
Analytical Plan for Incidence rates of scorpion accidents in Urban Planning Areas of Americana/SP in 1998–2018

DOCUMENT: SAP-2021-007-JB-v01

From: Felipe Figueiredo To: José Brites

2021-10-21

TABLE OF CONTENTS

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

Analytical Plan for Incidence rates of scorpion accidents in Urban Planning Areas of Americana/SP in 1998–2018

Document version

Version	Alterations
01	Initial version

1 ABBREVIATIONS

- IQR: interquartile range
- CI: confidence interval
- IRR: Incidence Rate Ratio
- SD: standard deviation
- UPA: Urban planning area

2 CONTEXT

2.1 Objectives

Describe the incidence rates of scorpion accidents in ten Urban Planning Areas of Americana/SP between 1998 and 2018.

2.2 Hypotheses

Incidence rates have increased in the UPAs during the study period.

2.3 Study design

Ecological study, with unit of analysis as the Urban Planning Areas (municipal geographic regions).

3 DATA

3.1 Raw data

Dataset with number of accidents and population for each UPA and year between 1998 – 2018.

Analytical Plan (SAP)

3.2 Analytical dataset

In order to enhance the numerical performance of the statistical models the year was recentered with a time index starting at 1, and covering all years in the study period.

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 5 variables were included in the analysis with 210 observations. Table 1 shows the structure of the analytical dataset.

Table 1 Analytical dataset structure after variable selection and cleaning.

upa	year	accidents	pop	time
1				
2				
3				
...				
210				

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

Primary outcome will be defined as the incidence rate of scorpion accidents in each APU. For each APU the incidence rates will be computed as the ratio between the number of incident accidents over the population. Incidence rates will be reported per 10000 persons.

Specification of outcome measures (Zarin, 2011):

1. (Domain) Scorpionism
2. (Specific measurement) Scorpion accidents
3. (Specific metric) Incidence rate (per 10000)
4. (Method of aggregation) Rates ratio

Primary outcome

Incidence rates of scorpion accidents in APUs of Americana/SP, per 10000 inhabitants.

4.2 Covariates

N/A

5 STATISTICAL METHODS

5.1 Statistical analyses

5.1.1 Descriptive analyses

The characteristics of the UPAs will be described at the year 2018, which will be hypothesized to be representative of static characteristics. All variables collected will be described as mean (SD), median (IQR) and range or as counts and proportions (%), as appropriate. Distributions will be summarized in tables and visualized in exploratory plots.

5.1.2 Inferential analyses

All inferences will be derived from the statistical model.

5.1.3 Statistical modeling

A multiple Poisson regression will be specified to explore the relations between the accident incidence rates and the UPAs between 1998 and 2018 (ten UPAs, 21 years, 210 observations). The unit of analysis for this study is the UPA, and the incidence rate of accidents will be estimated for each UPA and for each year. The population of the UPA was considered as a proxy for the exposure to accidents. This was incorporated in the model as the offset to compute the incidence rates and allowed to vary over time.

Two models will be adjusted to the data. The first model does not assume an increase in incidence rates in the UPAs. The second model will account for increases in incident accidents a second model by including an interaction term between year and UPA, to allow the incidence rates to increase independently for each UPA.

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

N/A

5.4 Statistical packages

This analysis will be performed using statistical software R version 4.1.2.

6 OBSERVATIONS AND LIMITATIONS

N/A

7 REFERENCES

- **SAR-2021-007-JB-v01** – Incidence rates of scorpion accidents in Urban Planning Areas of Americana/SP in 1998–2018
- Zarin DA, et al. The ClinicalTrials.gov results database — update and key issues. N Engl J Med 2011;364:852-60 (<https://doi.org/10.1056/NEJMsa1012065>).
- Gamble C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318(23):2337–2343 (<https://doi.org/10.1001/jama.2017.18556>).

8 APPENDIX

This document was elaborated following recommendations on the structure for Statistical Analysis Plans (Gamble, 2017) for better transparency and clarity.

8.1 Availability

Both this analytical plan and the corresponding analysis report (**SAR-2021-007-JB-v01**) can be downloaded in the following address:

<https://philsf-biostat.github.io/SAR-2021-007-JB/>