Synthesis
METHODS NOTE (E156 entry 348)

OutcomeSwitchDetector: Automated Detection of Primary Outcome Switching in ClinicalTrials.gov Protocols

Zaki Khan
Ahmadiyya Hospital, Kenema, Sierra Leone
Correspondence: zakikhan101@gmail.com

Primary estimand: Primary-outcome switching rate (proportion with any primary-outcome change between first and latest protocol version, 95% CI Wilson) across 1,000 interventional ClinicalTrials.gov trials.

Abstract

Can automated text comparison detect primary-outcome switching between the first and latest registered protocol versions on ClinicalTrials.gov? Using the ClinicalTrials.gov version-history interface, we sampled 1,000 interventional trials with at least two posted protocol versions and compared their first- and latest-version primary outcomes. A diff engine with semantic endpoint parsing matched outcomes by normalized text similarity, classifying primary-outcome changes as additions, removals, substitutions, or rank reversals. A pre-specified primary outcome had changed between the first and latest version in 26.0% of trials (95% CI 23.4 to 28.8), with outright substitution in 15.9% and demotion to secondary in 2.4%. Switching was far commoner among the 21% of trials that posted results (73.2%) than those without (13.5%), and highest in oncology (33.5%). Automated version-history screening therefore flags primary-outcome switching reproducibly and at scale, surfacing candidates for manual adjudication of clinical importance. Detection is limited to ClinicalTrials.gov structured outcome fields and to trials registering more than one protocol version.

FIGURES

Figure 1. Primary-outcome switching between the first and latest registered protocol version across 1,000 trials with ≥2 versions (Wilson 95% CIs). Switching far more frequent among trials posting results (73.2%) vs. without (13.5%).

Figure 2. Left: switching rate by specialty (highest in oncology 33.5%). Right: composition of primary-outcome changes (any change 26.0%; substitution 15.9%).

Visual abstract.

METHODS NOTE

Sample: 1,000 interventional trials on ClinicalTrials.gov with ≥2 posted protocol versions. Comparison: first-posted vs. latest-posted primary outcome text. Diff engine: normalized text similarity with semantic endpoint parsing. Change categories: addition, removal, substitution, rank reversal. Statistics: Wilson 95% CI for proportions; subgroups by results-posting status and specialty. Source data: ClinicalTrials.gov / AACT.

DATA & CODE AVAILABILITY

No patient-level data. Analysis derived from publicly available aggregate ClinicalTrials.gov / AACT records. Verification scripts and reproduced result files accompany this package (work/). Reporting checklist: methodological/software note. Funding: None. Competing interests: None declared.

REFERENCES

1. Zarin DA, Tse T, Williams RJ, Rajakannan T. Update on trial registration 11 years after the ICMJE policy was established. N Engl J Med. 2017;376(4):383-391.

2. Anderson ML, Chiswell K, Peterson ED, et al. Compliance with results reporting at ClinicalTrials.gov. N Engl J Med. 2015;372(11):1031-1039.

3. Chan AW, et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004;291(20):2457-2465.

4. Dwan K, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One. 2008;3(8):e3081.