

SAPI Checklist for Statistical Analysis Plans for Observational Studies

Section/Topic	Item	Checklist item
ADMINISTRATIVE INFORMATION		
Project title	1.1	State the title of the research project
Project documents	1.2	Provide links to any documents (e.g., protocol) describing the analysis project for this SAPI, if available
Ethics approval	1.3	Provide details of any ethics approval
Names and contact	1.4	Provide names, affiliations, and contacts of key project team members
PROJECT BACKGROUND		
Research aims and objectives	2.1	Describe the research aims and objectives of this analysis project
Target population	2.2	Describe the target population of interest for the research objectives
DESIGN AND DATA		
Data sources	3.1	Describe the sources of data (e.g., routinely collected electronic health records, registries, surveys)
Design	3.2	State the study design for this analysis project (e.g., cross-sectional, cohort, case-control, case-crossover, serial cross-sectional)
Eligibility criteria	3.3	Specify eligibility criteria for the research project based on information available at baseline
Sampling	3.4	Describe the sampling strategy for selecting eligible observation units from the data source for this analysis project
Data sets	3.5	Describe how the data is provided for this analysis project (e.g., the format and content of datasets) and summarize data preprocessing)
VARIABLES		
Variables used in the main data analysis	4.1	Define all variables and their roles in the analyses to answer the research objectives
Other variables	4.2	Define any variables that are not directly used to answer the research objectives, but which provide information about the observation units
METHODS: MAIN DATA ANALYSIS (MDA)		
Description of observation units	5.1	Describe the methods of analysis to summarize the characteristics of the observation units
Main data analysis methods	5.2	Describe the methods of analysis for each research objective
Assumptions and diagnostics	5.3	State any statistical assumptions of each analysis. Specify all measures and diagnostics used to evaluate statistical assumptions and appropriateness of analyses, including graphical tools
Sample size	5.4	Describe how the sample size was determined, including all assumptions supporting the sample size calculation
Software	5.5	Describe software used for all analyses, visualizations, data management, data archiving, or backups
METHODS: INITIAL DATA ANALYSIS (IDA)		
Data preparation	6.1	Describe any methods for preparing data for the analyses
Unit missingness	6.2	Describe any methods to identify the extent of unit missingness
Unit profile	6.3	Describe any methods to summarize the temporal or structural pattern of observations for each observation unit

Item missingness	6.4	Describe the methods to examine item missingness (i.e., missing values in variables)
Univariable descriptions	6.5	Describe the methods to summarize the distribution of each variable used in the analyses with numerical or graphical summaries
Multivariable descriptions	6.6	Describe any multivariable descriptive statistics and graphical summaries
EVALUATION AND UPDATES		
Evaluating the SAP	7.1	Indicate if an update of the SAP is needed after IDA
Updating the SAP, if applicable	7.2	List all SAP sections affected by updates and state the reasons for the updates
SUPPLEMENT		
Key references	8.1	Include references related to statistical methods, research aims, or other background information
Sustainable handling of analysis outputs	8.2	Describe the measures to ensure sustainable handling of analysis outputs aligned with FAIR principles, e.g., sharing or archiving analysis reports, code files, updated datasets, documentation for interoperability and reuse

Abbreviations: SAPI – statistical analysis plan with initial data analysis, MDA – main data analysis, IDA – initial data analysis, FAIR – findable accessible, interoperable, and reusable