

Statistical Analysis Plan for title

DOCUMENT: SAP-2021-004-TV-v01

From: Felipe Figueiredo To: Thiago Vivacqua

Date: 2021-08-24

TABLE OF CONTENTS

1 ABBREVIATIONS.....	2
2 INTRODUCTION.....	2
2.1 Context.....	2
2.2 Objectives.....	2
2.3 Hypotheses.....	2
3 DATA CLEANING.....	3
3.1 Raw data.....	3
3.2 Analytical dataset.....	3
4 STUDY VARIABLES.....	4
4.1 Primary and secondary outcomes.....	4
4.2 Covariates.....	4
5 STATISTICAL METHODS.....	4
5.1 Statistical analyses.....	4
5.1.1 Descriptive analyses.....	4
5.1.2 Inferential analyses.....	5
5.1.3 Statistical modeling.....	5
5.2 Significance and Confidence Intervals.....	5
5.3 Study size and Power.....	5
5.3.1 Statistical packages.....	5
6 EXCEPTIONS AND OBSERVATIONS.....	5
7 REFERENCES.....	5

Statistical Analysis Plan (SAP)

8 APPENDIX.....	6
8.1 Availability.....	6

Document version

Version	Alterations
01	Initial version

1 ABBREVIATIONS

- ACL: anterior cruciate ligament
- ACLR: ACL revision
- BMI: body mass index
- LCL: lateral collateral ligament
- LET: lateral extra-articular tenodesis
- MCL: medial collateral ligament
- PCL: posterior cruciate ligament

2 INTRODUCTION**2.1 Context****2.2 Objectives**

Compare the baseline characteristics of patients undergoing ACL revision surgery whether or not they undergone the LET technique. Compare the rates of surgical failure and functional scores of patients undergoing ACL revision surgery whether or not they undergone the LET technique.

2.3 Hypotheses

LET and no LET groups are comparable at baseline.

Lower ages are a risk factor.

3 DATA CLEANING

3.1 Raw data

The raw data was provided in an Excel file with two sheets, one for each group (LET and No LET). As a firsts processing step the raw data was copied into a new sheet with both groups, with one observation (patient) per line. The column structure of the resulting “clean” table corresponds to the original per-group tables.

In total there were 126 variables in the dataset received. Variables in the raw data sheet were grouped into sections. For this analysis only the baseline **demographic** characteristics and the variables in the **surgical data most recent revision ACLR** section were selected.

Column names were adjusted to avoid repeated names, using suffixes to identify the section each variable originally came from, as follows:

- section concomitant injuries primary ACLR – received the pri.inj suffix
- section Concomitant injuries most recent revision ACLR – received the inj suffix
- section surgical data primary ACLR – received the pri suffix
- section surgical data most recent revision ACLR – names were kept as the original

Some dates in Date of last clinical visit column (Follow-up notes data section) were typed with different separator symbols, and all values in this column were standardized to allow date format recognition.

There were two columns named grade (Slope data section), and they were both renamed grade1 and grade2.

The original raw data provided descriptive information for the encoding of each variable in the first line. A Data Dictionary was created from this information and was saved in a separate table for ease of consultation and processing.

The resulting “clean” raw data sheet was shared with the client for archival purposes.

3.2 Analytical dataset

All variables in the analytical set were labeled according to the raw data provided and values were labeled according to the data dictionary for the preparation of production-quality results tables and figures.

After the cleaning process 10 variables were included in the analysis with observations. Table 1 shows the structure of the analytical dataset.

Table 1 Analytical dataset structure after variable selection and cleaning.

id	age	sex	bmi	let	graft	graft_diameter	medial_meniscus	lateral_meniscus	cartilage
1									
2									
3									
...									
70									

The analytical dataset will be included in the private version of the report, and will be omitted from the public version of the report.

4 STUDY VARIABLES

4.1 Primary and secondary outcomes

The primary outcome of this analysis will be the difference in proportion of surgery failure between the two study groups.

The secondary outcome will be the comparison of functional scores (TO BE DEFINED).

4.2 Covariates

The distribution of the characteristics will be compared at baseline between the two study groups. Variables to be considered for baseline comparison are:

- Age
- Sex
- BMI
- Graft diameter
- Medial meniscus
- Lateral meniscus
- Cartilage

5 STATISTICAL METHODS

5.1 Statistical analyses

5.1.1 Descriptive analyses

The epidemiological profile of each study group (LET and no LET) will be described at baseline. Demographic and clinical variables will be described as mean and standard

deviation or as counts and proportions, as appropriate. Distributions will be summarized in tables and visualized in exploratory plots.

5.1.2 Inferential analyses

All comparisons will be performed as uni-variate analyses. Continuous variables will be compared between groups with the independent t test with Welch correction. Difference in distribution of categorical variables will be assessed with the Fisher exact test.

5.1.3 Statistical modeling

No adjustment for covariates will be performed.

5.2 Significance and Confidence Intervals

All analyses will be performed using the significance level of 5%. All significance hypothesis tests and confidence intervals computed will be two-tailed.

5.3 Study size and Power

Raw data received has 70 observations corresponding to patients assessed at primary ACL surgery and ACL revision. Surgeries that employed the LET technique include 38 patients and 32 patients undergone surgery with LET technique.

With these group sizes a t test can detect an effect size as large as $d = 0.68$ with 80% power and 5% significance level (Cohen, 1988).

5.3.1 Statistical packages

This analysis will be performed using statistical software R version 4.1.1.

6 EXCEPTIONS AND OBSERVATIONS

N/A.

7 REFERENCES

- **SAR-2021-004-TV-v01** – title
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd Ed.). New York: Routledge.

8 APPENDIX

8.1 Availability

Both this analytical plan and the corresponding analysis report (**SAR-2021-004-TV-v01**) can be downloaded in the following address:

<https://github.com/philsf-biostat/SAR-2021-004-TV/>

DRAFT