

## SOP 855 Clinical Data Management System - MANIPULATION OF DATA AFTER EXPORT

<b>For Use in:</b>	Research
<b>By:</b>	All staff
<b>For:</b>	All staff involved in the conduct of research
<b>Division responsible for document:</b>	Research & Development
<b>Key words:</b>	Data Management Emergency Unblinding
<b>Name of document author:</b>	Martin Pond
<b>Job title of document author:</b>	Head of Data Management, Norwich Clinical Trials Unit, UEA
<b>Name of document author's Line Manager:</b>	Prof Ann Marie Swat
<b>Job title of author's Line Manager:</b>	Director of Norwich Clinical Trials Unit
<b>Supported by:</b>	Julie Dawson NNUH Sarah Ruthven UEA
<b>Assessed and approved by:</b>	Julie Dawson: Research Services Manager NNUH Sarah Ruthven: Research Manager UEA
<b>Date of approval:</b>	27/10/2020
<b>To be reviewed before:</b> This document remains current after this date but will be under review	27/10/2023 (3 years, unless legislation or process changes)
<b>Reference and / or Trust Docs ID No:</b>	14286
<b>Version No:</b>	V3
<b>Description of changes:</b>	Updated to new SOP template

This Standard Operating Procedure (SOP) is available on the Research & Development pages on the NNUH website

Copies printed from the website are only valid on the day of printing.

# SOP 855 Clinical Data Management System - MANIPULATION OF DATA AFTER EXPORT

## 1. Contents

Section	Page
1. Contents	2
2. Definitions of Terms Used / Glossary	2
3. Objectives	2
4. Scope	3
5. Purpose	3
6. Rules	3
7. Procedure	3
8. Approval	4
9. Reason for Update & Training Implication	4

## 2. Definitions of Terms Used / Glossary

CDMS	Clinical Data Management System
CI	Chief Investigator
CTU	Clinical Trials Unit
DBM	Database Manager
GCP	Good Clinical Practice
ICH	International Conference for Harmonisation
NNUH	Norfolk and Norwich University Hospital
PI	Principal Investigator
R&D	Research and Development
SAS	Statistical Software for data management, advanced analytics
SM	Study Manager
SOP	Standard Operating Procedure
STATA	Software for Statistics and Data Science
TT	Tools and Templates
UEA	University of East Anglia

## 3. Objectives

It is recognised that substantial data manipulation will often be required, typically by the trial statistician once the trial database has been locked and data exported to statistical software for analysis. This may include the imputation of missing values, the creation of new variables, the exclusion of individuals from analysis, the changing of data values for a valid reason, etc.

This SOP has been written to outline necessary steps to be taken when manipulating the data after export.

# SOP 855 Clinical Data Management System - MANIPULATION OF DATA AFTER EXPORT

## 4. Scope

This SOP applies to all research managed by Norwich CTU however the principles contained in this SOP should be followed for all other trials.

## 5. Purpose

This SOP defines the procedure for recording data manipulation after export from the trials database.

## 6. Rules

### Data Manipulation

Creates the possibility for data errors and therefore should be carefully considered and conducted.

Changes should never be made 'by hand' in a manner that cannot be traced.

## 7. Procedure



- Any manipulation of data must be recorded in an appropriate manner. Most typically, this will be recorded in the code of programs written for analysis (e.g. in STATA, SAS).



- Manipulations should be highlighted, where possible, by comments in the code



- Data sets that contain manipulated data should be named in a manner that cannot be confused with any originally exported data.



- Documentation should be kept of dataset 'flow', i.e. input names, programme details, output names, to trace where data changes have taken place.



- Consistency checks should be carried out at each stage of data manipulation (which may include changes to one or more variables at a time) to highlight any possible errors.

- For example, if a new variable (NEWVAR) is created as the average of three existing variables (VAR1 – VAR3), then a proportion of individuals should be manually checked to ensure in each case that NEWVAR is indeed the average of VAR1-VAR3.

## SOP 855 Clinical Data Management System - MANIPULATION OF DATA AFTER EXPORT

### 8. Approval

<b>Author</b>	Martin Pond
<b>Role</b>	Head of Data Manager NCTU UEA
<b>Signature</b>	<i>Martin Pond</i>
<b>Date</b>	23/07/2020
<b>Approved &amp; Authorised NNUH</b>	Julie Dawson
<b>Role</b>	Research Services Manager
<b>Signature</b>	<i>Julie Dawson</i>
<b>Date</b>	27/10/2020
<b>Approved &amp; Authorised UEA</b>	Sarah Ruthven
<b>Role</b>	Research Manager
<b>Signature</b>	<i>Sarah Ruthven</i>
<b>Date</b>	01/07/2020

### 10. Reason for new version and Training Implication

This SOP replaces the previous version number V2.2

<b>Changes made</b>	What changes have been made to the contents of the document
<b>Reason</b>	<ul style="list-style-type: none"> <li>• New layout</li> <li>• Full Review</li> </ul>
<b>Training Implication</b>	<b>No</b>
<b>Actions required</b>	<ul style="list-style-type: none"> <li>• n/a</li> </ul>