Huebner M, le Cessie S, Schmidt C, Vach W on behalf of the Topic Group "Initial Data Analysis" of the STRATOS Initiative (2018): A Contemporary Conceptual Framework for Initial Data Analysis. Observational Studies 4 (2018):171-192.

Short summary

This paper proposes a contemporary conceptual framework for conducting Initial Data Analysis (IDA). IDA is crucial for ensuring transparent and reliable research outcomes to recognize IDA as an independent element of the research process. The framework consists of six steps: metadata setup, data cleaning, data screening, initial data reporting, refining and updating the analysis plan, and reporting of IDA in research papers. Each step is discussed, including types of errors to be addressed in the data cleaning step, topics and methods for data screening, and examples with potential influence on the statistical analysis plan. Practical challenges related to IDA are cost, IDA teamwork, and limits of manual inspection. The paper highlights the need for a comprehensive view of IDA and its role in the research process, and a clear definition of the aim of IDA.