The use of Stata `putdocx` for automating Data Safety Monitoring Committee reports

Alannah Rudkin

murdoch children's research institute | CEBU Clinical Epidemiology and Biostatistics Unit | The University of Melbourne

Acknowledgement of Country

Photo credit – Peter Glenane, HIVIS Pictures

About me

I work as a biostatistician and project officer at the Murdoch Children's Research Institute in Melbourne, Australia.

I work predominately in clinical trials.

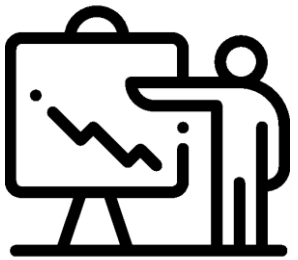
Producing lots of Data Safety Monitoring Committee (DSMC) reports…

What is a Data Safety Monitoring Committee/ Board (DSMC/DSMB) report?

A DSMC is a multidisciplinary group established by a trial sponsor to review, at regular intervals, accumulating trial data, in order to monitor the progress of a clinical trial.

Their role is to provide advice on safety and/or trial conduct issues by making recommendations to the sponsor or investigator, on whether to continue, modify or stop a trial for safety or ethical reasons (1).

Frequency of DSMC review depends on the needs of the trial but can commonly be one or twice a year for the life of the trial (generally several years).



1. National Health and Medical Research Council (2018), Data Safety Monitoring Boards (DSMBs)

What is a Data Safety Monitoring Committee (DSMC) report?

DSMC OPEN report

A trial statistician (part of the trial team) prepares a blinded (open) report for review by this committee.

Below is the commonly presented reported. This information would be provided across the whole study sample, not separated by treatment arm.

- Protocol modifications made by investigators (e.g., to inclusion criteria, trial endpoints, or sample size)
- Assessment of data quality, including completeness
- Screening and recruitment numbers and whether timelines are likely to be met, often by recruiting site
- Summary of loss to follow-up and adherence to study intervention
- Summary of demographics and baseline characteristics
- Summary of compliance with the protocol by investigators and clinical team
- Summary of relevant safety data

2. Australian Clinical Trial Alliance. Statistics in Trials Interest Group . The processes involved in a Data and Safety Monitoring Board/Committee with special reference to the role of the various statisticians

What is a Data Safety Monitoring Committee (DSMC) report?

DSMC CLOSED report

A reporting statistician (independent from the trial team) prepares a unblinded (closed) report.

This report will be separated by treatment arm.

- Same information as open report but separated by treatment arm.
- Can also include an interim efficacy analyses

2. Australian Clinical Trial Alliance. Statistics in Trials Interest Group . The processes involved in a Data and Safety Monitoring Board/Committee with special reference to the role of the various statisticians

Statisticians involved in a DSMC

TRIAL STATISTICIAN

- Involved in management and oversight of trial
- Responsible for statistical aspects of the trial
-
- Involved in writing protocol
- Conducts statistical analysis
- Can develop open DSMC report but can't prepare/ have access to closed report to stay blinded

DSMB STATISTICIAN

- Independent statistician who providing statistical expertise to the committee.
- Doesn't produce reports

REPORTING STATISTICIAN

- Prepares closed reports
- Can also contribute to open report
- Has access to unblinded data
- Couldn't be involved in oversight of trial
- Closed report

3. Australian Clinical Trial Alliance. . Statistical requirements for a Data Safety Monitoring Board (DSMB)

An example of how Stata `putdocx` can streamline the DSMC reports

Using Stata dataset cancer.dta

Patient survival in drug trial

Adding some other demographic and safety outcomes for the purposes of this presentation

Hypothetical trial cancer trial – 'TEST' trial

Primary outcome of complete response (Y/N)

Secondary outcome of IgG level (mg/dL)

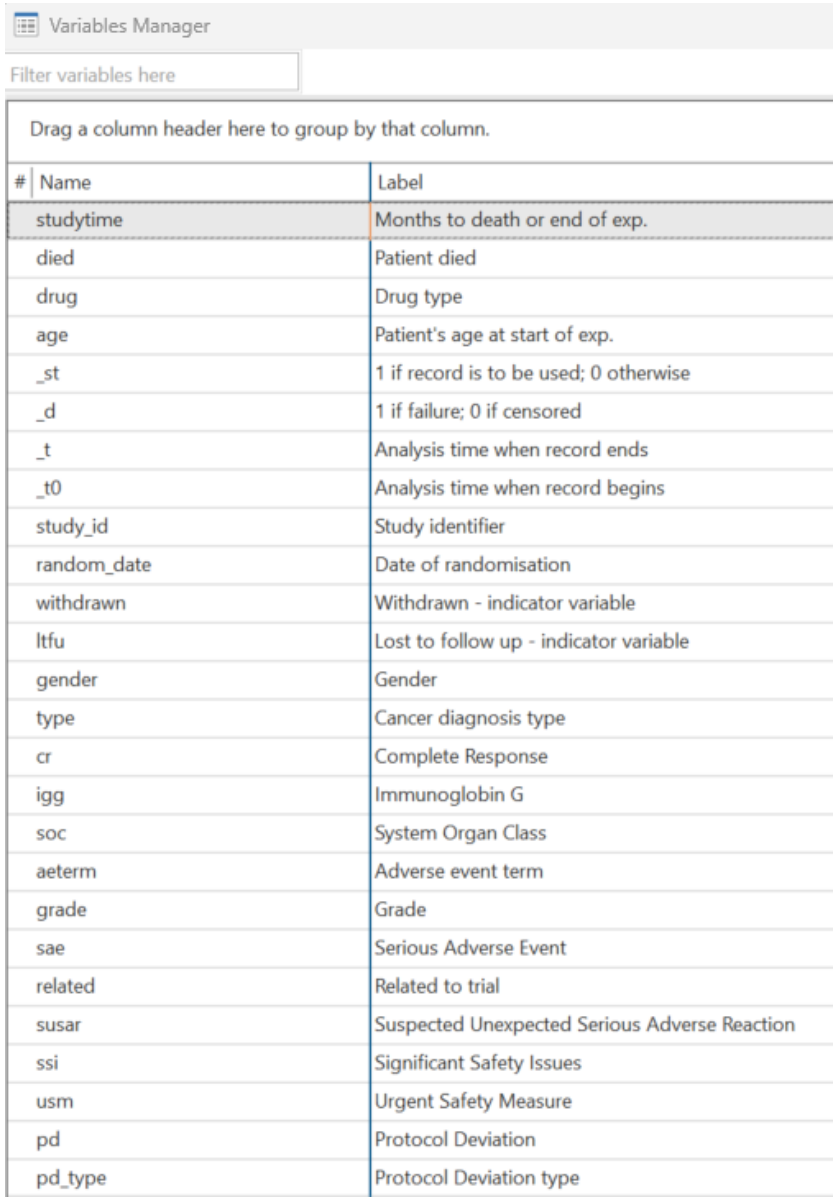
Safety outcomes of death and adverse events

Variables Manager

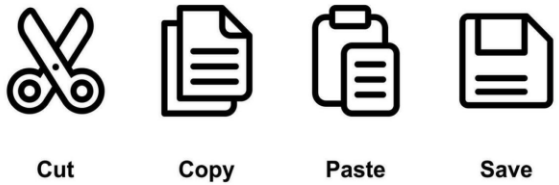
Filter variables here

Drag a column header here to group by that column.

#	Name	Label
	studytime	Months to death or end of exp.
	died	Patient died
	drug	Drug type
	age	Patient's age at start of exp.
	_st	1 if record is to be used; 0 otherwise
	_d	1 if failure; 0 if censored
	_t	Analysis time when record ends
	_t0	Analysis time when record begins



Usual approach vs streamlined `putdocx` approach



Usual approach (1st report)

- Create tables and graphs in a Stata do-file
  - Manually create a word document based on a template
    - Fill in report details in word
      - Paste in tables and graphs and format document

`putdocx` approach (1st report)

- Create tables and graphs in a Stata do-file
  - Create a word document in a Stata do-file
    - Add report details into do-file (using macros for totals etc)
      - Insert and format tables and graphs, and format document in do-file

Takes approximately the same amount of time ~10-20 hours dependent on trial

Usual approach vs streamlined `putdocx` approach

Usual approach (subsequent reports)

- Create tables and graphs, reusing code from last Stata do-file
  - Manually create a word document based on last report
    - Update report details in word
      - Paste in tables and graphs and format document

`putdocx` approach (subsequent reports)

- Create tables and graphs, reusing code from last Stata do-file
  - Create a word document re-using code from last Stata do-file
    - Update report details into do-file as needed

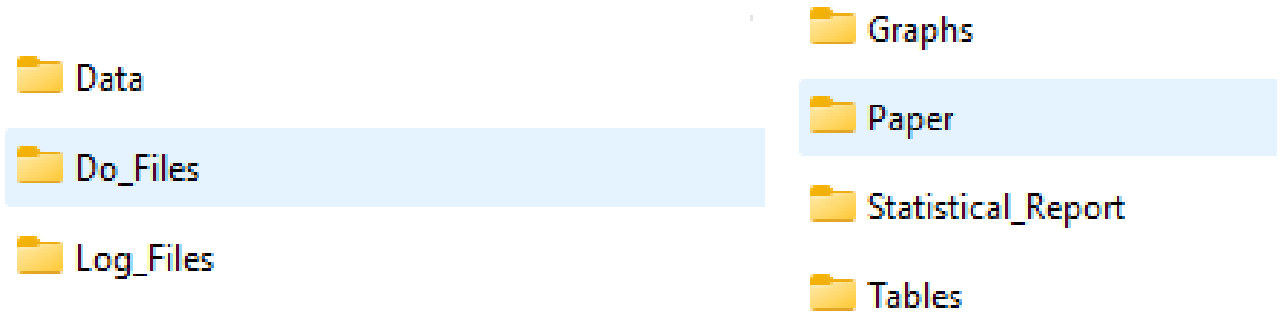
Subsequent reports is where the real benefit is seen. Time is saved by removing the copy/paste/ format step. Can typically save between ~5-15 hours

Step 0: Create do-files and outputs

You should be doing this for best practice regardless of `putdocx` use

I will only touch on briefly here,

- **Set up of files to keep do-files, data, outputs, and reports in a structured way (separate folder for each meeting timepoint)**
- **Utilise a master do-file to manage paths and do-file execution**
- **Ensures you can reproduce and easily update any data cleaning and tables etc**




```
*************************************************
* Study: TEST trial - fictional trial data used for DSMC demonstration purposes
* Program Name: Master
* Creation Date: 20 Jan 2026
* Purpose: MASTER programme
* Author: Alannah Rudkin
* Version: 1
* Note:  master and globals for DSMC

* Update: XXXXX, dd-mm-yyyy, Name Surname
*************************************************


clear all
version 19.5
set more off

global projectfolder "C:\Users\alannah.rudkin\OneDrive - Murdoch Children's Research Institute\Other\Contracts and PD\PD\Stata 2026\Analysis"
global output "$projectfolder\outputs"


*set directory
cd "$projectfolder"

*******************************************************************************
** Create dataset
*******************************************************************************
do "$projectfolder\making_dataset_cancer.do"

*******************************************************************************
***************OPEN report
*******************************************************************************
** Make required tables and graphs
do "$projectfolder\an_dsmc_tables_jan26.do"

** Make report
do "$projectfolder\an_dsmc_report_jan26.do"
```

Step 1: Specify data and macros

- The dataset used for each report will different (data updated from timepoint 1 to 2)
- Able to use the dataset to get counts and dates to reference throughout without having to manually update

```
*** Identify dataset
use test_trial_2.dta, clear

*** DEFINE MACROS
global today "`c(current_date)'"
di "$today"
global cut_date 14Jan2026

** Number enrolled at cut date
count if random_date <= td($cut_date)
global totno=r(N)

**last patient enrolled date_enroll
preserve
gen negdate_enrol=-(random_date)
sort negdate_enrol
gen lastenrol= random_date if _n==1
sum lastenrol if lastenrol!=.
local lastenrol :disp %tdDDMonCCYY r(mean)
di "`lastenrol'"
restore
```

Step 2 : Creating a word document

`putdocx begin` creates an Office Open XML (`.docx`) file. This is the active document that the remaining `putdocx` commands modify.

`putdocx describe` describes the active `.docx` file.

`putdocx save` saves and closes the `.docx` file.

`putdocx clear` closes the `.docx` file without saving.

`putdocx append` appends the contents of one or more `.docx` files to another `.docx` file.

begin_options	Description
`pagesize(`*psize*`)`	set document page size
`landscape`	change document orientation to landscape
`font(`*fspec*`)`	set font, font size, and font color for the document
`pagenum(`*pnspec*`)`	set page number format
`header(`*hname*`[ , `*header_opts*`])`	add a header
`footer(`*fname*`[ , `*footer_opts*`])`	add a footer
`margin(`*type*`, #[`*unit*`])`	set page margins for the document
`compmode(`*modetype*`)`	specify the compatibility mode to be used by Word when opening the document; default is `compmode(word2010)`

Step 2 : Creating a word document

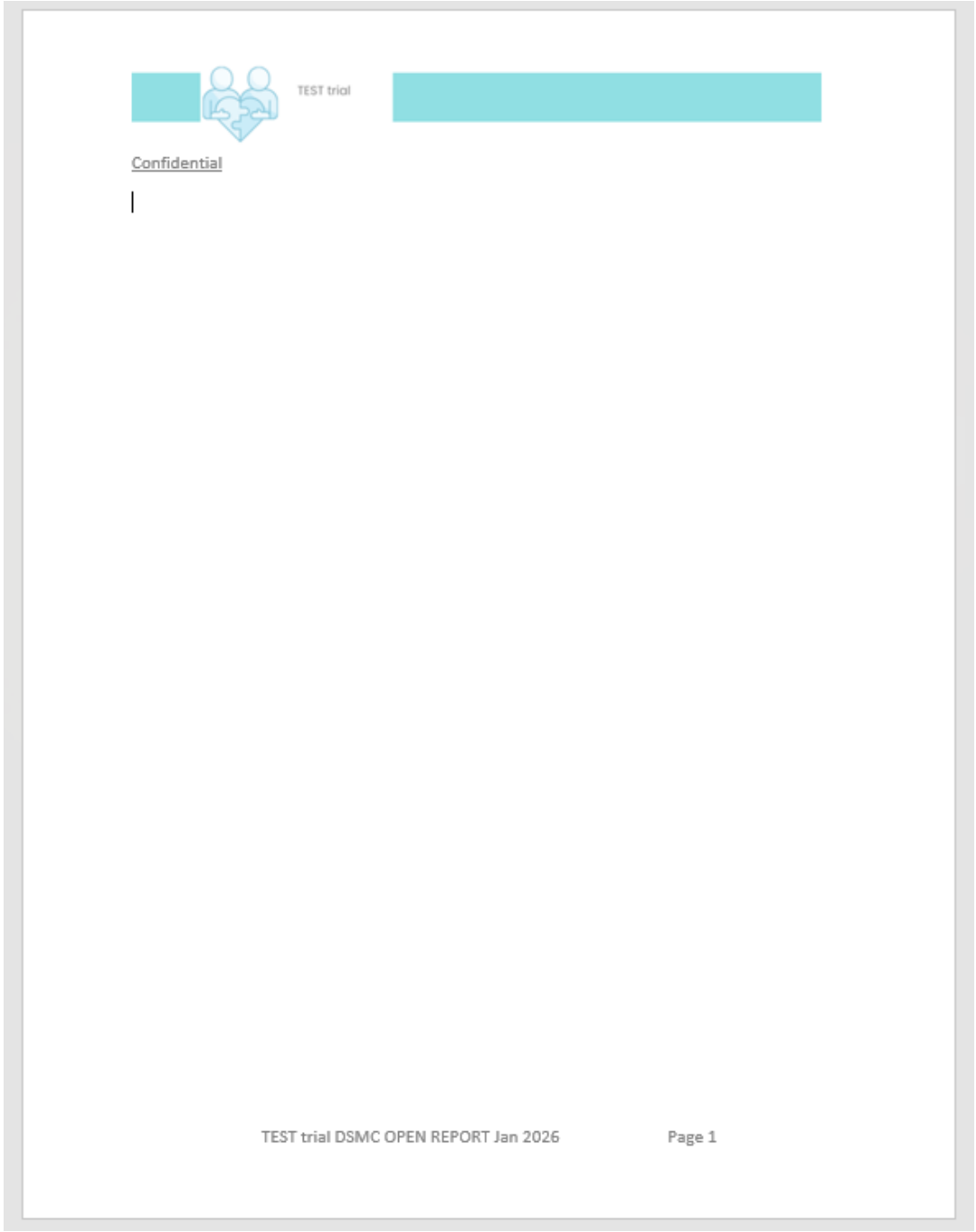
```
*******************************************************************************
* Start a Word document
set docx_paramode on
capture putdocx clear
putdocx begin, pagenum() header(header1) footer(footer1) font("Calibri", 12, black )

putdocx paragraph, toheader(header1)
putdocx image ("C:\Users\alannah.rudkin\OneDrive - Murdoch Children's Research Institute\Other\Contracts and PD\PD\Stata 2026\Picture1.png"), linebreak
putdocx text ("Confidential"), underline(single)

putdocx paragraph, tofooter(footer1) halign(center)
putdocx text ("TEST trial DSMC OPEN REPORT Jan 2026                    Page ")
putdocx pagenumber

*putdocx save "TEST_OPEN_report_$today.docx", replace
```

Step 2 : Creating a word document



Step 3: Adding a title page

Insert page breaks in a .docx file (see [RPT] **putdocx pagebreak**)

`putdocx pagebreak`	Adds a page break to the document
`putdocx sectionbreak`	Adds a new section to the document

Add paragraphs with text and images (see [RPT] **putdocx paragraph**)

`putdocx paragraph`	Adds a new paragraph to the active document
`putdocx text`	Adds text to the active paragraph
`putdocx textblock`	Adds a block of text to the active paragraph or to a new paragraph
`putdocx textfile`	Adds a block of preformatted text to a new paragraph with a predefined style

Step 3: Adding a title page

```
* TITLE PAGE Text
putdocx paragraph, halign(center)
putdocx text ("Data Safety Monitoring Committee (DSMC)"), bold font("Calibri", 14, black ) linebreak
putdocx text ("OPEN Report"), bold font("Calibri", 14, black ) linebreak
putdocx text ("Date of Report: $today"), bold font("Calibri", 14, black )
putdocx paragraph
putdocx paragraph
putdocx paragraph, halign(center)
putdocx text ("Name of trial:"), bold font("Calibri", 12, black )
putdocx text (" TEST trial - Randomised cancer treatment trial for example purposes"), font("Calibri", 12, black )
putdocx paragraph, halign(center)
putdocx text ("Meeting:"), bold font("Calibri", 12, black ) linebreak
putdocx text ("2nd meeting - Jan 2026"), font("Calibri", 12, black )
putdocx paragraph, halign(center)
putdocx text ("Trial registration number: Clinicaltrials.gov Identifier NCTXXXXXX"), font("Calibri", 12, black) linebreak
putdocx text ("Based on: Protocol Version 1.0, Dated 1 August 2022"), font("Calibri", 12, black)
putdocx paragraph,halign(left)
putdocx text ("Investigators: "), bold font("Calibri", 11, black ) linebreak
putdocx text ("1.   Dr Bob Smith"), font("Calibri", 11, black ) linebreak
putdocx text ("2.   Dr Penny Hudson"), font("Calibri", 11, black ) linebreak

putdocx paragraph,halign(left)
putdocx text ("Study biostatistician:"), bold font("Calibri", 11, black )
putdocx text (" Alannah Rudkin"), font("Calibri", 11, black )
putdocx paragraph,halign(left)
putdocx text ("Data and Safety Monitoring Committee Report produced by:"), bold font("Calibri", 11, black ) linebreak
putdocx text ("Alannah Rudkin"), font("Calibri", 11, black ) linebreak
putdocx text ("Murdoch Children's Research Institute"), font("Calibri", 11, black ) linebreak
putdocx text ("Melbourne"), font("Calibri", 11, black )
putdocx paragraph
putdocx paragraph
putdocx paragraph,halign(center)
putdocx text ("This report is based on data up to and including $cut_date."),  font("Calibri", 11, black )
putdocx pagebreak

....
```

Step 3: Adding a title page

Confidential

Data Safety Monitoring Committee (DSMC)
OPEN Report
Date of Report: 4 Feb 2026

Name of trial: TEST trial - Randomised cancer treatment trial for example purposes

Meeting:
2nd meeting - Jan 2026

Trial registration number: Clinicaltrials.gov Identifier NCTXXXXXX
Based on: Protocol Version 1.0, Dated 1 August 2022

Investigators:
1. Dr Bob Smith
2. Dr Penny Hudson

Study biostatistician: Alannah Rudkin

Data and Safety Monitoring Committee Report produced by:
Alannah Rudkin
Murdoch Children's Research Institute
Melbourne

This report is based on data up to and including 14Jan2026.

Step 4: Adding report content

Example of adding a table and graph

```
**** SECTION 1
putdocx paragraph, style(Heading2)
putdocx text ("1. Recruitment"), bold font("Calibri Light", 14, black)
putdocx paragraph, font("Calibri", 12, black )
putdocx text ("The planned sample size is 100 participants, of which 50 will be assigned to each of the 2 study arms."), linebreak
putdocx paragraph, font("Calibri", 12, black )
putdocx text ("Figure 1 shows the accrual of the first $totno participants into the study up to the $cut_date, and Table 1 shows a breakdown of these participants by randomisation stratum."), linebreak
putdocx text ("Since the beginning of the study, there has been 1 issue related to the randomisation procedure. Details are listed below.")
putdocx paragraph
putdocx text ("Electronic randomisation system was down for 24hours on 06Jan2025. Unable to randomise in this time. Back-up paper/phone randomisation system now available.")

***Insert Table 1
putdocx paragraph, font("Calibri", 12, black )
putdocx text ("Table 1: Recruitment by randomisation stratum"), italic
import excel using table1.xlsx , clear
putdocx table table1 =  data(A B) , border(insideH, nil) border(start) border(insideV,nil) border(end) headerrow(1) layout(autofitw)
putdocx table table1(1,.), bold
putdocx table table1(1,2) = ("Total"), bold
putdocx table table1(2,1) = ("Number of Participants Randomised"), bold
putdocx table table1(2,.), border(bottom)
putdocx table table1(.,2), halign(center)
putdocx table table1(6,.),  border(bottom)
putdocx paragraph

***Insert graph 1
putdocx paragraph, font("Calibri", 12, black )
putdocx text ("Figure 1: Recruitment Rate as of $cut_date"), bold linebreak
putdocx image ("$output\Graph1.png")
```

Step 4: Adding report content

Example of adding a table and graph

Protocol Amendments

Version 1.0, Dated 1 August 2022 is the most recent version of the protocol. No changes were made since the previous meeting. But if any changes were made this is where we'd put them.

1. Recruitment

The planned sample size is 100 participants, of which 50 will be assigned to each of the 2 study arms.

Figure 1 shows the accrual of the first 48 participants into the study up to the 14Jan2026, and Table 1 shows a breakdown of these participants by randomisation stratum.
Since the beginning of the study, there has been 1 issue related to the randomisation procedure. Details are listed below.

Electronic randomisation system was down for 24hours on 06Jan2025. Unable to randomise in this time. Back-up paper/phone randomisation system now available.

Table 1: Recruitment by randomisation stratum

Number of Participants Randomised	Total 48
Cancer diagnosis type	
GIT	14 (29.2%)
Brain	21 (43.8%)
Haematological	13 (27.1%)

TEST trial

Confidential

Figure 1: Recruitment Rate as of 14Jan2026

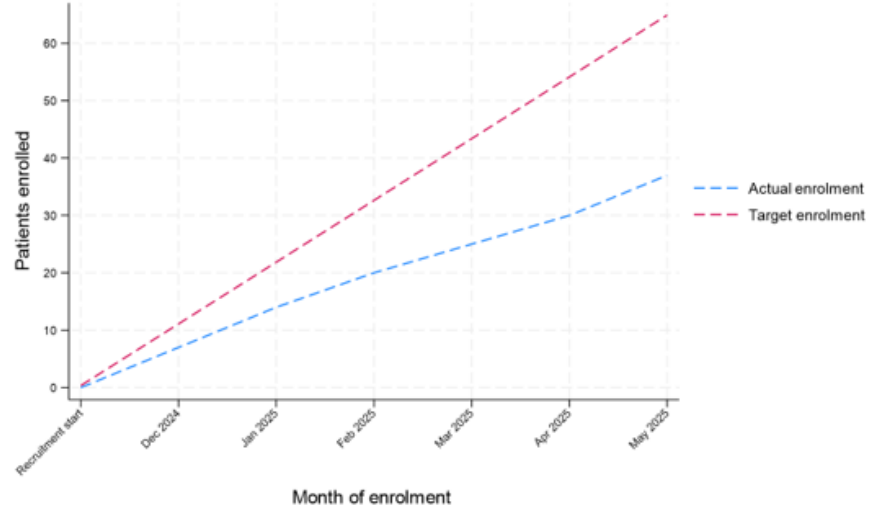


Table 2: Participant disposition

	Total N= 48
Baseline Data Available	48 (100.0%)
Primary Outcome Data Available	34 (70.8%)
Secondary Outcome Data Available	48 (100.0%)
Withdrawn from the study	3 (6.2%)
Lost to follow up	3 (6.2%)

Other benefits of using putdocx commands to create DSMC reports

- Any edits to formatting requested is easily updated. Rather than needing to remember and manually update at the next meeting
- Able to autopopulate numbers and dates into report
- Reproducibility – able to easily share code with other statisticians etc

  E.g. Able to draft code for a unblinded reporting statistician using a dummy randomisation variable. Saving them time in preparing a closed report
- DSMCs tend to follow a template and have a very similar content. Once used for one trial, a lot of the general code can help shorten the time it takes to use `putdocx` commands for other trials.

CLOSED do-file - tables with dummy arm

```
*******************************************************************************
*****CLOSED report tables
*******************************************************************************
/*
Note to statistician - please replace dummy_arm variable with actual randomised drug/treatment
*/
*******************************************************************************
rename drug dummy_arm  //changing the variable name from the default data to make this clear
label var dummy_arm "Trial arm"

*******************************************************************************
** Table 1: Recruitment by randomisation stratum - type
*******************************************************************************
**make table
count if random_date!=.
tab type if random_date!=.
dtable i.type if random_date!=. , by(dummy_arm)
collect preview
collect export table1_closed.xlsx, replace
```

Do-file example code available on GitHub