medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Large-scale empirical identification of candidate comparators for pharmacoepidemiological studies

Justin Bohn[1]
James P. Gilbert[1]
Christopher Knoll[1]
David M. Kern[1]
Patrick B. Ryan[1]

[1]Department of Global Epidemiology
Janssen Research and Development, LLC, Titusville, NJ, USA

Correspondence to:
Justin Bohn
1125 Trenton-Harbourton Road
Titusville, NJ, USA, 08560
jbohn2@its.jnj.com

Total word count	3,275
Abstract word count	248
Figures	4
Tables	3
References	34

Keywords (MeSH):

D017891	Pharmacoepidemiology
D015331	Cohort Studies
D057186	Comparative Effectiveness Research
D011358	Product Surveillance, Postmarketing
D015982	Bias
D015986	Confounding Factors, Epidemiologic
D000067575	Administrative Claims, Healthcare

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ABSTRACT

Objectives

The active comparator new user (ACNU) cohort design has emerged as a best practice for the estimation of drug effects from observational data. However, despite its advantages, this design requires the selection and evaluation of comparators for appropriateness, a process which can be challenging, especially in the context of many drugs. In this paper, we introduce an empirical approach to rank candidate comparators in terms of their similarity to a target drug in high-dimensional covariate space.

Methods

We generated new user cohorts for each RxNorm ingredient in five administrative claims databases, then extracted aggregated pre-treatment covariate data for each cohort across five clinically oriented domains. We formed all pairs of cohorts with ≥ 1,000 patients and computed a scalar similarity score defined as the average of cosine similarities computed within each domain for each pair. Ranked lists of candidate comparators were then generated for each cohort.

Results

Across up to 1,350 cohorts forming 922,761 comparisons, drugs that were more similar in the ATC hierarchy tended to have higher cohort similarity scores, and the most similar candidate comparators for each of six example drugs consistently corresponded to alternative treatments for the target drug's indication(s) that could be identified in the literature or publicly registered studies. 80%-95% of cohorts had at least one comparator with a cohort similarity score ≥ 0.95.

Conclusion

Empirical comparator recommendations may serve as a useful aid to investigators and could ultimately enable the automated generation of ACNU-derived evidence, a process that has previously been limited to self-controlled designs.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

BACKGROUND

Pharmacoepidemiologic studies routinely employ cohort designs to estimate the effect of a drug treatment on an outcome using observational data. And while many variations on the cohort design exist, the active comparator new user (ACNU) design[1] has emerged over the past several decades as a best practice. This design, which is intended to emulate a head-to-head randomized trial, involves identifying cohorts of patients at the time of initiation of one of two drugs: a target drug and a comparator. Patients are then followed for the occurrence of the study outcome and the incidence is compared between groups, usually with adjustment for confounding factors. The primary advantage to this approach is the mitigation of biases known to impact comparisons against an inactive – i.e., unexposed or untreated – group. In particular, confounding by indication, healthy user bias, and healthy adherer bias may all be reduced by the adoption of an ACNU design.[1]

However, while the ACNU design has many advantages, it requires the study designer to select an appropriate comparator, and the process of identifying and determining comparator adequacy can be challenging. The ideal comparator group represents a counterfactual determined by the research question of interest. This question may concern whether the incidence of an outcome in the presence of the target exposure is different than what would have occurred if some other exposure were substituted for the target, or alternatively whether that incidence differs from what would have occurred if the target exposure were simply removed. In the former case, which is termed comparative effectiveness or safety research, investigators have historically selected comparators on the basis of perceived clinical equipoise and empirical feasibility. In the latter case, which is common in the drug safety surveillance context, the use of an active comparator otherwise thought to be wholly unrelated to the study outcome has been recommended[2–5], and a theoretical framework for its choice developed[6]. While the counterfactual represented by the comparator thus varies according to the

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

research context, a common goal is to identify a comparator with a similar patient composition in terms of demographics and relevant medical characteristics to the target. And while several sets of authoritative recommendations[7] have emerged to guide investigators in the comparator selection process, it remains at least partially subjective and difficult to scale in a context with many target drugs. With thousands of drugs available on the market, each of which may be a viable comparator for another, the selection of comparators for all drugs is thus a large task, encompassing nearly one million target-comparator possibilities.

In this paper, we present an empirical approach for ranking potential comparator cohorts for pharmacoepidemiologic studies based on their similarity in high-dimensional covariate space. We propose a method to compute a scalar similarity score, the cohort similarity score, for all pairwise combinations of all active drug ingredients and rank candidate comparators for any given target drug based on this composite score. The cohort similarity score is based on the average of cosine similarity scores[8] computed within five separate covariate domains intended to emulate clinical thinking surrounding a treatment decision. These domains are demographics, medical history, prior medications, presentation, and visit context. We evaluate rankings based on the cohort similarity score in the context of existing drug classification schemes and assess whether they produce pairs of cohorts comprised of similar patient populations, based on having high proportions of covariates already in balance.

Throughout, we focus on six example drugs from three therapeutic areas: anticoagulation (warfarin and apixaban), rheumatoid arthritis (methotrexate and adalimumab), and multiple sclerosis (glatiramer and ocrelizumab). We demonstrate that this method produces ranked lists of comparators which broadly align with subject-matter knowledge, and which are already posited as good comparators in the literature or registered studies. Further, we propose that this approach may be of use both as a tool for researchers designing pharmacoepidemiologic studies and as a part of an eventual analytic pipeline to automate the evidence generation process. We provide the full suite of results, including

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

some from data sources not originally included in this analysis, in an interactive web application (data.ohdsi.org/ComparatorSelectionExplorer), which can be used to identify and diagnose potential comparators for a given target cohort, and the code necessary for others to implement our approach in their own data (https://github.com/OHDSI/ComparatorSelectionExplorer).

METHODS

Data sources

We utilized data from five sources of insurance claims: Merative™ MarketScan® Commercial Claims and Encounters (CCAE), Merative™ MarketScan® Multi-State Medicaid (MDCD), Merative™ MarketScan® Medicare Supplemental (MDCR), Japan Medical Data Center (JMDC), and Optum© De-Identified Clinformatics© Data Mart. All data were mapped to the Observational Medical Outcomes Partnership Common Data Model[9] (OMOP CDM). Episodes of continuous exposure to individual RxNorm drug ingredients were inferred from pharmacy claims[10]. ICD diagnosis codes were mapped to the Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT).

Cohort definitions

Prior to computing similarity scores, it was necessary to create cohorts for each active drug ingredient to be evaluated as either a target or comparator. As such, in each data source, we defined cohorts of new users for each RxNorm ingredient. While there are over 10,000 ingredient concepts in the RxNorm vocabulary, only a minority are used in medical practice and the number varies according to local usage patterns in each data source. To enter a given ingredient-specific cohort, patients were required to have 365 days of continuous observation in the database prior to the first observed exposure to that ingredient, which served as their index date. Patients were allowed to enter a given cohort only once, but there was no restriction on the number of distinct cohorts they could enter. In

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

order to ensure robust estimation of covariate prevalence, we restricted analysis to cohorts with at least 1,000 patients.

Extraction of covariate data

Within each cohort, covariate data was extracted in five domains intended to capture the clinical thinking surrounding a treatment decision. These domains were:

- **Demographics**: one covariate for female sex, one for male sex, and one covariate for each decade of age
- **Medical history**: one covariate for each SNOMED condition observed in the period from the start of patient observation to 31 days before the index date
- **Presentation**: one covariate for each SNOMED condition observed in the period from 30 days before the index date to the index date
- **Prior medications:** one covariate for each RxNorm ingredient with exposure start at least 30 days prior to the index date and exposure end on or before the index date
- **Visit context**: one covariate for the presence of an inpatient visit in the 30 days prior to the index date, and one for the presence of an emergency department visit in the same period

Information on the prevalence of each covariate in each cohort was stored in a separate table in the database from which it was extracted and used for the computation of similarity metrics. All covariates with prevalence < 1% were excluded to avoid distortion of similarity scores by extremely low prevalence estimates.

Computation of similarity metrics

Within each database, we formed all possible pairs of cohorts and computed similarity scores for each pair. We calculated the cosine similarity[8], the dot product of two vectors containing the

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

target and comparator cohorts' covariate prevalences divided by the product of their lengths, between the two cohorts, separately for each covariate domain, and then averaged across the five domains to get the final similarity score, termed the cohort similarity score. Informally, a cosine similarity score of 1 will occur if all covariates are the same between the two cohorts, while a score of 0 indicates that the cohorts have minimal agreement across the covariates. We then used the cohort similarity scores to generate ranked lists of comparators for each cohort, with scores closer to 1 indicating stronger comparator alignment. Additional information on cosine similarity can be found in the **Appendix.**

Under the assumption that cohorts with known relationships should have higher similarity than those without such relationships, we used the Anatomical Therapeutic Chemical (ATC) classification system[11] to compare the distribution of cohort similarity scores for cohort pairs related at different levels of the hierarchy. Cohort pairs were divided according to the most granular ATC level at which they co-occurred, if any. Pairs that co-occurred in a class at ATC level 5 were excluded from this analysis as this level is commonly used for individual drug products or combinations.

For a random subset of 100 cohorts within each data source, we also attempted to validate our ranked lists with a more familiar pharmacoepidemiological metric of covariate similarity, the standardized mean difference (SMD)[12]. For each of the randomly selected drugs, we identified their top comparator based on the cohort similarity score, calculated SMDs for all covariates present within the cohort pair and reported the proportion of covariates in various degrees of imbalance based on recommended cutoffs for the absolute SMD at 0.1, 0.2, 0.5, and 0.8 [13,14]. We then repeated this process for the comparator ranked 500th.

All code for cohort generation, covariate extraction, similarity calculation, and for our web application is publicly available at https://github.com/OHDSI/ComparatorSelectionExplorer.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

RESULTS

Between 1,729 and 2,409 RxNorm ingredient-level new-user cohorts were generated per data source. Of these, between 988 and 1,359 cohorts had at least 1,000 patients (**Table 1**). The mean number of covariates with prevalence ≥1% extracted per cohort was 934.5 - 1,479.9 overall, 65.4- 120.0 in the "Presentation" domain, 336.9- 1,016.4 in the "Medical History" domain, and 293.0 - 486.2 in the "Prior Medications" domain. The number of demographic and visit context covariates was roughly constant by design.

Between 487,578 and 922,761 pairs of cohorts with at least 1,000 patients were formed per data source. Among these, similarity was lowest in the "Presentation" domain (mean cosine similarity 0.213 to 0.490, depending on data source) and highest in the "Visit context" domain (mean cosine similarity 0.869 to 1.000, depending on data source). Between 90% and 95% of target cohorts with at least 1,000 patients had a comparator with an average cohort similarity score of ≥ 0.95. Among the six example drugs, the distribution of cohort similarity scores was consistently left-skewed, with most comparators lying between 0.7 and 0.9 (**Figure 1**). Similarity was lowest among pairs of cohorts that did not occur together in any ATC classes and tended to increase among pairs of cohorts that co-occurred within higher-level ATC classes (**Figure 2**).

The top five comparators for each of the six example drugs, by average rank across data sources (**Table 2**) and within data sources (**Table 3**), consistently corresponded to alternative treatments for each drug's indication(s). For example, the top five comparators for glatiramer (interferon beta-1a, interferon beta-1b, dimethyl fumarate, fingolimod, and natalizumab) are all other disease-modifying therapies for multiple sclerosis[15], and those ranked sixth through tenth (teriflunomide, dalfampridine, ocrelizumab, corticotropin, and modafinil) are either disease modifying therapies or are used to treat symptoms of MS or MS relapse. Ranked comparator lists for these example drugs varied according to

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

data source (**Table 3, Figure 3 panel A**), with the highest degree of correlation between CCAE, MDCD, and Optum. An example of between-data source heterogeneity is shown in **Figure 3 panel B** where, among 627 candidate comparators for warfarin observed in both CCAE and JMDC, rivaroxaban is identified as highly similar to warfarin in both data sources, whereas xylitol and vitamin A appear much more similar to warfarin in JMDC than in CCAE, while the reverse is true for hydralazine.

Among the 100 target drugs randomly selected within each data source and their top comparators, most covariates were well balanced. When the comparator ranked 500th was chosen instead, the proportion of imbalanced covariates was higher (**Figure 4**). The average proportion of covariates with |SMD| ≤ 0.1 was between 72% and 84% for the top comparator versus between 28% and 45% for the comparator ranked 500th.

DISCUSSION

Over the past several decades, the ACNU cohort design has emerged as a best practice for the estimation of drug effects from observational data. However, the implementation of this design may be mired by the difficulty of selecting an appropriate comparator for the target drug of interest, either because there are many potential alternatives to the target drug, or because the true target of inference is in fact a contrast between the target drug and its absence. The development of empirical comparator recommendations can aid investigator decision-making, simplifying the process of comparator selection and ultimately improving the quality of pharmacoepidemiological research. Furthermore, these recommendations may ultimately enable the automation of evidence generation using the ACNU design , something that previously relied on self-controlled designs[16,17] and temporal scans[18,19],.

In this paper, we demonstrated a method for ranking potential comparators based on similarity in high-dimensional covariate space. We showed that most target cohorts of reasonable size possessed at least one potential comparator with a high degree of similarity and that cohorts that were more

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

related in the ATC hierarchy also had higher cohort similarity scores. For a random sample of 100 cohorts within each data source, we also showed that selecting the top-most similar comparator would lead to balance on the majority of covariates before any adjustment took place.

We additionally demonstrated that, for six target cohorts from a diverse set of therapeutic areas, the top-ranked comparators aligned well with subject-matter knowledge. That is, all were used in the treatment of at least one of the target cohorts' indications. Indeed, many of these top-ranked comparators are found to be in use in head-to-head safety or effectiveness comparisons with the target cohort in the literature or in public registries of clinical and observational studies. For example, head-to-head studies exist comparing apixaban and its top comparator rivaroxaban[20,21], adalimumab and its top comparator certolizumab[22,23], and ocrelizumab and its top comparator teriflunomide[24]. While this is reassuring, it should also be noted that the correlation in comparator rankings across data sources was not consistently high (**Figure 3**). This may indicate that a given comparator is not uniquely fit-for-purpose across all data sources, suggesting that investigators take care to select a comparator using information from a source that at least resembles their intended analytic context.

This study complements a large body of work on drug similarity from within the bioinformatics community, including attempts to define similarity based on chemical structure, gene expression[25], genome-wide associations[26], co-occurrence of drugs with condition terms in the literature[27], individual patient data[28], or a combination of these[28]. However, our work is unique in its focus on the use of similarity indices for comparator selection rather than for identification of novel associations, and for its clinically oriented approach to feature extraction.

The empirical comparator recommendations provided here have several advantages. First, they are agnostic to investigator opinion and bias, both in terms of which drugs may serve as candidate comparators for a given target drug and which covariates should be used to determine empirical

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

feasibility. While these decisions should be guided by subject-matter knowledge[29], no investigator is omniscient and thus empirical recommendations may lessen investigator degrees of freedom and potentially enable less biased comparisons. Additionally, empirical recommendations may reveal previously unknown comparators that are not necessarily alternative treatments for any of the target drug's indications. They thus hold promise for the identification of appropriate negative control exposures, ultimately allowing for cohort studies that emulate user versus non-user comparisons within the less-biased ACNU framework. Additionally, these recommendations have been assessed across multiple large data sources comprised of diverse populations within and outside the United States. They have been observed, consistently across these data sources, to label drugs that are more closely related in the ATC hierarchy as more similar, to produce balance on a majority of covariates when the most similar comparator is chosen, and to identify top comparators that align with subject-matter knowledge.

A limitation of the approach introduced here is that the cohort similarity score is based on the *marginal* distributions of covariates in the two cohorts and as such is only a proxy for full covariate similarity. In contrast, propensity scores methods have the advantage of providing metrics (e.g., the c-statistic or the preference score[30]) that reflect the similarity in *joint* covariate distributions across cohorts, and such methods can furthermore be used to adjust for observed dissimilarity. However, unlike many propensity score models, which only incorporate a small number of covariates, the cohort similarity score dynamically defines a large set of covariates based on all available data elements. It should be noted that this data-adaptive approach has been incorporated into some propensity score methods like the high-dimensional propensity score (hdPS) [31] and, more recently, the large-scale propensity score (LSPS)[32]. However, the cohort similarity score remains the only approach that is scalable to millions of candidate target-comparator combinations, whereas PS models—particularly those using high-dimensional covariate data—can be computationally expensive, making them infeasible for screening large numbers of combinations. The cohort similarity score may therefore be

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

best suited as a tool to narrow the comparator search space before techniques like hdPS/LSPS are employed. Finally, the approach presented here provides a standardized framework for comparator ranking which can be readily implemented by the public. This standardization improves reproducibility and scalability and ultimately lowers the burden on domain expertise for the investigator.

There are several factors that may limit the usefulness of the comparator recommendations provided here. Because this analysis was limited to cohorts of new users of individual RxNorm ingredients, comparator recommendations for ingredients that are used differently depending on dose, form, or route of administration may be less coherent than recommendations for ingredients which are used more consistently regardless of these factors. The same limitation may apply to ingredients that are used in combination preparations or which possess multiple, possibly unrelated, indications. However, because the method presented here is extensible to cohorts defined in an arbitrary manner, these limitations do not preclude the future development of recommendations for such products. Additionally the assessment of these recommendations was limited by the lack of an objective metric for defining comparator "appropriateness". While we sought to benchmark appropriateness based on comparator location in the ATC hierarchy and appearance in published literature or registered studies, future work might attempt to go further by demonstrating that the use of empirical comparator recommendations can lead to less biased treatment effect estimates than traditional methods of comparator selection.

Finally, it should be clearly stated that empirical comparator recommendations are not a replacement for critical thinking in study design and analysis. Investigators may wish to use them as a "first pass" to identify and evaluate candidate comparators, after which more careful consideration and detailed analysis can determine true feasibility. Crucially, covariate adjustment is almost invariably required regardless of the quality of comparator chosen. Similarity between a target drug and comparator is no guarantee that any remaining covariate imbalance can be successfully adjusted via

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

established methods like restriction or propensity score adjustment. And, conversely, low similarity relative to other recommendations is not a guarantee that successful adjustment is impossible.

Our work leaves open many exciting opportunities for further exploration of empirical comparator recommendations. Among these are additional approaches to validation, including a comparison against more computationally expensive techniques like assessing propensity score overlap. Future work might also explore the utility of adding sources of metadata on known or suspected drug-outcome relationships, such as the OHDSI Common Evidence Model[33,34], in order to improve the ease of identifying negative control exposures.

CONCLUSION

We have developed and evaluated a method for recommending comparators that yields promising results and which may be of use to investigators wishing to conduct ACNU studies of drug safety and effectiveness. An open-source tool set is available for those wishing to identify candidate comparators using their own data, and the full suite of results, including the comparator rankings for up to 1,359 drugs across the five data sources included here, is available to the public. We hope these tools and the results presented here will accelerate and improve the quality of future pharmacoepidemiologic research.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

REFERENCES

1 Lund JL, Richardson DB, Stürmer T. The Active Comparator, New User Study Design in Pharmacoepidemiology: Historical Foundations and Contemporary Application. *Curr Epidemiol Rep* 2015;**2**:221–8. doi:10.1007/s40471-015-0053-5

2 Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies. *Epidemiology* 2010;**21**:383–8. doi:10.1097/EDE.0b013e3181d61eeb

3 Dusetzina SB, Brookhart MA, Maciejewski ML. Control Outcomes and Exposures for Improving Internal Validity of Nonrandomized Studies. *Health Serv Res* 2015;**50**:1432–51. doi:10.1111/1475-6773.12279

4 Setoguchi S, Glynn RJ, Avorn J, *et al.* Statins and the Risk of Lung, Breast, and Colorectal Cancer in the Elderly. *Circulation* 2007;**115**:27–33. doi:10.1161/CIRCULATIONAHA.106.650176

5 Solomon DH, Avorn J, Stürmer T, *et al.* Cardiovascular outcomes in new users of coxibs and nonsteroidal antiinflammatory drugs: High-risk subgroups and time course of risk. *Arthritis Rheum* 2006;**54**:1378–89. doi:10.1002/art.21887

6 Huitfeldt A, Hernan MA, Kalager M, *et al.* Comparative Effectiveness Research Using Observational Data: Active Comparators to Emulate Target Trials with Inactive Comparators. *EGEMs Gener Evid Methods Improve Patient Outcomes* 2016;**4**:20. doi:10.13063/2327-9214.1234

7 Setoguchi S, Gerhard T. Chapter 5: Comparator Selection. In: Velentgas P, Dreyer NA, Nourjah, P, *et al.*, eds. *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide*. Rockville (MD): : Agency for Healthcare Research and Quality (US) 2013. https://www.ncbi.nlm.nih.gov/books/NBK126184/

8 Tan P-N, Steinbach M, Kumar V. Chapter 8. In: *Introduction to data mining*. Boston: : Pearson Addison Wesley 2006. 500.

9 Overhage JM, Ryan PB, Reich CG, *et al.* Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc JAMIA* 2012;**19**:54–60. doi:10.1136/amiajnl-2011-000376

10 OMOP CDM v5.4. http://ohdsi.github.io/CommonDataModel/cdm54.html#Standardized_Derived_Elements (accessed 28 Nov 2022).

11 WHOCC - Structure and principles. https://www.whocc.no/atc/structure_and_principles/ (accessed 24 Nov 2022).

12 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;**28**:3083–107. doi:10.1002/sim.3697

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

13 Normand S-LT, Landrum MB, Guadagnoli E, *et al.* Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores. *J Clin Epidemiol* 2001;:12.

14 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, N.J: : L. Erlbaum Associates 1988.

15 Hauser SL, Cree BAC. Treatment of Multiple Sclerosis: A Review. *Am J Med* 2020;**133**:1380-1390.e2. doi:10.1016/j.amjmed.2020.05.049

16 Kulldorff M, Dashevsky I, Avery TR, *et al.* Drug safety data mining with a tree-based scan statistic: TREE-BASED SCAN STATISTIC DATA MINING METHOD. *Pharmacoepidemiol Drug Saf* 2013;**22**:517–23. doi:10.1002/pds.3423

17 Wang SV, Maro JC, Baro E, *et al.* Data Mining for Adverse Drug Events With a Propensity Score-matched Tree-based Scan Statistic. *Epidemiology* 2018;**29**:895–903. doi:10.1097/EDE.0000000000000907

18 Norén GN, Hopstadius J, Bate A, *et al.* Temporal pattern discovery in longitudinal electronic patient records. *Data Min Knowl Discov* 2010;**20**:361–87. doi:10.1007/s10618-009-0152-3

19 Lavallee M, Yu T, Evans L, *et al.* Evaluating the performance of temporal pattern discovery: new application using statins and rhabdomyolysis in OMOP databases. *BMC Med Inform Decis Mak* 2022;**22**:31. doi:10.1186/s12911-022-01765-1

20 Boehringer Ingelheim. Safety and Effectiveness Study Comparing Dabigatran, Rivaroxaban & Apixaban in Non-valvular Atrial Fibrillation Patients Enrolled in the US Department of Defense Military Health System. clinicaltrials.gov 2019. https://clinicaltrials.gov/ct2/show/NCT03026556 (accessed 21 Nov 2022).

21 Lip GYH, Keshishian A, Kamble S, *et al.* Real-world comparison of major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran, rivaroxaban, or warfarin: A propensity score matched analysis. *Thromb Haemost* 2016;**116**:975–86. doi:10.1160/TH16-05-0403

22 Smolen JS, Burmester G-R, Combe B, *et al.* Head-to-head comparison of certolizumab pegol versus adalimumab in rheumatoid arthritis: 2-year efficacy and safety results from the randomised EXXELERATE study. *The Lancet* 2016;**388**:2763–74. doi:10.1016/S0140-6736(16)31651-8

23 UCB Pharma SA. A Multicenter, Single-blind, Randomized Parallel-group Study to Assess the Short- and Long-term Efficacy of Certolizumab Pegol Plus Methotrexate Compared to Adalimumab Plus Methotrexate in Subjects With Moderate to Severe Rheumatoid Arthritis Responding Inadequately to Methotrexate. clinicaltrials.gov 2018. https://clinicaltrials.gov/ct2/show/NCT01500278 (accessed 21 Nov 2022).

24 Dirks P, Zingler V, Leemhuis J, *et al.* Design of a non-interventional post-marketing study to assess the long-term safety and effectiveness of ocrelizumab in German real world multiple sclerosis cohorts – the CONFIDENCE study protocol. *BMC Neurol* 2020;**20**:95. doi:10.1186/s12883-020-01667-7

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

25 Brown AS, Kong SW, Kohane IS, *et al.* ksRepo: a generalized platform for computational drug repositioning. *BMC Bioinformatics* 2016;**17**:78. doi:10.1186/s12859-016-0931-y

26 Grover MP, Ballouz S, Mohanasundaram KA, *et al.* Identification of novel therapeutics for complex diseases from genome-wide association data. *BMC Med Genomics* 2014;**7**:S8. doi:10.1186/1755-8794-7-S1-S8

27 Brown AS, Patel CJ. MeSHDD: Literature-based drug-drug similarity for drug repositioning. *J Am Med Inform Assoc* 2017;**24**:614–8. doi:10.1093/jamia/ocw142

28 Zeng X, Jia Z, He Z, *et al.* Measure clinical drug–drug similarity using Electronic Medical Records. *Int J Med Inf* 2019;**124**:97–103. doi:10.1016/j.ijmedinf.2019.02.003

29 Greenland S, Neutra R. Control of confounding in the assessment of medical technology. *Int J Epidemiol* 1980;**9**:361–7. doi:10.1093/ije/9.4.361

30 Walker A, Patrick, Lauer, *et al.* A tool for assessing the feasibility of comparative effectiveness research. *Comp Eff Res* 2013;:11. doi:10.2147/CER.S40357

31 Schneeweiss S, Rassen JA, Glynn RJ, *et al.* High-dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data. *Epidemiology* 2009;**20**:512–22. doi:10.1097/EDE.0b013e3181a663cc

32 Zhang L, Wang Y, Schuemie MJ, *et al.* Adjusting for indirectly measured confounding using large-scale propensity score. *J Biomed Inform* 2022;**134**:104204. doi:10.1016/j.jbi.2022.104204

33 Voss EA, Boyce RD, Ryan PB, *et al.* Accuracy of an automated knowledge base for identifying drug adverse reactions. *J Biomed Inform* 2017;**66**:72–81. doi:10.1016/j.jbi.2016.12.005

34 Home · OHDSI/CommonEvidenceModel Wiki. GitHub. https://github.com/OHDSI/CommonEvidenceModel (accessed 25 Nov 2022).

TABLES

Table 1: Summary of cohorts generated, and covariates extracted, by data source and covariate domain

	CCAE	MDCD	MDCR	JMDC	Optum
Total number of cohorts generated					
Overall	2,243	2,158	2,065	1,729	2,273
With at least 1,000 patients	1,359	1,112	988	1,082	1,332
Mean (SD) number of covariates with prevalence ≥ 1% extracted per cohort with ≥ 1,000 patients					
All	1,030.6 (324.65)	1,479.9 (432.59)	1,287.7 (279.28)	934.5 (186.76)	1,422.0 (441.61)
Demographics	7.8 (1.12)	9.6 (1.52)	6.2 (0.57)	8.9 (1.23)	9.3 (1.64)
Medical history	662.3 (234.73)	1,016.4 (319.43)	852.1 (207.44)	336.9 (74.25)	957.3 (309.51)
Presentation	65.4 (49.17)	120.0 (63.71)	88.3 (47.32)	101.5 (36.33)	119.2 (85.42)
Prior medications	293.0 (58.56)	331.9 (67.89)	339.1 (39.66)	486.2 (83.36)	334.3 (64.94)
Visit context	1.8 (0.36)	1.9 (0.25)	1.9 (0.08)	1.0 (0.00)	1.8 (0.34)
Number of pairs of cohorts with ≥ 1,000 patients	922,761	617,716	487,578	584,821	886,446
Mean (SD) similarity score among pairs of cohorts with ≥ 1,000 patients					
Cohort similarity score	0.721 (0.099)	0.774 (0.106)	0.847 (0.060)	0.793 (0.090)	0.743 (0.120)
Demographics	0.852 (0.139)	0.846 (0.128)	0.945 (0.075)	0.857 (0.134)	0.841 (0.147)
Medical history	0.781 (0.137)	0.798 (0.137)	0.913 (0.059)	0.879 (0.092)	0.798 (0.141)
Presentation	0.213 (0.173)	0.401 (0.216)	0.490 (0.190)	0.377 (0.189)	0.349 (0.212)
Prior medications	0.882 (0.085)	0.872 (0.090)	0.931 (0.059)	0.917 (0.071)	0.859 (0.108)
Visit context	0.875 (0.150)	0.955 (0.066)	0.957 (0.069)	1.000 (0.000)	0.869 (0.160)
Number (%) of target cohorts with at least one comparator with cohort similarity score of:					
At least 0.900	1,294 (95.2%)	1,094 (98.4%)	981 (99.3%)	1,067 (98.6%)	1,294 (97.1%)
At least 0.950	1,087 (80.0%)	961 (86.4%)	924 (93.5%)	1,024 (94.6%)	1,118 (83.9%)

Table 2: Top five comparators for six example drugs, according to average rank across data sources. Restricted to comparators occurring in at least two data sources. Entries are given as comparator name (average rank across data sources, mean cohort similarity score, number of data sources with information).

Rank	warfarin	apixaban	methotrexate	adalimumab	glatiramer	ocrelizumab
1	digoxin (4.6, 0.960, 5)	rivaroxaban (1.0, 0.977, 5)	sulfasalazine (1.2, 0.978, 5)	certolizumab pegol (2.2, 0.974, 4)	interferon beta-1a (1.3, 0.993, 3)	teriflunomide (1.0, 0.980, 2)
2	rivaroxaban (6.2, 0.968, 5)	diltiazem (7.0, 0.955, 5)	leflunomide (2.5, 0.970, 4)	sulfasalazine (3.4, 0.960, 5)	interferon beta-1b (1.5, 0.993, 2)	dimethyl fumarate (2.3, 0.969, 3)
3	diltiazem (6.4, 0.959, 5)	amiodarone (9.6, 0.953, 5)	etanercept (3.4, 0.966, 5)	infliximab (4.0, 0.958, 5)	dimethyl fumarate (3.3, 0.991, 3)	natalizumab (3.0, 0.967, 3)
4	carvedilol (9.0, 0.952, 5)	flecainide (14.2, 0.952, 5)	adalimumab (4.6, 0.953, 5)	methotrexate (4.2, 0.953, 5)	fingolimod (3.3, 0.989, 3)	fingolimod (3.7, 0.965, 3)
5	amiodarone (10.2, 0.953, 5)	torsemide (14.4, 0.947, 5)	certolizumab pegol (6.5, 0.953, 4)	leflunomide (6.0, 0.945, 4)	natalizumab (4.3, 0.988, 3)	dalfampridine (4.0, 0.967, 2)

Table 3: Top five comparators for six example drugs according to cohort similarity score, by data source. Entries are given as comparator name (cohort similarity score). Note that glatiramer and ocrelizumab did not meet the 1,000 patient sample size cutoff in MDCR or JMDC and thus results are not present for these data sources.

Example Drug	Data Source	Comparator #1	Comparator #2	Comparator #3	Comparator #4	Comparator #5
warfarin	CCAE	fondaparinux (0.970)	rivaroxaban (0.959)	dalteparin (0.957)	enoxaparin (0.932)	digoxin (0.915)
warfarin	MDCD	diltiazem (0.973)	carvedilol (0.970)	dabigatran etexilate (0.970)	rivaroxaban (0.968)	captopril (0.968)
warfarin	MDCR	diltiazem (0.975)	digoxin (0.972)	rivaroxaban (0.971)	amiodarone (0.971)	sotalol (0.970)
warfarin	Optum	rivaroxaban (0.970)	dabigatran etexilate (0.967)	captopril (0.967)	diltiazem (0.966)	digoxin (0.965)
warfarin	JMDC	protamines (0.988)	dexmedetomidine (0.986)	norepinephrine (0.985)	dobutamine (0.983)	digoxin (0.982)
apixaban	CCAE	rivaroxaban (0.956)	protamine sulfate (USP) (0.951)	digoxin antibodies Fab fragments (0.940)	diltiazem (0.940)	hydralazine (0.935)
apixaban	MDCD	rivaroxaban (0.985)	torsemide (0.967)	diltiazem (0.967)	bumetanide (0.966)	carvedilol (0.964)
apixaban	MDCR	rivaroxaban (0.974)	diltiazem (0.952)	hydralazine (0.952)	enoxaparin (0.952)	protamine sulfate (USP) (0.951)
apixaban	Optum	rivaroxaban (0.973)	digoxin antibodies Fab fragments (0.972)	amiodarone (0.971)	diltiazem (0.968)	torsemide (0.966)
apixaban	JMDC	rivaroxaban (0.998)	dabigatran etexilate (0.990)	digoxin (0.988)	edoxaban (0.985)	flecainide (0.985)
methotrexate	CCAE	sulfasalazine (0.972)	leflunomide (0.970)	etanercept (0.969)	hydroxychloroquine (0.956)	adalimumab (0.955)
methotrexate	MDCD	sulfasalazine (0.961)	etanercept (0.960)	hydroxychloroquine (0.955)	adalimumab (0.954)	leflunomide (0.950)
methotrexate	MDCR	leflunomide (0.984)	sulfasalazine (0.984)	hydroxychloroquine (0.981)	adalimumab (0.980)	etanercept (0.979)
methotrexate	Optum	sulfasalazine (0.983)	leflunomide (0.976)	hydroxychloroquine (0.974)	etanercept (0.958)	adalimumab (0.955)

Example Drug	Data Source	Comparator #1	Comparator #2	Comparator #3	Comparator #4	Comparator #5
methotrexate	JMDC	sulfasalazine (0.988)	tocilizumab (0.968)	etanercept (0.966)	golimumab (0.962)	adalimumab (0.922)
adalimumab	CCAE	certolizumab pegol (0.973)	etanercept (0.962)	sulfasalazine (0.957)	methotrexate (0.955)	infliximab (0.948)
adalimumab	MDCD	sulfasalazine (0.970)	certolizumab pegol (0.970)	etanercept (0.960)	methotrexate (0.954)	infliximab (0.942)
adalimumab	MDCR	etanercept (0.990)	leflunomide (0.983)	infliximab (0.982)	methotrexate (0.980)	certolizumab pegol (0.978)
adalimumab	Optum	certolizumab pegol (0.973)	etanercept (0.960)	sulfasalazine (0.959)	methotrexate (0.955)	golimumab (0.948)
adalimumab	JMDC	infliximab (0.982)	azathioprine (0.963)	golimumab (0.956)	sulfasalazine (0.939)	methotrexate (0.922)
glatiramer	CCAE	interferon beta-1a (0.994)	interferon beta-1b (0.993)	fingolimod (0.992)	natalizumab (0.992)	dimethyl fumarate (0.991)
glatiramer	MDCD	interferon beta-1a (0.994)	dimethyl fumarate (0.992)	fingolimod (0.987)	natalizumab (0.986)	ocrelizumab (0.962)
glatiramer	Optum	interferon beta-1b (0.992)	interferon beta-1a (0.991)	dimethyl fumarate (0.989)	fingolimod (0.989)	natalizumab (0.987)
ocrelizumab	CCAE	teriflunomide (0.981)	dimethyl fumarate (0.974)	fingolimod (0.969)	natalizumab (0.968)	peginterferon beta-1a (0.965)
ocrelizumab	MDCD	natalizumab (0.977)	dimethyl fumarate (0.975)	fingolimod (0.974)	glatiramer (0.962)	interferon beta-1a (0.953)
ocrelizumab	Optum	teriflunomide (0.979)	dalfampridine (0.971)	dimethyl fumarate (0.960)	natalizumab (0.957)	fingolimod (0.952)

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

FIGURE LEGENDS

Figure 1: Distribution of average cosine similarity scores computed across all available candidate comparators for six example drugs, by data source. Note that glatiramer and ocrelizumab did not meet the 1,000 patient sample size cutoff in MDCR or JMDC and thus results are not present for these data sources. The number of available comparators ranged from 988-1,359, depending on data source.

Figure 2: Distribution of cohort similarity scores (panel A) and number of comparisons (panel B), by relatedness in ATC hierarchy and data source. All target-comparator pairs were stratified by the distance between the two drugs in the ATC hierarchy ("4", indicating that both drugs are in same low-level class, through "0", indicating that the drugs do not share any common drug class ancestry). Cohort pairs in which the target and comparator are both members of an ATC level 5 class are omitted as these classes are frequently used for individual drugs or combination products. Whiskers are truncated at 10th and 90th percentiles. White circles indicate mean values.

Figure 3: Pairwise rank correlations in cohort similarity scores between data sources for six example drugs (A) and cosine similarity scores among 627 candidate comparators for warfarin observd in both CCAE and JMDC (B).

Figure 4: Proportion of covariates at various levels of imbalance according to standardized mean difference (SMD) among 100 randomly-selected target cohorts and their comparators ranked 1st and 500th , by data source.

COMPETING INTERESTS

The authors declare no competing interests. All authors are employees or contractors of Janssen Pharmaceuticals and may hold stock or stock options in Johnson & Johnson

FUNDING

This work was funded by Janssen Research and Development, a subsidiary of Johnson & Johnson. All authors are employees of Janssen Research and Development. All authors are employees or contractors of Janssen Pharmaceuticals and may hold stock or stock options in Johnson & Johnson

AUTHOR CONTRIBUTIONS

Author	Affiliation	Email Address	ORCID	Contributions
Justin Bohn	Janssen Research & Development, Titusville, NJ	Jbohn2@its.jnj.com	0000-0003-4430-798X	Designed the analysis, performed the analysis, drafted the manuscript
James P. Gilbert	Janssen Research & Development, Titusville, NJ	Jgilber2@its.jnj.com		Designed the analysis, performed the analysis, reviewed and contributed to the manuscript
Christopher Knoll	Janssen Research & Development, Titusville, NJ	Cknoll1@its.jnj.com		Designed the analysis, reviewed and contributed to the manuscript
David M. Kern	Janssen Research & Development, Titusville, NJ	Dkern2@its.jnj.com	0000-0001-5417-3925	Designed the analysis, reviewed and contributed to the manuscript

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Author	Affiliation	Email Address	ORCID	Contributions
Patrick B. Ryan	Janssen Research & Development, Titusville, NJ	Pryan4@its.jnj.com		Designed the analysis, reviewed and contributed to the manuscript

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

APPENDIX

Definition of cosine similarity

Cosine similarity is defined as

$$\frac{\boldsymbol{A} \cdot \boldsymbol{B}}{||\boldsymbol{A}||\,||\boldsymbol{B}||} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

where A_i and B_i are the ith components of vectors $\boldsymbol{A}$ and $\boldsymbol{B}$. In the context of this paper, $\boldsymbol{A}$ and $\boldsymbol{B}$ represent vectors of covariate prevalence estimates from the target and comparator cohorts, respectively. Because these prevalence values are always non-negative, cosine similarities based on them will be bounded in the interval [0,1]. Two cohorts with identical vectors of covariate prevalence will have a cosine similarity of 1, as would two cohorts will exactly proportional vectors, whereas two cohorts with orthogonal vectors of covariate prevalence will have a cosine similarity of 0.

Worked example

Below is a demonstration of the computation of the demographic cosine similarity between an example target cohort (ocrelizumab) and one of it's top-ranked comparator cohorts (teriflunomide) using data from the Merative™ MarksetScan Commerical Claims and Encounters Database. These data can be found and the result of the computation verified at https://data.ohdsi.org/ComparatorSelectionExplorer[1].

Description	i	In ocrelizumab cohort A_i	In teriflunomide cohort B_i
Proportion aged 20-29	1	0.070	0.025
Proportion aged 30-39	2	0.208	0.127
Proportion aged 40-49	3	0.315	0.305
Proportion aged 50-59	4	0.294	0.385
Proportion aged 60-69	5	0.107	0.156
Proportion male	6	0.709	0.785
Proportion female	7	0.291	0.215

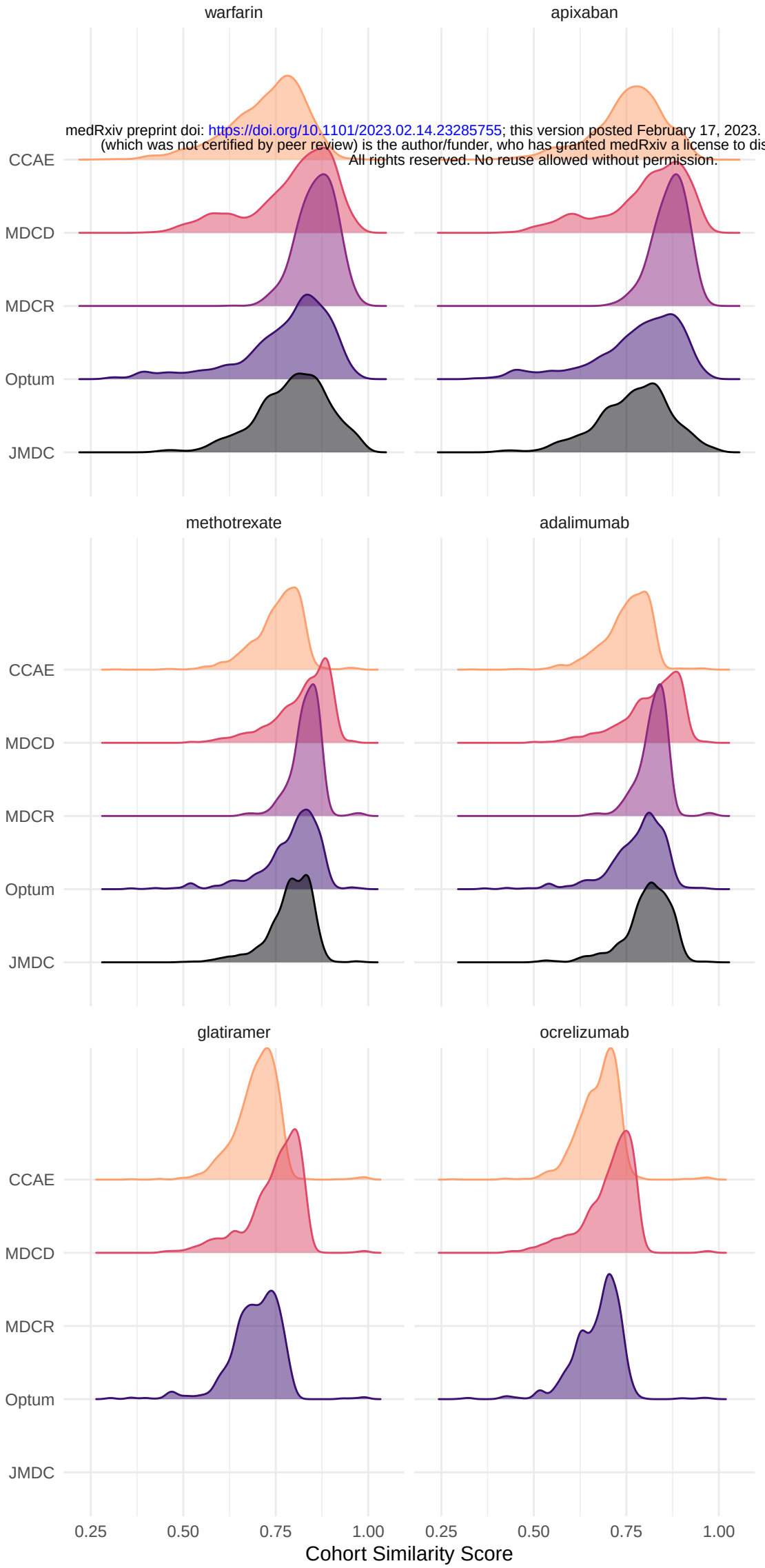
$$\text{cosine similarity} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} = \frac{0.873}{\sqrt{0.833}\sqrt{0.945}} = \frac{0.873}{0.887} = 0.985$$

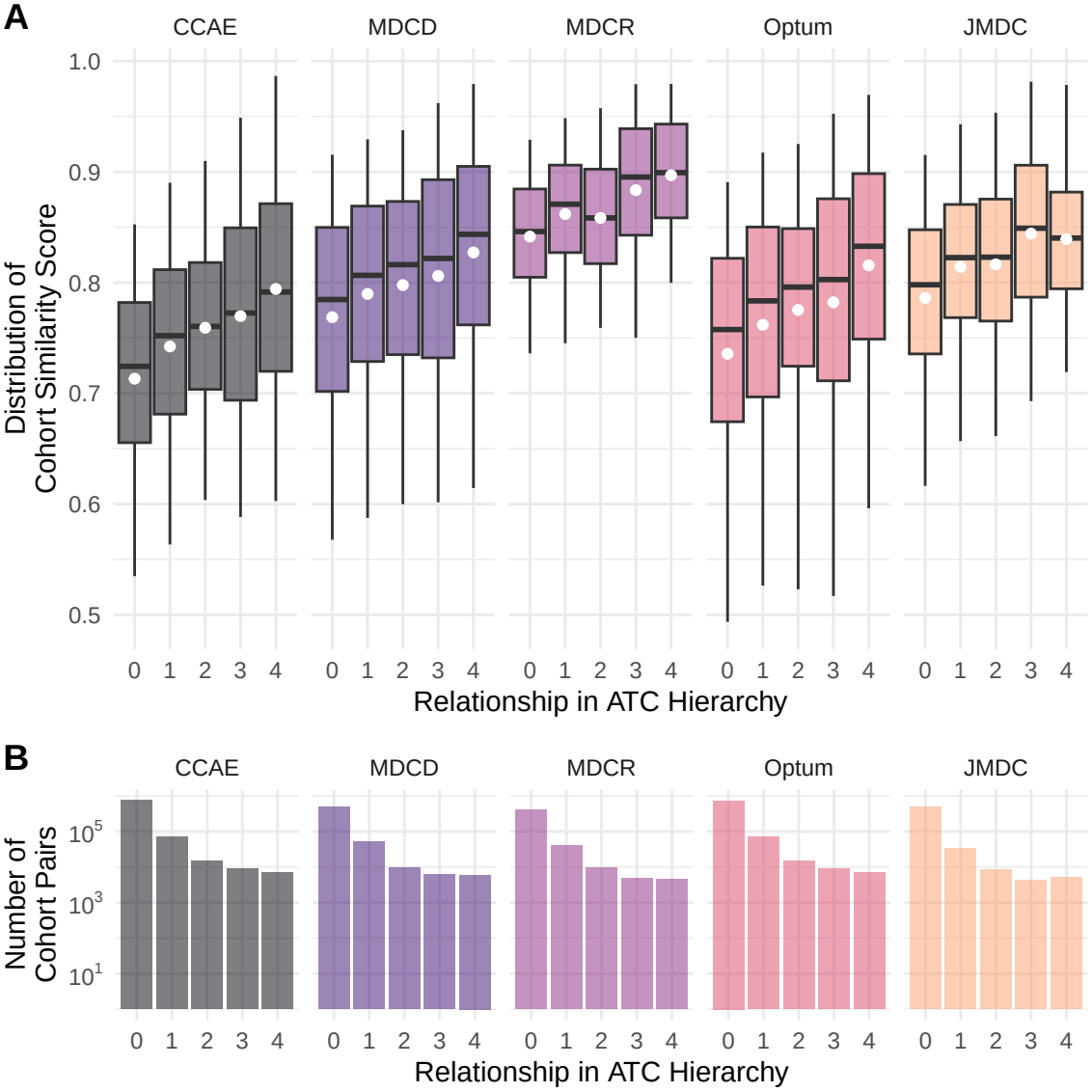
This process is repeated for the other four covariate domains (which are not demonstrated here) and the domain-specific cosine similarities are averaged to get the final similarity score as below.

$$\text{cohort similarity score} = \frac{1}{5}\sum_{j=1}^{5} \text{cosine similarity}_{\text{j}} = \frac{(0.985 + 0.989 + 0.953 + 0.980 + 1.0)}{5} = 0.981$$

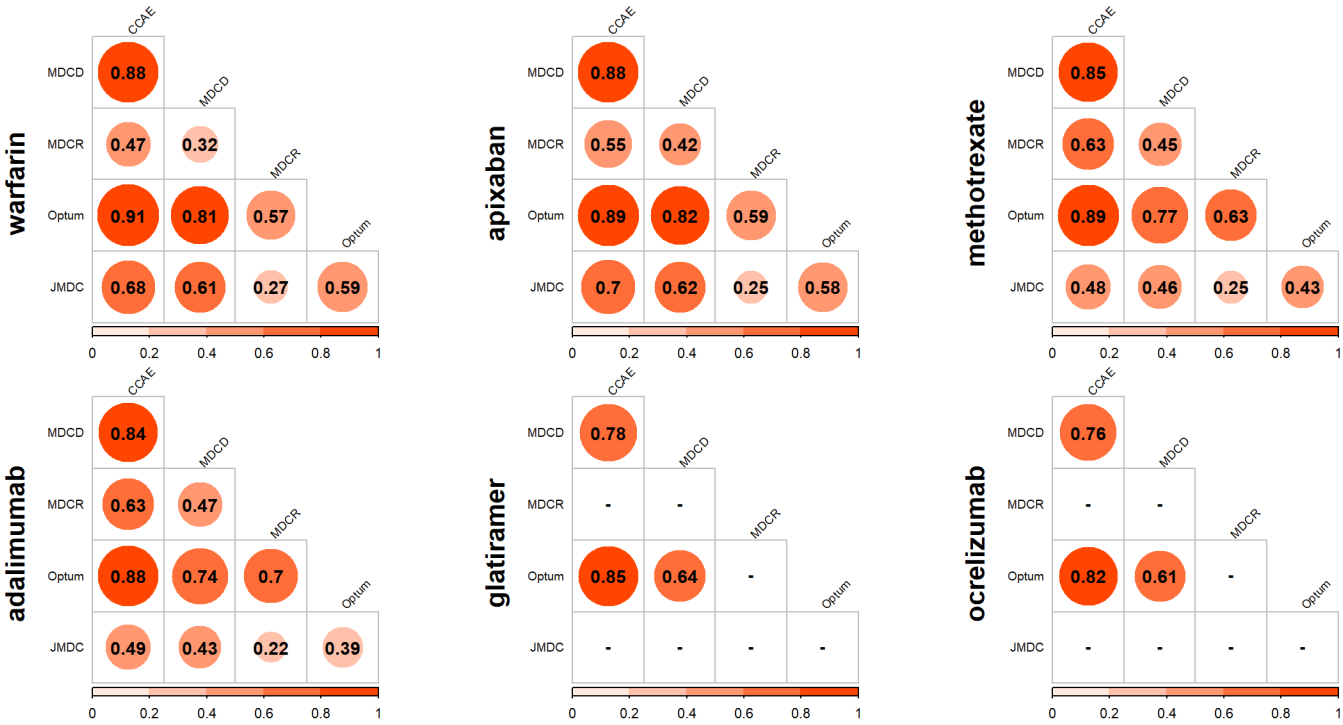
[1] Note that these numbers may change as data sources are updated

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285755; this version posted February 17, 2023.
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to dis
All rights reserved. No reuse allowed without permission.

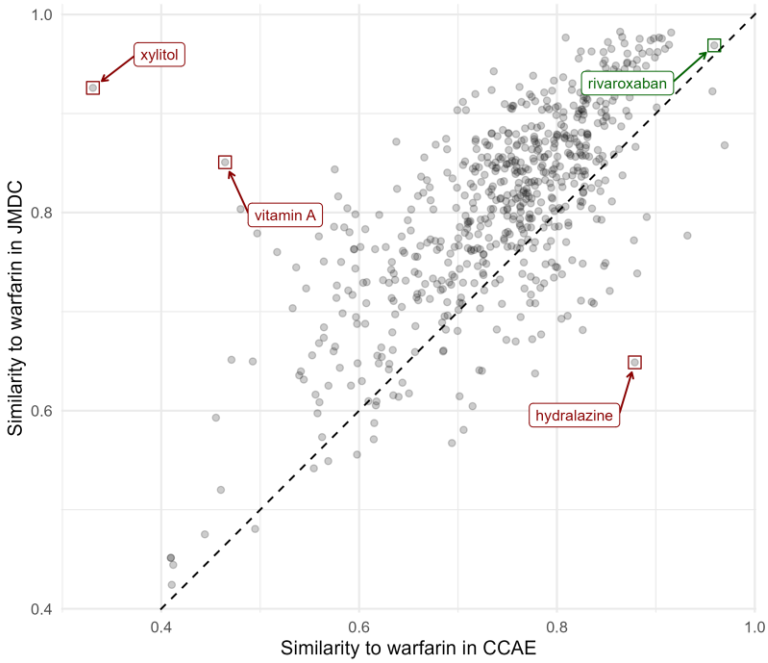


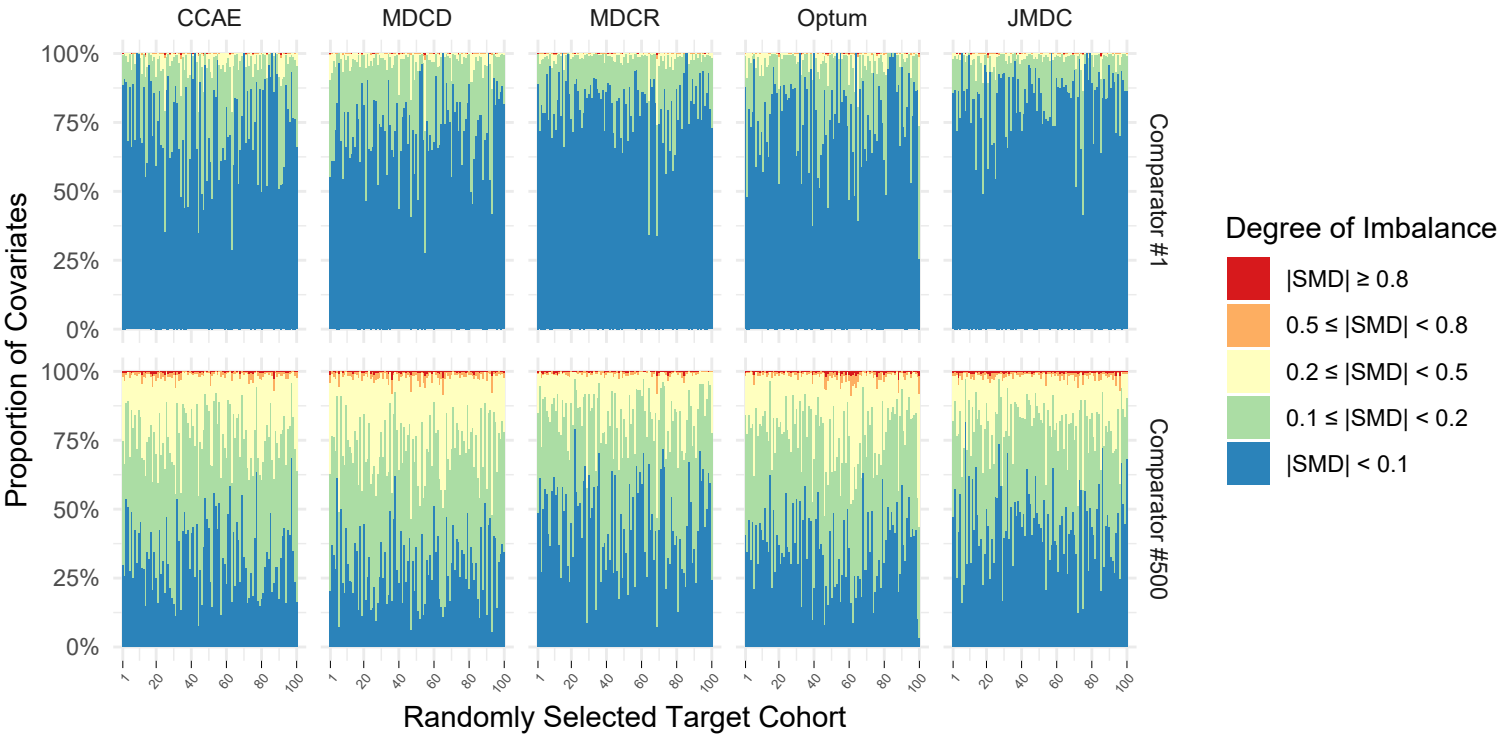

A
CCAE
MDCD
MDCR
Optum
JMDC
Distribution of Cohort Similarity Score
1.0
0.9
0.8
0.7
0.6
0.5
0 1 2 3 4
Relationship in ATC Hierarchy
B
Number of Cohort Pairs
10^5
10^3
10^1
Relationship in ATC Hierarchy

A


warfarin
CCAE
MDCD 0.88
MDCD
MDCR 0.47 0.32
MDCR
Optum 0.91 0.81 0.57
Optum
JMDC 0.68 0.61 0.27 0.59
0 0.2 0.4 0.6 0.8 1
apixaban
CCAE
MDCD 0.88
MDCD
MDCR 0.55 0.42
MDCR
Optum 0.89 0.82 0.59
Optum
JMDC 0.7 0.62 0.25 0.58
0 0.2 0.4 0.6 0.8 1
methotrexate
CCAE
MDCD 0.85
MDCD
MDCR 0.63 0.45
MDCR
Optum 0.89 0.77 0.63
Optum
JMDC 0.48 0.46 0.25 0.43
0 0.2 0.4 0.6 0.8 1
adalimumab
CCAE
MDCD 0.84
MDCD
MDCR 0.63 0.47
MDCR
Optum 0.88 0.74 0.7
Optum
JMDC 0.49 0.43 0.22 0.39
0 0.2 0.4 0.6 0.8 1
glatiramer
CCAE
MDCD 0.78
MDCD
MDCR - -
MDCR
Optum 0.85 0.64 -
Optum
JMDC - - - -
0 0.2 0.4 0.6 0.8 1
ocrelizumab
CCAE
MDCD 0.76
MDCD
MDCR - -
MDCR
Optum 0.82 0.61 -
Optum
JMDC - - - -
0 0.2 0.4 0.6 0.8 1

B


xylitol
vitamin A
rivaroxaban
hydralazine
Similarity to warfarin in JMDC
Similarity to warfarin in CCAE
0.4
0.6
0.8
1.0


CCAE
MDCD
MDCR
Optum
JMDC
Comparator #1
Comparator #500
Proportion of Covariates
100%
75%
50%
25%
0%
1
20
40
60
80
100
Randomly Selected Target Cohort
Degree of Imbalance
|SMD| ≥ 0.8
0.5 ≤ |SMD| < 0.8
0.2 ≤ |SMD| < 0.5
0.1 ≤ |SMD| < 0.2
|SMD| < 0.1