Methods for Variable Selection and Treatment Effect Estimation in Nonrandomized Studies with Few Outcome Events and Many Confounders

Principal investigator
Jessica M. Franklin, PhD

Organization
Brigham and Women's Hospital

December 2017

What was the research about?
Comparative effectiveness research compares two or more treatments to see which one works best for which patients. Information from health insurance claims could be useful for this type of research. These claims include data on how well patients respond to treatments. But many things—not just treatments—affect whether patients’ health improves. How well patients respond to treatments could depend on patients’ ages or medicines they take. It could also depend on how many health problems a patient has and how severe the problems are. Also, a doctor may suggest one treatment instead of another because of a patient’s personal situation and health. Researchers need ways to figure out whether changes in patients’ health result from treatment or something else.

Comparing treatments is hard in small studies with only a few patients. When there are few patients in a study, researchers can study only a few events. An event is an outcome related to the health problem or treatment researchers are studying. When there are few events and many things that could affect treatment results, it is hard to figure out what causes changes in patients’ health. To address this problem, researchers use different statistical methods to account for all the things that could affect treatment results. But researchers don’t know which methods might work best in studies with few events. In this study, the research team compared several methods to see which ones worked best.

What were the results?
The research team found that certain statistical methods worked better than others to account for all the things that could affect treatment results in studies with few events.

What did the research team do?
The research team wanted to see which statistical methods worked best to account for things that could affect treatment results. To do this, the team made a test set of health insurance claims using real patient data. The team made sure the set had only a few events and many things other than treatment that could affect the results. The team also made sure the test set had information on what happened after each patient got treatment. The test set made it possible to see which methods worked best.

During the study, patients gave input to the research team about the issues that are important to them in research that uses health insurance claims.

What were the limits of the study?
This study compared different statistical methods using data created by the research team. Studies using different data may have different results. Also, the results may not apply to all types of data.
How can people use the results?
This study can help researchers understand which statistical methods to use when doing a study with few events and many things that could affect treatment results. Knowing which methods work best can help researchers use health insurance claims to get information that patients can use to choose between treatment options.

To learn more about this project, visit pcori.org/Franklin049.