

Development of an automated comparator ranking algorithm for the REWARD initiative

Justin Bohn, James P. Gilbert, Christopher Knoll, David M. Kern

Background

The REWARD framework was developed to identify potential unknown benefits of marketed drugs in multiple large real-world data sources (1–3). This framework relies on the self-controlled cohort (SCC) design to efficiently generate effect estimates for large numbers of medication-outcome pairs, without the need for pre-specification of comparator medication cohorts. While the self-controlled cohort design is computationally simple and controls for confounding by time-invariant factors, it is susceptible to protopathic bias and other related temporal biases (4). To enable an alternative comparative cohort method within REWARD without the need for pre-specification or computationally-intensive multivariable approaches like propensity scores, we sought to develop an algorithm to select medications that could serve as reasonable comparators for target medications of interest based on covariate similarity. We then assessed the output of this algorithm for six medications from three disease areas: anticoagulation (warfarin, apixaban), multiple sclerosis (glatiramer acetate, ocrelizumab), and rheumatoid arthritis (methotrexate, adalimumab).

Methods

We utilized four US administrative claims databases described in (1) plus data from Japan Medical Data Center. Our proposed comparator ranking algorithm is as follows. First, within each data source, we identified cohorts of new users of all RxNorm ingredients. Second, feature extraction was performed on each cohort to generate covariates related to historical condition occurrence, medication use, healthcare utilization, and demographics. Third, all possible comparisons involving the six medications of interest were formed, and standardized mean differences were calculated for all covariates in each comparison. Fourth, lists of potential comparators were generated for each of the six medications of interest, ranked by the average absolute standardized mean difference (AASMD).

Results

Between 988 – 1,359 cohorts were created within each data source, with an average of 8,781 features extracted in each. An average of 9,294 comparisons involving the six medications of interest were considered. Figure 1 depicts how top- and bottom-ranked comparators differ in terms of covariate prevalence relative to the target cohort. Among the top comparators for each medication of interest (Table 1), several expected comparators can be found, including other anticoagulants (fondaparinux and rivaroxaban for warfarin and apixaban) and anti-rheumatic drugs (hydroxychloroquine and sulfasalazine for methotrexate and adalimumab). Across all comparisons, more than 75% of covariates had an absolute standardized mean difference < 0.1 (Figure 2).

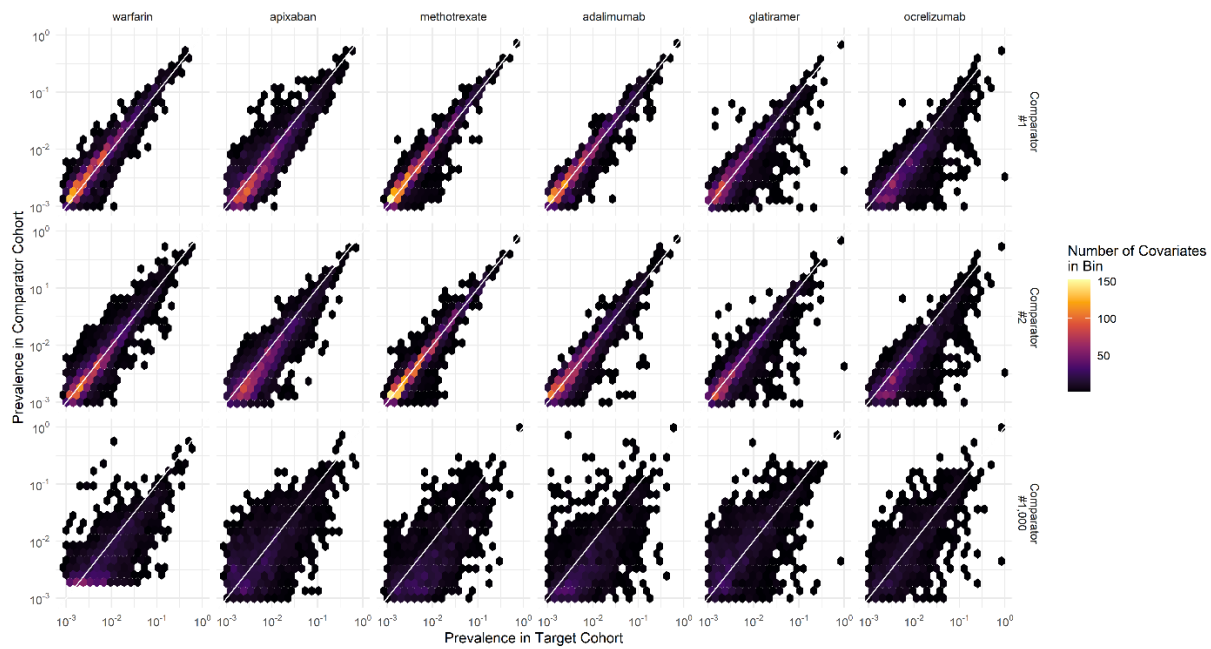


Figure 1: Distribution of covariate prevalence for six selected target cohorts and three selected comparators (the top 2 comparators and the 1,000th ranked comparator) for each target from the MarketScan CCAE database, demonstrating increasing degrees of divergence with increasing rank. Note that continuous covariates and binary covariates with prevalence below the limit of observation in either group were removed for plotting.

Comparator rank	warfarin	apixaban	methotrexate	adalimumab	glatiramer	ocrelizumab
1	fondaparinux [0.0114]	rivaroxaban [0.0170]	hydroxychloroquine [0.0082]	sulfasalazine [0.0088]	alprazolam [0.0131]	influenza B virus antigen [0.0224]
2	digoxin [0.0121]	perflutren [0.0182]	sulfasalazine [0.0083]	etanercept [0.0109]	citalopram [0.0135]	influenza A virus (H3N2) antigen [0.0224]
3	furosemide [0.0125]	umeclidinium [0.0194]	etanercept [0.0101]	methotrexate [0.0110]	diazepam [0.0137]	influenza A virus (H1N1) antigen [0.0224]
4	adenosine [0.0135]	protamine sulfate (USP) [0.0204]	adalimumab [0.0110]	infliximab [0.0113]	butalbital [0.0139]	propofol [0.0226]
5	diltiazem [0.0135]	propofol [0.0207]	lidocaine [0.0111]	hydroxychloroquine [0.0116]	caffeine [0.0139]	gadobutrol [0.0236]

Table 1: Top five comparators for six medications of interest. Values inside brackets are average absolute standardized differences.

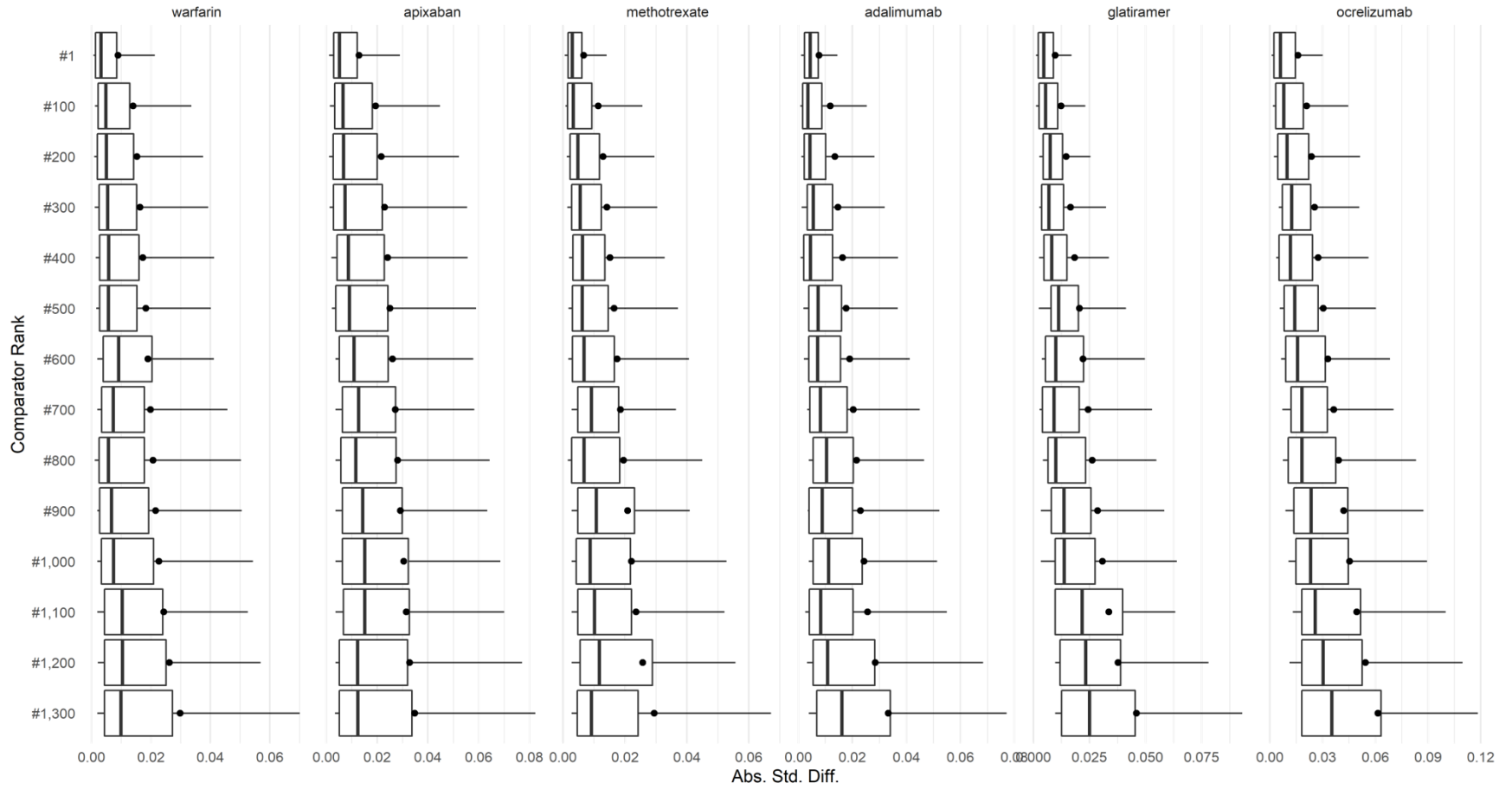


Figure 2: Distribution of absolute standardized mean differences for six target medications and selected comparators. Whiskers are set to percentiles 10 and 90, with the average plotted as a filled circle.

Conclusion

We demonstrated one approach to ranking comparator cohorts based on cohort similarity. Within comparisons, the average degree of covariate imbalance was well below established norms (i.e., $|SMD| \leq 0.1$). In some cases, top-ranked comparators corresponded with subject-matter expectation. Further research is needed to determine if highly ranked but unexpected comparators are in fact suitable for comparative inference in REWARD. Some of these comparators may in fact mimic non-user cohorts, but without their well-known pitfalls, and as such be attractive in the REWARD context, where the desired causal contrast is between a given drug and its absence. Alternative ranking methods and comparisons against rankings derived from balance on pre-specified covariates will be explored in future studies.

References/Citations

1. Kern DM, Cepeda MS, Flores CM, Wittenberg GM. Application of Real-World Data and the REWARD Framework to Detect Unknown Benefits of Memantine and Identify Potential Disease Targets for New NMDA Receptor Antagonists. *CNS Drugs*. 2021 Feb;35(2):243–51.
2. Cepeda MS, Kern DM, Seabrook GR, Lovestone S. Comprehensive Real-World Assessment of Marketed Medications to Guide Parkinson's Drug Discovery. *Clin Drug Investig*. 2019 Nov;39(11):1067–75.
3. Kern DM, Cepeda MS, Lovestone S, Seabrook GR. Aiding the discovery of new treatments for dementia by uncovering unknown benefits of existing medications. *Alzheimers Dement Transl Res Clin Interv*. 2019 Jan;5(1):862–70.
4. Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2020 Sep;29(9):1101–10.