

New Methods and Software to Determine the Impact of Missing Data in Clinical Trials

Principal investigator

Daniel O. Scharfstein, ScD

Organization

Johns Hopkins University

What was the research about?

Clinical trials study the effects of medical treatments, like how safe they are and how well they work. But most clinical trials don't get all the data they need from patients. Patients may not answer all questions on a survey, or they may drop out of a study after it has started. The missing data can affect researchers' ability to detect the effects of treatments.

To address the problem of missing data, researchers can make different guesses based on why and how data are missing. Then they can look at results for each guess. If results based on different guesses are similar, researchers can have more confidence that the study results are accurate. In this study, the research team created new methods to do these tests and developed software that runs these tests.

What were the results?

The research team created software and tested it by analyzing missing data from three clinical trials. For example, one clinical trial looked at a medicine for bipolar disorder. More than half of patients withdrew from the trial, creating missing data. Using the software to run tests, the team found that missing data didn't change the results of the trial.

What did the research team do?

The research team created new software to run the tests and made the software available online for free. The team also used this software on data from previous clinical trials to see if study results would differ depending on how the team handles missing data.

An advisory panel of 15 people in the fields of statistics, software development, and medicine helped solve technical problems when creating the software.

What were the limits of the study?

The software doesn't work with data that have outliers. Outliers fall far outside the normal range of other data within a study. In addition, data from studies may not be available to the public, which makes it hard to test the software further.

Future research could create software that works with data with outliers.

How can people use the results?

Researchers can consider using the software to see if missing data affect their results.

To learn more about this project, visit www.pcori.org/Scharfstein205.