

Baillie M, le Cessie S, Schmidt CO, Lusa L, Huebner M, for the Topic Group “Initial Data Analysis” of the STRATOS Initiative (2022): Ten simple rules for initial data analysis. PLoS Comput Biol 18(2): e1009819. <https://doi.org/10.1371/journal.pcbi.1009819>

Short summary

This paper presents ten simple rules for conducting initial data analysis (IDA) in research projects. IDA is an essential step to ensure responsible data analysis and avoid pitfalls. The rules cover various phases of IDA, including metadata setup, data cleaning, data screening, initial data reporting, refining the research analysis plan, and documenting and reporting IDA. The rules are based on extensive experience and are applicable to researchers in various fields, emphasizing the importance of considering data properties and context, highlighting the distinction between IDA and exploratory data analysis (EDA) and that IDA should be carefully designed based on the specific problem and context, rather than following a routine approach.