

# SOFFT

## API Recruitment SWAT

### Statistical Analysis Plan Final v1.0

**York Trials Unit**  
Department of Health Sciences  
University of York  
York, YO10 5DD

**Version date:** 25/03/2022  
**Author(s):** Danielle Podmore & Izzy Coleman  
**Chief Investigator:** Professor Adam C Watts  
**Trial Coordinator:** Liz Cook, Sophie James,  
Zohaib Akhter

## Contents

1. Definition of terms/acronyms .....	2
2. Introduction.....	2
3. Objectives .....	2
3.1. Primary objective .....	2
4. Outcomes .....	2
4.1. Primary Outcome .....	2
4.2. Secondary Outcomes .....	2
5. Study Methods; design, sample size and randomisation.....	3
5.1. Design .....	3
5.2. Sample Size.....	5
5.3. Randomisation .....	5
6. Data.....	5
6.1. Screening Data .....	5
6.2. Management Database .....	6
6.3. Data Storage.....	6
7. Analysis.....	6
7.1. Software .....	6
7.2. Blinding.....	6
7.3. Participant Flow .....	6
7.4. Baseline.....	6
7.5. Primary Analysis.....	6
7.6. Secondary Analysis .....	7
8. SAP Revisions .....	8
9. Roles and responsibilities.....	8
9.1. Signatures.....	8
10. References.....	8
11. Appendix.....	8
11.1 CONSORT Diagram.....	8
11.2 Template Tables.....	9

## 1. Definition of terms/acronyms

API	Associate Principal Investigator
CI	Chief Investigator
ITT	Intention to treat
SAP	Statistical Analysis Plan
SWAT	Study Within A Trial
TC	Trial Coordinator
YTU	York Trials Unit

## 2. Introduction

This statistical analysis plan (SAP) provides details on the analysis that is to be undertaken by any person analysing the results the recruitment Study within a Trial (SWAT) within the Simple Olecranon Fracture Fixation Trial (SOFFT).

The SAP has been developed and agreed prior to the end of data collection, and before any data is presented to the person responsible for the analysis.

## 3. Objectives

This SWAT will look at improving the recruitment rate for the host trial (SOFFT), by an enhanced associate principal investigator (API) training package and additional digital nudge delivered by a trial coordinator (TC).

In addition to this SWAT, a second SWAT focussed on the retention of participants will be embedded in the SOFFT trial and will investigate whether the inclusion of a social incentive text cover letter with postal questionnaires improves response rates to participant questionnaires; the analysis for this will be detailed in a separate SAP. It is not anticipated that the recruitment SWAT will impact on the retention of participants and therefore we will not explore the possibility of interaction between the two SWATs.

### 3.1. Primary objective

The primary objective of this SWAT is to assess the effect of an enhanced API package, a digital nudge, and a combination of the two compared to standard practice API and/or standard email alone on the number of patients recruited to the host trial during the six months that the API is in post at a recruiting site.

## 4. Outcomes

### 4.1. Primary Outcome

The primary outcome for this SWAT is the number of patients recruited. This is defined as the total number of patients recruited to SOFFT by a site in the six months that the API is in place (as identified from the SOFFT randomisation database).

### 4.2. Secondary Outcomes

Secondary outcomes of interest in this SWAT include:

- The total number of patients recruited to SOFFT by a site in months 7 - 12, the 6 months following the API being in place. This data will be collected from the SOFFT randomisation database.

- Total number of patients recruited to SOFFT by a site in the 12 months following the API being in place (identified from the SOFFT randomisation database).
- Proportion of eligible screened patients recruited to the trial; the numerator will be the number of randomisations (as recorded in the database), and the denominator will be the number screened (taken from the screening logs provided by the sites). Eligibility is defined by the inclusion and exclusion criteria of the host trial, which can be found in the SOFFT protocol.
- Estimated cost of implementing the interventions at a site. This will include the costs associated with the TC time taken to; deliver the intervention training, send intervention monthly emails, and send intervention nudge emails.
- The amount of time taken for the TC to deliver the interventions will be reported. Including the time taken to construct and send a digital nudge and the time taken to deliver the training for the enhanced API group.

## 5. Study Methods; design, sample size and randomisation

The study design, details of the recruitment strategies and population of interest can be found within the SWAT protocol for the host trial. Full details on the aspects such as treatment allocation, randomisation, blinding, selection of study population, and details on the specified follow-up time points are given in the protocol; a brief overview will be given here.

### 5.1. Design

This will be a multi-centre, 2x2 factorial randomised SWAT in the SOFFT randomised controlled trial to evaluate the effect on recruitment rates in a secondary care setting of two interventions: Enhanced API Package; and Digital Nudging (both are in addition to standard practice). This SWAT does not directly involve patient participation and sites are not aware of their participation in a SWAT.

The factorial design will allow for the evaluation of two interventions in one trial, there is potential for interaction between the two interventions and this will be explored in the analysis. There will be four groups in this SWAT (as shown in Table 1):

- Group A – Enhanced API Intervention and Standard Practice email
- Group B – Enhanced API Intervention and Digital Nudge Intervention
- Group C – Standard Practice API and Standard Practice email
- Group D – Standard Practice API and Digital Nudge Intervention

Table 1: Factorial design of the trial.

		API		
		(Intervention) Usual practice + Enhanced API	(Control) Usual Practice	
Digital Nudge	(Control) Usual Practice	Group A	Group C	<b>A+C</b>
	(Intervention) Usual practice + Digital Nudging	Group B	Group D	<b>B+D</b>
		<b>A+B</b>	<b>C+D</b>	

The interventions will run for 6 months in each recruitment site; this is due to the length of time an API surgical trainee will usually be in place at the site. Some APIs will not be surgical trainees, the actual time in post will be recorded for all APIs along with details of when the intervention is started. The collection of data for this SWAT will run for 12 months in each recruiting site. Any potential contamination, caused by APIs moving to a recruiting site that is not randomised to receive the enhanced intervention will be monitored using site delegation logs for movement of APIs and involvement across other participating sites.

The activity for each intervention is summarised below in comparison to current standard practice (Table 2), more detail is provided in the SWAT protocol.

Table 2: Summary of trial intervention activity.

ACTIVITY	STANDARD PRACTICE	ENHANCED API	DIGITAL NUDGE
<b>Identify API for the trial</b>	Local Principal Investigator	Local Principal Investigator	
<b>Training of API regarding how to perform their role</b>	Local Principal Investigator  SOFFT API Manual	Local Principal Investigator  1:1 telephone/videoconference induction by Trial coordinator  SOFFT API manual  Induction summary presentation	
<b>Training API regarding the SOFFT and consenting procedures</b>	Local Principal Investigator	Local Principal Investigator  1:1 telephone/videoconference induction by Trial Coordinator	
<b>Peer Support of API</b>		Monthly contact by Trial Coordinator working on SOFFT from YTU.	
<b>Digital information provided to API</b>	SOFFT API Manual	Induction agenda SOFFT API manual and new API checklist  Induction summary presentation  SOFFT consent flow diagram and protocol	

		API contact information consent form	
<b>Identifying patients for the trial</b>	A&E/Trauma meeting	A&E/Trauma meeting	
<b>Confirmation of randomisation</b>	Automated email to recruiting site team.		Automated email to recruiting site team.  Additional personalised email to randomiser to the trial sent by Trial Coordinator.

## 5.2. Sample Size

The sample size of this SWAT is driven by the size of the host trial, and as such no formal power calculation has been performed for this SWAT. The number of sites is constrained by the number of sites included in the host trial. Further to this, to be included in the SWAT the sites must have an API in post, and be open to recruitment for at least 6 months (i.e. those who open as a site with less than six months of the recruitment period left will not be eligible for inclusion in the SWAT). Due to this it is anticipated that a small number of sites will be included in the SWAT which are unable to provide secondary outcome data (that is where there is not 12 months left of the recruitment period to obtain this data). This decision was made to maximise the number of sites in the primary analysis. It is anticipated there will be at least 18 sites included in this SWAT. The site where the CI is based will not be included in the SWAT since this site may already have higher recruitment rates due to the involvement of the CI, which could result in bias and confound with the effect of the API intervention.

## 5.3. Randomisation

Sites (hospitals) will be the unit of randomisation for this SWAT. The SOFFT sites open to recruitment who have a confirmed API will be randomised 1:1:1:1 by minimisation to one of the four groups (as detailed in Section 5.1) to balance key baseline characteristics. A statistician at YTU will undertake the minimisation using the software MinimPY (1). The minimisation will be undertaken initially for all sites which are participating in the study as one batch in February 2021, with any additional sites or batches of sites randomised as they confirm that an API is in place. As there is no risk of subversion, or predictability due to the sites not being aware of the SWAT; naïve minimisation with base probability 1.0 will be used. The minimisation will include the following factors:

- Site size: large site (population > 500,000) vs. small site (population ≤ 500,000) based on population they serve.
- Previous recruiting time: site open for < 2 months vs. site open for ≥ 2 months; at the point the API is confirmed and the site is randomised into the SWAT.

## 6. Data

Methods detailing data collection are provided in the corresponding protocol and report.

### 6.1. Screening Data

Screening data will be collected from sites by TCs and stored in electronic format (e.g. excel files). The data will be transferred to the statistician in a secure manner and will be encrypted to ensure security (e.g. by using the YTU DropOff service).

## **6.2. Management Database**

Data collected for the SOFFT trial will be stored in a bespoke data management system. All participant-reported data is identified solely by the unique participant trial ID. CRFs are returned by post or electronically to YTU where they are scanned, using Teleform data capture software, into the data management system. This system is separate from the trial management system and contains no identifying details. All data are error checked and validated to ensure accuracy according to validation plans. The randomisation dates for each participant will be captured in the trial management system, these will be provided to the Statistician along with the corresponding participant trial IDs. Both the trial management system and the data management systems are held on secure University of York servers with access limited to specified members of YTU staff as detailed in the delegation log.

## **6.3. Data Storage**

Datasets will be stored in the SOFFT study folder (Y:\Project -- SOFFT - Statistics\1 Statistics) on a secure server at the Department of Health Sciences, University of York. Data will be stored in line with data protection procedures.

# **7. Analysis**

## **7.1. Software**

The analysis will be undertaken using Stata v17.0 (or later).

## **7.2. Blinding**

The statistician analysing the SWAT will not be blinded to the SWAT allocation. This SWAT does not directly involve patient participation. Sites are not aware of their participation in a SWAT; they cannot be blinded to the intervention they receive but will be unaware that is being evaluated.

## **7.3. Participant Flow**

A CONSORT flow diagram will be created to detail the flow of sites through the SWAT, including relevant details of the host trial wherever needed (see Appendix).

## **7.4. Baseline**

Baseline data relating to the sites (including the minimisation factors) will be summarised for the four groups (A to D as given in Section 5.1); and for the comparison groups (i.e., enhanced API (A+B) vs standard API (C+D), digital nudge (B+D) vs usual practice (A+C)). Continuous data will be presented using descriptive statistics (e.g., mean, standard deviation, median, minimum, maximum), while categorical data will be given as counts and percentages. No formal statistical comparison of baseline data will be undertaken between the groups.

The number of participants recruited will be summarised overall, as well as for each of the four groups (A to D), in addition to the comparison groups as described above.

## **7.5. Primary Analysis**

All analysis will be conducted on an ITT basis, where all sites are included in the group to which they were allocated. Statistical significance will be assessed using two-sided statistical tests at the 5% significance level unless specified otherwise.

A Poisson regression model will be fit with the number of patients recruited in the six months that the API is in place as the response variable, and with treatment groups for the two interventions (Enhanced API and Digital Nudge) and the minimisation factors (site size: small vs large, and recruiting time: <2 months vs  $\geq 2$  months) as the independent variables. Adjusted incidence rate ratios (IRRs), associated 95% confidence intervals (CIs) and p-values will be reported. The presence of an interaction between the two interventions will also be tested by re-running this model including an interaction term; this will be assessed at 10% significance level.

Model assumptions will be checked prior to use for all analysis. If the variance of the data is larger than the mean, this may give an indication that the data are over-dispersed. In this case, a likelihood ratio test of the over-dispersion parameter will be conducted. A significant p-value indicates that the data are over-dispersed and therefore Poisson regression is not appropriate. If there is evidence of over dispersion then negative binomial models will be used instead. Histograms will be produced to assess the distribution of the data, if an excess number of zeros are observed then zero-inflated Poisson and zero-inflated negative binomial models will be considered and compared to the Poisson and/or negative binomial models to determine best fit. Goodness-of-fit will be compared using Akaike's information criterion.

## **7.6. Secondary Analysis**

The total number of patients recruited to SOFFT by a site in the 12 months following the API being in place will be analysed in a similar way to in the primary analysis, using a Poisson regression model with the same covariates. Again, model assumptions will be checked and more appropriate models will be selected if required.

The proportion of eligible patients recruited will be analysed using a logistic regression model, adjusting for the same factors as in the primary analysis. All proportions will be presented as un-adjusted proportion by arm, and difference in proportions; as well as an adjusted odds ratio, with associated 95% confidence interval, and p-value.

The estimated cost of each additional participant recruited associated with the SWAT interventions (Enhanced API and Digital Nudge) will be calculated for each intervention, by calculating the total cost of the intervention, divided by the number of additional participants in the SWAT intervention arm. Alternatively, the cost per person recruited is calculated by the total number of recruited in the intervention arm, over the total cost of the intervention.

Feasibility outcomes, such as the TC time required to run the interventions and communication time and methods used for the peer support aspect of the intervention, will be reported descriptively. This will include the time taken to construct and send a digital nudge and the time taken to deliver the training for the enhanced API intervention.

Due to the possible contamination at sites as and when the APIs move on, the movement of the API will be monitored (i.e. recording of where and when an API moves to another site, and if this site is involved in recruiting for SOFFT). Should it be apparent that contamination may be present, a sensitivity analysis will be undertaken excluding possibly contaminated sites. Descriptive details will be given on the movement of APIs.

Additionally, as there will be sites in SOFFT that do not have an API in place, and thus are not included within the SWAT, a summary of recruitment at sites with no API will be provided, to allow for a comparison of sites with an API vs. sites without an API – this will be a visual comparison only. It may be theorised that sites which get an API are those that are more engaged in the trial, and thus would recruit better. This may be done by providing an



average recruitment over the most common six month period that APIs were in place for those sites involved in the SWAT.

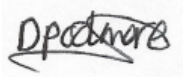

A meta-analysis combining the results from this SWAT, with any other evaluations of this intervention may be undertaken in the future.

## 8. SAP Revisions

Amendment/addition to SAP and reason for change	New version number, name and date

## 9. Roles and responsibilities

### 9.1. Signatures

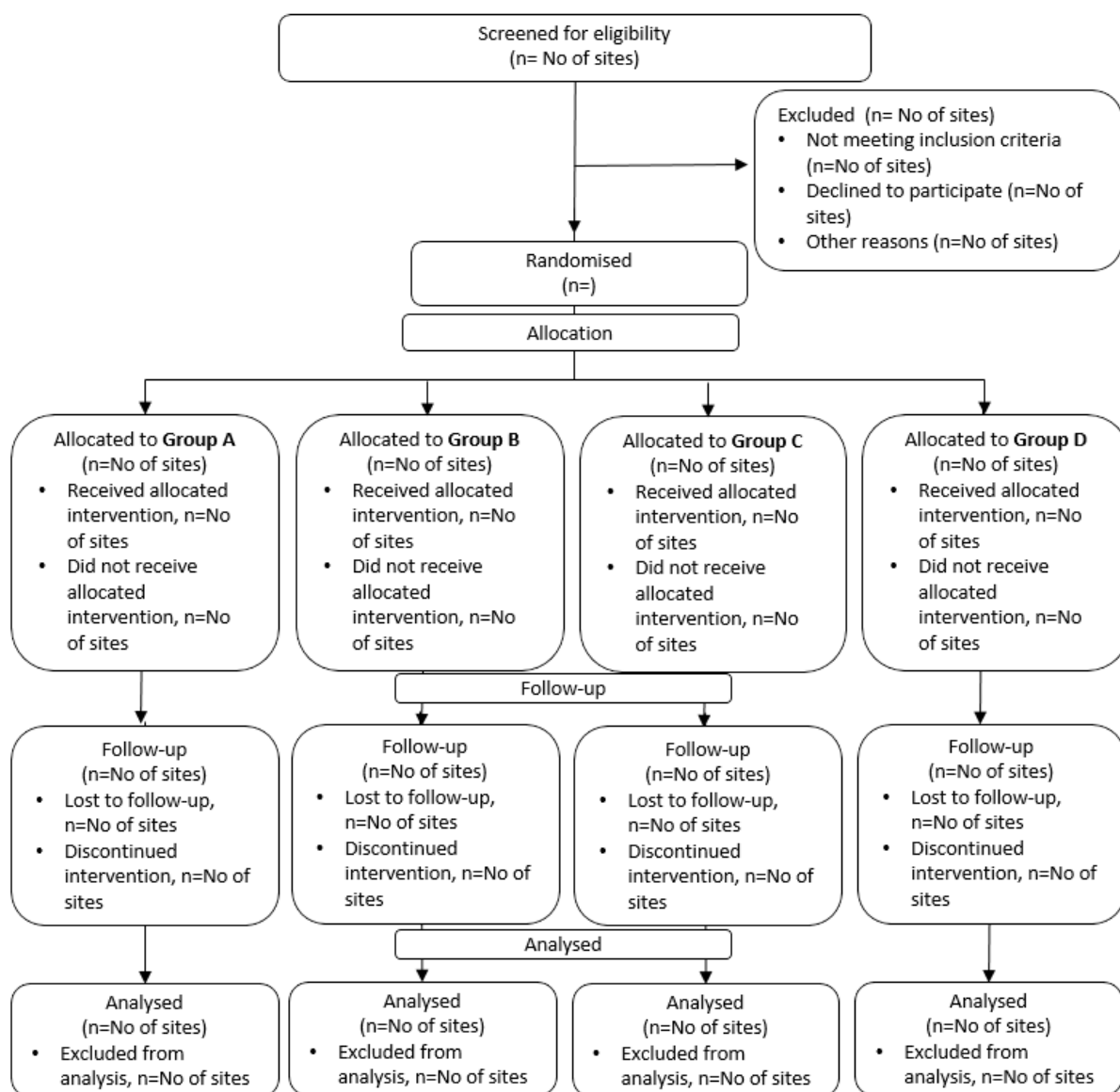
Name	Trial Role	Signature	Date
Danielle Podmore	Trial Statistician		25/03/2022
Izzy Coleman	Statistician		25/03/2022

## 10. References

1. Saghaei, Mahmoud & Saghaei, Sara. (2011). Implementation of an open-source customizable minimization program for allocation of patients to parallel groups in clinical trials. *Journal of Biomedical Science and Engineering*. 4. 734-739. 10.4236/jbise.2011.411090.

## 11. Appendix

### 11.1 CONSORT Diagram



## 11.2 Template Tables

**Table 1:** Site level baseline data presented by treatment allocation for each treatment group and overall, for the ITT population.

	Group A (n=Number of sites)	Group B (n=Number of sites)	Group C (n=Number of sites)	Group D (n=Number of sites)	Total
<b>Number of participants recruited</b> n (%) Mean (SD) Median (IQR) Min, Max					
<b>Site size, n (%)</b> Small site Large site					

<b>Recruiting time, n (%)</b> Short (<2 months) Long (≥2 months)					
...					

**Table 2:** Site level baseline data presented by API treatment allocation (enhanced API vs. standard API) and digital nudge treatment allocation (digital nudge vs. usual practice), for the ITT population.

	<b>Enhanced API (Group A+ Group B) (n=Number of sites)</b>	<b>Standard API (Group C+ Group D) (n=Number of sites)</b>	<b>Digital nudge (Group B+ Group D) (n=Number of sites)</b>	<b>Usual practice (Group A+ Group C) (n=Number of sites)</b>
<b>Number of participants recruited</b> n (%) Mean (SD) Median (IQR) Min, Max				
<b>Site size, n (%)</b> Small site Large site				
<b>Recruiting time, n (%)</b> Short (<2 months) Long (≥2 months)				
...				