TITLE: THE PERFORMANCE OF INSTRUMENTAL VARIABLES IN MEDICAL DEVICE EPIDEMIOLOGY: A SIMULATION STUDY OF NON-NORMALLY DISTRIBUTED CONTINOUS OUTCOME DATA BASED ON REAL WORLD POST-OPERATIVE PROMS

AUTHORS: ALBERT PRATS-URIBE, M SANNI ALI, VICTORIA Y STRAUSS, DANIEL PRIETO-ALHAMBRA

Background: Patient reported outcomes (PROMs) are often used in surgical RCTs for comparative efficacy studies. UTMOST, a cohort study assessing the comparative effectiveness of partial and total knee replacement was able to replicate an RCT using propensity scores. However, instrumental variables (IV) based on surgeon's preference failed to emulate the trial, yielding implausible estimates.

Objectives: To estimate bias and power of IV analyses of non-normally distributed continuous outcome data based on a PROM.

Methods: Different scenarios of outcome distributions and violation of IV assumptions were explored. We simulated data (n=130,000) using the distributions: gamma distributed (0.3,3) IV truncated at 1 and binarised at median to mimic UTMOST's IV; binary treatment based on a logit function of IV and 3 confounders; outcome as a linear function of IV, treatment effect of 2 (same as the RCT effect), and the 3 confounders. Different strengths of IVs (IV-Treatment correlation, with β ranging from 0 to 10^9, equivalent to OR 1 to 20), and distributions of outcome (rounded normal (30,8) and truncated rounded normal (30,8)) were considered. We tested different scenarios of patient eligibility for treatment (50% to 100%), treatment prevalence (10% to 50%), and IV violations, introducing a direct effect of the IV on the outcome (β 0 to 20). Two-stage least squares regression was used to estimate treatment effects, and relative bias and mean standard error (MSE) were calculated.

Results: In scenarios with low IV strength, IV estimates were extremely biased and inaccurate (high MSE and up to 1,000% bias). Conversely, bias was low (<2%) when IV was independent of outcome satisfying the assumption, and remained reliable with minor violations: bias was low (<5%) when IV strength β was >10,000-fold the IV-outcome correlation β. Exposure prevalence did not affect bias, nor did patient eligibility. Analyses of a truncated normal outcome yielded systematically lower estimates compared to non-truncated normal. Binarization of the IV did not introduce bias compared to continuous IV use.

Conclusions: IVs should be used with caution when dealing with non-normally distributed and/or truncated continuous outcomes, e.g. PROMs. Bias increased in the presence of weak IVs. Violation of IV assumptions introduced bias, except when the IV was very strong yet slightly correlated with outcome. More research including multi-level simulations are needed to further understand the reliability of IV methods for medical device epidemiology.