Analytical Plan for Effect of socioeconomic status of neighborhoods in mortality mortality rates after brain injury: retrospective cohort

DOCUMENT: SAP-2023-004-BH-v01

From: Felipe Figueiredo **To:** Brennan Hickson

2023-01-12

TABLE OF CONTENTS

FF Consultoria em Bioestatística e Epidemiologia
CNPJ: 42.154.074/0001-22
https://philsf-biostat.github.io/

SAP

Version	Year	Page
1	2023	1 / 6

Analytical Plan for Effect of socioeconomic status of neighborhoods in mortality mortality rates after brain injury: retrospective cohort

Document version

Version	Alterations
01	Initial version

1   ABBREVIATIONS

- FIM: Functional Independence Measure
- HR: hazards ratio
- SD: standard deviation
- SES: socioeconomic status

2   CONTEXT

2.1   Objectives

To determine the effect of socioeconomic status of the neighborhood on mortality of patients with brain injury.

2.2   Hypotheses

The neighborhood to which an individual is discharged from acute care affects the mortality rates of individuals that suffered a brain injury.

3   DATA

3.1   Raw data

The raw data was received in three distinct databases, one containing clinical and epidemiological data, a second one with follow up on the status of individuals and a third one containing SES information of each Zip code. The clinical and followup tables were merged by study ID, and this was joined with the SES table by the Zip code of the neighborhood of discharge, resulting in the original raw data base.

FF Consultoria em Bioestatística e Epidemiologia
CNPJ: 42.154.074/0001-22

https://philsf-biostat.github.io/

SAP

	Version	Year	Page
	1	2023	2 / 6

Before joining the tables by zip code, the missing locations at discharged were assumed to not have changed since injury, and those values were used to fill missing zip data where available. To maximize the availability of SES data, this was also assumed for each follow up location, where the preferred information used was the zip code at discharge, when available.

The original data base had 711 variables collected on 76665 observations from 19303 individuals.

3.2 Analytical dataset

The date of death of participants was available, and its presence was used as an indicator of the living status of the participant at each follow up. All variables in the raw dataset had varying missing data codes in the data dictionary made available by the researcher, which were used to attribute missingness status to each datum. Most of the categorical variables were measured with many levels that were condensed into fewer levels for analysis.

After the cleaning process 20 variables were included in the analysis. The total number of observations excluded due to incompleteness and exclusion criteria will be reported in the analysis.

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.

4 STUDY PARAMETERS

4.1 Study design

Retrospective cohort.

4.2 Inclusion and exclusion criteria

Inclusion criteria

- Participants with at most 10 years of follow up;
- Participants included in the cohort between 2010-01-01 and 2018-12-31.

Exclusion criteria

Observations after 2019-12-31 will be excluded in order to mitigate risk of confounding by COVID-19 related deaths. Observations prior to this date will still be considered for participants where such data is available.

FF Consultoria em Bioestatística e Epidemiologia
CNPJ: 42.154.074/0001-22

https://philsf-biostat.github.io/

SAP

Version	Year	Page
1	2023	3 / 6

4.3   Exposures

SES of the neighborhood to which the participant was discharged. The SES measure was stratified into its quintiles, and labelled according to the data dictionary to facilitate interpretation of the results.

4.4   Outcomes

Specification of outcome measures (Zarin, 2011):

1.   (Domain) Mortality
2.   (Specific measurement) Death
3.   (Specific metric) Time-to-event
4.   (Method of aggregation) Hazard ratio

Primary outcome

Death after a brain injury.

4.5   Covariates

- Sex
- Race
- Age at injury
- Substance Problem Use
- Education
- Employment status
- Rural area
- Previous seizure disorder diagnosis
- Spinal cord injury
- Cause of injury
- Primary rehabilitation payer
- Residence after rehab discharge
- Days From Injury to Rehab Discharge
- FIM Motor at Discharge
- FIM Cognitive at Discharge

FF Consultoria em Bioestatística e Epidemiologia
CNPJ: 42.154.074/0001-22

https://philsf-biostat.github.io/

SAP

Version	Year	Page
1	2023	4 / 6

5   STATISTICAL METHODS

5.1   Statistical analyses

5.1.1   Descriptive analyses

The epidemiological profile of the study participants will be described. Demographic and clinical variables will be described as mean (SD) or as counts and proportions (%), as appropriate. The distributions of participants' characteristics will be summarized in tables and visualized in exploratory plots.

5.1.2   Inferential analyses

All inferential analyses will be performed in the statistical models (described in the next section).

5.1.3   Statistical modeling

The hazard of mortality will be assessed with multivariate Cox regression models. In order to assess if there is an effect of the SES of the neighborhood to which the participant was discharged on mortality will be assessed with three models. A crude estimate of the HR between each SES quintiles and mortality will be calculated as the basis of interpretation of the effect. The best estimate of the true effect will be calculated adjusting for all covariates described in section 4.5. This full model will be evaluated for the proportional hazards assumption by testing the Schoenfeld residuals. Variables that are significantly associated with time will be removed before the final model is evaluated and described. Finally, the same model will be fitted to a filtered dataset that excludes all deaths happening within one year, to assess if there is an effect on late mortality in a sensitivity analysis.

5.1.4   Missing data

No missing data imputation will be performed. All evaluations will be performed as complete case analyses. Missing data counts and proportions will be reported in tables.

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

FF Consultoria em Bioestatística e Epidemiologia
CNPJ: 42.154.074/0001-22
https://philsf-biostat.github.io/

SAP

	Version	Year	Page
	1	2023	5 / 6

5.4   Statistical packages

This analysis will be performed using statistical software R version 4.2.1.

6   OBSERVATIONS AND LIMITATIONS

Recommended reporting guideline

The adoption of the EQUATOR network (http://www.equator-network.org/) reporting guidelines have seen increasing adoption by scientific journals. All observational studies are recommended to be reported following the STROBE guideline (von Elm et al, 2014).

7   REFERENCES

- **SAR-2023-004-BH-v01** – Effect of socioeconomic status of neighborhoods in mortality mortality rates after brain injury: retrospective cohort
- 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).
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg. 2014 Dec;12(12):1495-9 (https://doi.org/10.1016/j.ijsu.2014.07.013).

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

All documents from this consultation were included in the consultant's Portfolio.

The portfolio is available at:

https://philsf-biostat.github.io/SAR-2023-004-BH/

FF Consultoria em Bioestatística e Epidemiologia
CNPJ: 42.154.074/0001-22
https://philsf-biostat.github.io/

SAP

Version	Year	Page
1	2023	6 / 6