Data Studio

MSOB x399

DATE: February 20, 2019
TIME: 1:30 - 2:50 pm
TITLE: I wish we’d measured that at the start! Filling in missing baseline data based on a (biased) sample

PRESENTER: Mark Hlatky & Marc Søndergaard (Health Research and Policy)

Summary:

We are interested in the question of whether adverse pregnancy outcomes (APOs) in women, like gestational diabetes, are risk factors for developing heart disease in later life. We are working with the Women’s Health Initiative (WHI) data for the study. This is a large cohort of women (about 161,000) enrolled about 20 years ago, with extensive baseline data collection and scrupulous follow-up for detecting the development of heart disease. So the cohort is ideal for our study, BUT the investigators didn’t collect baseline data on APOs, the exposure of interest. So we did a survey in 2017 to collect data on APOs from about 51,000 women, and are trying to devise the best analysis strategy to relate APOs to heart disease.

Our "missing data problems" range from small (some of the 50,000 respondents did not answer all the survey questions, but did answer some of them) to very large (we are missing baseline APO data on about 110,000 women, many of whom have died from heart disease or other causes).