Name:	Chi Yuan
Affiliation:	Columbia University
Email:	cy2465@cumc.columbia.edu
Presentation type (s):	**Poster**

Generalizability of randomized trials' follow-up time to real-world practice

Chi Yuan[1], Patrick Ryan[1,2], Chunhua Weng[1]

1. Department of Biomedical Informatics, Columbia University, 2. Janssen Research and Development

Introduction

Randomized clinical trials (RCT) typically serve as a foundation for evidence about the effects of medical products. However, the generalizability of evidence from RCT to real-world clinical practice is often in question, as trials are often small and impose stringent inclusion criteria that many patients in routine care do not satisfy. In this work, we examine follow-up time as a novel component of generalizability. We seek to study how well the length of exposure in RCT studies reflects the typical length of exposure observed on clinical practice.

Methods

We developed a natural language processing procedure to extract the follow-up time frame, intervention and enrollments from clinical trial summaries posted at the public Clinicaltrials.gov. Interventions and enrollments are extracted from XML format of trial summaries using an XML parser, with an example shown below:

```
<enrollment type="Actual">1864</enrollment>
<intervention_browse>
        <mesh_term>Amlodipine</mesh_term>
        <mesh_term>Valsartan</mesh_term>
        <mesh_term>Amlodipine, Valsartan Drug Combination</mesh_term>
</intervention_browse>
```

In order to extract and standardize temporal information from clinical trial summaries, we developed a temporal tagger based on SUTime (https://nlp.stanford.edu/software/sutime.html). It extracts maximum time frame for primary outcomes, secondary outcomes and other outcomes respectively and calculate the longest time frame among them. With the standardization process, all time frames were converted to numbers by using "days" as their unit.

We restricted our analysis to Phase 3 trials. We used the Truven MarketScan Commercial Claims and Encounters (CCAE) database, standardized to the OMOP Common Data Model v5, to estimate a distribution of exposure time for each ingredient in the DRUG_ERA table. We created new user cohorts for each ingredient, requiring at least 1 year of observation pre- and post-index exposure, and followed patients on the target drug until they had at least 30 days with no record of exposure. For each drug, we computed the median and maximum length of follow-up in randomized clinical trials. We then estimated the proportion of new users of each drug from CCAE who had exposure length greater than the median and maximum RCT follow-up length. We summarized the results across the entire drug portfolio, and explored specific ingredients.

Results

We extracted follow-up time from 9.135 Phase 3 trials in clinicaltrials.gov, covering 1,670 distinct interventions. 1,324 of those interventions were able to be mapped into standard drug ingredient concepts, 914 of which had exposures in CCAE to yield an exposure length estimate. Across all drugs, 6% of persons are exposed to a drug longer than the longest RCT trial for that drug. 24% of drugs have >10% of persons with exposures longer than max RCT length. Figure 1 shows distribution of real-world exposure compared with RCT exposure-time for 6 commonly-used drugs. These drugs illustrate that the distribution of real-world exposure and distributions of trial lengths can vary substantially, which impacts the proportion of patients who have real-world use outside the bounds of what was studied

in RCTs. Figure 2 shows that most drugs with median follow-up time in RCTs > 1 year have <40% of patients with exposure time greater than median RCT exposure time.

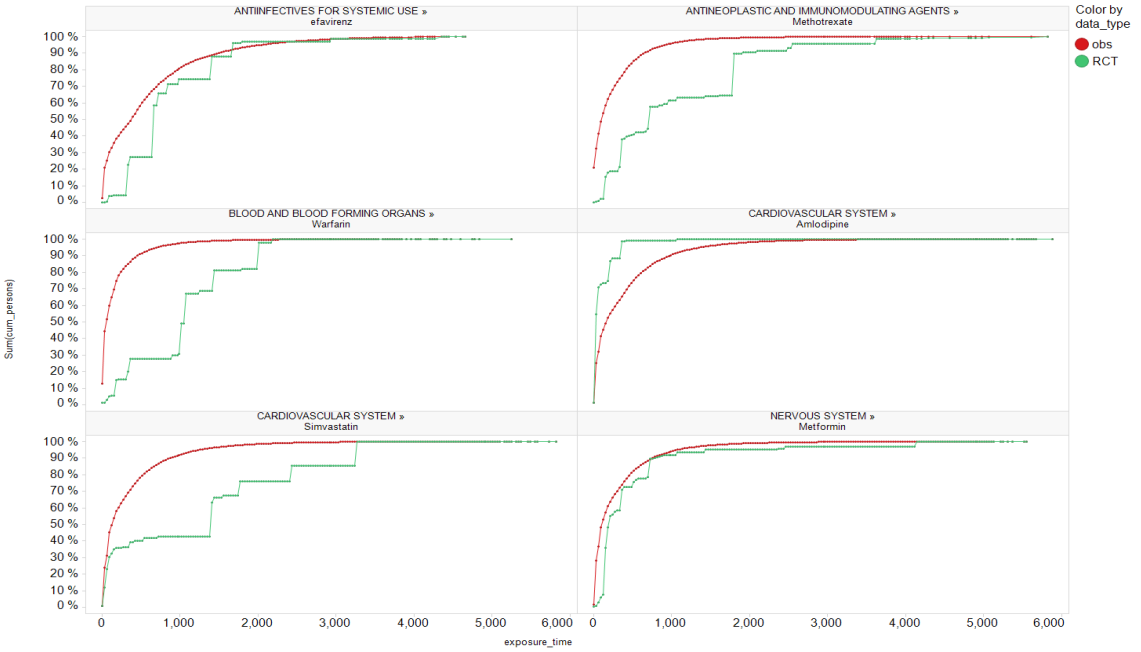

Figure 1. Comparison of distribution for follow-up time from RCT and claims data for select drugs

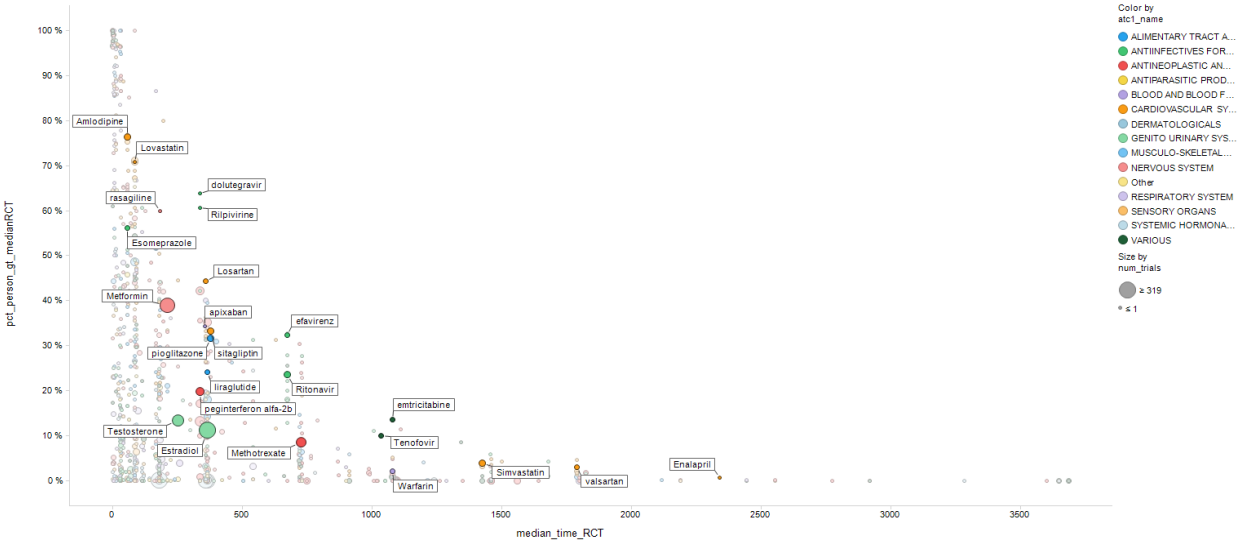

Figure 2. Summary of drugs by median follow-up time in RCTs (x-axis) vs. % of persons in CCAE whose exposure length is greater than the median follow-up time in RCTs

Conclusions

Most patients in clinical practice are exposed to a drug for a period that is shorter than the longest RCT for that drug, but many drugs are used by patients for a period longer than previously studied. Given the lack of generalizability from the RCT to these populations, further research is required to understand the long-term impact of exposure.