Sensitivity of mortality rates to the imputation of missing socioeconomic data: cohort study

DOCUMENT: SAR-2023-017-BH-v01

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

2023-09-29

TABLE OF CONTENTS

TABLE OF CONTENTS

Sensitivity of mortality rates to the imputation of missing socioeconomic data: cohort study

Document version

Version	Alterations
01	Initial version

1 ABBREVIATIONS

2 CONTEXT

2.1 Objectives

3 METHODS

The data procedures, design and analysis methods used in this report are fully described in the annex document **SAP-2023-017-BH-v01**.

This analysis was performed using statistical software R version 4.3.0.

4 RESULTS

4.1 Study population and follow up

- no changes to the id pool
- slight shift in SES props, but similar overall (also, still homogeneously distributed)

***Table 1** caption*

	Single observation per individual	Multiple observations per individual				
Characteristic	**CC**, N = 7,978[1]	**LOCF**, N = 7,978[1]	**LOCF+NOCB**, N = 7,978[1]	**CC**, N = 24,282[1]	**LOCF**, N = 24,282[1]	**LOCF+NOCB**, N = 24,282[1]
outcome, n	1,003	1,003	1,003	1,006	1,006	1,006
exposure, n (%)						

	Single observation per individual	Multiple observations per individual				
Characteristic	**CC**, N = 7,978[1]	**LOCF**, N = 7,978[1]	**LOCF+NOCB**, N = 7,978[1]	**CC**, N = 24,282[1]	**LOCF**, N = 24,282[1]	**LOCF+NOCB**, N = 24,282[1]
Prosperous	1,421 (22%)	1,428 (22%)	1,430 (22%)	4,363 (23%)	4,573 (23%)	4,577 (23%)
Comfortable	1,327 (20%)	1,339 (20%)	1,341 (20%)	3,862 (20%)	4,070 (20%)	4,072 (20%)
Mid-Tier	1,221 (19%)	1,237 (19%)	1,238 (19%)	3,573 (19%)	3,764 (19%)	3,765 (19%)
At-Risk	1,287 (20%)	1,296 (20%)	1,299 (20%)	3,782 (20%)	3,993 (20%)	3,997 (20%)
Distressed	1,285 (20%)	1,291 (20%)	1,294 (20%)	3,696 (19%)	3,905 (19%)	3,910 (19%)
Missing	1,437	1,387	1,376	5,006	3,977	3,961

[1]CC = complete-cases (not imputed), LOCF = last observation carried forward, NOCB = next observation carried backward

After dropping incomplete cases (data available to the model)

- few individuals lost after complete-cases deletion
- Single: all Single datasets are equal
- multiple:: most deaths are dropped in multiple CC (ZIP not recorded -> no SES)
- multiple: LOCF is equal to LOCF+NOCB
- single death added

dataset	n
sing_cc	693
sing_locf	693
sing_locf+nocb	693
mult_cc	2
mult_locf	694
mult_locf+nocb	694

4.2 Inferential analysis

- FIMCOG4 is included in single CC (table 4)
- FIMMOTD4 and FIMCOG4 are included in multiple LOCF
- Days to rehab is included in multiple LOCF
- multiple: time-varying exposure is associated with the outcome up to model 5 (clinical+FIM scores)

***Table 3** caption*

	Single CC*			Multiple LOCF		
Characteristic	**HR**[12]	**95% CI**[2]	**p-value**	**HR**[2]	**95% CI**[2]	**p-value**
SES quintiles						
Prosperous	—	—		—	—	
Comfortable	0.98	0.78 to 1.25	0.893	1.06	0.83 to 1.35	0.623
Mid-Tier	1.09	0.84 to 1.41	0.515	1.18	0.91 to 1.52	0.207
At-Risk	1.12	0.87 to 1.43	0.386	1.11	0.87 to 1.43	0.400
Distressed	1.21	0.95 to 1.56	0.129	1.33	1.03 to 1.72	**0.027**

[1]* FIM scores violate PH assumption

[2]CC = complete-cases (not imputed), CI = Confidence Interval, LOCF = last observation carried forward

	Single CC*			Multiple LOCF		
Characteristic	**HR[12]**	**95% CI[2]**	**p-value**	**HR[2]**	**95% CI[2]**	**p-value**
term	cc					locf
exposure	0.5					0.5
SexF	0.2					0.2
Race	0.3					0.4
AGE	0.8					0.8
EDUCATION	>0.9					0.7
EMPLOYMENT	0.3					0.2
RehabPay1	0.6					0.7
SCI	0.2					0.9
DAYStoREHABdc	0.055					0.2
PROBLEMUse	0.4					0.8
ResDis	0.4					0.6
RURALdc	0.4					0.3
FIMMOTD4	0.047					0.055
FIMCOGD4	0.2					0.13
GLOBAL	0.3					0.2

5 OBSERVATIONS AND LIMITATIONS

6 CONCLUSIONS

7 REFERENCES

- **SAP-2023-017-BH-v02** – Analytical Plan for Sensitivity of mortality rates to the imputation of missing socioeconomic data: cohort study

8 APPENDIX

8.1 Exploratory data analysis

N/A

8.2 Availability

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

The portfolio is available at:

https://philsf-biostat.github.io/SAR-2023-017-BH/

8.3 Associated analyses

This analysis is part of a larger project and is supported by other analyses, linked below.

Effect of socioeconomic status in mortality rates after brain injury: cohort study

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

Time-adjusted effect of socioeconomic status in mortality rates after brain injury: cohort study

https://philsf-biostat.github.io/SAR-2023-016-BH/

8.4 Analytical dataset

Table A1 shows the structure of the analytical dataset.

Table A1 *Analytical dataset structure*

id	exposure	outcome	Time	SexF	Race	Mar	AGE	PROBLEMUse	EDUCATION	EMPLOYMENT	RURALdc	PriorSeiz	SCI	Cause	RehabPay1	ResDis	DAYStoREHABdc	FIMMOTD	FIMCOGD	FollowUpPeriod	FIMMOTD4	FIMCOGD4
1																						
2																						
3																						
…																						
N																						

Due to confidentiality the data-set used in this analysis cannot be shared online in the public version of this report.