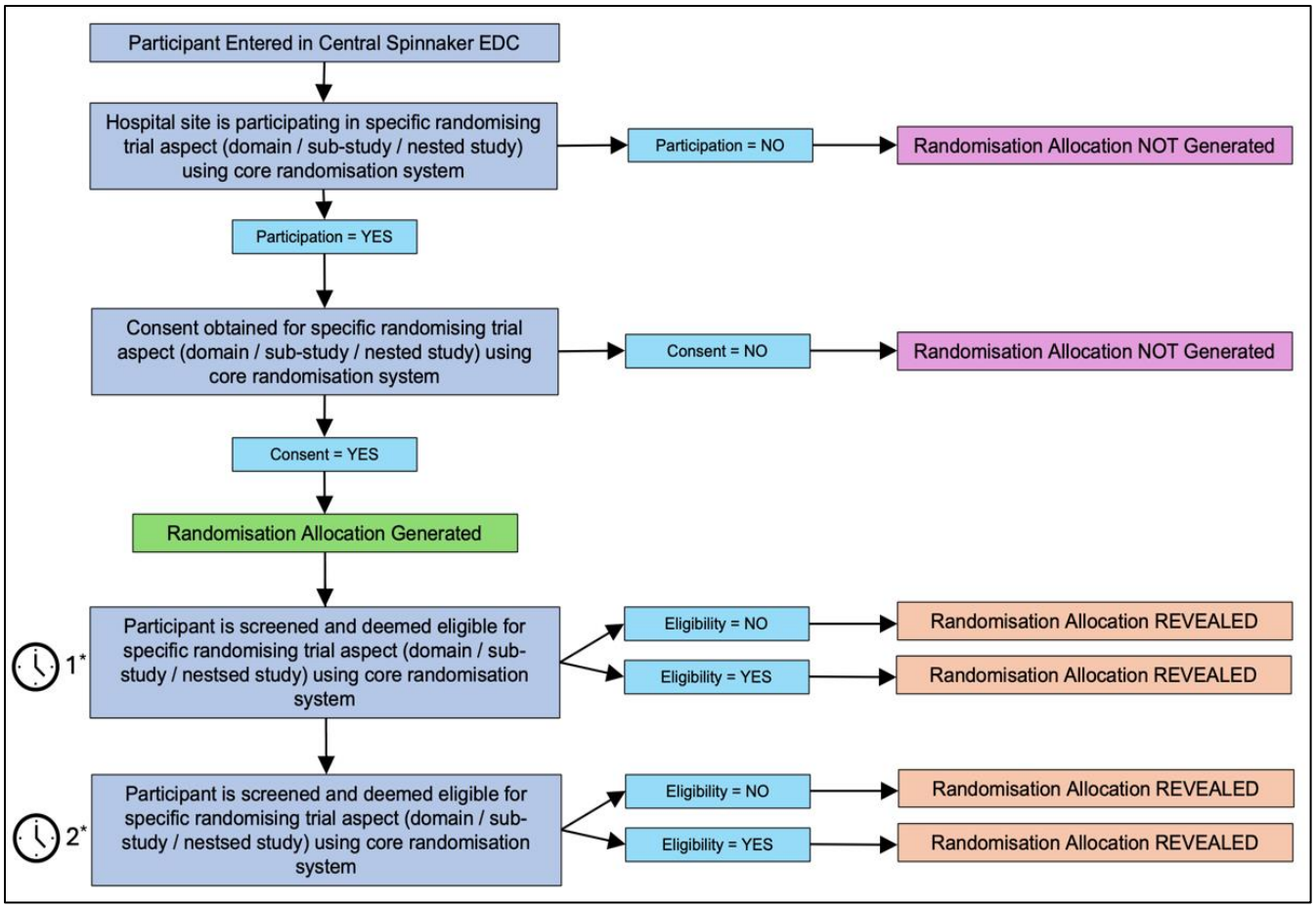

Figure 2 Timing of Core Randomisation Generation & Reveal

**It is important to note that not all core randomisation allocations will be revealed at the same time as they are generated, as allocation reveal is a staggered process based on the timing of eligibility assessments across the different domain / nested trials.*

8.1.2 CORE RANDOMISATION ALGORITHMS

The randomisation algorithm, developed by the Spinnaker EDC Provider, uses a "double-seed" strategy to generate random allocations. The seed combines the time and the unique eligibility ID.

Different randomisation algorithms may be employed for each trial aspect, based on anticipated size:

- **For larger domains/studies with high recruitment (e.g. >300),** simple randomisation is preferred. The large sample size is expected to mitigate any significant imbalances in group size.
- **For smaller domains/studies with lower recruitment (e.g. ≤300)**, block randomisation (with the potential to stratify by country or site) may be used to maintain balance in group size across groups.

The following table outlines the randomisation algorithms for each domain/nested trial including possible allocations.

Table 5 Randomisation Strategies for the Core Randomisation System

Domain	Silo	Randomisation Algorithm	Possible Allocations
Antibiotic Backbone Domain	PSSA	Simple randomisation	(Flu)cloxacillin
			Penicillin
			Cefazolin

(Domain A)	MSSA	Simple randomisation	(Flu)cloxacillin
			Cefazolin
	MRSA	Simple randomisation	Vancomycin / Daptomycin (clinician's choice)
			Vancomycin / Daptomycin (clinician's choice) + Cefazolin
Adjunctive Domain (Domain B)	-	Simple randomisation	No Adjunctive Clindamycin
			Adjunctive Clindamycin
Early Oral Switch Domain (Domain C)	-	Simple randomisation	Continued IV
			Switch to Orals
PET/CT Domain (Domain D)	-	Permuted block randomisation stratified by country	No PET/CT Scan
			PET/CT Scan
SIMPLY-SNAP (Nested Trial)	-	Permuted block randomisation stratified by language and decision-maker (surrogate vs patient)	Simplified Consent
			Conventional Consent

8.1.3 IMPLEMENTING RESPONSE ADAPTIVE RANDOMISATION (RAR)

At some point in the future, the trial may elect to implement Response Adaptive Randomisation (RAR). If this is approved by the GTSC, weighted ratios will be provided by the Stats-WG to the EDC Provider to implement within the EDC randomisation algorithm.

The format of these updated ratios will be discussed at the time and a suitable format agreed upon.

8.2 NON-CORE RANDOMISATION SYSTEMS

Any randomisation systems that randomise participants outside of the Central Spinnaker EDC will be considered a non-core randomisation system.

If non-core randomisation systems are used, this needs to be justified by the EDC-RP and approved by the GTMG, and the randomisation system validated as per **Section 8.3** below.

8.2.1 TIMING OF NON-CORE RANDOMISATIONS

Non-core randomisation may occur simultaneously with core randomisation or at a different time point. As per the core randomisation, randomised allocations are only revealed once the participant meets all the inclusion criteria and none of the exclusion criteria for that aspect.

The participant is only considered 'enrolled' in a trial aspect once reveal has occurred. If the participant is deemed 'not eligible', their randomisation allocation will never be revealed.

- **Site NOT participating in specific trial aspect or participant NO consent** = allocation not generated
- **Site IS participating in specific trial aspect and participant IS eligible** ***(due to meeting all inclusion criteria and being assessed within the required timeframe)*** = allocation IS revealed
- **Site IS participating in specific trial aspect and participant NOT eligible** ***(due to meeting exclusion criteria and/or not being assessed within the required timeframe)*** = allocation NOT revealed

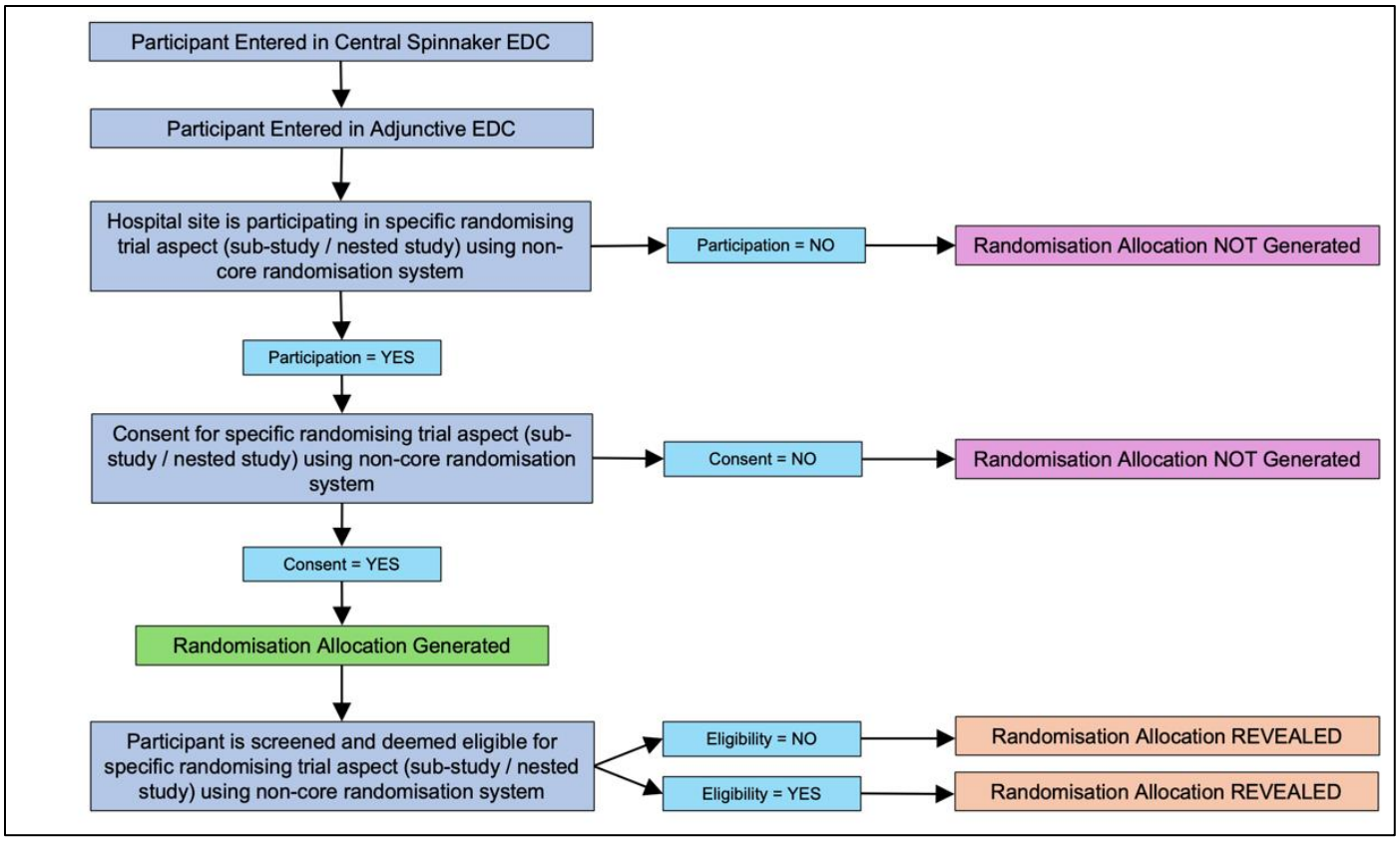

Figure 3 Timing of Non-Core Randomisation Generation & Reveal

8.2.2 NON-CORE RANDOMISATION ALGORITHMS

If a separate randomisation algorithm is used, the strategy for the algorithm must be documented by the lead investigator and this documentation provided to the GTMG prior to implementation.

8.3 VALIDATION OF THE RANDOMISATION ALGORITHMS

All randomisation algorithms used in the SNAP Trial will undergo validation by a qualified statistician via the following processes:

1. **Simulation of Randomisation:** Test randomisations will be generated using the randomisation software to simulate the trial and identify potential errors, imbalances, or other issues.

   AND

2. **Validation of Randomisation Algorithm (where code is available*):** The randomisation algorithm code should be provided for validation wherever possible. At a minimum, the algorithm should be checked by a statistician independent of the code development.

   OR

3. **Validation of Randomisation Output (where code is not available*):** Where the randomisation algorithm code is not available, the randomisation output will be reviewed to ensure that the allocations are balanced in accordance with the specified strategy (e.g., proper stratification and block sizes)

**If the code cannot be provided (e.g. Intellectual Property (IP) reasons), point 3 will be performed instead of point 2.*

Checks will be conducted by the EDC-RP to ensure that allocation concealment is maintained until the point of allocation reveal, and that the allocation assignments are only available to unblinded members with permission to view intervention allocations.

8.3.1 VALIDATION OF THE CORE RANDOMISATION ALGORITHM

Validation of the core randomisation algorithm will be undertaken by statistician(s) on the SNAP Trial Stats-WG and AT. Validation of the randomisation strategy will be conducted each time a new aspect that makes use of the core randomisation system is opened or when updates to the availability of domains and/or interventions are updated. Additionally, the DSMC reviews the randomisation numbers and outputs at each interim analysis as part of their ongoing oversight of the trial.

Further ad hoc validations can occur as requested by the GTSC, Stats-WG, or AT.

If a SNAP proposed trial elects to use the core randomisation module, but it has no impact on the primary statistical model, it will be the responsibility of the lead(s) to enlist a qualified statistician to conduct validation of the core randomisation module and provide documented evidence of this validation processing to the GTMG & Stats-WG prior to implementation.

The core randomisation algorithm was validated by the EDC Provider, Stats-WG and sub-study leads (where relevant) at the following timepoints:

Table 6 Validation Events of the Core Randomisation Module

Validation Event	Date	Validated by
Beginning of Trial including Adjunctive, Backbone and Early Oral Switch Domains	October 2021	Stats-WG
After 100 participants randomised	July 2022	Stats-WG
New Nested Trial – SIMPLY-SNAP	November 2023	Lead
New Domain – PET/CT	May 2025	Stats-WG
New Interventions – Cefazolin introduced within the PSSA Silo of the Backbone Domain	May 2025	Stats-WG
**This table will be updated when a new version of the DMP is released.*		

Documentation of each validation event will be stored in the GTMG TMF.

8.3.2 VALIDATION OF NON-CORE RANDOMISATION ALGORITHMS

The EDC-RP will be responsible for engaging a qualified statistician to perform validation of any non-core randomisation algorithms. Validation of the randomisation strategy should be conducted prior to EDC implementation at a minimum, and when the availability of the interventions is updated. Documentation of this validation must be kept according to local and regional guidelines and regulations and stored in the TMF.

Where the EDC will be used in more than 1 participating SNAP region and/or the data analyses will utilise the core statistical model, the validation documentation must be provided to the GTMG prior to the implementation of the algorithm or algorithm updates. The GTMG will store this documentation in the GTMG TMF.

9 DATA ENTRY, HANDLING & EXTRACTION PROCEDURES

Core data collection and randomisation will be maintained in the Central Spinnaker EDC. When new trial aspects are integrated, data collection may be incorporated into an adjunctive EDC (such as REDCap, OpenClinica, etc.)

The following sections outline the minimum requirements for data entry, handling and extraction across the SNAP Trial EDCs.

*See **Section 7** for more information on SNAP Trial EDCs.*

9.1 DATA ENTRY PROCEDURES & TIMELINES

Data is entered directly into the trial EDCs by approved users (see **Section 7.3.2).**

Data entry is required for each participant throughout their study duration, in accordance with the timelines outlined in the protocols and appendices. The CRFs are designed to be pragmatic – streamlining the validity of the entered data by utilising prespecified validity checks with on screen alerts and confirmations to notify the user when data is unexpected or missing.

Events such as protocol deviations, safety events, and consent and withdrawal events should be entered in real-time or as soon as possible to ensure that the EDCs are current and accurate.

If the usual data entry staff are not able to enter data as per the data entry timelines outlined in the protocols and/or EDC-specific documentation, the situation will be monitored by the RTMG, and a decision made regarding whether the site should pause recruitment until data entry issues can be resolved. The GTMG and RTMG will be in communication with the trial sites to ensure that the data points of highest priority are completed, and that any pauses to recruitment are logged and documented.

A **CRF Summary Guide for the Central Spinnaker EDC** will be provided and maintained by the GTMG; this guide outlines the EDC data fields and timepoints for collection for use as a reference tool. It will be the responsibility of each EDC-RP to provide similar documentation, if deemed necessary or upon request.

9.2 DATA COMPLETION & MISSINGNESS

CRFs will only be considered complete once all required data fields have been entered and completion status checked by the RTMG.

Each CRF will be checked to confirm that:

- The required data fields are complete. If unable to be completed, it is documented by the RTMG in collaboration with EDC-RP and the site that the missing data is unable to be obtained and field(s) is/are left blank on the EDC
- That all data alerts and queries are resolved, either by updating the data or responding to data queries

Built-in completion logic will be implemented in the EDCs, where possible, to auto-mark CRFs as complete when all data fields have been entered.

Data missingness in the Spinnaker EDC will be checked as part of the central monitoring process **(See Section 10).**

9.3 DATA VALIDATIONS

Built-in validity, consistency, and logic checks will be implemented within the EDCs (e.g., range checks for data points, consistency of dates, numeric/alphabetic restrictions) to assist sites in entering correct

data. Warnings may appear when a data entry contradicts another field or what is reasonably expected for that field, or to alert sites of a potential protocol deviation or safety event.

Additional data validity checks will be performed by the GTMG and/or RTMGs as part of the central monitoring process (See **Section 10.1**).

9.4 DATA CORRECTIONS

Data correction is the process of identifying and rectifying errors, inconsistencies, and inaccuracies within the dataset. This involves finding and fixing or removing incorrect, incomplete, duplicate, or improperly formatted data. Data corrections are crucial for ensuring data accuracy, reliability, and usability.

Data corrections should be applied by the site directly within the EDC wherever possible. If data corrections are unable to be made to the EDC (e.g. due to restricted fields), a data corrections process will be developed, maintained, and implemented by the relevant EDC-RP. This process should be documented.

Please consult the **EDC-specific documentation** for more information regarding data corrections.

9.5 DATA COLLECTION OUTSIDE OF THE EDCS

In rare occasions, participant data may need to be collected outside of an EDC. Some examples of this could include:

- Paper CRFs where CRF in the EDC is not (yet) available
- Separate data collection tool used where errors in EDC logic or EDC bugs are discovered which prevent data from being collected accurately

If data is collected outside of the EDC, data will be entered into the relevant EDC system as soon as possible and a Note to File (NTF) written by the EDC-RP to document the processing, handling and archiving of this data.

9.6 SOURCE DATA

This trial will not in usual practice mandate the use of paper CRFs for data collection, therefore trial documents considered "source data" will be minimal; the participant's medical records will be the primary source of data unless otherwise specified in the monitoring plan.

- PDF downloads of blank CRFs may be available from the EDCs – if available, these will be considered a data collection tool only.
- Where documents *are considered source data,* these should be documented and stored in accordance with and for the duration of the trial as specified by regional and local regulations.

Where data is collected outside of the EDC (See **Section 9.5** above), the NTF will document if the data collection tool is considered source data.

9.7 DATA EXTRACTIONS (DATA CUTS)

A data extraction (data cut) refers to a specific extraction event of the entire participant data set from the relevant EDCs that is not modified/updated. It will serve as a snapshot in time and, therefore, it is expected that some participant data may be incomplete or inaccurate, due to the ongoing nature of the adaptive platform trial. As such, participant data may differ between data cuts.

There will be numerous data extractions that occur throughout the life of the SNAP Trial. Events that may trigger a data cut include, and not be limited, the following:

1. Data Quality Control & Monitoring Procedures - *see* ***Section 10***
2. Scheduled Interim Analyses (as defined in the protocol documents) – *see* ***Section 13.1***
3. Final Analyses (upon the last participant of a particular cohort reaching completion) - *see* ***Section 13.2***

It will be the responsibility of the EDC-RP to ensure that the data extractions are performed according to the timelines specified in the protocols and Statistical Analysis Plans, and that appropriate quality control checks are undertaken prior to the extraction event.

9.8 STANDARDISED DERIVED VARIABLES

A derived variable is a variable that is calculated or categorised from one or more other variables in a dataset.

Due to the pragmatic design, many trial endpoints will be based on derived variables. The methods used to derive these variables will be documented in a Standardised Derived Variables document by the EDC-RP and made publicly available on the SNAP Trial website for transparency and reproducibility.

10 DATA QUALITY CONTROL & MONITORING PROCEDURES

Data quality control and monitoring will be a continuous process throughout the lifetime of the SNAP Trial. It remains the primary responsibility of each RTMG to maintain their trial data and ensure accuracy and validity to the best of their ability.

To assist with data quality control, the following monthly summaries will be generated by the GTMG for EDC's where the GTMG is the EDC-RP. These reports will be provided to the GTSC, RTSCs and RTMGs in the Data Cleaning and GTSC Reports. A master copy will be stored in the GTMG TMF by the GTMG and local copies provided to RTMGs should be stored in the RTMG TMF.

Table 7 Data Quality Control Summaries

Summary Name	Detail	Report Name	Type of Report	Timescale
Data Cleaning Summary	Summary of data inconsistencies and errors for correction	Data Cleaning Report	Regional	Every 4 weeks *Will be made available to the RTMGs via the Data Cleaning Report and can be shared with the RTSCs.*
Data Completion Summary	Summary of data completion and completion timeliness for CRFs by: • Site (Regional only) • Region • Whole trial (Global only)	Data Cleaning Report	Regional	Every 4 weeks *Will be made available to the RTMGs via the Data Cleaning Report and can be shared with the RTSCs.*
		GTSC Report	Global	Every 4 weeks *Will be made available to the GTSC via the GTSC Report.*
Data Missingness Summary	Summary of data missingness for EDC fields by:	Data Cleaning Report	Regional	Every 4 weeks *Will be made available to the RTMGs via the Data Cleaning*

	• Site (Regional only) • Region • Whole trial (Global only)			*Report and can be shared with the RTSCs.*
		GTSC Report	Global	Every 4 weeks *Will be made available to the GTSC via the GTSC Report.*

10.1 CENTRAL DATA MONITORING & REPORTS

There are many embedded validities checks in the EDC tools that will encourage sites to enter correct and accurate data. In addition to these automatic validity checks, sites will undergo regular data cleaning and monitoring processes to ensure quality and integrity of the data. Details of the overall monitoring for the trial can be found in the **Global & Region-Specific Monitoring Plans;** the aspects discussed here relate to data management only (data cleaning and monitoring).

The GTMG will conduct regular and routine central data monitoring, summarised in monthly **Data Cleaning Reports** and **GTSC Reports**:

- A monthly data extraction will be undertaken to perform these checks and summaries, and will be conducted using scripts developed by the GTMG in R (programming software)
- The reports will summarise data monitoring conducted on priority data points (See **Section 10.1.1**) and other data points (see **Section 10.1.2**)
- The reports will contain aggregated summaries of data completion and missingness at the global, regional and site level
- The **Data Cleaning Report** only will contain line listings of relevant regional and site data inconsistencies and cleaning checks

Throughout the lifetime of the trial, new checks and summaries may be incorporated into the report templates as required, and items that are no longer required removed. The reports are a working document that continue to change and grow as the trial and its EDCs evolve.

For more information on what checks and summaries are being currently conducted, please see the **Data Cleaning Report Template** and **GTSC Report Template.**

10.1.1 CENTRAL DATA CLEANING OF PRIORITY DATA POINTS

Central data cleaning will be focused on priority data points to ensure the integrity of the scheduled interim analyses.

Priority data points are ones which:

- Will affect the participant's randomisation within the nested trials and progression through the trial.
- Are imperative for the conduct of the nested scheduled analyses
- Are indicative of a protocol deviation or safety event (or other noteworthy event as defined in the trial protocols)

The table below shows the priority data points for the SNAP Trial, and the type of priority they are given within the EDCs. Please consult the **Monitoring Plan** and **Data Cleaning Report Template** for further information.

Some priority data points are mandated by the EDCs through built-in verification as described in **Table 8**; these include data points that are necessary for screening, enrolment and randomisation into nested

trials. Other priority data points will have data checks run and/or Source Data Verification (SDV) performed to confirm the data is complete, within an expected range, and/or have an associated safety or deviation event logged.

- SDV will be performed by RTMGs as per the Region-Specific Monitoring Plan
- Data cleaning checks are led by the GTMG to ensure that any omissions or discrepancies in the data can be resolved as soon as possible, and prior to the scheduled analyses, as applicable.

Any SDV/data cleaning checks performed on these variables should be documented in the EDC system, unless otherwise outlined in the EDC-specific documentation.

Table 8 Priority Data Points for the SNAP Trial

Priority Data Points	Type of Priority		
	Mandated by EDC	Data Check Conducted[1]	SDV Performed[2]
Participant Consent	X	X	X
Platform Eligibility Criteria	X	X	X
Domain/Nested Trial Eligibility Criteria	X	X	X
Date and time of Index Blood Culture Collection	X	X	X
Participant Initials[3], DOB[3], Age & Sex at Birth	X		X
Vital status at Platform Day 90		X	X
Date of death		X	X
Participant Silo		X	X
Day 2 and Day 5 Blood Culture Results		X	X
Acute Discharge – date and discharged to (if relevant)		X	X
Total Discharge – date (if relevant)		X	
Acute Kidney Injury (AKI) Safety Event – Derived Variable		X	
Acute Liver Injury (ALI / hepatotoxicity) Safety Event – Derived Variable		X	
C. difficile Safety Event – Derived Variable		X	
Change in Treatment Safety Event – Derived Variable		X	
Protocol Deviations		X	

Serious Adverse Reactions (SARs)[4]		X	
Storage of Index Isolate & Lab Number			X

[1] for 1 or more data points within the category; see **Data Cleaning Report Template** for more information

[2] for 1 or more data points within the category; see **Global and Region-Specific Monitoring Plans** for more information

[3] if collected in that region, as per regional regulatory and ethical approvals

[4] Serious Adverse Reactions (SARs) will be collected and monitored in all regions. Some regions will collect additional safety data such as SAEs or AEs; these will not be considered part of the priority data points

10.1.2 OTHER CENTRAL DATA CLEANING ACTIVITIES

Routine checks of common potential data inconsistencies and missing data will also be performed by the EDC-RP on an ongoing basis and summarised for the RTMGs and will be specified in the EDC specific documentation; where possible, these will be integrated into the Data Cleaning Reports. These checks relate to data fields that are of lower priority as they do not relate directly to the scheduled analyses.

When a final analysis or other data extraction event is scheduled, additional comprehensive checks pertaining to the domain/nested trial-specific eligibility and endpoints will be incorporated into the central monitoring activities to ensure that the data is as accurate and complete as possible for the extraction event.

For further information on the current checks being conducted, please see the **Data Cleaning Report Template.**

10.2 DATA QUALITY QUERY & RESOLUTION PROCESS

Where possible, all data querying should be performed using a query system inbuilt within the EDC. If a query cannot be entered within the EDC, it should be emailed to the site, with specific reference to the EDC and CRFs it relates to. A copy of the sent email should be kept as part of the RTMG TMF, and where possible, a corresponding query should also be added to the EDC to confirm the validity of the datapoint.

It is expected that data cleaning queries are responded to and/or resolved as soon as possible and within 14 days by the site, in communication with the RTMGs.

Aggregated reports on data completion and timeliness of data completion will be provided by the GTMG to the GTSC and RTMGs as per **Table 7**. Line listing of relevant regional and site data completion will be provided to RTMGs.

10.3 OTHER MONITORING ACTIVITIES

All other monitoring activities will be led and conducted by the RMTGs on behalf of the ISO.

For more information on the trial global monitoring requirements, please see the **Global Monitoring Plan.**

10.4 SIGN OFF ON DATA ENTRY

Some event CRFs (e.g. protocol deviations, safety events, etc) will require signatures from specific parties. If sign off is required, a digital signature can be added to these CRFs within the EDC to show that they have been reviewed and confirmed by the user; these types of sign offs are documented in EDC-specific documentation.

Permissions for sign off will only be provided to specific user profiles to ensure that the sign off has been made by the authorised personnel.

Bulk electronic sign off on all EDC CRFs within the EDC system is required at the following timepoints as a minimum global requirement, commencing from the timepoint that electronic sign-off is available within the EDC:

- At each final analysis (prior to each final analysis data cut)
- At site closure

Additional electronic sign off within the EDC system by site PIs and/or RTMGs will not be required as part of the global minimum requirement but may be implemented at a regional level as required.

11 DATA PRIVACY & CONFIDENTIALITY

In accordance with ethical guidelines, regulatory requirements, and institutional policies, this clinical trial will implement robust data privacy and confidentiality measures to protect participants' personal health information (or personal data). 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 in participating regions.

The below sections outline the trial strategies used to ensure the privacy and confidentiality of data throughout the trial.

11.1 MAINTAINING A ROBUST INFORMED CONSENT PROCESS

The informed consent process ensures that participants are fully aware of the nature of the study, the types of data being collected, how their data will be used, and their rights regarding privacy and confidentiality. This process will be conducted in compliance with ethical guidelines and regulatory requirements to protect participant autonomy and ensure transparency in data usage.

Informed consent will be obtained from all participants prior to their enrolment in the SNAP Trial, and before any data is collected, unless otherwise specified in the region-specific documentation and approvals.

Prior to enrolment, each participant will receive detailed information explicitly outlining the following key points about their data:

- **Types of data collected**: Participants will be informed about the types of personal and health-related data that will be collected, such as age, comorbidities, lab results, etc.
- **Purpose of data use**: Participants will be informed if their data will be used exclusively for the purposes of the SNAP Trial, sub-study(s) and/or future or other research.
- **Data security and confidentiality**: Participants will be reassured that their data will be kept confidential and stored securely, in compliance with data protection regulations. They will be informed about the measures in place to protect their privacy, such as encryption and access controls.
- **Right to withdraw**: in the case of withdrawal, data collected up to the point of withdrawal will be handled according to the terms outlined in the informed consent form.
- **Potential for data sharing**: If applicable, participants will be informed of any potential sharing of pseudonymised data with third parties (e.g., regulatory bodies, academic collaborators, or external researchers), and the steps taken to protect their privacy in such cases.

- **Retention Period:** Participants will be informed as to how long their data will be kept and what will be done with their data after the retention period is complete.

If there are any changes to the data use or the purpose of the trial, participants will be informed, and updated consent will be obtained, if necessary. This ensures that participants remain fully informed about how their data is being used throughout the course of the trial.

11.2 PSEUDONYMISATION OF PARTICIPANT DATA

The trial EDC(s) contain personal data and will therefore be treated as confidential information and stored securely. To protect participants' personal information, all study data will be pseudonymised prior to data analysis, reporting or sharing. A unique participant identifier (Trial ID) will be assigned to each participant (see **Section 11.3** below).

11.2.1 PARTICIPANT IDENTIFYING INFORMATION (PII)

Participant-Identifying Information (PII) is a subset of personal data that refers to any data or combination of data elements that can be used to directly identify, contact, or locate a specific individual (for example name, date of birth, health care number, address, etc).

Collection of and access to PII is highly restricted in the SNAP Trial and is only collected and accessed where it is specifically required to undertake approved trial processes (such as data linkage or biobanking) and if allowed by local regulations.

A summary of allowable access to PII is provided in in the table below.

Table 9 Access to Participant-Identifying Information (PII)

Role	Access to PII Collected in the study EDC(s)[1]	Access to PII Collected in the Data Linkage EDC(s)[1]	Access to PII Collected in the Biobanking EDC[1]
Relevant EDC Provider	**YES**	**YES**	**YES**
Analytic Team (AT)	NO	NO	NO
Statistical Subcommittee (Stats-WG)	NO	NO	NO
Data Safety Monitoring Committee (DSMC)	NO	NO	NO
Global Trial Management Group (GTMG)	Visible On-Screen Only *(encrypted in the exports)*	NO	NO
GTMG/RTMG – Designated Biobank Coordinator(s)	Visible On-Screen Only *(encrypted in the exports[2])*	NO	**YES[3]**
Regional Trial Management Groups (RTMGs)	Visible On-Screen Only *(encrypted in the exports)*	Visible On-Screen Only *(encrypted in the exports)*	NO
RTMG – Designated Data Coordinator(s)	Visible On-Screen Only *(encrypted in the exports[4])*	**YES[5]**	NO
Site Staff	Visible On-Screen Only *(no access to exports)*	Visible On-Screen Only *(no access to exports)*	NO

Global Trial Steering Committee (GTSC)	NO	NO	NO
Other Working Groups & Committees	NO	NO	NO
Sub-Study Leads	NO	NO	NO

1 if approved for collection in the region

2 only designated biobank coordinator(s) from the RTMG & GTMG will have access to PII in an unencrypted manner, for the purposes of conducting the biobanking.

3 Only designated biobank coordinator(s) from the RTMG & GTMG will have access to PII for the purposes of conducting the biobanking.

4 only designated data coordinator(s) from the RTMG will have access to PII in an unencrypted manner, for the purposes of conducting data linkage

5 Only designated data coordinator(s) from the RTMG will have access to PII for the purposes of conducting data linkage. Please refer to the **Region-Specific Appendix** for further information about data linkage processes in each region.

11.2.2 PARTICIPANT IDENTIFYING INFORMATION (PII) COLLECTED IN THE STUDY EDC(S)

Participant Identifying Information (PII) is collected in the study EDCs for the specific purposes of:

1. **Implementing EDC logic for specific cohorts:** PII collected during screening (such as date of birth/sex at birth/initials, if applicable in the region) will be used to show/hide relevant interventions and data points which are specific to that cohort (paediatric / potential for pregnancy / etc).
2. **Verifying data entry of the data linkage EDC:** PII collected during screening (date of birth/sex at birth/initials, if applicable in the region) will be used to verify the data entry for the data linkage component of the trial
3. **Verifying data entry of the Central Biobank EDC:** the laboratory ID for each participant's index isolate sample will be collected in the study EDC for regions participating in the Central Biobank and will be used to verify the laboratory isolates received as part of the central biobanking process.

PII are encrypted in the study data exports as per **Table 9**. Please see **EDC-specific documentation** for more information.

11.2.3 PARTICIPANT IDENTIFYING INFORMATION (PII) COLLECTED IN THE DATA LINKAGE EDC

Additional identifiable data (including full name, date of birth, health care number; and if allowed by local regulation) may be collected in the data linkage EDC for the purposes of data linkage, in regions that have approval to conduct data linkage once specific consent is obtained, if required in the region.

If collected, this data will be kept on a separate secure EDC and will only be accessed by the site trial staff at that site, and designated data coordinator(s) from the RTMG for the purposes of providing the data for linkage, the data linkage unit who will perform the linkage, and regulatory authorities who may want to check that the trial is being carried out correctly.

11.2.4 PARTICIPANT IDENTIFYING INFORMATION (PII) COLLECTED IN THE BIOBANKING EDC

The laboratory ID associated with the participant's index isolate sample will be collected in the Central Biobank EDC for the purposes of identifying the isolates that are received for the Central Biobank, in

regions that have approval to conduct data linkage once specific consent is obtained, if required in the region.

If collected, this data will be kept on a separate secure EDC and will only be accessed by designated biobank coordinator(s) from the GTMG for the purposes of maintaining and updating the biobank and regulatory authorities who may want to check that the trial is being carried out correctly.

11.3 USE OF PARTICIPANT IDENTIFICATION (ID) CODES

Each participant is assigned two codes when entered into the Central Spinnaker EDC, which will be used to identify the participant throughout the trial and between EDCs.

1. **Participant Screening ID:** a participant screening ID is assigned to each participant that is screened for the trial, including screen failures.
2. **Participant Trial ID:** an additional trial ID number when the participant is enrolled into the trial (either platform or registry).

Additional identification codes may be assigned to participants who are participating in additional sub-studies / nested studies / adjacent studies.

11.3.1 PARTICIPANT SCREENING ID

During platform screening, each participant is given a randomly generated 6-letter alphabetic ID. This identification number can be used to identify each unique participant, regardless of their inclusion or exclusion into the trial.

11.3.2 PARTICIPANT TRIAL ID

After platform eligibility screening has been completed and the participant is included in either the platform or registry, a new 8- or 9-digit Participant Trial ID number is assigned.

The Participant Trial ID comprises three letters designating the site code of the enrolling site, and 5 numbers, allocated sequentially per site, in the order of enrolment into the trial. Registry-only participants are intercalated within this sequence and are designated with the prefix "R-" before the site code. If a patient is removed from the EDC for whatever reason, their unique number will be archived and cannot be reused. Thus, a sequential list of participants at a site may be as follows:

- ABC00001
- ABC00002
- R-ABC00003
- ABC00005*

**R-ABC00004 skipped as participant was removed from the EDC*

The 3-letter site codes used are listed in a shared tracker and are assigned by the RTMGs. The RTMGs will be responsible for checking with the GTMG to ensure the site code has not already been used in another region. All site codes are unique and reflect the name of the site in some manner (for example John Hunter Hospital is denoted by the site code 'JHH').

11.3.3 OTHER PARTICIPANT IDS AND LINKING PARTICIPANTS BETWEEN EDCS

If using adjunctive EDCs, participants may be assigned additional participant IDs within the Adjunctive EDCs.

It is important that an ID from the Central Spinnaker EDC (either the Screening ID or Participant Trial ID) is collected in the adjunctive EDCs to ensure that participants can be linked.

- The **Screening ID** should be used when linking participants to EDCs that contain Personal Identifying Information (PII; see **Section 11.2** [e.g Biobanking EDC or Data Linkage EDC]) to protect participant re-identifiability.
- The **Trial ID** should be used when linking participants to EDCs that do not contain PII (See **Section 11.2**; e.g adjunctive EDCs) as this ID is more widely used and easier to identify in the Central Spinnaker EDC.
    - The use of the Trial ID is discouraged when linking participant records across EDCs that contain PII, as the Trial ID contains information about the participant's recruitment site.

12 MANAGING BLINDING TO DATA & ANALYSIS RESULTS IN THE ONGOING TRIAL

Blinding in this section refers to blinding of the data and analysis results in the ongoing trial. This is separate to blinding described in the core protocol, which refers to blinding of treatment allocation.

Measures to maintain objectivity and minimise potential biases within the analyses of the ongoing trial are essential for the integrity and validity of the study results.

Aggregate participant data will remain blinded to everyone except to specific nominated unblinded members, until a trial aspect reaches certain milestones (such as domain conclusion reached/ substudy complete, etc.). Once this occurs, the results from that aspect are published and become unblinded to everyone.

Please refer to the figure below for an overview of how blinding to trial data and results is arranged within the SNAP Trial.

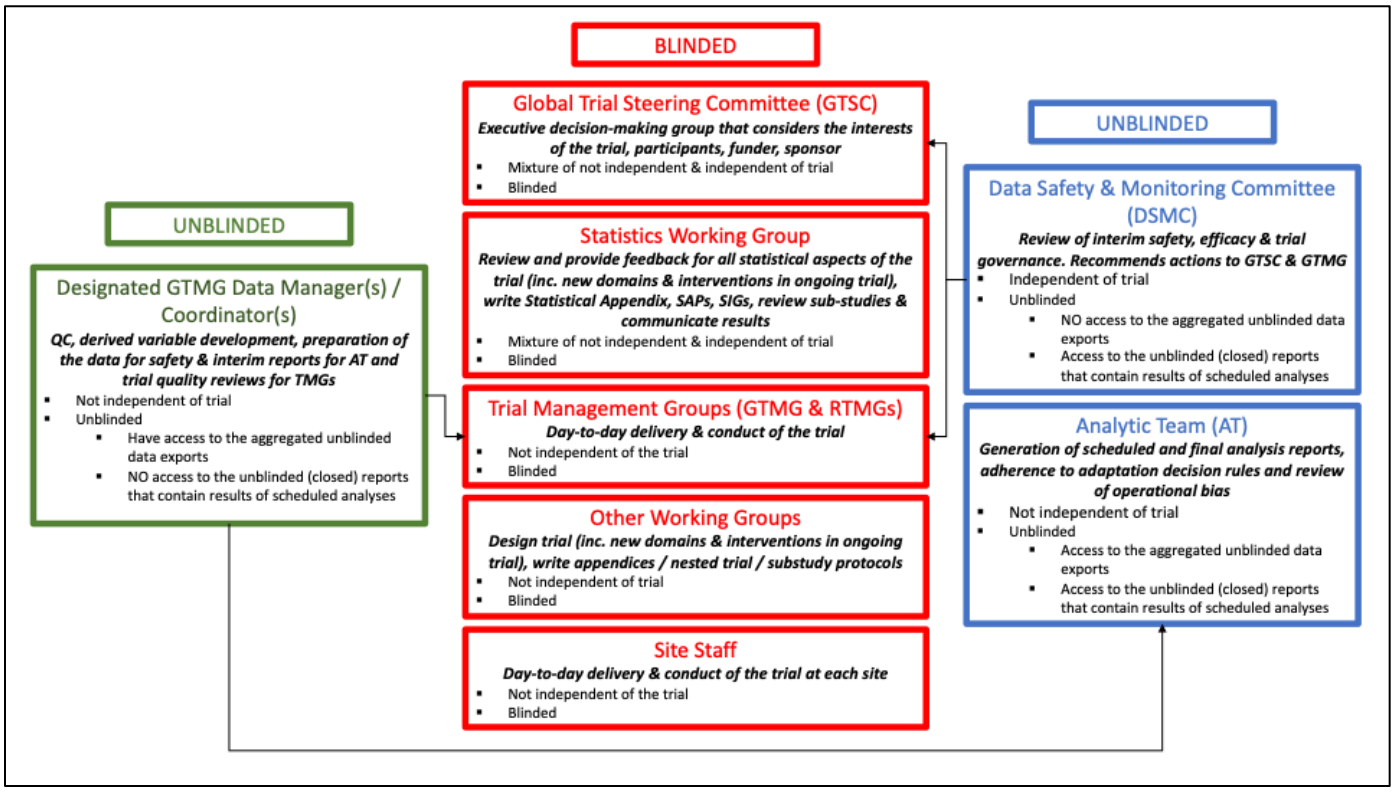

Figure 4 Simplified Flowchart of Unblinding to Data in the SNAP Trial

The blinding permissions for the ongoing study aspects is as follows:

1. **Blinded** = blinded members do not have access to aggregated unblinded data exports or the unblinded (closed) reports that contain the results of the scheduled analyses.
2. **Unblinded** = unblinded members have access to aggregated unblinded data exports (*See **Section 12.1***), and/or have access to the closed reports which contain the results of the scheduled analyses (*See **Section 12.2***).
    a. **Designated GTMG Data Manager(s)/Coordinator(s)** = have access to aggregated unblinded data exports, but do not access the unblinded (closed) reports that contain the results of the scheduled analyses.
    b. **Analytic Team (AT)** = have access to aggregated unblinded data exports, as per the Designated GTMG Data Manager(s) / Coordinator(s) and additionally access the unblinded (closed) reports that contain the results of the scheduled analyses.
    c. **Data Safety Monitoring Committee (DSMC)** = do NOT have access to aggregated unblinded data exports but do access the unblinded (closed) reports that contain the results of the scheduled analyses.

    Unblinded members with access to the closed reports are prevented from participating in any working groups or committees in which decisions are made relating to trial activities.

No groups have access to individual treatment allocations prior to the point of randomisation allocation reveal.

The table below provides a summary of blinding status for each of the core roles.

Table 10 Blinding Status within the SNAP Trial

Role	Blinding Status	Access to Aggregated Unblinded Data Exports	Access to Scheduled Analysis Results and Unblinded Reports [1]
Analytic Team (AT)	**Unblinded**	**YES**	**YES**
Data Safety Monitoring Committee (DSMC)	**Unblinded**	NO	**YES**
GTMG – Designated Data Manager(s) / Coordinator(s)[2]	**Unblinded**	**YES**	NO
Global Trial Steering Committee (GTSC)	**Blinded**	NO	NO
Global & Regional Trial Management Groups (GTMG & RTMG)	**Blinded**	NO	NO
Statistical Subcommittee (Stats-WG)	**Blinded**	NO	NO
Other Working Groups & Committees	**Blinded**	NO	NO
Site Staff	**Blinded**	NO	NO
Nested Trial / Sub-Study Leads	**Unblinded[3]**	**YES[3]**	NO

[1] Only the AT and DSMC will have access results of the scheduled analysis and can view the unblinded (closed) reports

[2] Only designated data manager(s) / coordinator(s) from the GTMG will have access to the aggregated unblinded data exports for the purposes of conducting cleaning and validation checks, and for liaising with the AT as required.

[3] Nested Trial / Sub-study leads may have access to unblinded data pertaining to their sub-study / nested study but not to the SNAP Trial in general

12.1 ACCESS TO AGGREGATED UNBLINDED DATA EXPORTS

Aggregated unblinded data exports contain post-randomisation data stratified by allocated treatment(s).

Prior to any domain, cell or trial conclusions, the only users with access to the aggregated unblinded data exports will be the AT and the designated data manager (s) / coordinator(s) from the GTMG.

- The GTMG designated data manager(s) / coordinator(s) will require access to the unblinded data exports for the purposes of providing data, data cleaning, and data field derivation for the scheduled and final statistical analyses.
- The AT will require access to the unblinded data exports for the purposes of conducting the statistical analyses and producing the scheduled and final analysis reports.

Following the public release of any trial conclusions, access to unblinded data (including data included in the publication of the results) must be carefully considered to maintain trial integrity for the ongoing trial aspects.

12.2 ACCESS TO UNBLINDED RESULTS/REPORTS

Although the designated data manager(s) / data coordinator(s) from the GTMG will have access to the aggregated unblinded data exports, they will *not* have access to the results of the scheduled analyses or the unblinded (closed) reports produced by the AT as part of the scheduled analyses – only the AT and DSMC will have access to this level of unblinding.

Access to the scheduled analysis report(s) will be protected while the trial is in progress - trial integrity demands clear delineation between those who have access to blinded and unblinded scheduled reports to protect the decision-making processes of the adaptive platform trial.

Following the public release of any trial conclusions, access to final analysis report(s) from that trial aspect conclusion will be protected to maintain trial integrity – only those involved in the writing group for that domain will be allowed access to the report.

12.3 COMMUNICATION BETWEEN BLINDED AND UNBLINDED PARTIES

Communications between blinded members and unblinded members with access to the closed reports (e.g. Analytic Team or DSMC) will be limited wherever possible. If communications between these parties is required to take place, the chairs of the AT, DSMC, Stats-WG and GTSC will be included in the correspondence for oversight of the communication.

Communication between unblinded members with access to the closed reports and unblinded members who do not (e.g. Analytic Team or DSMC and Unblinded Data Coordinator(s)) will be limited to queries pertaining directly to the data. If communications between these members is required to take place, a copy of this communication will be kept in the unblinded data drive and archived as part of the TMF at trial closure. If any files are shared between unblinded parties containing unblinded trial data, they will be shared according to the Data Sharing process outlined in ***Section 14.***

13 STATISTICAL ANALYSES & REPORTING

Statistical analyses will be performed using the unblinded data exports, extracted from the trial EDC(s) by the Designated Data Manager(s) / Coordinator(s) and provided to the nominated analysts. Any unblinded data exports that are extracted from the EDCs will only be stored on a SNAP Research-NAS for the duration of the trial.

Data extraction for the statistical analyses will occur when:

- **Core Scheduled Analyses:** Each 500th participant has reached day 90 and a scheduled analysis occurs (See **Section 13.1**)
- **Domain Final Analyses:** When a domain is completed, once data quality control procedures have been undertaken (See **Section 13.2/13.3**)
- **Sub-Study/ Nested Trial Scheduled & Final Analyses:** When a Sub-Study / Nested Trial reaches the trigger for an interim analysis or when the trial is completed, and once data quality control procedures have been undertaken (See **Section 13.2/13.3**)

Prior to analysis, data quality will be improved using the methods outlined in **Section 10**. Comprehensive efforts will be made to ensure that data is as clean as possible for each analysis conducted.

The data extraction (data cut; as defined in **Section 9.7**) used for each analysis will serve as the locked analysis dataset.

- Data for the statistical analyses will be downloaded from the relevant EDCs by the designated Data Managers (or delegates) and provided to the nominated analysts.
- A copy of this data extraction (data cut) is securely retained by both the nominated analysts and the designated Data Managers (or delegates).
- Email correspondence is sent by the designated Data Managers (or delegates) to the EDC-RP and nominated analysts to confirm successful data transfer of the data for these analyses.

As a pragmatic adaptive platform trial, participant data may contribute to more than one analysis. Therefore, in general, the data extraction (data cut) used for an interim or final analysis will serve as the "locked analysis dataset" (see **Section 9.7**). This allows for the data pertaining to that particular analysis to remain unchanged, whilst allowing the full dataset to undergo continuous data quality control procedures throughout the lifetime of the trial.

For example, the same data fields may be collected for the analysis of the Domain X as Domain Y; however, Domain X closed 1 year prior to Domain Y. During that year, some data fields have been updated due to usual data cleaning and update processes. As such, the same data fields for certain participants may differ between the locked dataset for Domain X and Domain Y.

See **Sections 13.1-13.3** for more information on the scheduled interim and final analyses within the SNAP Trial.

13.1 CORE SCHEDULED ANALYSES

Scheduled analyses, which occur when each 500th participant has reached day 90 and a scheduled analysis occurs will be performed using the unblinded data exports, extracted from the trial EDC(s) by the Designated Data Manager(s) / Coordinator(s) at the timepoints outlined in the AT TOR and provided to the AT via a secure, access-limited drive.

Data extraction for the scheduled analyses will occur when each 500th participant has reached day 90 and a scheduled analysis occurs.

The schedule of the key extraction events for the analyses will occur as specified in the statistical documentation and the DSMC and AT Terms of Reference documentation. The GTMG will ensure that the AT, DSMC, RTMGs and GTSC are informed of the timelines for each scheduled analysis. See **Table 13** for a summary of these key extraction events; this table will be updated with each review of the DMP.

- The data exports will be downloaded by the Data Manager to a local file temporarily before being immediately transferred to the SNAP Analytic Team Secure, 1TB University of Melbourne Research Data Drive (Research-NAS(CIFS)) located at the following address: \\research-cifs.unimelb.edu.au\5050-snap.
- It is a standard network file share that can be mounted onto a computer
- It is only accessible via the University network or via UoM VPN connection by university staff (or honorary appointees)
- The storage infrastructure is physically located within the University's data centres in Parkville and Noble Park, Victoria, Australia.
- Access to SNAP Research-NAS drive is limited to the members of the AT and the unblinded Data Manager(s) / Data Coordinator(s). The access approver for this drive is the University of Melbourne appointed member of the AT, Robert Mahar.

The *a priori* statistical instructions for each scheduled analysis will be documented in the Statistical Implementation Guides (SIGs) produced by the blinded members of the Stats-WG. The SIG will be version controlled, with a new version produced for each scheduled analysis, incorporating any changes or new information that the AT will need to know to conduct the analysis. The SIGs will be made publicly available via the SNAP website prior to each scheduled analysis

13.1.1 DATA QUALITY CONTROL & QUERY RESOLUTION

As per the timelines specified in the AT ToR, there will be a period of time between the participant reaching day 90 and the extraction date, where the GTMG and RTMGs will finalise outstanding data quality control checks relating to the Priority Data Points (see **Table 8**).

As part of the scheduled analysis, the AT will also perform routine quality checks on the data and communicate any queries about the data to the unblinded Data Manager(s) / Data Coordinator(s) from the GTMG.

The unblinded Data Manager(s) / Data Coordinator(s) from the GTMG will work with the AT to resolve any concerns or issues with the data. If communications between the data manager(s) / data coordinator(s) and the AT are required to take place, a copy of this communication will be kept archived in the access-limited drive and archived at trial closure.

The data cut used for each scheduled analysis will serve as the locked scheduled analysis dataset. As these participants may also form part of other analyses, the data in the EDC may continue to be updated as part of subsequent data quality control.

See **Section 15.1** for more information on data locking.

13.1.2 PRODUCTION OF THE OPEN AND CLOSED REPORTS

The AT will be responsible for producing both the open (blinded) and closed (unblinded) scheduled analysis reports, as per the timelines specified in the AT ToR.

- **Closed Report (unblinded):** will only be reviewed by the Data Safety Monitoring Committee (DSMC)

- **Open Report (blinded):** will be reviewed by the GTSC and GTMG and can be accessed by the RTMGs in order to complete region-specific reporting requirements. These reports will not be shared outside of these groups, unless specified or requested by the GTMG.

13.1.3 DSMC LETTER OF RECOMMENDATION

After each scheduled analysis, the DSMC reports its recommendations in writing, usually within two weeks of the meeting.

Two letters will be produced after each DSMC meeting – a simple letter and a more detailed letter:

- **Simple Letter** – this letter should include a general statement that can be provided to the wider SNAP community.
- **Detailed Letter** – this should provide a general statement and will remain within the GTSC, GTMG and RTMGs. This letter should contain all information from the simple letter, but will also include any suggestions, information, and questions the DSMC would like to communicate.

13.2 FINAL ANALYSES THAT USE CORE STATISTICAL MODEL

Final analyses that make use of the core statistical model will be performed by the AT when a domain, cell, nested trial, or the trial as a whole has reached a conclusion, as per the prespecified stopping rules in the appendices, upon recommendation by the DSMC or other possible reasons including but not limited to participant safety, futility or funding availability.

The *a priori* statistical instructions for the final analyses will be documented in the Statistical Analysis Plans (SAPs), produced by the blinded members of the Stats-WG. The SAP will be version controlled, if any changes are made prior to the final report being produced. The SAPs will be made publicly available via the SNAP website prior to each final analysis.

These analyses will be performed using the unblinded data exports, extracted from the trial EDC(s) by the Designated Data Manager(s) / Coordinator(s) at the timepoints outlined in the AT ToR and provided to the AT via a secure, access-limited drive.

If other aspects of the trial are still open, the trial will continue to accrue data, however, further enrolment into the closed interventions of a domain, silo, etc.) that is undergoing final analysis will be prohibited.

13.2.1 DATA QUALITY CONTROL & QUERY RESOLUTION

Final analyses will be conducted when the GTMG have confirmed that all queries are resolved, and the data is verified as complete for the relevant data points. As part of the final analysis, the AT will also perform the routine quality checks on the data outlined in **Section 13.1.1**. There will be a period of time between the final participant reaching day 90 and the extraction date, where the GTMG and RTMGs will finalise outstanding data quality control checks.

The procedure for data extraction of the participant records will commence once the following data quality conditions have been met:

- Participant eligibility screening data is complete
- Priority data points (see **Table 8**) have been validated
- Data cleaning checks are performed and raised queries are resolved
- CRFs/data fields are complete or have a data override performed
- Known Protocol Deviations and Safety Events have been recorded

- Data corrections have been applied

The unblinded Data Manager(s) / Data Coordinator(s) from the GTMG will work with the AT to resolve any concerns or issues with the data. If communications between the data manager(s) / data coordinator(s) and the AT are required to take place, a copy of this communication will be kept archived in the access-limited drive and archived at trial closure.

The data cut used for the final analysis will serve as the locked analysis dataset. As these participants may also form part of another final analysis, the data in the EDC may continue to be updated as part of subsequent data quality control.

See **Section 15** for more information on data locking.

13.2.2 PRODUCTION OF THE REPORTS

The AT will be responsible for producing the final reports. These reports will be shared with the GTMG, GTSC, Stats-WG and any other relevant working groups who may be involved in producing the final manuscript. These reports will only contain primary, secondary and safety outcome data relevant to that analysis; whole-trial outcome data will remain blinded until the trial as a whole is concluded.

13.3 FINAL ANALYSES THAT DO NOT USE CORE STATISTICAL MODEL

Various domains / sub-studies and nested trials will be incorporated into the SNAP Trial throughout its lifetime that may not use the core statistical model.

These analyses may be performed on data already collected for the SNAP Trial and/or additional data collected either within the Central Spinnaker EDC or in an adjunctive EDC. Unless otherwise agreed upon, these analyses typically will not be performed by the SNAP AT; but instead by the lead, working group, or an external nominated analyst. The *a priori* statistical instructions for the sub-study analyses will be documented in the study protocols and/or the study Statistical Analysis Plan (SAP).

If data is requested from the Central Spinnaker EDC, the GTSC will provide approval of this data extraction, via the **Data Access Request Process** (See **Data Sharing Policy**). As a general rule, these analyses will not occur prior to the analysis and publication of the SNAP Trial results; exceptions may be made by the GTSC on a case-by-case basis, if the publication will not compromise trial integrity.

Where data extraction has been approved, the GTMG will liaise with the lead to organise relevant data export access. The AT will also be notified of the analysis.

A full record of all Data Extractions will be maintained by the GTMG.

See the **Authorship & Publication Policy** and **Data Sharing Policy** for further detail.

13.3.1 DATA QUALITY CONTROL & QUERY RESOLUTION

Scheduled and final analyses will be conducted when the lead(s) have confirmed that all queries are resolved, and the data is verified as complete for the relevant data points; it will be the full responsibility of the lead(s) to ensure that all relevant quality control checks have been completed prior to analysis.

If the trial is using data from the Central Spinnaker EDC, the unblinded Data Manager(s) / Data Coordinator(s) from the GTMG will ensure that the cleaning checks outlined in **Section 10** are completed and will work with the lead(s) to resolve any concerns or issues with the data. If communications between unblinded and blinded members are required to take place, a copy of this communication will be kept in the secure drive and archived at trial closure. If any files are shared that contain unblinded trial data, they will be password protected.

The data cut used for the sub-study / nested trial will serve as the locked analysis dataset. As these participants may also form part of another analysis, the data in the EDC(s) may continue to be updated as part of subsequent data quality control procedures.

See **Section 15** for more information on data locking.

13.3.2 PRODUCTION OF THE REPORTS

The nominated analysts will be responsible for producing the sub-study / nested trial analysis reports.

13.4 AGGREGATE REPORTING OF SAFETY EVENTS

To protect trial integrity, it is vital that the aggregated reports of safety events does not inadvertently unblind the investigators and/or the wider public to the primary or secondary safety outcomes of the SNAP Trial.

The primary endpoint of the trial is 90-day mortality, and the objective of the trial is to identify differences in the primary endpoint that can be attributed to treatment allocation, which will often include treatments that are believed to be or known to be safe, and effective but for which it is not known whether some treatments are more effective than others.

Therefore, details regarding trial intervention and outcome of death will be omitted from aggregate reports provided to the blinded community – this includes open reports from the interim analyses as well as region-specific reports (such as DSURs). It is recognised that while certain safety events can be omitted from the line listings and summary tabulations, these omissions should be explained.

See **Global Safety Reporting Plan** for more information.

14 DATA SHARING

The SNAP Trial is committed to responsible sharing of clinical trial data to serve the dual goals of (1) ensuring the privacy and informed consent of trial participants and (2) advancing scientific knowledge to improve public health.

A goal of the SNAP Trial is to enable other researchers and members of the community to access the SNAP dataset to maximise the benefits that can be derived from the data. The SNAP Trial acknowledges the importance of making research data publicly accessible, and as such, has developed a comprehensive data sharing policy and procedure to encourage the reuse of the data, whilst maintaining trial integrity and ensuring appropriate acknowledgments of the dataset contributors, funders and collaborators.

Data sharing includes but is not limited to providing access to SNAP Trial data for the purposes of:

- Maintaining oversight of trial progress and management of the trial
- Conducting scheduled interim and final analyses of the trial data
- Providing regional and site data to regional and site data owners as part of contractual arrangements
- Providing data to regulatory or legal authorities, upon request, to facilitate compliance and/or support investigations
- Providing data to external researchers / research organisations, upon request and relevant approval, to encourage collaboration and enhance value of the dataset

See the **Data Sharing Policy** for more information.

15 DATA LOCK, ARCHIVING & RETENTION

As a pragmatic adaptive platform trial, participant data may contribute to more than one analysis. Therefore, in general, the data extraction (data cut) used for an interim or final analysis will serve as the "locked analysis dataset" (see **Section 9.7**).

Therefore, in the context of this trial, "data lock" will refer to the procedures that are required to take place prior to data archiving.

15.1 DATA LOCK PROCEDURES

Data lock is the process by which all modifications to the data within the EDC systems are halted.

Full data lock procedures will be undertaken following closure/completion of:

- A sub-study / nested trial
- The trial (including all domains and sub-studies) at a trial site
- The entire trial (including all domains and sub-studies) at all trial sites in a region

Data lock procedures are not required following closure/completion of a domain or nested study as data fields within a domain/nested study are often used across multiple domains and are embedded within shared CRFs. It is important for the data integrity that the fields are continuously updated throughout the duration that these fields are used within these domain / nested study analyses.

Data lock occurs in two distinct stages, "soft lock" and "hard lock".

- **Soft Lock (see section 13.1.1):** A temporary or preliminary lock placed on data. It restricts further edits but can be reversed if necessary
- **Hard Lock (see Section 13.1.2):** A final, irreversible lock on the data by the EDC Provider. Once this lock is applied, no changes can be made without undergoing a formal unlock process (See **Section 13.1.3**).

Specific data lock procedures will be outlined in the **EDC-specific documentation** as required.

15.1.1 SOFT LOCK

The soft lock is a temporary or preliminary lock on the data that will be performed by the appropriately delegated RTMGs. It restricts further edits but can be reversed if necessary. This type of lock is applied when data cleaning is complete yet still allows for adjustments if errors or discrepancies are later identified.

At a minimum, soft lock should be applied as part of the full data lock procedures. RTMGs can implement additional soft lock requirements above the minimum requirements as per local procedures and regulations

Depending on the EDC, soft lock may be applied to individual CRFs and/or to individual participants and/or to individual recruiting sites. The procedure for soft data lock should commence once the following data quality conditions have been met as relevant:

- Participant eligibility screening data is complete
- Participant consent has been Source Data Verified (SDV'd) as per the Region-Specific Monitoring Plans.
- Monitoring visits have been completed and queries resolved

- Priority datapoints (see **Table 8**) have been validated as per **Section 10**
- Data cleaning checks and raised queries are resolved
- CRFs/data fields are complete or have a data override performed
- Known Protocol Deviations and Safety Events have been recorded
- Data corrections have been applied

Data quality will be improved using the methods outlined in **Section 10** and repeated until the above conditions are met. Once the data quality conditions have been met and documented, the soft data lock will occur for the site(s).

15.1.2 HARD LOCK

A hard lock is a final, irreversible lock on the data by, approved by the EDC-RP in consultation with the GTMG and actioned by the EDC Provider. Once this lock is applied, no changes can be made without undergoing a formal unlock process. A hard lock is implemented after all data cleaning, reviews, and verifications are complete, and the dataset is ready for archiving.

Once all sites have undergone soft data lock, the EDC will undergo a final round of cleaning and quality control to ensure no further changes are required. As locked data should only be updated in extraordinary circumstances, comprehensive efforts will be made to ensure that data is as clean as possible prior to hard lock.

15.1.3 FORMAL UNLOCK PROCESS FOR LOCKED DATA

There may be legitimate situations where previously hard-locked data needs to be unlocked for corrections or updates; as such, this process below applies to all hard-locked data in the SNAP Trial EDCs. It is the responsibility of the GTMG, in collaboration with the EDC-RPs and RTMGs, to ensure that the unlocking process is followed and documented appropriately.

The steps for performing a data unlock on hard-locked data is as follows:

1. **Unlock Request Submission:** Any request to unlock hard-locked data must be formally submitted to the GTSC via the GTMG.

   The request must include the following information:

   - Subject ID(s) and visit(s) for which data unlock is required
   - Specific form(s) or field(s) to be unlocked
   - Justification for the unlock request
   - Any supporting documentation (e.g., data clarification forms, medical records)
2. **Review and Approval:** the GTSC will review the unlock request and supporting documentation.
   - If the request is deemed valid and justified, the GTSC will approve the request.
   - If the request is not approved, the GTMG will provide feedback to the requestor, and the process will end.
3. **Data Unlock:** Upon approval by the GTSC, the GTMG will coordinate with the relevant EDC-RP(s) to unlock the specified data in EDC(s).
4. **Data Update and Relock:** Once the data has been unlocked, the requestor (or nominated delegate) will make the necessary or corrections to the data. After the updates are complete, the EDC-RP(s) in collaboration with the GTMG will review the changes and ensure data integrity.

If the changes are approved, the EDC-RP(s) will coordinate with the EDC Provider(s) to re-lock the updated data.

5. **Documentation and Reporting:** The unlock process will be formally documented in a Note to File (NTF) that is maintained in the Global TMF. All unlock requests, approvals, data updates, and relocks will be thoroughly documented by the EDC-RP(s) in collaboration with the GTMG in a NTF that will be kept in the Global TMF. Any significant changes to the data should be reported to the relevant sponsors, DSMC, AT and RTMGs.

Full documentation of the above steps should be compiled and saved in the Trial Master File by the GTMG.

15.2 DATA ARCHIVING AT CLOSURE

The GTMG in collaboration with the RTMGs will ensure the following data management tasks are completed in accordance with ICH-GCP, where relevant, at archiving:

- Perform data lock (see **Section 15.1**)
- Undertake data cleaning and quality control procedures (see **Section 10.0**)
- Work with the EDC Provider(s) to arrange the archiving of the EDC(s) and ensure the full package of data exports are saved to the SNAP secure drive as the global coordinating centre is the custodian of this complete data set
- Ensure a complete copy of each region's data is provided to the relevant RTMG for archiving in accordance with their regional archiving procedures
- Provide, upon request, a copy of each site's data to the requesting site for archiving in accordance with the site's regional and local archiving procedures.
- Work with the AT to ensure all data exports and code/methods used in the analyses are saved to the SNAP secure drive.

Bulk electronic sign off on all EDC CRFs within the EDC system is required at the following timepoints as a minimum global requirement, commencing from the timepoint that electronic sign-off is available within the EDC:

- At each final analysis (prior to each final analysis data cut)
- At site closure

Electronic sign off on all EDC CRFs within the EDC system is required at site closure as a minimum global requirement, commencing from the timepoint that electronic sign-off is available within the EDC. If electronic sign-off is not available at the time of site closure, sign off on all EDC CRFs within the EDC system is expected at site closure by documentation on the Site Closure Visit Report and/or by application of signature to a PDF version of the CRFs.

Additional electronic sign off within the EDC system by site PIs and/or RTMGs will not be required as part of the global minimum requirement but may be implemented at a regional level as required.

15.3 DATA RETENTION

Appropriate retention of trial data is critical for ensuring regulatory compliance, enabling auditing, and allowing for potential re-analysis or secondary use of data after trial completion.

15.3.1 DURING THE TRIAL

Data will be retained for the duration of the trial and retention period from trial initiation through data lock. Data must be retained by all parties involved in a secure, accessible manner and remain available for monitoring, auditing, and inspection by regulatory authorities.

15.3.2 AFTER REGIONAL TRIAL COMPLETION

After the trial is complete in a region and the data is locked, all trial data (electronic, paper, and multimedia) will be retained by the RMTGs on behalf of the International Sponsor Organisation (ISO) at a minimum according to the retention period required to comply with regional requirements for data retention. The retention period will commence from the recruitment of the final participant to the SNAP platform or registry in that region.

The complete trial data package will be archived securely according to ISO requirements; each ISO will retain their region's data for at least the minimum period according to their regional requirements.

15.3.3 AFTER GLOBAL TRIAL COMPLETION

After the trial is complete and the data is locked, all trial data (electronic, paper, and multimedia) will be retained by the Global Coordinating Centre at a minimum according to the longest retention period of the regions involved in the trial to comply with all regional requirements for data retention; at present, this is 25 years. The retention period will commence from the recruitment of the final participant to the SNAP platform or registry.

The complete trial data package will be archived securely in an electronic archive managed by University of Melbourne according to The University of Melbourne Records Management Policy (MPF1106).

15.3.4 LONG-TERM RETENTION

After the retention period, archived data may be further retained indefinitely if still required for further research purposes. Data archiving systems will be validated to ensure data integrity is maintained for the full retention period and procedures will be implemented to ensure archived data can be retrieved and migrated to new archival systems as needed.

15.4 DISPOSAL AND DESTRUCTION OF DATA

After the suitable retention period and if the data are no longer useful or cannot be or are unlikely to be used in further research, the sponsor will destroy it according to ICH-GCP and International Sponsor Organisation institutional procedures.

Approval from the GTSC, Coordinating Investigators or the Global Coordinating Centre Sponsorship Office will be obtained prior to destruction of the data. The destruction will be formally documented in a NTF and saved in the Trial Master File (TMF).

16 DATA STORAGE & SECURITY

All trial data will be stored in secure, limited-access environments to protect against unauthorized access, loss, or alteration. The following procedures will be followed for data storage:

16.1 PROTECTING DATA DURING TRANSMISSION

Participant data provided from clinical trial sites will be entered directly into the secure EDC systems used within the SNAP Trial, accessible only to authorised users with appropriate credentials.

If other data transfer is required between clinical sites, sponsors, and other stakeholders, this will be performed using password-protected secure file transfer services to maintain data confidentiality and security.

16.2 DATA STORAGE LOCATIONS

The data storage locations of the data in the SNAP Trial will be outlined in the section or in the corresponding EDC-specific documentation and will be available upon request from the EDC Provider. It will be the responsibility of the EDC-RP to confirm that the data storage locations used by the EDC are compliant and documented.

- **EDC data** will be stored on encrypted, password-protected servers maintained by the EDC Provider. The details of the storage location will be documented in the EDC-specific documentation.
- **Paper source data** will be stored in a locked, access-controlled room at the trial site, maintained by the site investigator team. Access to the paper source data will be limited to authorised trial site staff members and trial monitors.
- **Data exports containing unblinded data** will be stored on a secure, access-limited drive for the duration of the trial.
- **Data exports containing blinded data** will be stored on a secure, access-restricted drive, managed by the GTMG.

16.3 DATA ENCRYPTION

The data encryption methods implemented in the SNAP Trial will be outlined in the corresponding EDC-specific documentation and will be available upon request from the EDC Provider. It will be the responsibility of the EDC-RP to confirm that appropriate data encryption methods are implemented in the EDC and are compliant and documented.

16.4 BACKUP AND RECOVERY

Regular data backups will be performed to ensure data is not lost in the event of system failure or natural disaster. A recovery plan will be in place to restore data to a secure environment if needed.

All trial EDCs and electronic data files will be fully backed up daily at a minimum.

The data backup and recovery methods implemented in the SNAP Trial will be outlined in the corresponding EDC-specific documentation and will be available upon request from the EDC Provider. It will be the responsibility of the EDC-RP to confirm that appropriate data backup and recovery methods are implemented in the EDC and are compliant and documented.

16.5 SECURITY MONITORING, AUDITING & REPORTING

The security monitoring and auditing procedures implemented in the SNAP Trial will be outlined in the corresponding EDC-specific documentation and will be available upon request from the EDC Provider. It will be the responsibility of the EDC-RP to confirm that appropriate security methods are implemented in the EDC and are compliant and documented.

Reports on security events, audit findings, vulnerabilities, and overall risk posture will be generated by the EDC Provider. Any suspected incidents or data breaches will be reported immediately by the EDC Provider to the EDC-RP and communicated to the GTMG, so that the incident can be reported to the RTMG, GTSC, ethical review bodies and regulatory authorities.

16.6 RESPONDING TO DATA BREACHES AND SECURITY INCIDENTS

A data breach or security incident is defined as an event that results in the unauthorised access, disclosure, alteration, or destruction of data, including personal data. Breaches can occur in various forms, such as:

- Unauthorised access to data (e.g., a security loophole or compromised login credentials)
- Accidental exposure of data (e.g., sending data to the wrong recipient)
- Physical theft or loss of devices containing sensitive information
- Data corruption due to system failure or cyberattacks

Any suspected incidents or data breaches will be reported immediately and within 24 hours to the GTMG so that the incident can be reported to the RTMG, GTSC, ethical review bodies and regulatory authorities within 72 hours, if applicable.

16.6.1 DATA BREACH OR SECURITY INCIDENT PROTOCOL

In the event a data breach or security incident is suspected or identified, the following protocol will be immediately initiated by the EDC-RP in collaboration with the relevant RTMGs, EDC Provider, and/or trial site team as required, in order to identify the cause, mitigate the impact, and notify affected participants if necessary.

It is the responsibility of the EDC-RP, in collaboration with the GTMG and RTMGs to ensure that the data breach or security incident protocol followed aligns with the Global Coordinating Centre's guidelines as described in this Data Management Plan.

16.6.1.1 IMMEDIATE CONTAINMENT AND INVESTIGATION

The first step in responding to a data breach is to contain the breach and prevent further unauthorised access. This includes:

- **Securing Data**: The affected systems will be isolated, and access to the compromised data will be restricted to prevent further exposure.
- **Investigating the Incident**: A thorough investigation will be conducted to determine the nature, scope, and cause of the breach. This investigation will involve technical personnel to assess whether data was accessed, altered, or lost, and the identity of the individuals or systems involved.
- **Performing a Root Cause Analysis**: A root cause analysis will be carried out to identify how the breach occurred (e.g., human error, technical vulnerabilities, or malicious activity) and whether additional preventive measures are necessary to avoid recurrence.

16.6.1.2 SECURITY AND RISK ASSESSMENT

Once the breach is contained, the potential impact on participants' privacy must be evaluated, and the potential risks assessed. This includes:

- **Assessing the Severity**: Evaluating the type of data involved in the breach (e.g., personally identifiable information, medical records, etc.) and the extent of the exposure (e.g., how many individuals were affected and whether sensitive information was compromised).
- **Assessing Impact on Participants**: Assessing the potential harm to affected participants, including whether they are at risk of identity theft, fraud, or other adverse consequences.

16.6.1.3 NOTIFY REGULATORY AUTHORITIES & ETHICAL REVIEW BODIES

Review whether the breach must be reported to relevant regulatory authorities and ethical review bodies. If determined that reporting is required, ensure that reporting is done within the specified timelines. The notification will include:

- **Details of the Breach**: A clear description of the breach, including the nature of the data compromised, how it occurred, and the potential impact on participants.
- **Actions To Be Taken**: Information on the steps that will be taken to mitigate the breach and prevent future incidents, such as changes to security measures or system fixes.
- **Notification of Participants**: If notification to participants is required, and the manner by which this will be done. A copy of the notification to participants should be provided for review/approval prior to distribution.

If the breach involves personal data, the relevant ISO's Privacy and Data Protection Officer may be required to be notified.

16.6.1.4 NOTIFY AFFECTED PARTICIPANTS

If the breach involves personal data and poses a risk to participants' rights and/or safety, affected participants will be notified promptly. The notification will include:

- **Details of the Breach**: A clear description of the breach, including the nature of the data compromised, how it occurred, and the potential impact on participants.
- **Action Taken**: Information on the steps taken to mitigate the breach and prevent future incidents, such as changes to security measures or system fixes.
- **Advice and Support**: Guidance for affected individuals on steps they can take to protect themselves (e.g., monitoring credit reports, changing passwords, or reporting suspicious activity). Participants may also be offered free credit monitoring or other protective services, depending on the severity of the breach.
- **Contact Information**: Contact details for the trial site team or Privacy and Data Protection Officer (PDPO) for any further questions or concerns.

16.6.1.5 CORRECTIVE ACTIONS AND PREVENTION

Following a data breach, corrective actions will be implemented to minimise the risk of similar incidents occurring in the future. These may include:

- **Improved Security Measures**: Updating or strengthening technical and organisational security measures, such as encryption, access controls, and multi-factor authentication, to protect sensitive data.
- **Staff Training**: Conducting refresher training for all trial personnel on data security best practices, recognising potential security threats, and following proper protocols for data handling and breach prevention.
- **System Updates**: Installing patches or updates to address any software vulnerabilities or weaknesses identified during the breach investigation.
- **Review of Audit and Monitoring Procedures**: Enhancing monitoring systems to detect any unusual activity or signs of potential breaches in real time. Regular audits will be conducted to ensure continued compliance with data protection policies and protocols.

16.6.1.6 ONGOING MONITORING AND REPORTING

After the breach has been resolved, ongoing monitoring will be implemented to assess the effectiveness of the corrective actions. Any new risks or vulnerabilities identified will be promptly addressed. The data management plan will be reviewed and updated as necessary to reflect lessons learned and improve future responses to similar incidents.

16.6.1.7 FULL DOCUMENTATION OF THE INCIDENT

Full documentation of the above steps should be compiled by the EDC-RP and saved in the Trial Master File by the GTMG and relevant RTMGs.

17 APPENDICES

17.1 APPENDIX 1: RECORD OF CHANGES

VERSION NUMBER	VERSION DATE	SECTIONS AFFECTED	SUMMARY OF CHANGES	LEAD AUTHOR
1.0	27-Jun-23	-	-	Jay Keatley / Jocelyn Mora
2.0	31-Mar-26	All	Substantial overhaul to all sections to outline the overarching data management plan for all EDCs used within the SNAP Trial	Lauren Barina
2.1	05-Jun-26	10.4	Administrative update to correct textual error where part of the sentence was missing.	Lauren Barina

17.2 APPENDIX 2: SUMMARY OF DATA MANAGEMENT RESPONSIBILITIES

Key:

- **R:** Primary Responsibility
- **S:** Support or Feedback responsibility
- **A:** Approval responsibility

1 = if the EDC is implemented in more than 1 region

2= if the EDC will make use of the core statistical model for analyses

3 = if the EDC is implemented in a single region and will not make use of the core statistical model for analyses

4 = a qualified and designated safety reviewer will be delegated to review, confirm and add MEDdra code

5 = pertaining to the sub-study / nested trial

6 = data cleaning reports will be produced by the central team to assist, but the primary responsibility for oversight of data quality rests with the RTMGs

Table 11 Data Management Responsibilities

Category	Data Management Responsibilities	EDC Provider	EDC-RP	GTMG	RTMGs	GSTC	DSMC	Other WGs	Sites	Stats-WG / Nominated Statistician	AT / Nominated Analysts	Sub-Study or Nested Trial Lead
Documentation & Processes	Review of Trial Integrity Documentation pertaining to data management			R	S	A				A	A	

Category	Data Management Responsibilities	EDC Provider	EDC-RP	GTMG	RTMGs	GSTC	DSMC	Other WGs	Sites	Stats-WG / Nominated Statistician	AT / Nominated Analysts	Sub-Study or Nested Trial Lead
	Creation and update of the EDC completion guidance, access request, and other EDC operational document(s)	S	R	A[1,2]	A[3]							
	Establishment and management of an EDC change process during the life of the trial	S	R	A[1,2]	A[3]							
Electronic Data Capture (EDC) Systems	Validation of EDC randomisation algorithm	S	R	A[1,2]	A[3]					S		
	Design of the EDC including the data dictionary (DD), user settings, content, user interface, etc.	S	R	A[1,2]	A[3]						S	
	Build, validation and testing of the initial EDC and subsequent amendments/changes prior to releasing to live, including completion of validation statement	S	R	A[1,2]	A[3]					S		

Category	Data Management Responsibilities	EDC Provider	EDC-RP	GTMG	RTMGs	GSTC	DSMC	Other WGs	Sites	Stats-WG / Nominated Statistician	AT / Nominated Analysts	Sub-Study or Nested Trial Lead
	Maintenance of data integrity through the data life cycle	R	R	R	R	R	R		R	R	R	R
	Establishment of EDC site settings and management of subsequent changes to the site settings	S	R	S	R							
	Establishment of EDC users and management of subsequent changes to the site users	S	R	S	R							
	Provision of EDC support to sites and monitors	R	R	S[1,2]	S[1,3]							
Participant CRFs / Records	Timely completion of EDC CRFs	S	S	S[1,2]	S[1,3]				R			
	Enrolment of participants and data entry	S	S	S[1,2]	S[1,3]				R			
	Patient record deletion	S	R	S[1,2]	A				A			
	Mark records as complete (where data cannot be completed)	S	R	S[1,2]	A				A			

Category	Data Management Responsibilities	EDC Provider	EDC-RP	GTMG	RTMGs	GSTC	DSMC	Other WGs	Sites	Stats-WG / Nominated Statistician	AT / Nominated Analysts	Sub-Study or Nested Trial Lead
	Review of participant safety data and MedDRA coding of confirmed safety events			S	R[4]				S			
Data Validation & Quality	Data Quality Control Summaries[7]			R								R[5]
	Data Monitoring			S	R				S			R[5]
	Timely data query generation and resolution			S	R				R			S[5]
	Data correction, editable fields	S			S				R			
	Data correction, fields that are locked	S	R	S[1,2]	S[1,3]				A			
Locking and Closure	Soft lock of CRFs / participant records	S	S	S	R				A			
	Hard lock of CRFs / participant records	R	R	A[1,2]	A[1,3]				A			
	Locking sites upon closure (includes deactivating enrolment and removing access users)	S	R	S	A							S[5]

Category	Data Management Responsibilities	EDC Provider	EDC-RP	GTMG	RTMGs	GSTC	DSMC	Other WGs	Sites	Stats-WG / Nominated Statistician	AT / Nominated Analysts	Sub-Study or Nested Trial Lead
Analysis	Data extraction from the EDC and provision of extracts for analysis	S	R	S[1,2]		S	A				A	S[5]
	Data analysis, preparation, and circulation of the interim and/or final analysis reports			A[1,2]						S	R	A[5]
	Data review, recommendations and circulation of the DSMC letter of recommendations			S	S	A	R				S	S[5]
	Responding to DSMC queries and recommendations			R[1,2]		S	A			S		R[5]
	Management of the secure location for containing trial data for analysis and access approval			R[1,2]		S					A	R[5]
EDC Support	EDC Access Request	S	R	S	A							
	EDC Support Request	S	R	S	S							S[5]

Category	Data Management Responsibilities	EDC Provider	EDC-RP	GTMG	RTMGs	GSTC	DSMC	Other WGs	Sites	Stats-WG / Nominated Statistician	AT / Nominated Analysts	Sub-Study or Nested Trial Lead

KEY:

- R: Primary Responsibility
- S: Support or Feedback responsibility
- A: Approval responsibility

1 = if the EDC is implemented in more than 1 region

2= if the EDC will make use of the core statistical model for analyses

3 = if the EDC is implemented in a single region and will not make use of the core statistical model for analyses

4 = a qualified and designated safety reviewer will be delegated to review, confirm and add MEDdra code

5 = pertaining to the sub-study / nested trial

6 = data cleaning reports will be produced by the GTMG/sub-study/nested trial leads, but the primary responsibility for oversight of data quality rests with the RTMGs

17.3 APPENDIX 3: TIMESCALE OF DATA MANAGEMENT ACTIVITIES

Table 12 Timescale of Data Management Activities

Category & Detail	Primary Responsibility (R)	Timescale
Documentation & Processes		
Review of the Trial Integrity Documentation pertaining to data management	GTMG	Annually at a minimum
Creation and update of the EDC completion guidance, access request, and other EDC operational document(s)	EDC-RP	Ad hoc as required
Establishment and management of an EDC change process during the life of the trial	EDC-RP	At EDC implementation
Electronic Data Capture (EDC) Systems		
Validation of EDC randomisation algorithm	EDC-RP	At EDC implementation, at a minimum Ad hoc as changes to the interventions are made
Design of the EDC including the data dictionary (DD), user settings, content, user interface, etc.	EDC-RP	At EDC implementation
Build, validation and testing of the initial EDC and subsequent amendments/changes prior to releasing to live, including completion of validation statement	EDC-RP	At EDC implementation, at a minimum Ad hoc as changes to the interventions are made
Maintenance of data integrity through the data life cycle	All	Ongoing
Establishment of EDC site settings and management of subsequent changes to the site settings	EDC-RP & RTMG	At site initiation and subsequent site updates
Establishment of EDC users and management of subsequent changes to the site users	EDC-RP & RTMG	At site initiation and subsequent site staff updates
Provision of EDC support to sites and monitors	EDC-RP	Ongoing
Participant CRFs / Records		
Timely completion of EDC CRFs	Sites	As per the timelines outlined in the protocols and/or EDC-specific documentation.
Enrolment of participants and data entry	Sites	As per the timelines outlined in the protocols
Patient record deletion	EDC-RP	Ad hoc as required
Mark records as complete (where data cannot be completed)	EDC-RP	Ad hoc as required
Review of participant safety data and MedDRA coding of confirmed safety events	RTMG	As per Safety Management Plan

		The MedDRA coding is required to be added as per timelines outlined in region-specific safety management documentation.
Data Validation & Quality		
Data Quality Control Summaries	GTMG & Sub-Study / Nested Trial Lead	Ongoing
Data Monitoring	RTMG	Ongoing
Timely data query generation and resolution	RTMGs & Sites	Within 14 days of query raised
Data correction, editable fields	Sites	Within 14 days of query raised
Data correction, fields that are locked	EDC-RP	Prior to final analyses
Locking and Closure		
Soft lock of CRFs / participant records	RTMG	Upon site closure at a minimum *Sites will be locked, and access removed at the point that the site has been closed to recruitment and there are no outstanding data queries.*
Hard lock of CRFs / participant records	EDC-RP	Prior to final archiving of the EDC
Locking sites upon closure (includes deactivating enrolment and removing access users)	EDC-RP	Upon site closure *Sites will be locked, and access removed at the point that the site has been closed to recruitment and there are no outstanding data queries.*
Analysis		
Data extraction from the EDC and provision of extracts for analysis	EDC-RP	At each analysis *As per the timelines outlined in the protocols and/or Statistical Analysis Plan and/or Terms of Reference documents.*
Data analysis, preparation, and circulation of the interim and/or final analysis reports	AT / Nominated Analysts	At each analysis *As per the timelines outlined in the protocols and/or Statistical Analysis Plan and/or Terms of Reference documents.*
Data review, recommendations and circulation of the DSMC letter of recommendations	DSMC	At each analysis

		As per the timelines outlined in the protocols and/or Terms of Reference documents.
Responding to DSMC queries and recommendations	GTMG & Sub-Study / Nested Trial Lead	At each analysis and prior to the next analysis.
Management of the secure location for containing trial data for analysis and access approval	GTMG & Sub-Study / Nested Trial Lead	Ongoing
EDC Support		
EDC Access Request	EDC-RP	Within 5 working days of receipt of the request
EDC Support Request	EDC-RP	Within 5 working days of receipt of the request or as detailed in the EDC-specific documentation.
Urgent EDC Support Request	EDC-RP	Within 2 working days.

17.4 APPENDIX 4: SCHEDULE OF KEY EXTRACTION EVENTS

The schedule of the key extraction events for the analyses will occur as specified in the statistical documentation.

This table will be updated with each version of the Data Management Plan.

Table 13 Key Extraction Events

Scheduled Interim Analyses					
Analysis	**Milestone**	**Date Recruited**	**Date Reached Day 90**	**Date Extraction**	**Date of DSMC Meeting**
Analysis 1	500th participant	03-Mar-2023	31-May-2023	14-Jun-2023	28-Jun-2023
Analysis 2	1,000th participant	04-Jul-2023	02-Oct-2023	16-Oct-2023	01-Nov-2023
Analysis 3	1,500th participant	13-Oct-2023	10-Jan-2024	24-Jan-2024	07-Feb-2024
Analysis 4	2,000th participant	25-Jan-24	28-Apr-2023	07-May-2024	04-Jun-2024
Unscheduled	-	-	-	07-May-2024	05-Aug-2024
Analysis 5	2,500th participant	24-Apr-2024	22-Jul-2024	05-Aug-24	26-Aug-2024
Analysis 6	3,000th participant	24-Jul-2024	21-Oct-2024	04-Nov-24	02-Dec-2024
Analysis 7	3,500th participant	25-Oct-2024	22-Jan-2025	05-Feb-25	13/May-2025
Analysis 8	4,000th participant	7-Feb-2025	7-May-2025	21-May-25	18-Jun-2025
Analysis 9	4,500th participant	16-May-2025	13-Aug-2025	27-Aug-25	08-Oct-2025
Analysis 10	5,000th participant	19-Aug-2025	16-Nov-2025	30-Nov-25	07-Jan-26
Analysis 11	5,500th participant	12-Dec-2026	11-Mar-2026	25-Mar-2026	29-Apr-2026

Final Analyses				
Analysis	**Milestone**	**Domain/Sub-Study /Platform Details**	**Last Participant Recruited**	**Data Extraction**
Domain Analysis	Last participant recruited to a domain	Backbone Domain – MSSA Silo (Adults): Cefazolin vs (flu)cloxacillin	21-Jun-2024	29-Nov-2024
		Backbone Domain – PSSA Silo (Adults & Paediatrics): Penicillin vs (flu)cloxacillin	21-Jun-2024	29-Nov-2024
		Adjunctive Domain (Adults & Paediatrics): No adjunctive cefazolin vs adjunctive cefazolin	17-Jul-2025	14-Jan-2026

17.5 APPENDIX 5: VALIDATION STATEMENT TEMPLATE

SNAP Trial Electronic Data Capture (EDC) Validation Statement

EDC Release Specifications

EDC Name	
EDC Provider	
EDC Responsible Party (EDC-RP)	
Release Name / Version	
Release Date	
Testing Dates	
Description of Release	

EDC Randomisation Algorithm [Remove Section if Not Relevant to EDC]

[Provide a description of the randomization algorithm used. Reference the relevant section(s) of protocol/documentation that specify the randomization requirements.]

Validation of the Randomisation Algorithm

[Provide a description of the validation steps performed to ensure the algorithm is working correctly including:

1. **Simulation of Randomisation:** Test randomisations will be generated using the randomisation software to simulate the trial and identify potential errors, imbalances, or other issues.
2. **Validation of Randomisation Algorithm (where code is available*):** The randomisation algorithm code should be provided for validation wherever possible. At a minimum, the algorithm should be checked by a statistician independent of the code development.

OR

3. **Validation of Randomisation Output (where code is not available*):** Where the randomisation algorithm code is not available, the randomisation output will be reviewed to ensure that the allocations are balanced in accordance with the specified strategy (e.g., proper stratification and block sizes)]

**If the code cannot be provided (e.g. IP reasons), point 3 will be performed instead of point 2.*

Randomisation Allocation Concealment

[Provide a description of the checks performed to ensure that allocation concealment is maintained until the point of allocation reveal, and that the allocation assignments are only available to unblinded members with permission to view intervention allocations.]

EDC Specifications

[Specifications of the release are outlined in the following documents: / insert specifications] or detail using sections below]

EDC Password Policy

[Describe the password policy requirements implemented for the database, including criteria like length, complexity, expiration, etc. Alternatively, reference the document that details the password policy.]

User Profiles and Permissions

User profiles and permissions have been specified as follows:

[Provide list or table of the different user roles/profiles and a high-level description of the permissions/access granted to each]

Alerts [Remove Section if Not Relevant to EDC]

The following on-screen and email alerts have been configured: [Provide list or table of on-screen or email alerts and emails that will be implemented within the EDC]

Other [Remove Section if Not Relevant to EDC]

[insert as required]

EDC Testing

[insert details of when the testing of the EDC was completed and when and by who]

User Profiles and Permissions

The user profiles and permissions in the EDC has been tested and confirmed to be working as expected based on the requirements.

[Elaborate on the testing performed, where applicable]

Data Field Mapping to Protocol

The data fields in the EDC adequately collect the outcomes and other relevant data as specified in [PROTOCOL/APPENDIX NAME & VERSION].

[Elaborate on the testing performed, where applicable. Optionally, include a table mapping the protocol outcomes and data collection specifications to the corresponding data fields.]

Show/Hide Logic

The show/hide logic for data fields has been tested and confirmed to be working as expected based on the requirements.

[Elaborate on the testing performed, where applicable]

Built-In Verifications for Data Entry

The built-in verifications for data entry have been tested and confirmed to be working as expected based on the requirements.

[Elaborate on the testing performed, where applicable]

Alerts [Remove Section if Not Relevant to EDC]

The On-Screen and Email alerts have been tested and confirmed to be working as expected based on the requirements.

[Elaborate on the testing performed, where applicable]

Data Exports

The data exports generated from the EDC have been tested and confirmed to be working as expected based on the requirements.

[Elaborate on the testing performed, where applicable]

Other [Remove Section if Not Relevant to EDC]

[insert additional aspects as required]

Post-Release Validation [Remove Section if Not Relevant to EDC]

[insert any testing that took place on the live environment post-release]

EDC Alignment with Data Dictionary

It is confirmed by [NAME/ROLE] that the EDC aligns with the with the specifications outlined in the Data Dictionary [VERSION NUMBER AND DATE] for the following aspects:

- [insert]

It is confirmed by [NAME/ROLE] that the following aspects do not align with the EDC specifications, and that these discrepancies will be addressed in the following database release [remove if not relevant to EDC]

- [insert]

Outstanding Issues at Release

[Outstanding issues are described here]

Training & Updated Documentation

[document any training or updated documentation that is done as part of the release]

Back Up of the Exports

	Folder Name	File Size
Before release		
After release		

EDC-RP Validation Statement Sign Off

I confirm that the validation activities described in this document have been completed, and that the EDC has been built and tested in accordance with the current protocol, data dictionary, and applicable regulatory requirements.

I confirm that the EDC complies with the SNAP Trial Data Management Plan and the system is qualified for use within the SNAP Trial.

Signature			
Name		**Date of Signature**	