

USING META-ANALYSIS TO EVALUATE HETEROGENEITY ACROSS A NETWORK OF HEALTH DATA PROVIDERS

Cindy Liu, Aongus O’Murchadha, Richard Rast
TriNetX, Inc., Cambridge, MA, USA | richard.rast@trinetx.com



TriNetX

OBJECTIVES

It is often desirable to use multiple data sources in outcomes research to increase statistical power. However, it cannot be assumed that all are homogeneous. This study compares pooled versus meta-analysis estimates of the relative risk (RR) of treatment outcomes expected to be reasonably consistent across health data providers, to determine if unobserved heterogeneity exists, and to compare methods of dealing with it if it does.

METHODS

Using the EMRs of 32 independent healthcare organizations (HCOs) who are TriNetX data network contributors, patients aged 70+ with a diagnosis of atrial fibrillation within the past 5 years were selected for study. Patients were split into two cohorts, those prescribed apixaban within a year of their initial diagnosis and those prescribed warfarin. Patients were identified and categorized based on diagnosis codes and prescribed medications found in their EMR.

The RR and 95% confidence interval (CI) of stroke within 3 years of starting treatment was calculated for the pooled population and each HCO separately. Meta-analysis with fixed- and random-effects were used to assess the impact of unobserved heterogeneity across the HCOs on the estimate of RR.

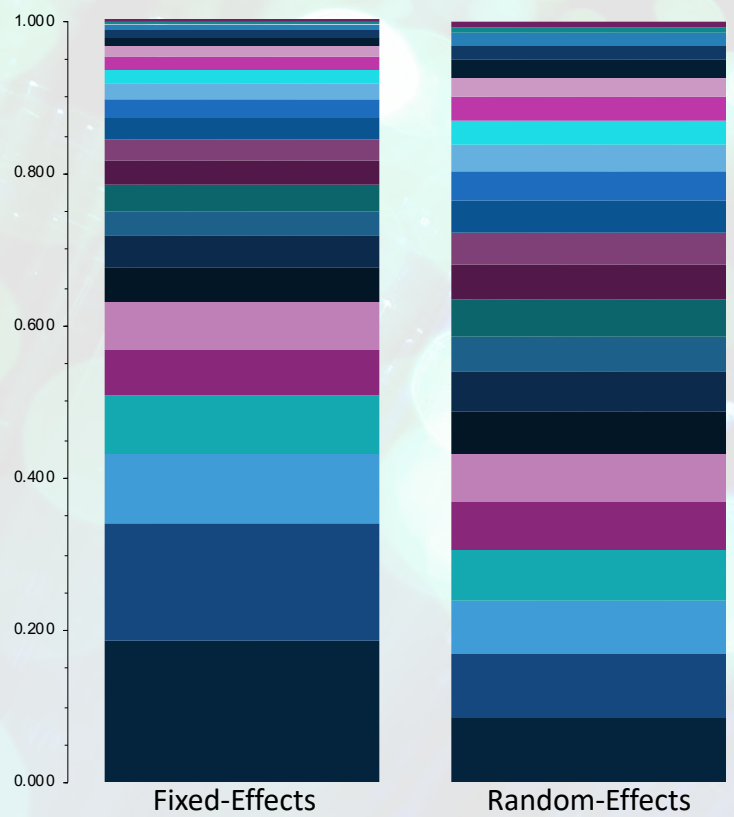


Figure 2. Relative weights by data source (fixed and random effects)

The fixed effects model gives much more weight to more powerful estimates. The random effects model gives slightly more weight to all estimates, to compensate more heavily for between-source heterogeneity.

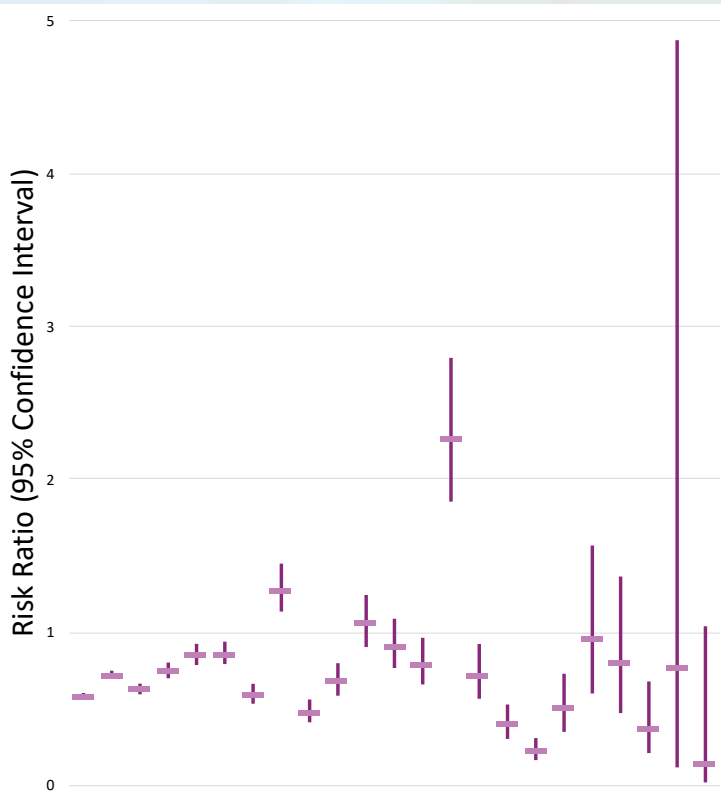


Figure 1. Risk ratio by data source, excluding sources with undefined weight

Many data sources have extremely high statistical power, but the confidence intervals do not overlap. This indicates high levels of source-level heterogeneity, and cautions against using any individual data source for the entire study.

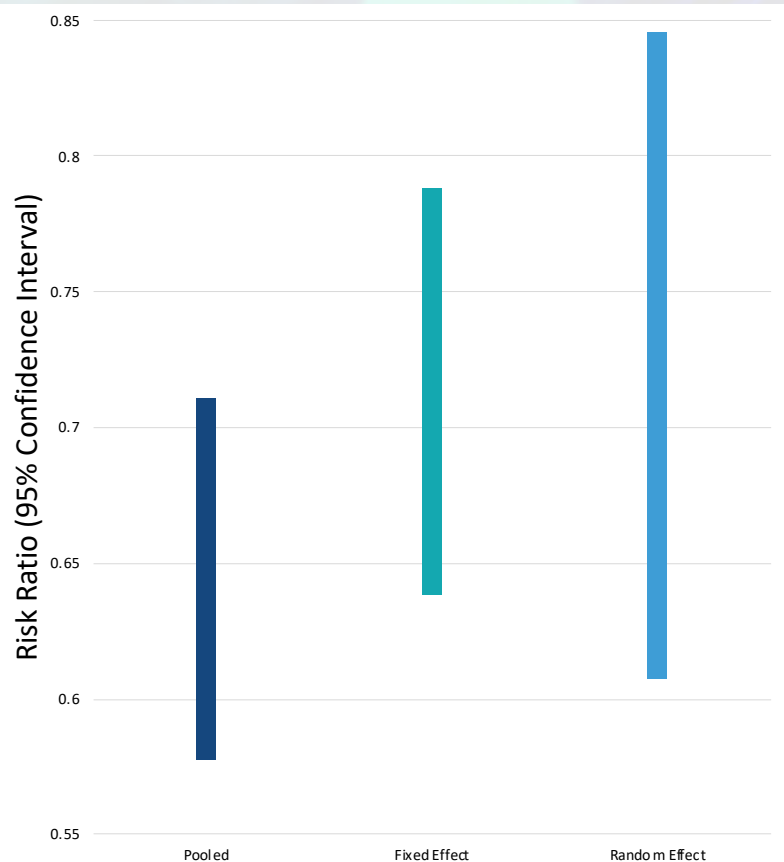


Figure 3. Risk ratio by pooling method. Meta analysis methods (fixed effects and random effects) exclude sources with undefined weight

Despite source-level variation, the combined confidence intervals all overlap and give similar center estimates.

RESULTS

The study cohorts consisted of 13,636 apixaban and 14,943 warfarin patients across 32 data sources, of which 23 had sufficient cohort size to be included for meta-analysis.

Pooling patients across the HCOs within each cohort, 4.404% and 6.866% of patients respectively had a stroke event [RR= 0.641; 95% CI (0.581, 0.707)]. Computing the RR via a fixed-effects meta-analysis model yielded a RR= 0.709 [95% CI (0.642,0.784)] and using a random-effects model yielded a RR= 0.718 [95% CI (0.611,0.842)].

CONCLUSIONS

Within this HCO data network, both fixed- and random-effects meta-analysis yielded a slightly attenuated, but still statistically significant RR of stroke for apixaban versus warfarin compared to the pooled RR. This analysis demonstrated that meta-analysis can be used to determine if unobserved heterogeneity within an HCO data network leads to a conclusion different than estimates from the pooled population.