

# Researchers propose new standards to fix what's wrong in statistics

April 26 2023

---



Credit: CC0 Public Domain

A new paper in the *Journal of Survey Statistics and Methodology* indicates that the methods researchers use to report on analyses of survey

data vary widely, and frequently contain mistakes. Publications containing these incorrect analyses yield results that can misinform policymakers, researchers, and practitioners. The researchers here propose new standards to improve the reporting of analyses using complex sample survey data.

For decades, researchers have documented methodological problems and analytic errors that are common in papers that use complex sample [survey](#) data of populations. These surveys employ sampling design features that—when used appropriately—can produce unbiased estimates of a population. For example, population samples routinely use complex design features to improve statistical efficiency, reduce costs, and increase sample sizes of underrepresented populations. However, complex samples deviate from simple random samples. This has important implications for analyzing and reporting the results.

By default, most statistical software programs assume that data come from simple random samples. But not all survey data are collected using a simple random sample. It is therefore essential that investigators use the correct software procedures to account for complex sample design features when analyzing such data. Failing to account for complex design features can yield biased estimates and incorrect interpretations of the material.

A 2016 paper analyzed data from the Scientists and Engineers Statistical Data System and found that only 7.6% correctly accounted for sampling in variance estimation. The same paper found that a little more than half (54.5%) of papers correctly accounted for the sampling weights in [analyses](#) and only 10.7% of papers used appropriate subpopulation estimation. A separate review of publications analyzing data from the National Inpatient Sample found that some 80% of papers did not account for the sample's clustering and stratification. Another analysis found that less than half of papers analyzing data from the Medicare

Current Beneficiary Survey described appropriate weighting or variance estimation.

The researchers here propose an itemized [checklist](#) to guide researchers in publishing analyses using complex sample survey data. The checklist, which they call the Preferred Reporting Items for Complex Sample Survey Analysis (or PRICSSA), consists of 17 important items to report for any analyses conducted on complex survey data, including sample sizes for all estimates, missing data rates and imputation methods, information about any data deleted, and an explanation about survey weighting and variance estimation. In addition to the checklist, the investigators here propose that researchers using complex [survey data](#) make all corresponding software code available.

The authors believe that such reforms could greatly increase transparency and make analytic mistakes easier to spot. This, in turn, would make academics or other researchers less likely to commit them. The researchers here emphasize that they modeled their checklist after other checklists, such as the PRISMA checklist, widely used for systematic reviews and meta-analyses, and the CONSORT guidelines, which are standard in randomized trials.

Scholars and institutions have invested tremendous resources into survey design and data collection to try to produce accurate population estimates. Analyzing such data correctly necessitates that researchers incorporate certain complex survey design features into their work. The authors of this paper want to ensure that results reported in peer-reviewed publications do not misinform policymakers, practitioners, and researchers. They argue that their proposed checklist has the potential to increase the rigor and reproducibility of survey research by improving the quality of analysis and increasing transparency.

"It's a problem when papers get published and the analyses were

performed incorrectly or cannot be reproduced," said the [paper](#)'s lead author, Andrew Seidenberg. "We created this checklist to help prevent that from happening."

**More information:** Andrew B Seidenberg et al, Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA), *Journal of Survey Statistics and Methodology* (2023). [DOI: 10.1093/jssam/smac040](https://doi.org/10.1093/jssam/smac040)

Provided by Oxford University Press

Citation: Researchers propose new standards to fix what's wrong in statistics (2023, April 26) retrieved 5 May 2024 from <https://phys.org/news/2023-04-standards-wrong-statistics.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.