Empirical comparator identification for pharmacoepidemiological studies

1	Large-scale empirical identification of candidate comparators for						
2	pharmacoepidemiological studies						
3 4	Justin Bohn ¹ James P. Gilbert ¹						
5			Christopr				
6			David N	1. Kern ¹			
7			Patrick I	B. Ryan ¹			
8							
9		¹ Glo	bal Epidemio	logy Organization			
10	Jansse	en Research	and Develop	oment, LLC, Titusville, NJ	I, USA		
11			Correspor	ndence to:			
12			Justin	Bohn			
13		112	25 Trenton-H	arbourton Road			
14			Titusville, NJ	, USA, 08560			
15			jbohn2@i	<u>ts.jnj.com</u>			
16			+1 (857) 🕻	772-2015			
		10tal	word count	4,098			
	Abstract			250			
			Figures	4			
			Tables	3 20 in main tout			
			References	39 in main text			
				60 in literature			
		Cupplonenterr		2 de eurocente			
		Sup :	plementary	2 documents			
17		I	mormation	1 spreadsneet			
1/							
18			Keywords	s (MeSH):			
D017891 Pharmacoepidemiology							
		D015331 Cohort Stud		lies			
		D011358	Product Sur	veillance, Postmarketing	5		
		D015982	Bias				
		D015986	Confoundin	g Factors, Epidemiologi			
	DO	00067575	Administrat	tive Claims, Healthcare			
19							

Empirical comparator identification for pharmacoepidemiological studies

1 ABSTRACT

2 **Objectives**

3 The new user cohort design has emerged as a best practice for the estimation of drug effects

- 4 from observational data. However, despite its advantages, this design requires the selection and
- 5 evaluation of comparators for appropriateness, a process which can be challenging. In this paper, we
- 6 introduce an empirical approach to rank candidate comparators in terms of their similarity to a target
- 7 drug in high-dimensional covariate space.

8 Methods

9 We generated new user cohorts for each RxNorm ingredient in five administrative claims
10 databases, then extracted aggregated pre-treatment covariate data for each cohort across five clinically

oriented domains. We formed all pairs of cohorts with \geq 1,000 patients and computed a scalar similarity

12 score, defined as the average of cosine similarities computed within each domain, for each pair. Ranked

13 lists of candidate comparators were then generated for each cohort.

14 Results

Across up to 1,350 cohorts forming 922,761 comparisons, drugs that were more similar in the ATC hierarchy had higher cohort similarity scores. The most similar candidate comparators for each of six example drugs corresponded to alternative treatments used in the target drug's indication(s), and choosing the top-ranked comparator for randomly selected drugs tended to produce balance on a majority of covariates. This approach also ranked highly those comparators chosen in high quality published new user cohort design studies.

21 Conclusion

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

Empirical comparator identification for pharmacoepidemiological studies

1 KEY POINTS

Empirical comparator recommendations based on similarity to a target cohort's covariate profile can aid
investigators in study design and align with subject matter knowledge and published literature.

4

5 1 INTRODUCTION

6 Pharmacoepidemiologic studies routinely employ cohort designs to estimate the effect of a drug 7 treatment on an outcome using observational data. While many variations on the cohort design exist, 8 the new user (NU, sometimes called *incident* new user or *initiator*) design[1,2]has emerged over the 9 past two decades as a best practice. This design, which is intended to emulate a randomized trial in its 10 well-defined, exposure-based timescale, involves identifying cohorts of patients at the time of initiation 11 of one of two drugs: a target drug and a comparator. Patients are then followed for the occurrence of 12 the study outcome and the incidence is compared between groups, usually with adjustment for 13 confounding factors. The NU design can be further divided based on whether an active comparator (i.e., 14 an alternative treatment for the indication of interest) or an *inactive* comparator (i.e., something not 15 used for the indication of interest and which is believed to be unrelated to the outcome of interest) is 16 chosen[3]. The primary advantage to this approach is the mitigation of biases known to impact 17 comparisons against nonuser – i.e., unexposed or untreated – groups. In particular, healthy user bias, 18 healthy adherer bias, and immortal time bias may all be reduced by the adoption of an NU design, with 19 particular reductions in confounding by indication expected if an active comparator is chosen[4,5]

However, while the NU 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 outcome experience determined

Empirical comparator identification for pharmacoepidemiological studies

1	by the research question of interest. This question may concern whether the incidence of an outcome in
2	the presence of the target exposure is different than what would have occurred if some other exposure
3	were substituted for the target, or alternatively whether that incidence differs from what would have
4	occurred if the target exposure were simply removed. In the former case, which is termed comparative
5	effectiveness or safety research, investigators have historically selected comparators on the basis of
6	perceived clinical equipoise and empirical feasibility. In the latter case, which is common in the drug
7	safety surveillance context, the use of a comparator otherwise thought to be wholly unrelated to the
8	study outcome has been recommended[6–9], and a theoretical framework for its choice developed[10].
9	While the counterfactual represented by the comparator thus varies according to the research context,
10	a common goal is to identify a comparator with a similar patient composition in terms of demographics
11	and relevant medical characteristics to the target. While several sets of recommendations[11] have
12	emerged to guide investigators in the comparator selection process, it remains at least partially
13	subjective and difficult to scale in a context with many target drugs. With thousands of drugs available
14	on the market, each of which could theoretically be a viable comparator for another, the selection of
15	comparators for all drugs is thus a large task, encompassing nearly one million target-comparator
16	possibilities.

17 In this paper, we present an empirical approach for ranking potential comparator cohorts for 18 pharmacoepidemiologic studies based on their similarity in high-dimensional covariate space. We 19 propose a method to compute a scalar similarity score, the *cohort similarity score*, for all pairwise 20 combinations of all active drug ingredients and rank candidate comparators for any given target drug 21 based on this composite score. The cohort similarity score is based on the average of cosine similarity 22 scores[12] computed within five separate covariate domains intended to emulate clinical thinking 23 surrounding a treatment decision. These domains are demographics, medical history, prior medications, 24 presentation, and visit context. We evaluate rankings based on the cohort similarity score in the context

Empirical comparator identification for pharmacoepidemiological studies

1	of existing drug classification schemes and assess whether they produce pairs of cohorts comprised of
2	similar patient populations, based on having high proportions of covariates already in balance.
3	Throughout, we focus on six example drugs from three therapeutic areas: anticoagulation
4	(warfarin and apixaban), rheumatoid arthritis (methotrexate and adalimumab), and multiple sclerosis
5	(glatiramer and ocrelizumab). We demonstrate that this method produces ranked lists of comparators
6	which broadly align with subject-matter knowledge, and which are already posited as good comparators
7	in the literature or registered studies. Further, we propose that this approach may be of use both as a
8	tool for researchers designing pharmacoepidemiologic studies and as a part of an eventual analytic
9	pipeline to automate the evidence generation process. We provide the full suite of results in an
10	interactive web application (<u>data.ohdsi.org/ComparatorSelectionExplorer</u>), which can be used to identify
11	and diagnose potential comparators for a given target cohort, and the code necessary for others to
12	implement our approach in their own data (<u>https://github.com/OHDSI/ComparatorSelectionExplorer</u>).

13 2 METHODS

14 2.1 Data Sources

We utilized data from five sources of insurance claims: Merative[™] MarketScan[®] Commercial 15 Claims and Encounters (CCAE), Merative[™] MarketScan[®] Multi-State Medicaid (MDCD), Merative[™] 16 MarketScan®Medicare Supplemental (MDCR); Japan Medical Data Center (JMDC); and Optum© De-17 Identified Clinformatics[©] Data Mart. All data were mapped to the Observational Medical Outcomes 18 19 Partnership Common Data Model[13] (OMOP CDM). Episodes of continuous exposure to individual 20 RxNorm drug ingredients were inferred from pharmacy claims[14]. International Classification of 21 Diseases (ICD) diagnosis codes were mapped to the Systematized Nomenclature of Medicine Clinical 22 Terms (SNOMED-CT).

Empirical comparator identification for pharmacoepidemiological studies

1 2.2 Cohort Definitions

2	Prior to computing similarity scores, it was necessary to create cohorts for each active drug
3	ingredient to be evaluated as either a target or comparator. As such, in each data source, we defined
4	cohorts of new users for each RxNorm ingredient. While there are over 10,000 ingredient concepts in
5	the RxNorm vocabulary, only a minority are used in medical practice and the number varies according to
6	local usage patterns in each data source. To enter a given ingredient-specific cohort, patients were
7	required to have 365 days of continuous observation in the database prior to the first observed
8	exposure to that ingredient, which served as their index date. Patients were allowed to enter a given
9	cohort only once, but there was no restriction on the number of distinct cohorts they could enter. In
10	order to ensure robust estimation of covariate prevalence, we restricted analysis to cohorts with at least
11	1,000 patients.
12	2.3 Extraction of Covariate Data
13	Within each cohort, covariate data was extracted in five domains intended to capture the clinical
14	thinking surrounding a treatment decision. These domains were:
15	• Demographics : one covariate for female sex, one for male sex, and one covariate for each
16	decade of age
17	• Medical history: one covariate for each SNOMED condition observed in the period from the
18	start of patient observation to 31 days before the index date
19	• Presentation : one covariate for each SNOMED condition observed in the period from 30 days
20	before the index date to the index date
21	
	• Prior medications: one covariate for each RxNorm ingredient with exposure start at least 30

Empirical comparator identification for pharmacoepidemiological studies

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.

7 2.4 Computation of Similarity Scores

8 Within each database, we formed all possible pairs of cohorts and computed similarity scores 9 for each pair. We calculated the cosine similarity [12], the dot product of two vectors containing the 10 target and comparator cohorts' covariate prevalences divided by the product of their lengths, between 11 the two cohorts, separately for each covariate domain, and then averaged across the five domains to get 12 the final similarity score, hereafter called the *cohort similarity score*. While other measures of 13 multivariable similarity/distance such as the average absolute standardized difference and the 14 Mahalanobis distance were available, the computational efficiency of the cosine similarity made it an 15 attractive choice for our large-scale evaluation. A cohort similarity score of 1 will occur if all covariates 16 are the same between two cohorts, while a score of 0 indicates that the cohorts have no agreement across the covariates. We then used these scores to generate ranked lists of comparators for each 17 18 cohort, with scores closer to 1 indicating stronger comparator alignment.

Under the assumption that cohorts with known clinical or mechanistic relationships should have
higher similarity than those without such relationships, we used the Anatomical Therapeutic Chemical
(ATC) classification system[15] to compare the distribution of cohort similarity scores for cohort pairs
with relationships at different levels of the hierarchy. Cohort pairs were divided according to the most

Empirical comparator identification for pharmacoepidemiological studies

1	granular ATC level at which they co-occurred, if any. Pairs that co-occurred in a class at ATC level 5 were
2	excluded from this analysis as this level is commonly used for individual drug products or combinations.
3	For a random subset of 100 cohorts within each data source, we also attempted to validate our
4	ranked lists with a more familiar pharmacoepidemiological metric of covariate similarity, the
5	standardized mean difference (SMD)[16]. For each of the randomly selected target drugs, we identified
6	their top comparator based on the cohort similarity score, calculated SMDs for all covariates present
7	within the cohort pair and reported the proportion of covariates in various degrees of imbalance based
8	on recommended cutoffs for absolute SMD of 0.1, 0.2, 0.5, and 0.8 [17,18]. We then repeated this
9	process for the comparator ranked 500 th . All code for cohort generation, covariate extraction, similarity
10	calculation, and for our web application is publicly available at

11 https://github.com/OHDSI/ComparatorSelectionExplorer.

12 2.5 Review of NU Design Studies Published in High-Quality Journals

13 We also searched the literature to identify high-quality observational studies employing NU 14 designs and compared their comparator choices to our empirical estimates of similarity. We limited our 15 search to papers published in the following journals between January 2021 and May 2023: Science, 16 Nature, The New England Journal of Medicine, JAMA, JAMA Internal Medicine, JAMA Surgery, JAMA 17 Psychiatry, JAMA Cardiology, JAMA Pediatrics, JAMA Neurology, The Lancet, BMJ, Diabetes Care, 18 Hypertension, Journal of the American College of Cardiology, Pharmacoepidemiology and Drug Safety, 19 Drug Safety, Annals of Internal Medicine, The Lancet Digital Health, The Lancet Infectious diseases, The 20 Lancet Diabetes & Endocrinology, The Lancet Neurology, The Lancet Oncology, The Lancet Respiratory 21 Medicine, The Lancet Psychiatry, and The Lancet Global Health. Full details on the search parameters 22 used and the results of the review can be found in the supplementary material. As some of the 23 identified studies defined target/comparator cohorts at the drug *class* level, we additionally generated

Empirical comparator identification for pharmacoepidemiological studies

new-user cohorts for each Anatomical Therapeutic Chemical (ATC) level 4 class and re-ran our similarity
 pipeline with these cohorts included.

3 3 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.

10 Between 487.578 and 922.761 pairs of cohorts with at least 1.000 patients were formed per 11 data source. Among these, similarity was lowest in the "Presentation" domain (mean cosine similarity 12 0.213 to 0.490, depending on data source) and highest in the "Visit context" domain (mean cosine 13 similarity 0.869 to 1.000, depending on data source). Between 90% and 95% of target cohorts with at 14 least 1,000 patients had a comparator with an average cohort similarity score of ≥ 0.950 . Among the six 15 example drugs, the distribution of cohort similarity scores was consistently left-skewed, with most 16 comparators lying between 0.7 and 0.9 (Figure 1). Similarity was lowest among pairs of cohorts that did 17 not occur together in any ATC classes and tended to increase among pairs of cohorts that co-occurred 18 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[19], and those ranked sixth through tenth (teriflunomide, dalfampridine, ocrelizumab,

Empirical comparator identification for pharmacoepidemiological studies

1	corticotropin, and modafinil) are either disease modifying therapies or are used to treat symptoms of
2	MS or MS relapse. Ranked comparator lists for these example drugs varied according to data source
3	(Figure 1, Figure 3 panel A), with the highest degree of correlation between CCAE, MDCD, and Optum.
4	An example of between-data source heterogeneity is shown in Figure 3 panel B where, among 627
5	candidate comparators for warfarin observed in both CCAE and JMDC, rivaroxaban is identified as highly
6	similar to warfarin in both data sources, whereas xylitol and vitamin A appear much more similar to
7	warfarin in JMDC than in CCAE, while the reverse is true for hydralazine.
8	
9	Among the 100 target drugs randomly selected within each data source and their top
10	comparators, most covariates were well balanced. When the comparator ranked 500 th was chosen
11	instead, the proportion of imbalanced covariates was higher (Figure 4). The average proportion of
12	covariates with $ SMD \leq 0.1$ was between 72% and 84% for the top comparator versus between 28%
13	and 45% for the comparator ranked 500 th .
14	Our literature search returned 60 studies, of which 24 employed an NU design. 56 distinct
15	target-comparator pairs were identified, of which 23 could be mapped directly to RxNorm ingredients or
16	ATC level 4 classes. Nine such pairs were comparisons among anticoagulants, and five were among
17	antidepressants. Among these 20 pairs, the comparator selected by the authors was in the top 5% of
18	computed similarity for its target in at least one of our data sources in 96% of cases, and in 91% of cases
19	it was ranked among the top 10. Some pairs with notably low computed similarity were apixaban vs.
20	warfarin (cohort similarity score 0.879, rank 136 in CCAE), beta blocking agents vs. aldosterone

21 antagonists (cohort similarity score 0.757, rank 359 in CCAE), and topiramate vs. lamotrigine (cohort

22 similarity score 0.846, rank 167 in CCAE). See supplementary file for similarity scores from reviewed

23 literature.

Empirical comparator identification for pharmacoepidemiological studies

1 4 DISCUSSION

23

2	Over the past two decades, the NU cohort design has emerged as a best practice for the
3	estimation of drug effects from observational data. However, the implementation of this design may be
4	hampered by the difficulty of selecting an appropriate comparator for the target drug of interest, either
5	because there are many potential alternatives to the target drug, or because the true target of inference
6	is in fact a contrast between the target drug and its absence. The development of empirical comparator
7	recommendations can aid investigator decision-making, simplifying the process of comparator selection
8	and improving the quality of pharmacoepidemiological research and the speed of postmarketing
9	surveillance. Ultimately, these approaches may even enable the automation of evidence generation
10	using the NU design, something that was previously possible only with self-controlled designs[20,21]
11	and temporal scans[22,23].
12	We demonstrated a method for ranking potential comparators based on similarity in high-
13	dimensional covariate space. We showed that most target cohorts of reasonable size possessed at least
14	one potential comparator with a high degree of similarity and that cohorts that were more related in the
15	ATC hierarchy also had higher cohort similarity scores. For a random sample of 100 cohorts within each
16	data source, we also showed that selecting the top-most similar comparator identified a cohort with
17	balance on the majority of covariates before any adjustment took place.
18	We additionally demonstrated that, for six target cohorts from a diverse set of therapeutic
19	areas, the top-ranked comparators aligned well with subject-matter knowledge. That is, all were used in
20	the treatment of at least one of the target cohorts' indications. Indeed, many of these top-ranked
21	comparators are found to be in use in head-to-head safety or effectiveness comparisons with the target
22	cohort in the literature or in public registries of clinical and observational studies. For example, head-to-

11

head studies exist comparing apixaban and its top comparator rivaroxaban[24,25], adalimumab and its

Empirical comparator identification for pharmacoepidemiological studies

1	top comparator certolizumab[26,27], and ocrelizumab and its top comparator teriflunomide[28]. This
2	was also borne out in our formal literature review, which demonstrated that NU studies published in
3	high quality journals tended to select comparators that our method also ranked highly. While this is
4	reassuring, it should also be noted that the correlation in comparator rankings across data sources was
5	not consistently high (Figure 3). This may indicate that a given comparator is not uniquely fit-for-purpose
6	across all data sources, suggesting that investigators take care to select a comparator using information
7	from a source that at least resembles their intended analytic context.
8	This study complements a large body of work on drug similarity from the bioinformatics
9	community, including attempts to define similarity based on chemical structure, gene expression[29],
10	genome-wide associations[30], co-occurrence of drugs with condition terms in the literature[31],
11	individual patient data[32], and a combination of these. However, our work is unique in its focus on the
12	use of similarity indices for comparator selection rather than for identification of novel associations, and
13	for its clinically oriented approach to covariate data extraction. The empirical comparator
14	recommendations provided here have several advantages. First, they can reduce subjectivity, both in
15	terms of which drugs may serve as candidate comparators for a given target drug and which covariates
16	should be used to determine empirical feasibility. While comparator choice should be guided by subject-
17	matter knowledge[33], no investigator is omniscient and thus the availability of empirical
18	recommendations may promote more informed decisions and could potentially enable less biased
19	comparisons. Additionally, empirical recommendations may reveal previously unknown comparators
20	that are not necessarily alternative treatments for any of the target drug's indications. They thus hold
21	promise for the identification of appropriate inactive comparators/negative control exposures,
22	ultimately allowing for cohort studies that emulate user versus non-user comparisons within the NU
23	framework. Additionally, these recommendations have been assessed across multiple large data sources
24	comprising diverse populations within and outside the United States. They have been observed,

Empirical comparator identification for pharmacoepidemiological studies

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 and high quality
published research. Finally, the approach presented here provides a standardized framework for
comparator ranking which can be readily implemented by investigators. This standardization can
improve reproducibility and scalability and ultimately lowers the burden on domain expertise for the
investigator.

8 In an era of rapid advancement in deep learning methods, the simplicity of our approach may 9 appear to be a limitation. For example, emerging research on large language models (LLMs) [34] may 10 suggest their potential utility as a source of comparator recommendations, and much has been written 11 on the utility of various similarity-based deep learning approaches to predict drug-drug and drug-target 12 interactions. Unfortunately, the lack of a reference corpus explicitly identifying "good" and "bad" 13 comparators prevents us from trialing these methods here. And while we look forward to future 14 assessments of their utility, LLM approaches provide little empirical insight as to the origin of their 15 responses, can suffer from hallucinations, and are naturally biased towards pre-existing information, 16 including non-refereed online context, reducing their utility in identifying negative control exposures or 17 previously unused but otherwise suitable comparators.

Compared to traditional approaches, a limitation of the approach introduced here is that the cohort similarity score is based on the *marginal* distributions of covariates in 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[35]) that reflect the similarity in *joint* covariate distributions across cohorts. Furthermore, such methods can be used to adjust for observed dissimilarity when it occurs. 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

Empirical comparator identification for pharmacoepidemiological studies

1	based on all available data elements. It should be noted that this data-adaptive approach has been
2	incorporated into some propensity score methods like the high-dimensional propensity score (hdPS)
3	[36] and, more recently, the large-scale propensity score (LSPS)[37]. However, the cohort similarity score
4	remains the only approach that is scalable to millions of candidate target-comparator combinations,
5	whereas PS models—particularly those using high-dimensional covariate data—can be computationally
6	expensive, making them infeasible for screening large numbers of candidate comparators. For example,
7	on an Amazon EC2 instance running Windows Server 2019 with a 2.2 GHz processor and 64 GB RAM and
8	accessing data stored in a Redshift database, our entire pipeline (including generating cohorts,
9	extracting covariate data, and computing cohort similarity scores for all drugs) took between 33 and 172
10	minutes per database, whereas an LSPS pipeline (including generating cohorts, extracting covariate data,
11	and fitting models) took between 3 and 40 hours for a single target-comparator pair. The cohort
12	similarity score may therefore be best suited as a tool to narrow the comparator search space before
13	techniques like hdPS/LSPS are employed. It should also be noted that, like PS methods and assessments
14	of standardized differences in covariates, our approach only addresses half of the confounding picture.
15	That is, it summarizes covariate imbalance without addressing covariates' associations with study
16	outcomes. And while investigators must keep this limitation in mind, they can address it <i>post-hoc</i> by
17	searching our web application to identify the degree of balance on important risk factors for their study
18	outcome of interest.
19	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.

Empirical comparator identification for pharmacoepidemiological studies

1	However, this can be ameliorated by running the entire analytic pipeline in the subset of the population
2	with a specific indication of interest. A demonstration of this for the target spironolactone, which is used
3	both in the treatment of acne and heart failure, showed that such restriction both reduces the
4	comparator search space and improves the coherence of recommendations (see supplementary
5	material for details). While such restriction is not feasible at the scale of all indications, it may be
6	attractive to investigators whose work routinely focuses on a small set of indications. Additionally, the
7	assessment of these recommendations was limited by the lack of an objective metric for defining
8	comparator "appropriateness". While we sought to benchmark appropriateness based on comparator
9	location in the ATC hierarchy and appearance in published literature or registered studies, future work
10	might attempt to go further by demonstrating that the use of empirical comparator recommendations
11	can lead to less biased treatment effect estimates than traditional methods of comparator selection.
12	Finally, it should be clearly stated that empirical comparator recommendations are not a
13	replacement for critical thinking in study design and analysis. Investigators may wish to use them as a
14	"first pass" to identify and evaluate candidate comparators, after which more careful consideration and
15	detailed analysis can determine true feasibility. Crucially, investigators must remain aware that
16	covariate adjustment is almost invariably required regardless of the quality of the comparator chosen.
17	Similarity between a target drug and comparator is also no guarantee that any remaining covariate
18	imbalance can be successfully adjusted via established methods like restriction or propensity score
19	adjustment. And, conversely, low similarity relative to other recommendations is not a guarantee that
20	successful adjustment is impossible.
21	Our work leaves open many exciting opportunities for further exploration of empirical
22	comparator recommendations. Among these are additional approaches to validation, including a
23	comparison against more computationally expensive techniques like assessing propensity score overlap.

24 Future work might also explore the utility of adding sources of metadata on known or suspected drug-

Empirical comparator identification for pharmacoepidemiological studies

- 1 outcome relationships, such as the OHDSI Common Evidence Model[38,39], in order to improve the
- 2 ease of identifying inactive comparators/negative control exposures.

3 **5 CONCLUSION**

- 4 We have developed and evaluated a method for recommending comparators that yields
- 5 promising results and which may be of use to investigators wishing to conduct NU studies of drug safety
- 6 and effectiveness. An open-source tool set is available for those wishing to identify candidate
- 7 comparators using their own data, and the full suite of results, including comparator rankings for
- 8 thousands of drugs across multiple data sources, is available to the public. We hope these tools and the
- 9 results presented here will accelerate and improve the quality of future pharmacoepidemiologic
- 10 research.
- 11

12 **REFERENCES**

Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol.
 2003;158:915–20.

Johnson ES, Bartman BA, Briesacher BA, Fleming NS, Gerhard T, Kornegay CJ, et al. The incident user
 design in comparative effectiveness research. Pharmacoepidemiol Drug Saf. 2013;

- D'Arcy M, Stürmer T, Lund JL. The importance and implications of comparator selection in
 pharmacoepidemiologic research. Curr Epidemiol Rep. 2018;5:272–83.
- 19 4. 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.
- 5. Sendor R, Stürmer T. Core concepts in pharmacoepidemiology: Confounding by indication and the
 role of active comparators. Pharmacoepidemiol Drug Saf. 2022;31:261–9.
- 6. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative Controls: A Tool for Detecting Confounding and
 Bias in Observational Studies. Epidemiology. 2010;21:383–8.
- 26 7. Dusetzina SB, Brookhart MA, Maciejewski ML. Control Outcomes and Exposures for Improving
- 27 Internal Validity of Nonrandomized Studies. Health Serv Res. 2015;50:1432–51.

Empirical comparator identification for pharmacoepidemiological studies

- 1 8. Setoguchi S, Glynn RJ, Avorn J, Mogun H, Schneeweiss S. Statins and the Risk of Lung, Breast, and
- 2 Colorectal Cancer in the Elderly. Circulation. 2007;115:27–33.
- 3 9. Solomon DH, Avorn J, Stürmer T, Glynn RJ, Mogun H, Schneeweiss S. Cardiovascular outcomes in new
- 4 users of coxibs and nonsteroidal antiinflammatory drugs: High-risk subgroups and time course of risk.
- 5 Arthritis Rheum. 2006;54:1378–89.
- 6 10. Huitfeldt A, Hernan MA, Kalager M, Robins JM. 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.
- 9 11. Setoguchi S, Gerhard T. Chapter 5: Comparator Selection. In: Velentgas P, Dreyer NA, Nourjah, P,
- 10 Smith, SR, Torchia MM, editors. Dev Protoc Obs Comp Eff Res Users Guide [Internet]. Rockville (MD):
- 11 Agency for Healthcare Research and Quality (US); 2013. Available from:
- 12 https://www.ncbi.nlm.nih.gov/books/NBK126184/
- 12. Tan P-N, Steinbach M, Kumar V. Chapter 8. Introd Data Min. 1st ed. Boston: Pearson Addison
 Wesley; 2006. p. 500.
- 15 13. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for
- active safety surveillance research. J Am Med Inform Assoc JAMIA. 2012;19:54–60.
- 17 14. OMOP CDM v5.4 [Internet]. [cited 2022 Nov 28]. Available from:
- 18 http://ohdsi.github.io/CommonDataModel/cdm54.html#Standardized_Derived_Elements
- 19 15. WHOCC Structure and principles [Internet]. [cited 2022 Nov 24]. Available from:
- 20 https://www.whocc.no/atc/structure_and_principles/
- 21 16. 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.
- 23 17. Normand S-LT, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, et al. Validating
- 24 recommendations for coronary angiography following acute myocardial infarction in the elderly: A
- 25 matched analysis using propensity scores. J Clin Epidemiol. 2001;12.
- 18. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, N.J: L. Erlbaum
 Associates; 1988.
- 19. Hauser SL, Cree BAC. Treatment of Multiple Sclerosis: A Review. Am J Med. 2020;133:1380-1390.e2.
- 29 20. Kulldorff M, Dashevsky I, Avery TR, Chan AK, Davis RL, Graham D, et al. Drug safety data mining with
- a tree-based scan statistic: TREE-BASED SCAN STATISTIC DATA MINING METHOD. Pharmacoepidemiol
- 31 Drug Saf. 2013;22:517–23.
- Wang SV, Maro JC, Baro E, Izem R, Dashevsky I, Rogers JR, et al. Data Mining for Adverse Drug Events
 With a Propensity Score-matched Tree-based Scan Statistic. Epidemiology. 2018;29:895–903.
- 34 22. Norén GN, Hopstadius J, Bate A, Star K, Edwards IR. Temporal pattern discovery in longitudinal
- 35 electronic patient records. Data Min Knowl Discov. 2010;20:361–87.

Empirical comparator identification for pharmacoepidemiological studies

- 1 23. Lavallee M, Yu T, Evans L, Van Hemelrijck M, Bosco C, Golozar A, et al. Evaluating the performance of
- 2 temporal pattern discovery: new application using statins and rhabdomyolysis in OMOP databases. BMC
- 3 Med Inform Decis Mak. 2022;22:31.
- 4 24. Boehringer Ingelheim. Safety and Effectiveness Study Comparing Dabigatran, Rivaroxaban &
- 5 Apixaban in Non-valvular Atrial Fibrillation Patients Enrolled in the US Department of Defense Military
- 6 Health System [Internet]. clinicaltrials.gov; 2019 Feb. Report No.: NCT03026556. Available from:
- 7 https://clinicaltrials.gov/ct2/show/NCT03026556
- 8 25. Lip GYH, Keshishian A, Kamble S, Pan X, Mardekian J, Horblyuk R, et al. Real-world comparison of
- 9 major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran,
- 10 rivaroxaban, or warfarin: A propensity score matched analysis. Thromb Haemost. 2016;116:975–86.
- 12 26. Smolen JS, Burmester G-R, Combe B, Curtis JR, Hall S, Haraoui B, et al. Head-to-head comparison of
- 12 certolizumab pegol versus adalimumab in rheumatoid arthritis: 2-year efficacy and safety results from
- 13 the randomised EXXELERATE study. The Lancet. 2016;388:2763–74.
- 14 27. UCB Pharma SA. A Multicenter, Single-blind, Randomized Parallel-group Study to Assess the Short-
- 15 and Long-term Efficacy of Certolizumab Pegol Plus Methotrexate Compared to Adalimumab Plus
- 16 Methotrexate in Subjects With Moderate to Severe Rheumatoid Arthritis Responding Inadequately to
- 17 Methotrexate [Internet]. clinicaltrials.gov; 2018 Jul. Report No.: NCT01500278. Available from:
- 18 https://clinicaltrials.gov/ct2/show/NCT01500278
- 19 28. Dirks P, Zingler V, Leemhuis J, Berthold H, Hieke-Schulz S, Wormser D, et al. Design of a non-
- 20 interventional post-marketing study to assess the long-term safety and effectiveness of ocrelizumab in
- 21 German real world multiple sclerosis cohorts the CONFIDENCE study protocol. BMC Neurol.
- 22 2020;20:95.
- 23 29. Brown AS, Kong SW, Kohane IS, Patel CJ. ksRepo: a generalized platform for computational drug
 24 repositioning. BMC Bioinformatics. 2016;17:78.
- 25 30. Grover MP, Ballouz S, Mohanasundaram KA, George RA, H Sherman CD, Crowley TM, et al.
- Identification of novel therapeutics for complex diseases from genome-wide association data. BMC MedGenomics. 2014;7:S8.
- 31. Brown AS, Patel CJ. MeSHDD: Literature-based drug-drug similarity for drug repositioning. J Am Med
 Inform Assoc. 2017;24:614–8.
- 30 32. Zeng X, Jia Z, He Z, Chen W, Lu X, Duan H, et al. Measure clinical drug–drug similarity using Electronic
 31 Medical Records. Int J Med Inf. 2019;124:97–103.
- 32 33. Greenland S, Neutra R. Control of confounding in the assessment of medical technology. Int J
 33 Epidemiol. 1980;9:361–7.
- 34. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical
 knowledge. Nature. 2023;620:172–80.
- 35. Walker A, Patrick, Lauer, Hornbrook, Marin, Platt, et al. A tool for assessing the feasibility of
- 37 comparative effectiveness research. Comp Eff Res. 2013;11.

Empirical comparator identification for pharmacoepidemiological studies

- 1 36. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional Propensity
- 2 Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data. Epidemiology.
- 3 2009;20:512–22.
- 37. Zhang L, Wang Y, Schuemie MJ, Blei DM, Hripcsak G. Adjusting for indirectly measured confounding
 using large-scale propensity score. J Biomed Inform. 2022;134:104204.
- 6 38. Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. Accuracy of an automated
- 7 knowledge base for identifying drug adverse reactions. J Biomed Inform. 2017;66:72–81.
- 8 39. Home · OHDSI/CommonEvidenceModel Wiki [Internet]. GitHub. [cited 2022 Nov 25]. Available from:
- 9 https://github.com/OHDSI/CommonEvidenceModel

1 **TABLES**

2

3

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 pre	evalence $\geq 1\%$ ext	racted per cohort w	ith \geq 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 \geq	922,761	617,716	487,578	584,821	886,446
1,000 patients					
Mean (SD) similarity score among pairs of cohorts with \geq 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 cosine similarity	0.852 (0.139)	0.846 (0.128)	0.945 (0.075)	0.857 (0.134)	0.841 (0.147)
Medical history cosine similarity	0.781 (0.137)	0.798 (0.137)	0.913 (0.059)	0.879 (0.092)	0.798 (0.141)
Presentation cosine similarity	0.213 (0.173)	0.401 (0.216)	0.490 (0.190)	0.377 (0.189)	0.349 (0.212)
Prior medications cosine similarity	0.882 (0.085)	0.872 (0.090)	0.931 (0.059)	0.917 (0.071)	0.859 (0.108)
Visit context cosine similarity	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%)

4 CCAE: Merative[™] MarketScan[®] Commercial Claims and Encounters; MDCD: Merative[™] MarketScan[®] Multi-State Medicaid; MDCR: Merative[™]

5 MarketScan[®]Medicare Supplemental; JMDC: Japan Medical Data Center; Optum: Optum© De-Identified Clinformatics© Data Mart. Cohort

6 similarity score is average of domain-specific cosine similarities.

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

1 Table 2: Top five comparators for six example drugs, according to average rank across five data sources.

2 Restricted to comparators occurring in at least two data sources. Entries are given as comparator name (average rank across data sources, mean

3 cohort similarity score, number of data sources with information on comparison).

4

5

6

Table 3: Top five comparators for six example drugs according to cohort similarity score, by data source.

Example	Data					
Drug	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)

Example Drug	Data Source	Comparator #1	Comparator #2	Comparator #3	Comparator #4	Comparator #5
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)
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)

Example	Data					
Drug	Source	Comparator #1	Comparator #2	Comparator #3	Comparator #4	Comparator #5
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)

1 Entries are given as comparator name (cohort similarity score). Note that glatiramer and ocrelizumab did not meet the 1,000 patient sample size

2 cutoff in MDCR or JMDC and thus results are not present for these data sources.

Empirical comparator identification for pharmacoepidemiological studies

1

2 FIGURES

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



Empirical comparator identification for pharmacoepidemiological studies

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.

7 White circles indicate mean values.



Empirical comparator identification for pharmacoepidemiological studies

1 Figure 3: Pairwise rank correlations in cohort similarity scores between data sources for six example drugs (A) and cosine

2 similarity scores among 627 candidate comparators for warfarin observed in both CCAE and JMDC (B).



Empirical comparator identification for pharmacoepidemiological studies





4

5 **DECLARATIONS**

6 Funding

- 7 This work was funded by Janssen Research and Development, a subsidiary of Johnson & Johnson. All
- 8 authors are employees of Janssen Research and Development.

9 **Conflicts of interest**

10 The authors declare no competing interests.

11 Data availability

- 12 The full suite of results described here is available as an interactive web application
- 13 (<u>data.ohdsi.org/ComparatorSelectionExplorer</u>), which can be used to identify and diagnose potential
- 14 comparators for a given target cohort. Note that this application undergoes periodic updates, including
- 15 the updating of existing data sources to newer versions, the addition of new data sources, and the
- 16 modification of existing functionality. As such, some results derived from the application may differ from
- 17 those presented here.

18 Code availability

- 19 All code necessary to run the analytic pipeline described here and to generate the web application is
- 20 available via GitHub (<u>https://github.com/OHDSI/ComparatorSelectionExplorer</u>).

Empirical comparator identification for pharmacoepidemiological studies

Author	Affiliation	Email Address	ORCID	Contributions
Justin Bohn	Janssen	Jbohn2@its.jnj.com	0000-0003-	Designed the analysis,
	Research &		4430-798X	performed the analysis,
	Development,			conducted literature review,
	Titusville, NJ			drafted the manuscript
James P.	Janssen	Jgilber2@its.jnj.com	0000-0003-	Designed the analysis,
Gilbert	Research &		2294-3459	performed the analysis,
	Development,			reviewed and contributed to
	Titusville, NJ			the manuscript
Christopher	Janssen	Cknoll1@its.jnj.com		Designed the analysis,
Knoll	Research &			reviewed and contributed to
	Development,			the manuscript
	Titusville, NJ			
David M.	Janssen	Dkern2@its.jnj.com	0000-0001-	Designed the analysis,
Kern	Research &		5417-3925	reviewed and contributed to
	Development,			the manuscript
	Titusville, NJ			
Patrick B.	Janssen	Pryan4@its.jnj.com		Designed the analysis,
Ryan	Research &			conducted literature review,
	Development,			reviewed and contributed to
	Titusville, NJ			the manuscript

1 Author contributions