SENTINEL METHODS PROTOCOL

Development and Evaluation of a Global Propensity Score for Data Mining with Tree-Based Scan Statistics

Prepared by: Shirley V Wang[1], Joshua J Gagne[1], Judith C Maro[2], Efe Eworuke[3], Sushama Kattinakere[1], Martin Kulldorff[1], Elande Baro[4], Rima Izem[4], Michael Nguyen[3], Rita Ouellet-Hellstrom[3], Sandra DeLuccia[2], Ella Pestine[2], Danijela Stojanovic[3]

Author Affiliations: 1. Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; 2. Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA; 3. Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD; 4. Office of Biostatistics, Center for Drug Evaluation and Research, FDA, Silver Spring, MD

Version 1: Aug 10, 2018
Version 2: Oct 1, 2018

The Sentinel System is sponsored by the U.S. Food and Drug Administration (FDA) to proactively monitor the safety of FDA-regulated medical products and complements other existing FDA safety surveillance capabilities. The Sentinel System is one piece of FDA's Sentinel Initiative, a long-term, multi-faceted effort to develop a national electronic system. Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I.

History of Modifications

Version	Date	Modification	By
V2	Oct 1, 2018	• Responses to public comments detailed in Appendix B were incorporated to protocol text. • Maximum follow up for case studies 1 and 2 were changed from 30 days to 183 days • For all examples, patients will be censored if they switch between compared therapies (defined as a dispensation for a drug in the other group during follow up)	Shirley V. Wang

Sentinel

Sentinel Methods Protocol

Development and Evaluation of a Global Propensity Score for Data Mining with Tree-Based Scan Statistics

Table of Contents

I. INTRODUCTION 1

II. SPECIFIC AIMS 2

III. EXPERT ENGAGEMENT 3

IV. PROPENSITY SCORE MATCHED COHORT DESIGN 3

A. CANDIDATE GLOBAL PROPENSITY SCORES 3

V. TREESCAN 5

A. HIERARCHICAL TREE (MULTI-LEVEL CLINICAL CLASSIFICATION SOFTWARE) 5
B. DEFINING INCIDENT OUTCOMES 6
C. UNCONDITIONAL BINOMIAL TREE SCAN STATISTIC 6
D. SELECTED OUTCOMES AND ALERTS TO EVALUATE 7

VI. DRUG SAFETY SIGNAL DETECTION CASE STUDIES 8

A. DATA 8
B. PERFORMANCE METRICS 8
C. SELECTED EXAMPLES 9

VII. REFERENCES 15

VIII. APPENDICES 20

A. CURATION OF THE MLCCS TREE 20
B. RESPONSES TO PUBLIC COMMENT 20
C. DETAILED PROTOCOL SPECIFICATIONS FOR EXAMPLES 22

I. INTRODUCTION

TreeScan™ (http://www.treescan.org) is a statistical data mining tool that is compatible with multiple study designs and addresses many methodological challenges[1] with making secondary use of healthcare databases for detection of potential signals related to marketed medical products. It uses a hierarchical outcome tree to group related codes together and applies tree-based scan statistics to adjust for multiple testing when screening across thousands of potential adverse events.[2]

TreeScan has previously been used to screen for adverse events when conducting signal detection in vaccine safety studies that used a self-controlled design.[3-5] Studies of childhood vaccine safety are well suited for self-controlled designs because these are often administered on an age-based schedule[6] rather than in response to a change in clinical condition. In contrast, the clinical context behind decisions of whether, when and for how long to treat patients with specific drugs can make issues of confounding related to timing of exposure more complex than typically found in vaccine studies.[7] While self-controlled designs cannot be confounded by time-invariant characteristics due to the within-person comparisons, they are susceptible to bias from time-varying characteristics (e.g. healthy user bias[8,9], trends in exposure probability in the population[10] or within individual[11,12], decline in overall health status, protopathic bias[13]). Recent work evaluating TreeScan for adverse event screening using drug examples with a self-controlled risk interval design resulted in many alerts for outcomes that were related to underlying changes in health condition that prompted the initiation of therapy (Maro et al, draft in progress). For example, when screening for adverse events after antibiotics, there were numerous alerts for conditions related to the underlying infection that prompted the need for antibiotic.

For drug safety evaluations, a cohort study of new initiators is a powerful design that can better address unmeasured time-varying characteristics associated with initiation of treatment that are also related to outcomes of interest (e.g. confounders) through judicious selection of an appropriate active comparator group.[14,15] While selection of an appropriate comparator in the design phase is crucial, cohort studies often require additional adjustment of measurable confounders and/or their proxies. A common method for adjusting for confounding is via a propensity score (PS). The PS is the predicted probability of exposure (versus comparator) conditional on covariates and is typically estimated via a logistic regression model.[16] When the PS is used to conduct matched, weighted or stratified analyses, the results are adjusted for confounders that are included in the PS model as well as closely related proxies. The covariates selected for inclusion in the PS can be predefined by the investigator or empirically defined via machine learning algorithms.[17-21]

A recent analysis involved simulation to evaluate the performance of tree-based scan statistics with PS matched analysis of new initiator cohorts. These analyses showed promise for TreeScan with PS matching as a method for screening and prioritization of potential adverse events (Wang et al, in press, Epidemiology). In the simulations, the true confounding structure was known to the investigators and the scenarios were limited to under 30 potential confounders for selected outcomes. However, when applying TreeScan for a real drug safety screening using a PS matched active-comparator cohort design, special consideration will be necessary to address confounding. Risk factors, and therefore confounders, will vary by outcome and TreeScan scans across thousands of possible outcomes. Therefore, it would not be feasible to identify potential confounders for every possible adverse event that is screened.

FDA is interested in development of a "global" PS that could be applied generically across different active-comparator cohort studies in varied populations to adjust for many of the major confounders associated with the different outcomes being scanned. A global PS would include a set of pre-defined variables that are risk factors for a variety of outcomes and/or empirically selected variables related to exposure selection. After scanning for potential adverse events using a global PS, outcomes that alert can be further explored and refined in analyses where confounders are carefully selected and tailored to the specific adverse events of interest.

II. SPECIFIC AIMS

This methods project aims to develop and evaluate candidate global propensity scores for application with the propensity score cohort matched design and tree-based scan statistics. The examples used in this project are not part of a regulatory evaluation.

Aim #1: Develop and compare the relative performance of candidate global propensity scores which could be applied generally in cohort studies involving tree-based scan statistics by evaluating 4 case studies for drug safety signal detection, where the selected drugs have well characterized safety profiles.

We will evaluate tree-based scan statistics with crude, age-sex matched and 1:1 global propensity score matched cohorts. The workgroup will identify a set of candidate global propensity scores, then conduct analyses to evaluate performance of 1:1 matched cohorts for each drug safety signal detection case study using a proprietary database licensed by Brigham and Women's Hospital, Truven Health MarketScan® Research Database. Evaluation of performance will focus on a subset of selected outcomes in the hierarchical tree for each example and will include assessment of balance on known risk factors as well as impact of including additional adjustment for known risk factors on alerting (see section VIB for more detail).

Aim #2: Using the *a priori* specified primary global propensity score to adjust for confounding, we will conduct clinical and epidemiologic review of TreeScan alerts.

The workgroup will review alerts from one or more case studies after running TreeScan on a cohort 1:1 matched on the pre-specified primary global propensity score to adjust for confounding. We will review alerts identified at a pre-specified threshold of $p \leq 0.01$ from both clinical and epidemiological perspectives. The interpretation of alerts will include (but not necessarily be limited to) consideration of biologic pathways, confounding by indication, and reverse causality. If there are known signals, the WG will review and interpret the pattern of alerting for these known adverse events. Because the outcome nodes in the hierarchical tree may not reflect the most sensitive or specific algorithms for any particular outcome, this may influence alerting. For example, consider the situation where a drug increases the risk of Type II diabetic ketoacidosis (a very rare outcome), and incident outcomes are defined at level 3 of the Multilevel Clinical Classifications Software (MLCCS) tree hierarchy. At level 3 of the tree, diagnosis codes for diabetic ketoacidosis are part of a node that includes more prevalent general diabetes codes. Thus, when scanning, the effect of the drug on diabetic ketoacidosis may be diluted due to the lack of specificity of the outcome definition. The WG will learn from observing patterns of alerts for known adverse events. After the workgroup considers clinical context and methodological issues to assess whether there may be clear, non-causal explanations for unanticipated alerts, at the FDA's discretion, a separately scoped project may be initiated to distribute relevant Cohort Identification and Descriptive Analyses (CIDA) + Propensity Score Analysis Tool and TreeExtraction software packages to the Sentinel

Distributed Database (SDD). [22,23] If similar unknown or unexplained alerts are observed, the workgroup may follow up with clinical review of Patient Episode Profile Retrieval (PEPR)[24], and design of a protocol based safety assessment with confounding control targeted to the outcome of interest.

III. EXPERT ENGAGEMENT

To increase opportunities for input around signal detection and TreeScan methods, the workgroup will invite members of the Methods Core to review and provide input on the project protocols. The workgroup will engage the Methods Core for feedback at three points during the project: Project initiation/protocol review; presentation of interim findings and presentation of project results.

IV. PROPENSITY SCORE MATCHED COHORT DESIGN

The propensity score matched cohort design is currently being used by the FDA Sentinel Program in active surveillance activities.[25-31] In this context, a cohort is generally defined by incident exposure to a new drug and an appropriate comparator drug. The initiators of each drug are then matched based on a propensity score, which is a summary measure capturing the probability of being an initiator of the new drug of interest rather than the comparator drug based on multiple baseline characteristics.

There are many variants on how a cohort may be extracted from longitudinal healthcare claims data (e.g. defining cohort entry, washout period for incidence), choices regarding what goes into the propensity score (e.g. which covariates, defined over what period) as well as variants on how to match patients (e.g. matching caliper, 1:1 versus variable ratio). These design and implementation decisions are captured as options that investigators may alter when using the Sentinel Program's CIDA + Propensity Score analysis tool.[32]

A. CANDIDATE GLOBAL PROPENSITY SCORES

For Aim 1, candidate global PS will include a combination of demographics, comorbidity score, frailty, screening measures, healthcare utilization measures, exposure based high dimensional propensity score (hdPS)[18] and investigator selected risk factors that may influence choice of therapy (**Table 1**).

Table 1. Evaluation of TreeScan with variants of a global propensity score to adjust for confounding

	Predefined global[1]	Empirically selected[2]	Predefined tailored[3]
1	Yes	No	No
2	No	Yes	No
3	Yes	Yes	Yes
4	Yes	Yes	No
5	Yes	No	Yes

[1] Demographics, comorbidity score components, frailty score components, screening, healthcare utilization

[2] Exposure-based high dimensional propensity score selection

[3] Investigator selected confounders tailored to each example

Predefined global covariates can be applied "out of the box" without tailoring for different drug evaluations. While the global predefined covariates may be the same in each application, the coefficients to derive the estimated PS will depend on the covariate relationships with the exposure and comparator. Including covariates selected empirically via the exposure based hdPS algorithm increases computational complexity, while including tailored potential confounders increases staff time and effort. Staff time and effort may be the most difficult to scale up if there are numerous screening activities being launched for active surveillance. The candidate PS models in Table 1 were chosen to allow comparison of models that:

1. use only predefined global covariates to those that use only empirical covariates (model 1 vs 2)
2. models that include predefined global covariates with or without inclusion of covariates tailored to the exposure-comparator under investigation (models 1 vs 5)
3. models that include predefined global covariates and empirically selected covariates with or without inclusion of covariates tailored to the exposure-comparator under investigation (model 3 vs 4)

Algorithms to define independent variables in the global propensity score will be based on previously published papers that evaluated performance of the algorithm whenever available. For example, prior papers have evaluated a claims based combined comorbidity score[33] as well as a claims based frailty index.[34] When published and/or evaluated algorithms are not available, the workgroup will create algorithms based on content validity from selected code descriptions and knowledge of coding for billing purposes. Algorithms for pre-specified covariates included in the candidate global propensity scores (other than exposure based hdPS empirically identified covariates) are provided in **Appendix C**.

We chose to use hdPS rather than other machine learning to select variables such as LASSO[19] or Elastic Net[35] or hybrid approaches[36,37] that combine the two to further reduce the dimensionality of variables included. In the context of variable selection based on potential for bias when evaluating a single outcome, these empirical variable selection approaches have performed similarly in prior evaluations. However, we will be scanning across thousands of potential outcomes. It would not be feasible to apply a machine learning or hybrid approach which selects variables based on association with outcome. Furthermore, it may be helpful in our scanning context to include a slightly broader base of variables to provide proxy adjustment for confounders on a wider range of outcomes.

The hdPS software program creates and selects baseline covariates using diagnosis and procedure codes occurring within a user defined covariate assessment window.[18] The user can provide the data as dimensions representing different aspects of care (e.g. inpatient diagnoses, outpatient diagnoses, procedure codes, drug codes). The algorithm will identify the top n (default 200) most prevalent codes from each dimension and create binary variables for each, as well as a variable measuring frequency of occurrence (once, sporadically, frequently).[38] The exposure based selection option will select k covariates for inclusion (default = 500) based on strength of the association with exposure.

Exposure-based selection is not the default hdPS option.[39] The default selection criterion is based on potential to bias an exposure-outcome relationship, as defined by the Bross formula.[40] The bias based selection option is not appropriate for this signal detection context, because the Bross formula is based on relationships between exposure and covariates with a single outcome. Because we will be screening across thousands of potential outcomes simultaneously rather than focusing on one outcome of interest, we focus on empirically identifying covariates using selection based on relationships with exposure. This may result in inclusion of covariates that are strongly predictive of exposure but not

necessarily risk factors for all evaluated outcomes. However, on the whole we expect that there will be little harm to including covariates that are not strong risk factors for a particular outcome relative to the anticipated reduction in confounding for other outcomes.[38,41]

For Aim 2, the pre-specified primary candidate global propensity score to adjust for confounding will include demographic, comorbidity score, frailty, screening and exposure based hdPS selected variables (**#3 from** Table 1). This PS was selected as primary because it includes all covariates that do not require additional investigator input, making it easier to scale up screening activities. Further investigation of potential alerts could then use PS tailored to refine understanding for specific outcomes of interest.

V. TREESCAN

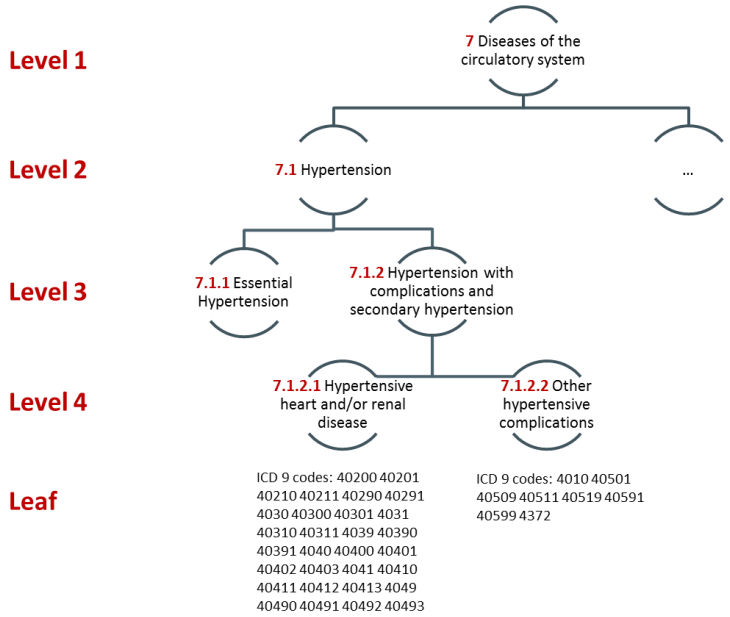
A. HIERARCHICAL TREE (MULTI-LEVEL CLINICAL CLASSIFICATION SOFTWARE)

The tree structure that we will use for TreeScan will be based on the MLCCS (https://www.hcup-us.ahrq.gov/toolssoftware/ccs/CCSUsersGuide.pdf). The MLCCS tree we will use includes International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes and is grouped into 4 hierarchical levels representing increasingly specific clinical concepts. At the top level, there are 18 categories representing different body systems.

The increasing specificity of the hierarchical levels is depicted in

Figure 1, where at the top level the category specifies only that the person has a disease of the circulatory system. Moving to level 2, one might see that the patient has hypertension. At level 3, essential versus secondary hypertension or hypertension with complications can be differentiated. At level 4, the type of secondary or hypertensive complications are specified. Finally, each node in level 4 of the hierarchy is based on specific ICD-9-CM codes.

Figure 1. Example from Multi-Level Clinical Classification



Decimal points in ICD-9-CM codes at the leaf level have been removed.

The MLCCS based tree that we will use is a curated tree based on the 2015 version of ICD-9-CM codes. The tree has been independently curated by 2 members of FDA (**Appendix A**). Disagreements were adjudicated after discussion between the curators. Curation of the tree involved removal of conditions that were 1) congenital/hereditary, 2) unlikely to be caused by drugs (e.g. pregnancy, flu, well-visits) as well as 3) conditions with long induction times such as cancer (details on curation in the appendix). The diagnostic codes and their classification into different levels are not based on validated algorithms and could misclassify outcomes. Nevertheless, MLCCS classification of ICD-9-CM codes into clinical concepts can be useful as part of a screening tool for potential adverse events, followed by more rigorous and targeted protocol-based investigations. While an ICD-10-CM version of the MLCCS tree is available, our case studies will be conducted using data prior to widespread use of ICD-10-CM codes in the United States. We may elect to use trees that have been further curated for specific examples.

B. DEFINING INCIDENT OUTCOMES

We will define incident outcomes based on level 3 nodes across the MLCCS tree hierarchy. Incident outcomes will be defined by the first diagnosis from the node that occurs in the emergency department (ED) or inpatient (IP) setting; without any diagnoses in the same MLCCS level 3 node in the prior 183 days in any care settings. Multiple incident outcomes may be contributed by each patient as long as they meet the incidence criteria at MLCCS level 3 nodes.

Each patient will be allowed to enter the cohort only one time, after the first qualifying incident use of either the exposure or comparator of interest. Patients will be censored at death, disenrollment, or maximum days follow up for the example. If one member of a 1:1 propensity score matched set is censored, the other member will also be censored at the same time. Incident outcomes occurring during the patient's follow up after treatment initiation will be included in outcome counts for TreeScan.

C. UNCONDITIONAL BINOMIAL TREE SCAN STATISTIC

We will use the unconditional Bernoulli version of the tree-based scan statistic. This statistic conditions on the number of cases in the node but does not fix the total number of cases across the tree for each exposure group to be the same in the observed and randomly permuted data. The threshold for alerting in both Aims 1 and 2 will be $p \leq 0.01$ (1-sided).

The distribution of the test statistic *T* below is unknown. However, a Monte Carlo based p-value can be obtained by generating random datasets under the null hypothesis that every outcome occurs, independently of other outcomes, with the same probability among in the treatment group versus the comparator group.[4]

The log likelihood ratio (LLR) based test statistic *T* can be calculated as:

$$LLR(G) = ln\left(\frac{\left(\frac{c_G}{c_G + n_G}\right)^{c_G}\left(\frac{n_G}{c_G + n_G}\right)^{n_G}}{(p)^{c_G}(1-p)^{n_G}}\right) I\left(\frac{c_G}{c_G + n_G} > p\right)$$

$$T = \max_G LLR(G)$$

Where: T = unconditional Bernoulli tree scan statistic
c_G = cases in the treatment group for a given node G
n_G = cases in the reference group for a given node G

p = probability of being in the treatment group (for 1:1 matched this is 0.5)
G = node of interest

Random datasets can be generated under the null hypothesis by creating replicates of the original data where each node contains the same number of events as observed in the original data, however the events within each node are assigned to exposure based on a binomial draw with the expected proportion based on the null hypothesis. In our 1:1 matched setting, this proportion is 0.5. For these Monte Carlo generated data sets, the outcomes in each node are assigned randomly. If 9,999 random replicates are generated, and ranked according to *T,* then the Monte Carlo based p-value = Rank of the observed data/(9999+1). When the type 1 error alpha for alerting is set to a threshold of 0.01, then only nodes with rank within the top 1% of the real and random replicates will constitute a statistical alert. If the null hypothesis is true, the probability that all p-values are larger than 0.01 is 99%.

Hypothesis tests will be performed at level 3 and all more finely specified nodes at numerically higher levels of the MLCCS, including level 4 and the leaf level (specific ICD-9-CM codes). A LLR will be computed at every node where a hypothesis test is performed.

Some of the major strengths of tree-based scan statistics include:

i. They were developed based on scan statistical theory
ii. They use a hierarchical diagnosis tree to simultaneously evaluate outcomes at different levels of granularity (including specific diagnoses and groups of related diagnoses)
iii. They use a frequentist method to formally adjust for the multiple testing inherent in evaluation of thousands of potential adverse events that accounts for correlation between tests of related hypotheses (unlike traditional frequentist methods which are too conservative)
iv. They can be useful when screening for unanticipated safety signals where there is no informative prior

Some of the major limitations of tree-based scan statistics include:

v. Bias is adjusted by design, not inherent in the scan statistic
vi. The hierarchical classification system for outcomes used are generally not based on validated algorithms
vii. Adjusting for multiplicity when scanning across outcomes will decrease power compared to evaluating a single pre-specified hypothesis

D. SELECTED OUTCOMES AND ALERTS TO EVALUATE

Realistically, most of the potential outcomes scanned will be unrelated to the evaluated drugs. We will focus on doing a deeper dive for evaluate a subset of outcomes with and without alerts in a variety of examples where we have prior knowledge of where true signals may or may not be present. By focusing on areas where there are known signals or unanticipated alerts, we target high yield areas for learning about the method and its performance. Our case studies are intended to be diverse with respect to study populations as well as types of outcomes with true signals. That said, the performance of global PS in these case studies will not necessarily be generalizable to all contexts.

For Aim 1, we will focus on a subset of up to 5 outcomes with and up to 5 outcomes without alerts for each example. These outcomes will be chosen based on clinical knowledge of the safety profiles for the selected examples and may include outcomes where elevation in risk with exposure relative to the comparator is anticipated, where the risk is expected to be no different for exposure and comparator or where the effect of exposure is unknown (**Figure 2**). These outcomes may be the specific MLCCS nodes that signaled, or a validated algorithm may be used to capture an underlying clinical concept for related nodes. These follow on evaluations will provide insight regarding the relative performance of the candidate global propensity scores.

Figure 2. Sampling specific outcomes to evaluate relative performance of candidate generic propensity scores

		True effect of exposure		
		Known effect	Known null	Unknown ?
	Alert	True positive	False positive	True/False positive?
	No alert	False negative	True negative	True/False negative?

For Aim 2, all alerts from the pre-specified primary global propensity score will be reviewed by the workgroup.

VI. DRUG SAFETY SIGNAL DETECTION CASE STUDIES

A. DATA

The new initiator cohorts for each example will be created from a commercial insurance claims data licensed and housed by the Brigham and Women's Hospital (BWH), Truven Health MarketScan® Research Database. No use of the SDD is planned for this project. While MarketScan is not part of the SDD, there is a small, but unknown amount of overlap between Optum (which is a part of the SDD) and MarketScan data.

B. PERFORMANCE METRICS

Performance metrics will include measures of balance on known risk factors for a sample of outcomes in the tree (absolute and standardized differences for individual covariates, average standardized absolute mean difference (ASAMD) across all covariates) and evaluation of how additional adjustment for known risk factors of those outcomes affects alerting. If some candidate global PS tend to have greater imbalance on known risk factors for sampled outcomes, this would suggest that they are missing important dimensions that should be included in a global confounding adjustment score. We will describe what happens to alerts when we do or do not include specific, known risk factors for a selected sample of outcomes in the PS.

In each case study, we have known signals that have previously been identified. As in a real surveillance activity, we may not necessarily have adequate power to find signals of interest at stringent pre-specified alpha levels. However, we will be looking at patterns of alerting in these examples to observe how signal detection using the method could play out in a real scenario. Outcomes that don't alert at the pre-specified threshold may still have relatively low likelihood under the null. The method can play an important part in screening and prioritization even if there is not sufficient power to alert at a pre-

specified threshold by painting a clinical picture of the pattern of outcomes that are unlikely to be observed if there was no relationship with exposure.

C. SELECTED EXAMPLES

In consultation with the FDA, the work group chose four case studies to evaluate. We describe these case studies here in brief. Additional details of the specifications for these case studies can be found in **Appendix C**. We deliberately chose older examples with known safety profiles. Given the delays in data refreshes, we have limited years of data available after Oct 2015. Although an ICD10 based tree is available, if we focused only on the ICD-10 era, we would have lower power to detect known signals in our examples. Running hdPS in a mixed ICD9-10 era would require incorporation of forward and backward mapping and is beyond the scope of this project.

We will use Sentinel's routine query tools [22,23] supplemented with *de novo* coding and other software as necessary to extract new initiator cohorts and counts of outcomes across the hierarchical tree. After matching on each candidate global propensity score, the 1:1 matched cohorts will be used as inputs for TreeScan to run scan statistics.

Example 1: New initiators of macrolides versus fluoroquinolones for community acquired pneumonia (a between class comparison)

Macrolides and fluoroquinolones are two classes of antibiotic drugs that are used to treat community acquired pneumonia. There is no clear evidence of a strong benefit or safety risks for one class versus the other. A meta-analysis and review of randomized clinical trials that primarily recruited mild to moderate outpatient cases of community acquired pneumonia reported that macrolides had more gastrointestinal side effects (e.g. diarrhea, nausea, vomiting) and rate of treatment failure (as defined by persistent signs and symptoms of pneumonia requiring treatment modification).[42]

We will identify patients who newly initiate a macrolide or fluoroquinolone after at least 183 days without dispensations of either class. Patients must be enrolled in the database with medical and drug coverage for at least 183 days prior to the initiation date (index date for cohort entry) and have diagnoses for community acquired pneumonia in an outpatient setting as well as codes for chest radiography within 14 days prior to or on the date of initiation (see design diagram in appendix). This algorithm for outpatient community acquired pneumonia is adapted from one developed and validated using Group Health data, which had a positive predictive value of around 71%.[43,44] We will exclude patients who initiate both a macrolide and a fluoroquinolone on the same day, are younger than 18 or over 65 on the date of initiation, or were hospitalized for any reason within 90 days prior to or on the date of initiation of macrolide or fluoroquinolone.

When defining covariates for candidate global propensity scores (see **Table 1** above), we will use diagnosis and procedure codes from claims occurring in the 183 days prior to but not including the date of initiation. We will exclude the index date from the covariate assessment window and follow up window because we cannot distinguish the timing of exposure and health events on the same day (e.g. could be a baseline condition or an outcome). The propensity scores that use investigator specified covariates tailored to the example will include variables such as those listed in **Table 2**.

The study period will include Jan 1, 2003to Sep 30, 2015. The follow up window when scanning for potential adverse events will begin one day after initiation. The maximum follow up will be 30 days after initiation. Follow up will be censored at Sep 30, 2015 due to transition to ICD-10-CM coding. Patients in

matched sets will additionally be censored at death or disenrollment of either member of the set or switching of exposure (defined by dispensation for a drug in the other group). Because there are not *a priori* anticipated safety signals, our evaluation will include selection of up to 5 incident outcome nodes that do and up to 5 nodes that do not alert at the 0.01 threshold. We will select nodes based on our ability to identify known risk factors for those outcomes.

Example 2: New initiators of azithromycin versus clarithromycin for community acquired pneumonia (a within class comparison)

Azithromycin versus clarithromycin are macrolide antibiotics that are used to treat community acquired pneumonia. Prior trials have found them to be similar in terms of both efficacy and safety.[45,46] The specification for this cohort will be identical to that for the between class comparison of macrolides and fluoroquinolone, with the exceptions that initiation will be defined with respect to macrolides as a class and co-prescription or prior use of fluoroquinolones will be included as indication related covariates.

Example 3: New initiators of meloxicam versus celecoxib for osteoarthritis (two nonsteroidal anti-inflammatory drug (NSAID) options)

Meloxicam and celecoxib are two non-steroidal anti-inflammatory drugs that are used to treat osteoarthritis pain. Although potential effect sizes are small, prior research has suggested there could be more gastrointestinal adverse events[47] and fewer vascular and renal events[48,49] with meloxicam versus celecoxib.

We will identify patients who newly initiate meloxicam or celecoxib after at least 183 days without dispensations of any non-steroidal anti-inflammatory drug (see design diagram in appendix). Patients must be enrolled in the database with medical and drug coverage for at least 183 days prior to the initiation date (index date for cohort entry) and have at least one diagnosis for osteoarthritis (ICD-9 715*) in an inpatient or outpatient setting during this time. We will exclude patients who initiate both meloxicam and celecoxib on the same day or are younger than 18 on the date of initiation.

When defining covariates for candidate global propensity scores (**Table 1**), we will use diagnosis and procedure codes from claims occurring in the 183 days prior to but not including the date of initiation. In addition to global demographic, comorbidity, frailty, screening and healthcare utilization covariates, we will consider numerous investigator specified covariates for this example such as those listed in **Table 2**.

The study period will include Jan 1, 2003 to Sep 30, 2015. When scanning for potential adverse events, follow up will begin one day after initiation. The maximum follow up will be 183 days after initiation. Follow up will be censored at Sep 30, 2015 due to transition to ICD-10-CM coding. Patients in matched sets will additionally be censored at death or disenrollment for either member of the set, or switching of exposure (defined by dispensation for a drug in the other group). The outcomes that we select for evaluation of balance on known risk factors will include symptomatic upper GI events such as acid/peptic ulcer related, upper GI conditions such as perforations and bleeding, as well as myocardial infarction and cerebrovascular events. Candidate global PS that only includes un-tailored pre-specified covariates will not include the indication related risk factors listed above, but may adjust for them by proxy via correlation with the variables that are included. We will evaluate balance in nodes where alerts are anticipated. Our evaluation will also include selection of up to 5 incident outcome nodes that do and up to 5 nodes that do not alert at the 0.01 threshold. We will select nodes based on our ability to identify known risk factors for those outcomes.

Example 4: New initiators of valproic acid and lamotrigine for epilepsy

Valproic acid and lamotrigine have been used for decades as therapies for treatment of epilepsy.[50-52] There is not a clear consensus regarding superiority in effectiveness for any of the numerous currently available therapies for epilepsy, each of which has different known adverse effects and risk profiles.[53] For example, valproic acid has black box warnings for hepatotoxicity, teratogenicity, and pancreatitis[54] whereas lamotrigine can cause serious rashes (Stevens-Johnson syndrome) within 8 weeks of initiation[52] and serious immune system reactions (hemophagocytic lymphohistiocytosis (HLH))[55].

We will identify patients who newly initiate valproic acid or lamotrigine after at least 183 days without dispensations of drugs used for seizures (e.g. valproic acid, lamotrigine, topiramate, carbamazepine, phenytoin, oxcarbazepine, ethusuximide, zonisamide, phenytoind, phenobarbital, primidone, felbamate, levetiracetam, tiagabine, zonisamide, gabapentin, esclicarbazepine, vigabatrin, lacosamide, pregabalin, pregabalin, rufinamide, clonazepam). Patients must be enrolled in the database with medical and drug coverage for at least 183 days prior to the initiation date (index date for cohort entry) and have at least one diagnosis for epilepsy (ICD9: 345.xx) and one diagnosis code for convulsions (ICD9: 779.0, 780.3, and 780.39) in an inpatient or outpatient setting during this time (algorithm adapted from Wilner et al ENREF 53[56]). We will exclude patients who initiate both valproic acid and lamotrigine on the same day, are younger than 2 or older than 65 on the date of initiation, or have claims indicative of traumatic brain injury or stroke during the 183 days prior to the index date (see design diagram in appendix). We will conduct separate analyses for subgroups of patients aged 2 to 18 years and patients 19-65 years old at the time of initiation.

The study period will include Jan 1, 2003 to Sep 30, 2015. When scanning for potential adverse events, follow up will begin one day after initiation. The maximum follow up will be 183 days after initiation. Follow up will be censored at Sep 30, 2015 due to transition to ICD-10-CM coding. Patients in matched sets will additionally be censored at death or disenrollment for either member of the set, or switching of exposure (defined by dispensation for a drug in the other group). The investigator specified covariates tailored to this example may include risk factors for known adverse events such as those listed in **Table 2**. In addition to evaluation of balance in nodes where alerts are anticipated, our evaluation will include selection of up to 5 incident outcome nodes that do and up to 5 nodes that do not alert at the 0.01 threshold. We will select nodes based on our ability to identify known risk factors for those outcomes.

Table 2. Known adverse events, risk factors and investigator defined covariates for case studies

Example	Known adverse events	Investigator defined measurable risk factors for known adverse events	Other investigator defined covariates
1 Macrolides and fluoroquinolones (between class comparison)	• Diarrhea • Nausea • Vomiting • Treatment failure	n/a	• co-prescription of beta-lactam with macrolide or fluoroquinolone • prior prescription of other antibiotic classes: penicillins, cephalosporins, sulfonamides, tetracyclines, animoglycosides • pregnancy at time of initiation (class B vs class C drug)
2 Azithromycin and clarithromycin (within class comparison)	n/a	n/a	same as above
3 Meloxicam and celecoxib	• Upper GI events (e.g. acid/peptic ulcer, perforations, bleeding) • Myocardial infarction • Cerebrovascular events • Renal failure, acute kidney injury, hyperkalemia	• Anticoagulants • Obesity • Smoking • Angina • Coronary Revascularization • Statins • Hormone Replacement Therapy • Prior Nonselective Nsaid Use • Stroke/Tia • History Of Peptic Ulcer Disease Or Gi Bleeding • Antiplatelets	• Potential drug interactions: ○ Antidepressants ○ Fluconazole ○ Lithium ○ blood pressure medications ○ cyclosporine ○ methotrexate ○ steroids • Potential contraindications: ○ pregnancy at the time of initiation • Potential other indications: ○ Fibromyalgia ○ rheumatoid arthritis

Example	Known adverse events	Investigator defined measurable risk factors for known adverse events	Other investigator defined covariates
4 Valproic acid and lamotrigine	• hepatotoxicity • teratogenicity (spina bifida, atrial septal defect, cleft palate, hypospadias, polydactyly) • pancreatitis • serious rashes (Stevens-Johnson syndrome) • serious immune reactions (hemophagocytic lymphohistiocytosis - HLH)	• HIV infection • Other viral infections (mumps, flu, herpes, Coxsackie, Epstein-Barr, cytomegalovirus, parvovirus B19, pneumocystosis and histoplasmosis) • Bacterial infection • Organ transplant • Autoimmune diseases (e.g. Lupus, type 1 diabetes, rheumatoid arthritis, psoriatic arthritis, multiple sclerosis, inflammatory bowel disease, Addisons disease, Grave's disease, Hashimoto's thyroiditis, myasthenia gravis, vasculitis, celiac disease) • Chemotherapy • Cancer • Acute b-lymphoblastic leukemia • Medicines for gout (e.g. Allopurinol) • Sulfa antibiotics (e.g. Bactrim, septra), • Sertraline • "oxicam" anti-inflammatory drugs (e.g. Meloxicam, piroxicam) • Alcohol use disorders • Liver disease • Gallstones • Cystic fibrosis • Kawasaki disease • Reye's syndrome	• Potential other indications: ○ migraine, bipolar disorder

Example	Known adverse events	Investigator defined measurable risk factors for known adverse events	Other investigator defined covariates
		• Hemolytic uremic syndrome (HUS) • Thrombotic thrombocytopenic purpura (TTP) • Hyperparathyroidism • Obstructions in the biliary system • Peptic ulcer • Azathioprine • 6-mercaptopurine (e.g., Imuran®) • Diuretics • Didanosine • Estrogens • Pentamidine • Tetracycline	

VII. REFERENCES

1. Nelson JC, Shortreed SM, Yu O, et al. Integrating database knowledge and epidemiological design to improve the implementation of data mining methods that evaluate vaccine safety in large healthcare databases. Statistical Analysis and Data Mining: The ASA Data Science Journal 2014;7:337-51.
2. TreeScan User Guide v1.3. (Accessed June 12, 2017, at www.treescan.org.)
3. Yih WK, Maro JC, Nguyen M, et al. Assessment of Quadrivalent Human Papillomavirus Vaccine Safety Using the Self-Controlled Tree-Temporal Scan Statistic Signal-Detection Method in the Sentinel System. Am J Epidemiol 2018;187:1269-76.
4. TreeScan Power (PRISM). US Food and Drug Administration (FDA). (Accessed June 26, 2017, at https://www.sentinelinitiative.org/sentinel/methods/336.)
5. Pilot Of Self-Controlled Tree-Temporal Scan Analysis For Gardasil Vaccine. US Food and Drug Administration (FDA). (Accessed June 12, 2017, at https://www.sentinelinitiative.org/sentinel/methods/339.)
6. The childhood immunization schedule and safety: Stakeholders concerns, scientific evidence, and future studies. Washington, DC: The National Academies Press 2013.
7. Taxonomy for monitoring methods within a medical product safety surveillance system: Report of the Mini-Sentinel Taxonomy Project Work Group. 2010.
8. Fine PE, Chen RT. Confounding in studies of adverse reactions to vaccines. Am J Epidemiol 1992;136:121-35.
9. Glynn RJ, Knight EL, Levin R, Avorn J. Paradoxical relations of drug treatment with mortality in older persons. Epidemiology 2001;12:682-9.
10. Suissa S. The case-time-control design. Epidemiology 1995;6:248-53.
11. Wang S, Linkletter C, Maclure M, et al. Future cases as present controls to adjust for exposure trend bias in case-only studies. Epidemiology 2011;22:568-74.
12. Wang SV, Gagne JJ, Glynn RJ, Schneeweiss S. Case-crossover Studies of Therapeutics: Design Approaches to Addressing Time-varying Prognosis in Elderly Populations. Epidemiology 2013;24:375-8.
13. Horwitz RI, Feinstein AR. The problem of "protopathic bias" in case-control studies. Am J Med 1980;68:255-8.

14. Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. Pharmacoepidemiol Drug Saf 2010;19:858-68.
15. Ray WA. Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. American journal of epidemiology 2003;158:915-20.
16. Brookhart MA, Wyss R, Layton JB, Sturmer T. Propensity score methods for confounding control in nonexperimental research. Circ Cardiovasc Qual Outcomes 2013;6:604-11.
17. Schneeweiss S, Eddings W, Glynn RJ, Patorno E, Rassen J, Franklin JM. Variable Selection for Confounding Adjustment in High-dimensional Covariate Spaces When Analyzing Healthcare Databases. Epidemiology 2017;28:237-48.
18. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology 2009;20:512-22.
19. Tibshirani R. The Lasso Method for Variable Selection in the Cox Model. Statistics in Medicine 1997;16.
20. Wyss R, Ellis AR, Brookhart MA, et al. The role of prediction modeling in propensity score estimation: an evaluation of logistic regression, bCART, and the covariate-balancing propensity score. Am J Epidemiol 2014;180:645-55.
21. Ju CC, M; Lendle, SD; Franklin, JM; Wyss, R; Schneeweiss, S; van der Laan, MJ. Propensity score prediction for electronic healthcare databases using Super Learner and High-dimensional Propensity Score Methods. Cornell University Library; 2017.
22. Routine Querying Tools (Modular Programs). 2014. 2016, at http://mini-sentinel.org/data_activities/modular_programs/details.aspx?ID=166.)
23. Gagne JW, SV; Schneeweiss, S. . FDA Mini-Sentinel Prospective Routine Observational Monitoring Program Tool: Cohort Matching. Technical Users' Guide version: 1.0 . January 2014.
24. Cole D, Kulldorff M, Baker M, et al. Infrastructure for Evaluation of Statistical Alerts Arising from Vaccine Safety Data Mining Activities in Mini-Sentinel. 2016 July 22, 2016.
25. Gagne JJ HX, Hennessy S, Leonard CE, Chrischilles EA, Carnahan RM, Wang SV, Fuller C, Iyer A, Katcoff H, Woodworth TS, Archdeacon P, Meyer TE, Schneeweiss S, Toh S. . Successful comparison of US Food and Drug Administration Sentinel analysis tools to traditional pharmacoepidemiologic approaches. . Clinical Pharmacology and Therapeutics

26. Mini-Sentinel Product Assessment: A Protocol for Assessment of Dabigatran Version 3. Food and Drug Administration, Sentinel Program., 2015. (Accessed June 12, 2017, at http://www.mini-sentinel.org/work_products/Assessments/Mini-Sentinel_Protocol-for-Assessment-of-Dabigatran.pdf.)
27. Mini-Sentinel Prospective Routine Observational Monitoring Program Tools (PROMPT): Rivaroxaban Surveillance Version 3. Food and Drug Administration, Sentinel Program, 2015. (Accessed June 12, 2017, at http://www.mini-sentinel.org/work_products/Assessments/Mini-Sentinel_PROMPT_Rivaroxaban-Surveillance-Plan.pdf.)
28. Toh S, Reichman ME, Houstoun M, et al. Comparative risk for angioedema associated with the use of drugs that target the renin-angiotensin-aldosterone system. Arch Intern Med 2012;172:1582-9.
29. Leonard CR, ME; Toh, D; Kulldorff, M; Nelson, JC; Gagne, JJ; Ouellet-Hellstrom, RP; Moeny, DG; Mott, KA; By, K; Wang, SV; Hennessy, S. Mini-Sentinel Prospective Surveillance Plan: Prospective Routine Observational Monitoring of Mirabegron
30. Toh S, Avorn J, D'Agostino RB, Sr., et al. Re-using Mini-Sentinel data following rapid assessments of potential safety signals via modular analytic programs. Pharmacoepidemiol Drug Saf 2013;22:1036-45.
31. Taxonomy for monitoring methods within a medical product safety surveillance system: year two report of the Mini-Sentinel Taxonomy Project Workgroup. 2012. (Accessed June 12, 2017, at http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Taxonomy-Year-2-Report.pdf.)
32. Routine Querying Tools (Modular Programs). 2014. (Accessed June 12, 2017, at http://mini-sentinel.org/data_activities/modular_programs/details.aspx?ID=166.)
33. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol 2011;64:749-59.
34. Kim DH, Schneeweiss S. Measuring frailty using claims data for pharmacoepidemiologic studies of mortality in older adults: evidence and recommendations. Pharmacoepidemiol Drug Saf 2014;23:891-901.
35. Zhou T, Tao D, Wu X. Manifold elastic net: a unified framework for sparse dimension reduction. Data Mining and Knowledge Discovery 2011;22:340-71.

36. Franklin JM, Eddings W, Glynn RJ, Schneeweiss S. Regularized Regression Versus the High-Dimensional Propensity Score for Confounding Adjustment in Secondary Database Analyses. Am J Epidemiol 2015;182:651-9.
37. Karim ME, Pang M, Platt RW. Can We Train Machine Learning Methods to Outperform the High-dimensional Propensity Score Algorithm? Epidemiology 2018;29:191-8.
38. Myers JA, Rassen JA, Gagne JJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. American journal of epidemiology 2011;174:1213-22.
39. Using the Pharmacoepidemiology Toolbox in SAS. (Accessed June 12, 2017, at http://www.drugepi.org/wp-content/uploads/2013/10/Using_the_Pharmacoepi_Toolbox_in_SAS_2.4.15.pdf.)
40. Bross ID. Spurious effects from an extraneous variable. J Chronic Dis 1966;19:637-47.
41. Myers JA, Rassen JA, Gagne JJ, et al. Myers et al. Respond to "Understanding Bias Amplification". American journal of epidemiology 2011;174:1228-9.
42. Skalsky K, Yahav D, Lador A, Eliakim-Raz N, Leibovici L, Paul M. Macrolides vs. quinolones for community-acquired pneumonia: meta-analysis of randomized controlled trials. Clinical Microbiology and Infection 2013;19:370-8.
43. Jackson LA, Neuzil KM, Yu O, et al. Effectiveness of Pneumococcal Polysaccharide Vaccine in Older Adults. New England Journal of Medicine 2003;348:1747-55.
44. Jackson ML, Neuzil KM, Thompson WW, et al. The Burden of Community-Acquired Pneumonia in Seniors: Results of a Population-Based Study. Clinical Infectious Diseases 2004;39:1642-50.
45. Drehobl MA, De Salvo MC, Lewis DE, Breen JD. Single-Dose Azithromycin Microspheres vs Clarithromycin Extended Release for the Treatment of Mild-to-Moderate Community-Acquired Pneumonia in Adults. CHEST 2005;128:2230-7.
46. O'Doherty B, Muller O. Randomized, Multicentre Study of the Efficacy and Tolerance of Azithromycin versus Clarithromycin in the Treatment of Adults with Mild to Moderate Community-Acquired Pneumonia. European Journal of Clinical Microbiology and Infectious Diseases 1998;17:828-33.
47. Layton D, Hughes K, Harris S, Shakir SA. Comparison of the incidence rates of selected gastrointestinal events reported for patients prescribed celecoxib and

meloxicam in general practice in England using prescription-event monitoring (PEM) data. Rheumatology (Oxford) 2003;42:1332-41.
48. Asghar W, Jamali F. The effect of COX-2-selective meloxicam on the myocardial, vascular and renal risks: a systematic review. Inflammopharmacology 2015;23:1-16.
49. Layton D, Hughes K, Harris S, Shakir SAW. Comparison of the incidence rates of thromboembolic events reported for patients prescribed celecoxib and meloxicam in general practice in England using Prescription-Event Monitoring (PEM) data. Rheumatology 2003;42:1354-64.
50. Valproic acid and sodium valproate approved for use in epilepsy. FDA Drug Bull 1978;8:14-5.
51. Goldenberg MM. Overview of Drugs Used For Epilepsy and Seizures: Etiology, Diagnosis, and Treatment. Pharmacy and Therapeutics 2010;35:392-415.
52. Highlights of Prescribing Information. (Accessed July 6, 2018, at https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/022251,020764s029,020241s036lbl.pdf.)
53. Karceski SG, P; Dashe, JF. Initial treatment of epilepsy in adults. UpToDate. Jun 2018 ed.
54. Depakote Tablets (divalproex sodium). (Accessed July 6, 2018, at https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/018723s039lbl.pdf.)
55. Lamictal (lamotrigine): Drug Safety Communication - Serious Immune System Reaction. (Accessed 7/6/2018, at https://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm605628.htm.)
56. Wilner AN, Sharma BK, Soucy A, Thompson A, Krueger A. Common comorbidities in women and men with epilepsy and the relationship between number of comorbidities and health plan paid costs in 2010. Epilepsy Behav 2014;32:15-20.

VIII. APPENDICES

A. CURATION OF THE MLCCS TREE

Two members of the FDA independently reviewed the MLCCS tree based on 2015 ICD-9-CM codes. Each reviewer flagged codes that were:

1. of known etiology (e.g. pregnancy, congenital condition),
2. unlikely to be an adverse reaction caused by drugs (e.g. gingival recession, recurrent dislocation of shoulder),
3. not an incident diagnosis (e.g. alcoholism in family, social maladjustment),
4. for conditions with long latency/induction periods (e.g. cancer, osteoporosis)

After adjudication of disagreement, 7,078 of 15,075 (47%) ICD-9-CM diagnostic codes were excluded.

Source:

https://www.sentinelinitiative.org/sentinel/surveillance-tools/software-toolkits/treeextraction-documentation (supporting tree file)

B. RESPONSES TO PUBLIC COMMENT

This study was posted for public comment from August 10, 2018 to September 7, 2018. We received one set of comments and have responded to each question individually below. We have minimally rephrased the questions slightly to a more general format but the content and intent remains the same.

Question 1: Given the large number of available empirical approaches for model selection, it could be helpful to provide motivation for why high dimensional propensity score (hdPS) was chosen for this evaluation as opposed to other options. For instance, Karim et al. (Epidemiology 2018 Mar; 29(2): 191-198) recently showed that a machine learning with hdps hybrid often outperforms hdps alone.

> Response: This paper found that machine learning based approaches such as LASSO and ElasticNet in combination with hdPS performed marginally better than hdPS alone in the context of selection based on potential for bias for a single outcome. The machine learning component of the hybrid empirical variable selection methods worked to further reduce the dimensionality of variables identified with hdPS.
>
> In our context, we are scanning across thousands of potential outcomes. It would not be feasible to apply a hybrid approach which selects variables based on association with outcome. Furthermore, it may be helpful in our scanning context to include a slightly broader base of variables to provide proxy adjustment for confounders on a wider range of outcomes.
>
> We will include this citation and a brief explanation as above in the background.

Question 2: Similarly, it could be useful to motivate why the TreeScan methodology was selected as opposed to other scan statistics (or, minimally, to provide its major advantages and limitations in this specific setting).

> Response: We will list some of the major strengths and limitations of TreeScan in the protocol. A comparison of different signal detection methods is currently underway in another ongoing project entitled, “Evaluation of Three Self-Controlled Methods for Signal Detection: TreeScan, Sequence Symmetry Analysis, and Information Component

Temporal Pattern Discovery." Further details on this project can be found here on the Sentinel website.

Strengths:

i. Developed based on scan statistical theory

ii. Uses a hierarchical diagnosis tree to simultaneously evaluate outcomes at different levels of granularity (including specific diagnoses and groups of related diagnoses)

iii. Uses a frequentist method to formally adjust for the multiple testing inherent in evaluation of thousands of potential adverse events that accounts for correlation between tests of related hypotheses (unlike traditional frequentist methods which are too conservative)

iv. Useful when screening for unanticipated safety signals where there is no informative prior

Weaknesses:

v. Bias is adjusted by design, not inherent in the scan statistic

vi. Hierarchical classification system for outcomes are not based on validated algorithms

vii. Adjusting for multiplicity when scanning across outcomes will decrease power compared to evaluating a single pre-specified hypothesis

Question 3: Please explain why you chose to implement only an ICD-9-CM based outcome tree, rather than an ICD-10-CM based outcome tree. Do you believe that this choice will impact generalizability of this study?

Response: Although different hierarchical trees may have different properties, the method of TreeScan with PS-matching is not tied to a particular coding system and should be extensible.

We deliberately chose older examples with known safety profiles. Given the delays in data refreshes, we have limited years of data available after Oct 2015. Although an ICD10 based tree is available, if we focused only on the ICD-10 era, we would have lower power to detect known signals in our examples. Using high dimensional propensity score methods in a mixed ICD9-10 era would require incorporation of mapping and is beyond the scope of this project.

Question 4: It could be helpful to explain why you chose to focus your Aim 1 evaluation on a subset of outcomes (feasibility?) and whether/how this may limit the generalizability of the evaluation of method performance. For instance, why would we expect performance metrics for the selected set of outcomes to carry-over to the other 100's of outcomes one might evaluate? Are there any performance metrics that could be feasibly evaluated for 'all' outcomes to avoid issues of selection and uncertain generalizability?

Response: The general performance metric of balance on predefined and empirically selected covariates used in the PS will apply to all outcomes. However, the covariates in the PS will not necessarily be relevant or optimal for all outcomes in all drug comparison

scenarios. It would be infeasible to identify the best set of covariates for each outcome and evaluate balance on each.

That is why we will be focusing our deeper dive on balance for known risk factors on a subset of outcomes with and without alerts in a variety of examples where we have prior knowledge of where true signals may or may not be present. These examples are intended to be diverse with respect to study populations as well as types of outcomes with true signals. That said, the performance in these examples will not necessarily be generalizable to all contexts.

Realistically, most of the potential outcomes scanned will be unrelated to the evaluated drugs. By focusing on areas where there are known signals or unanticipated alerts, we target high yield areas for learning about the method and its performance.

We will include discussion to that effect in the protocol.

Question 5: How will you ensure that you have identified example scenarios with adequate power to find signals of interest?

Response: As in a real surveillance activity, we may not necessarily have adequate power to find signals of interest at stringent pre-specified alpha levels. In each of our example scenarios, we have known signals that were previously identified. We will be looking at patterns of alerting in these examples to observe how signal detection using the method could play out in a real scenario. Outcomes that don't alert at the pre-specified threshold may still have relatively low likelihood under the null. The method can play an important part in screening and prioritization even if there is not sufficient power to alert at a pre-specified threshold by painting a clinical picture of the pattern of outcomes that are unlikely to be observed if there was no relationship with exposure.

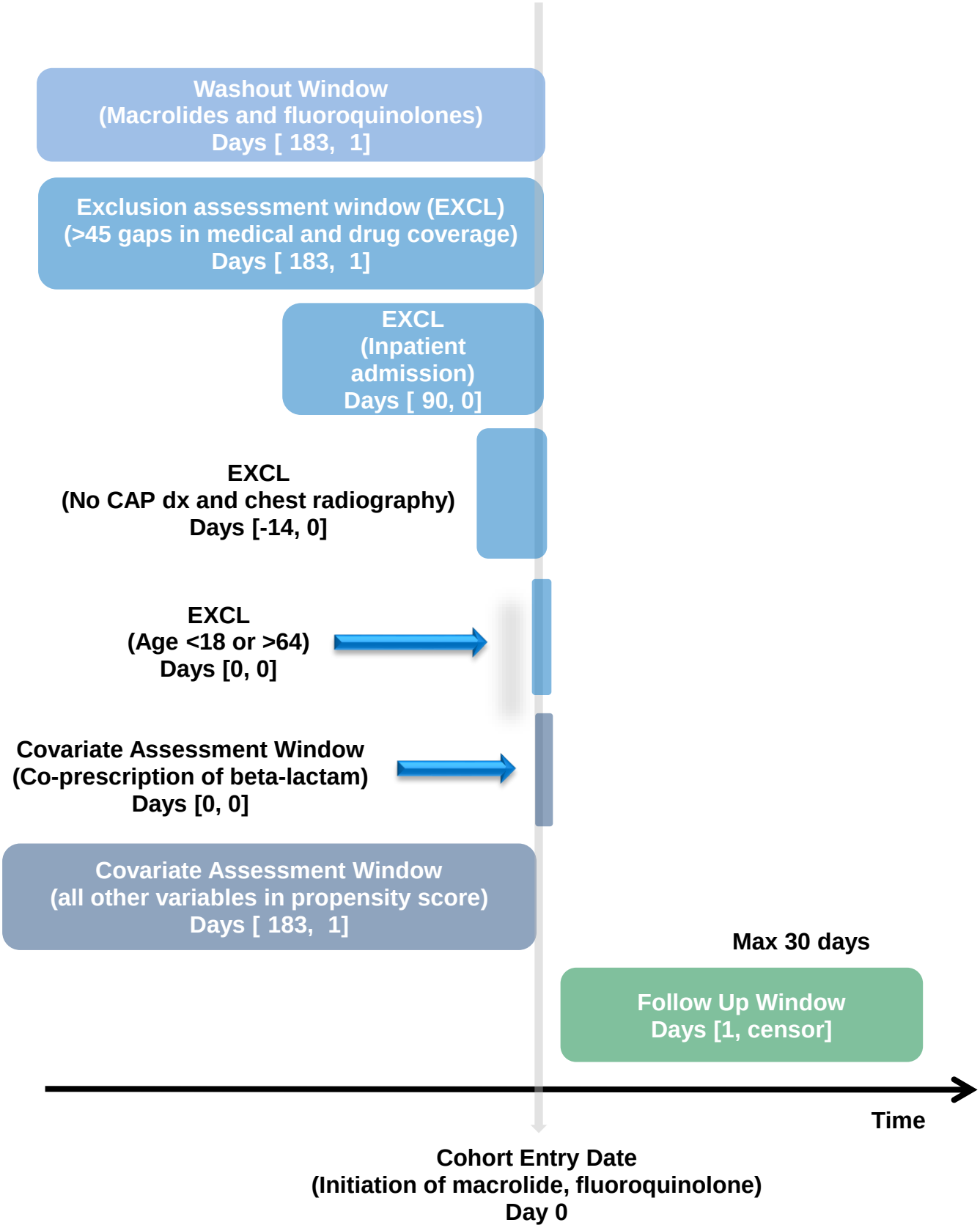
Question 6: Is there an existing 'standard' data mining method with which you can compare your new global PS-based options? This may help give some evidence about the level of improvement your methods may provide beyond what a more basic approach that a researcher might do typically in practice now.

Response: There does not appear to be a 'standard' cohort based approach. The workgroup will be comparing three signal detection methods that use a self-controlled design, which are more frequently used in pharmacovigilance activities.

C. DETAILED PROTOCOL SPECIFICATIONS FOR EXAMPLES

DESIGN DIAGRAM

Example 1 Macrolides vs fluoroquinolones for community acquired pneumonia (CAP)



Specifications for protocol: *Example 1 Macrolides vs fluoroquinolones for community acquired pneumonia*

Study Period: January 1, 2001 - September 30, 2015
Data Source: Truven MarketScan Commercial and Medicare Supplemental
Data Extraction Date/Version: January 1, 2018
Data sampling/extraction criteria: All enrollees in data source between January, 2001 - September 30, 2015
Type of data: Commercial claims
Data linkage: None
Data conversion: Sentinel Common Data Model
Software to create study population: SAS 9.4, CIDA, TreeExtraction, Pharmacoepi Toolbox hdPS macro (http://www.drugepi.org/dope-downloads/), Aetion Platform

Study Entry Criteria (Index Date)	Description	Number of entries	Type of entry	Washout window	Incident w.r.t.	Index date (day 0)	
Exposure	Macrolide	Single	Incident	[-183, 0]	Macrolide and fluoroquinolone	Date of incident dispensation	
Comparator	Fluoroquinolone	Single	Incident	[-183, 0]	Macrolide and fluoroquinolone	Date of incident dispensation	
Exclusion Criteria	**Description**		**Order of application**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Applied to:**
Enrollment/coverage							
	Medical and drug coverage		Before selection of index date	[-183, 0]	n/a	n/a	Exposure, comparator
Max. enrollment gap allowed	45 days						
Demographics							
	Age 18-64 yrs		Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Male, female		Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Race (any)		Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Drug dispensation	Days supply > 0		Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Pre-Existing Condition							
	Do not have recent community acquired pneumonia diagnosis and chest radiogr		Before selection of index date	[-14, 0]	Any	n/a	Exposure, comparator
Other	Initiate macrolide and fluoroquinolone on the same day		Before selection of index date	[0, 0]	Any	n/a	Exposure, comparator
	Inpatient hospital admission		Before selection of index date	[-90, 0]	IP	n/a	Exposure, comparator
Predefined Covariates	**Description**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
Metastatic cancer	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Tumor	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Arrhythmia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Congestive heart failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Dementia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Renal failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Weight loss	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Hemiplegia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Alcohol abuse	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Coagulopathy	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator

Complicated diabetes	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Anemia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Fluid and electrolyte disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Liver disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Peripheral vascular disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Psychosis	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary circulation disorders	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
HIV/AIDS	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Hypertension	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Degenerative disease of central nervous system	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Durable medical equipments	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Vaccine administration	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Screening examinations and disease management training	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Pap smear	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
HPV DNA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Mammogram	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Fecal occult blood test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Colonoscopy	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
PSA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Number of inpatient hospitalizations	Healthcare utilization measure	Include	The number of unique CASEIDs where the admission date (inclusive) is within lookback window	[-183, -1]	IP	n/a	Exposure, comparator
Number of outpatient visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one OP claim	[-183, -1]	OP	n/a	Exposure, comparator
Number of emergency department visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one ED claim	[-183, -1]	ED	n/a	Exposure, comparator
Number of unique generics	Healthcare utilization measure	Include	The number of unique generics (based on generic name) dispensed within the lookback window	[-183, -1]	Any	n/a	Exposure, comparator
Co-prescription of beta-lactam	Covariate tailored to example	Include	Binary, 1 if dispensed on same day, otherwise 0	[0, 0]	n/a	n/a	Exposure, comparator
Prior prescription of penicillins	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator

Prior prescription of cephalosporins	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Prior prescription of sulfonamides	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Prior prescription of tetracyclines	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Prior prescription of aminoglycosides	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Pregnancy at time of initiation	Covariate tailored to example	Include	Binary, 1 if codes relating to pregnancy occur within assessment window	[-183, -1]	Any	n/a	Exposure, comparator

Empirically Defined Covariates	**Settings**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
high dimensional propensity score (hdPS) - exposure based selection	**Dimensions:** IP/OP ICD9 diagnoses ICD9/HCPCS/CPT procedures FDB generic names **Top_N** = 200 **k** = 50 **Ranking_method** = exp_assoc **Outcome_zero_cell_corr** = 1 **Default settings for all other macro parameters**	Include	Binary and categories of frequency	[-183, -1]	Any	n/a	Exposure, comparator

Outcome		**Type of entry**	**Washout window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
Event/Outcome	MLCCS tree (incidence defined at level 3 nodes)	Incident	[-183, 0]	IP, ED	Yes	Exposure, comparator

Follow up			
Begins	Day 1		
Ends	**Check all that apply**	**Specify**	
Day X *(specify day)*	X	30	
End of study period	X	30-Sep-15	
Outcome			
Death	X		
Disenrollment	X		
End of exposure *(specify grace and/or stockpiling algorithm)*			
Add/switch exposure *(specify)*			
Other *(specify)*			

Code algorithms for cohort entry date are provided in Appendix A, study entry criteria in Appendix B, covariates in Appendix C and D, outcomes in Appendix E (a * in code algorithm indicates use of a wildcard)

All temporal windows anchored on study population entry date (Day 0) unless otherwise specified.

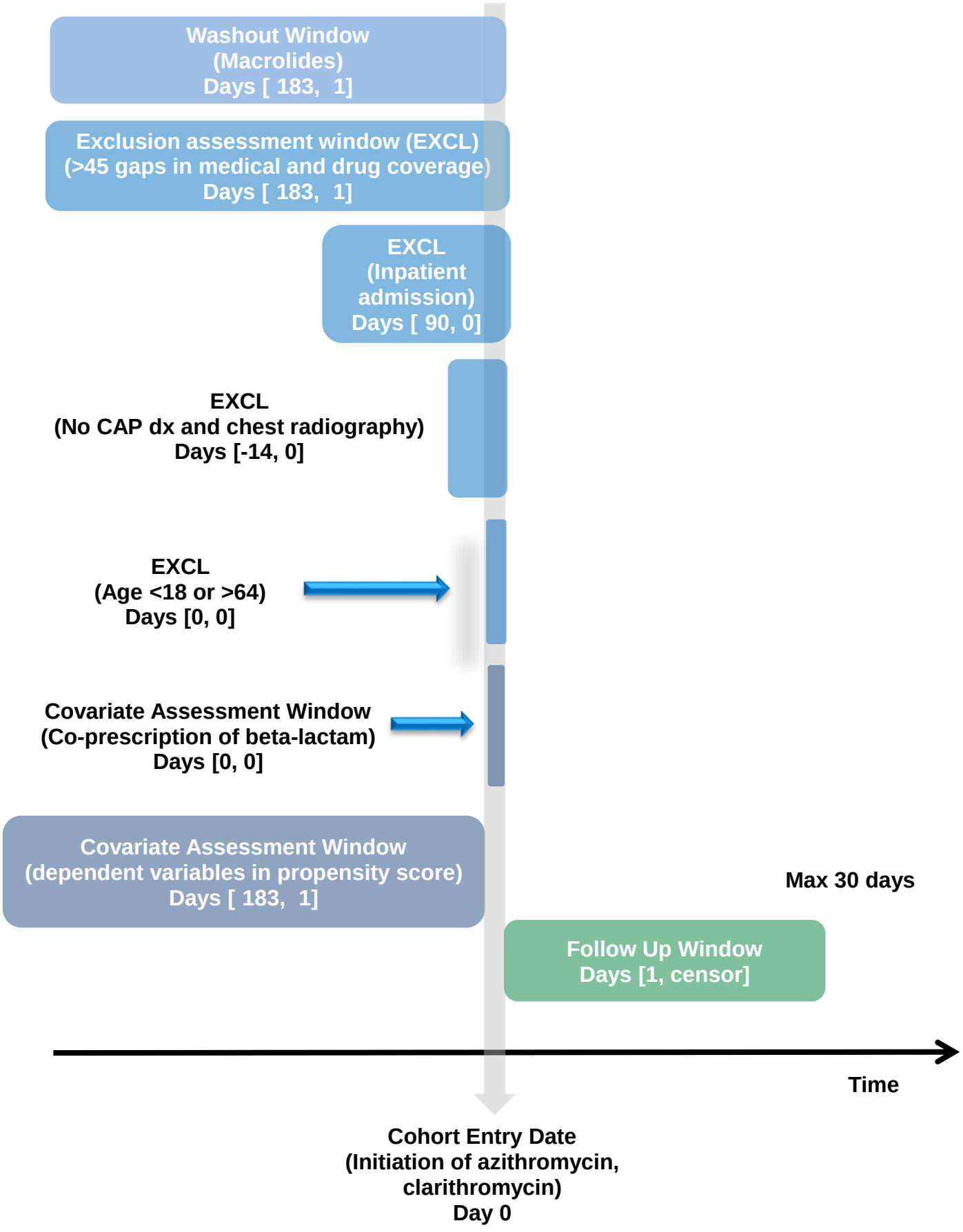
( ) represent open intervals that do not include the end points

[ ] represent closed intervals that do include the end points

[1] Please enter all that apply. Valid entries: IP = inpatient, OP = outpatient, ED = emergency department, any, other, n/a = not applicable. See Appendix F for details on how care setting is defined

DESIGN DIAGRAM

Example 2 azithromycin versus clarithromycin for community acquired pneumonia



Specifications for protocol: *Example 2 azithromycin versus clarithromycin for community acquired pneumonia*

Study Period:	January 1, 2003 - September 30, 2015
Data Source:	Truven MarketScan Commercial and Medicare Supplemental
Data Extraction Date/Version:	Sep 2016
Data sampling/extraction criteria:	All enrollees in data source between January, 2003 - September 30, 2015
Type of data:	Commercial claims
Data linkage:	None
Data conversion:	Sentinel Common Data Model
Software to create study population:	SAS 9.4, CIDA, TreeExtraction, Pharmacoepi Toolbox hdPS macro (http://www.drugepi.org/dope-downloads/), Aetion Platform

Study Entry Criteria (Index Date)	**Description**	**Number of entries**	**Type of entry**	**Washout window**	**Incident w.r.t.**	**Index date (day 0)**
Exposure	Azithromycin	Single	Incident	[-183, 0]	Azithromycin and clarithomycin	Date of incident dispensation
Comparator	Clarithomycin	Single	Incident	[-183, 0]	Azithromycin and clarithomycin	Date of incident dispensation

Exclusion Criteria	**Description**	**Order of application**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Applied to:**
Enrollment/coverage						
	Medical and drug coverage	Before selection of index date	[-183, 0]	n/a	n/a	Exposure, comparator
Max. enrollment gap allowed	45 days					
Demographics						
	Age 18-64 yrs	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Male, female	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Race (any)	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Drug dispensation	Days supply > 0	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Pre-Existing Condition						
	Do not have recent community acquired pneumonia diagnosis and chest radiogr	Before selection of index date	[-14, 0]	Any	n/a	Exposure, comparator
Other	Initiate azithromycin and clarithromycin on the same day	Before selection of index date	[0, 0]	Any	n/a	Exposure, comparator
	Inpatient hospital admission	Before selection of index date	[-90, 0]	IP	n/a	Exposure, comparator

Predefined Covariates	**Description**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
Metastatic cancer	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Tumor	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Arrhythmia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Congestive heart failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Dementia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Renal failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Weight loss	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Hemiplegia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Alcohol abuse	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Coagulopathy	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator

Complicated diabetes	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Anemia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Fluid and electrolyte disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Liver disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Peripheral vascular disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Psychosis	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary circulation disorders	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
HIV/AIDS	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Hypertension	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Degenerative disease of central nervous system	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Durable medical equipments	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Vaccine administration	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Screening examinations and disease management training	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Pap smear	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
HPV DNA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Mammogram	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Fecal occult blood test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Colonoscopy	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
PSA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Number of inpatient hospitalizations	Healthcare utilization measure	Include	The number of unique CASEIDs where the admission date (inclusive) is within lookback window	[-183, -1]	IP	n/a	Exposure, comparator
Number of outpatient visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one OP claim	[-183, -1]	OP	n/a	Exposure, comparator
Number of emergency department visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one ED claim	[-183, -1]	ED	n/a	Exposure, comparator
Number of unique generics	Healthcare utilization measure	Include	The number of unique generics (based on generic name) dispensed within the lookback window	[-183, -1]	Any	n/a	Exposure, comparator
Co-prescription of beta-lactam	Covariate tailored to example	Include	Binary, 1 if dispensed on same day, otherwise 0	[0, 0]	n/a	n/a	Exposure, comparator

Prior prescription of penicillins	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Prior prescription of cephalosporins	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Prior prescription of sulfonamides	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Prior prescription of tetracyclines	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Prior prescription of aminoglycosides	Covariate tailored to example	Include	Binary, 1 if dispensed within lookback window, otherwise 0	[-183, -1]	n/a	n/a	Exposure, comparator
Pregnancy at time of initiation	Covariate tailored to example	Include	Binary, 1 if codes relating to pregnancy occur within assessment window	[-183, -1]	Any	n/a	Exposure, comparator
Empirically Defined Covariates	**Settings**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
high dimensional propensity score (hdPS) exposure based selection	**Dimensions:** IP/OP ICD9 diagnoses ICD9/HCPCS/CPT procedures FDB generic names **Top_N** = 200 **k** = 50 **Ranking_method** = exp_assoc **Outcome_zero_cell_corr** = 1 **Default settings for all other macro parameters**	Include	Binary and categories of frequency	[-183, -1]	Any	n/a	Exposure, comparator
Outcome			**Type of entry**	**Washout window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
Event/Outcome	MLCCS tree (incidence defined at level 3 nodes)		Incident	[-183, 0]	IP, ED	Yes	Exposure, comparator
Follow up							
Begins	Day 1						
Ends	**Check all that apply**	**Specify**					
Day X *(specify day)*	X	30					
End of study period	X	30-Sep-15					
Outcome							
Death	X						
Disenrollment	X						
End of exposure *(specify grace and/or stockpiling algorithm)*							
Add/switch exposure *(specify)*							
Other *(specify)*							

Code algorithms for cohort entry date are provided in Appendix A, study entry criteria in Appendix B, covariates in Appendix C and D, outcomes in Appendix E (a * in code algorithm indicates use of a wildcard)

All temporal windows anchored on study population entry date (Day 0) unless otherwise specified.

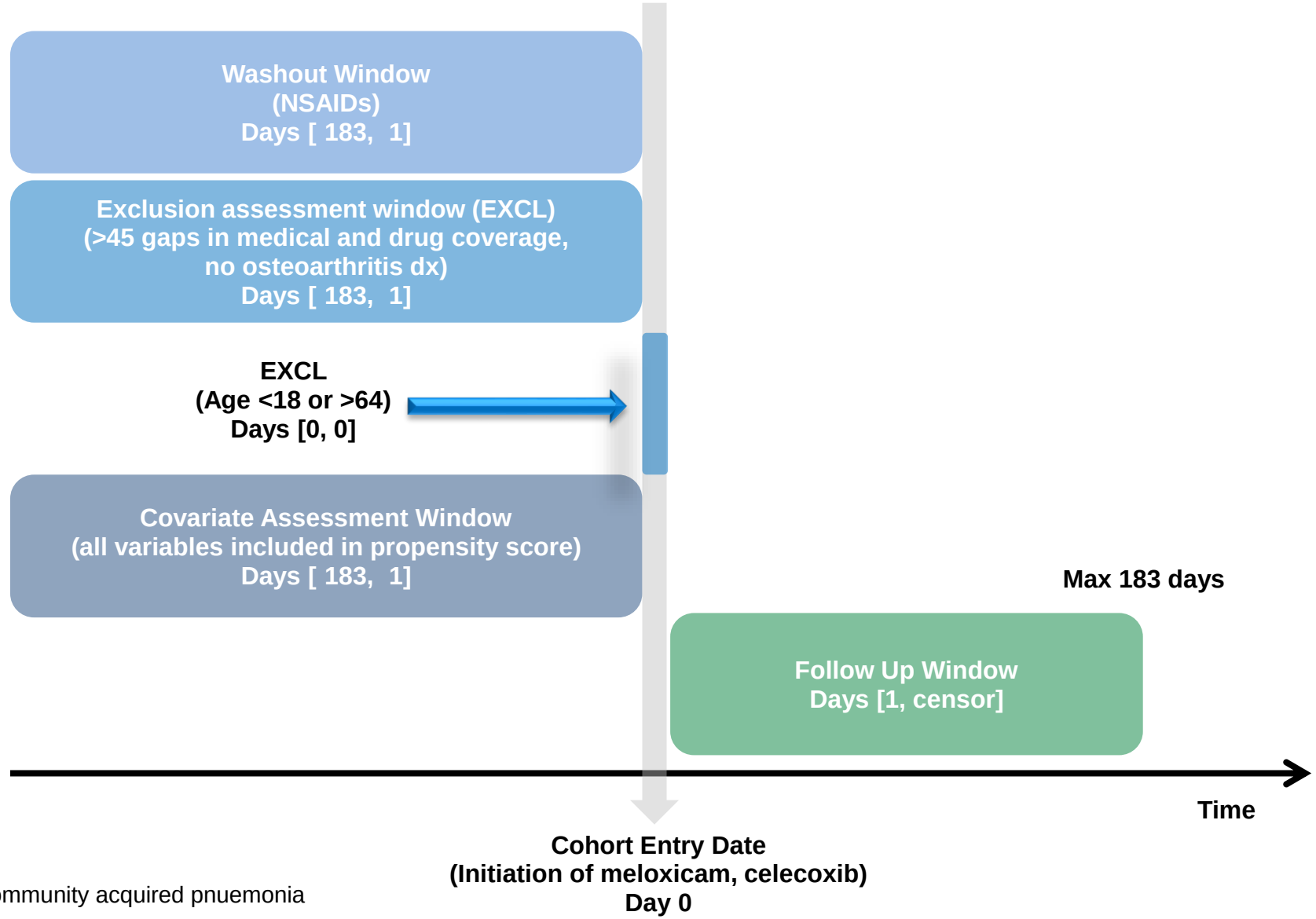
( ) represent open intervals that do not include the end points

[ ] represent closed intervals that do include the end points

[1] Please enter all that apply. Valid entries: IP = inpatient, OP = outpatient, ED = emergency department, any, other, n/a = not applicable. See Appendix F for details on how care setting is defined

DESIGN DIAGRAM

Example 3 Meloxicam vs celecoxib for osteoarthritis



CAP = community acquired pnuemonia

Specifications for protocol: *Example 3 Meloxicam vs celecoxib for osteoarthritis*

Study Period:	January 1, 2003 - September 30, 2015
Data Source:	Truven MarketScan Commercial and Medicare Supplemental
Data Extraction Date/Version:	Sep 2016
Data sampling/extraction criteria:	All enrollees in data source between January, 2003 - September 30, 2015
Type of data:	Commercial claims
Data linkage:	None
Data conversion:	Sentinel Common Data Model
Software to create study population:	SAS 9.4, CIDA, TreeExtraction, Pharmacoepi Toolbox hdPS macro (http://www.drugepi.org/dope-downloads/), Aetion Platform

Study Entry Criteria (Index Date)	**Description**	**Number of entries**	**Type of entry**	**Washout window**	**Incident w.r.t.**	**Index date (day 0)**
Exposure	meloxicam	Single	Incident	[-183, 0]	all NSAIDs	Date of incident dispensation
Comparator	celecoxib	Single	Incident	[-183, 0]	all NSAIDs	Date of incident dispensation

Exclusion Criteria	**Description**	**Order of application**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Applied to:**
Enrollment/coverage						
	Medical and drug coverage	Before selection of index date	[-183, 0]	n/a	n/a	Exposure, comparator
Max. enrollment gap allowed	45 days					
Demographics						
	Age 18-64 yrs	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Male, female	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Race (any)	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Drug dispensation	Days supply > 0	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Pre-Existing Condition						
	Do not have diagnosis of osteoarthritis	Before selection of index date	[-183, 0]	Any	n/a	Exposure, comparator
Other	Initiate meloxicam and celecoxib on the same day	Before selection of index date	[0, 0]	Any	n/a	Exposure, comparator

Predefined Covariates	**Description**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
Metastatic cancer	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Tumor	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Arrhythmia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Congestive heart failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Dementia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Renal failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Weight loss	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Hemiplegia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Alcohol abuse	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Coagulopathy	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Complicated diabetes	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator

Anemia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Fluid and electrolyte disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Liver disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Peripheral vascular disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Psychosis	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary circulation disorders	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
HIV/AIDS	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Hypertension	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Degenerative disease of central nervous system	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Durable medical equipments	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Vaccine administration	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Screening examinations and disease management training	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Pap smear	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
HPV DNA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Mammogram	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Fecal occult blood test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Colonoscopy	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
PSA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Number of inpatient hospitalizations	Healthcare utilization measure	Include	The number of unique CASEIDs where the admission date (inclusive) is within lookback window	[-183, -1]	IP	n/a	Exposure, comparator
Number of outpatient visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one OP claim	[-183, -1]	OP	n/a	Exposure, comparator
Number of emergency department visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one ED claim	[-183, -1]	ED	n/a	Exposure, comparator
Number of unique generics	Healthcare utilization measure	Include	The number of unique generics (based on generic name) dispensed within the lookback window	[-183, -1]	Any	n/a	Exposure, comparator
Anticoagulants	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Obesity	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Smoking	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator

Angina	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Coronary Revascularization	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Statins	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Hormone Replacement Therapy	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Prior Nonselective Nsaid Use	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Stroke/TIA	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
History Of Peptic Ulcer Disease Or Gi B	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Antiplatelets	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Antidepressants	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Fluconazole	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Lithium	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Anti-hypertensives	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Cyclosporine	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Methotrexate	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Steroids	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Pregnancy at the time of initiation	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Fibromyalgia	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Rheumatoid arthritis	Covariate tailored to example	Include	Binary 1 if code in lookback window, 0 otherwise	[-183, -1]	Any	n/a	Exposure, comparator
Empirically Defined Covariates	**Settings**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
high dimensional propensity score (hdPS) - exposure based selection	**Dimensions:** IP/OP ICD9 diagnoses ICD9/HCPCS/CPT procedures FDB generic names **Top_N** = 200 **k** = 50 **Ranking_method** = exp_assoc **Outcome_zero_cell_corr** = 1 **Default settings for all other macro parameters**	Include	Binary and categories of frequency	[-183, -1]	Any	n/a	Exposure, comparator

Outcome		Type of entry	Washout window	Care Settings[1]	Primary Dx	Measured for:
Event/Outcome	MLCCS tree (incidence defined at level 3 nodes)	Incident	[-183, 0]	IP, ED	Yes	Exposure, comparator
Follow up						
Begins	Day 1					
Ends	**Check all that apply**	**Specify**				
Day X *(specify day)*	X	183				
End of study period	X	30-Sep-15				
Outcome						
Death	X					
Disenrollment	X					
End of exposure *(specify grace and/or stockpiling algorithm)*						
Add/switch exposure *(specify)*						
Other *(specify)*						

Code algorithms for cohort entry date are provided in Appendix A, study entry criteria in Appendix B, covariates in Appendix C and D, outcomes in Appendix E (a * in code algorithm indicates use of a wildcard)

All temporal windows anchored on study population entry date (Day 0) unless otherwise specified.

( ) represent open intervals that do not include the end points

[ ] represent closed intervals that do include the end points

[1] Please enter all that apply. Valid entries: IP = inpatient, OP = outpatient, ED = emergency department, any, other, n/a = not applicable. See Appendix F for details on how care setting is defined

Specifications for protocol: *Example 4 Valproic acid versus lamotrigine for epilepsy*

Study Period:	January 1, 2001 - September 30, 2015
Data Source:	Truven MarketScan Commercial and Medicare Supplemental
Data Extraction Date/Version:	January 1, 2018
Data sampling/extraction criteria:	All enrollees in data source between January, 2001 - September 30, 2015
Type of data:	Commercial claims
Data linkage:	None
Data conversion:	Sentinel Common Data Model
Software to create study population:	SAS 9.4, CIDA, TreeExtraction, Pharmacoepi Toolbox hdPS macro (http://www.drugepi.org/dope-downloads/), Aetion Platform

Study Entry Criteria (Index Date)	**Description**	**Number of entries**	**Type of entry**	**Washout window**	**Incident w.r.t.**	**Index date (day 0)**
Exposure	Vaproic acid	Single	Incident	[-183, 0]	Valproic acid, lamotrigine and other anticonvulsants	Date of incident dispensation
Comparator	Lamotrigine	Single	Incident	[-183, 0]	Valproic acid, lamotrigine and other anticonvulsants	Date of incident dispensation

Exclusion Criteria	**Description**	**Order of application**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Applied to:**
Enrollment/coverage						
	Medical and drug coverage	Before selection of index date	[-183, 0]	n/a	n/a	Exposure, comparator
Max. enrollment gap allowed	45 days					
Demographics						
	Age 18-64 yrs	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Male, female	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
	Race (any)	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Drug dispensation	Days supply > 0	Before selection of index date	[0, 0]	n/a	n/a	Exposure, comparator
Pre-Existing Condition						
	Do not have at least 1 diagnosis of epilepsy	Before selection of index date	[-183, 0]	Any	n/a	Exposure, comparator
	Do not have at least 1 diagnosis of convulsions	Before selection of index date	[-183, 0]	Any	n/a	Exposure, comparator
Other	Initiate valproic acid and lamotrigine on the same day	Before selection of index date	[0, 0]	Any	n/a	Exposure, comparator

Predefined Covariates	**Description**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
Metastatic cancer	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Tumor	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Arrhythmia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 5	[-183, -1]	Any	n/a	Exposure, comparator
Congestive heart failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Dementia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Renal failure	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Weight loss	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 2	[-183, -1]	Any	n/a	Exposure, comparator
Hemiplegia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Alcohol abuse	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator

Coagulopathy	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Complicated diabetes	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Anemia	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Fluid and electrolyte disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Liver disease	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Peripheral vascular disorder	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Psychosis	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
Pulmonary circulation disorders	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = 1	[-183, -1]	Any	n/a	Exposure, comparator
HIV/AIDS	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Hypertension	Component of combined comorbidity score	Include	Binary, 1 if diagnosis on a claim, otherwise 0, weight = -1	[-183, -1]	Any	n/a	Exposure, comparator
Degenerative disease of central nervous system	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Durable medical equipments	Component of frailty measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Vaccine administration	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Screening examinations and disease management training	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Pap smear	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
HPV DNA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Mammogram	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Fecal occult blood test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Colonoscopy	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
PSA test	Screening measure	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Number of inpatient hospitalizations	Healthcare utilization measure	Include	The number of unique CASEIDs where the admission date (inclusive) is within lookback window	[-183, -1]	IP	n/a	Exposure, comparator
Number of outpatient visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one OP claim	[-183, -1]	OP	n/a	Exposure, comparator
Number of emergency department visits	Healthcare utilization measure	Include	The number of unique dates in the lookback window with at least one ED claim	[-183, -1]	ED	n/a	Exposure, comparator
Number of unique generics	Healthcare utilization measure	Include	The number of unique generics (based on generic name) dispensed within the lookback window	[-183, -1]	n/a	n/a	Exposure, comparator
HIV infection	Covariate tailored to example	Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator

Other viral infections (mumps, flu, herpes, Coxsackie, Epstein-Barr, cytomegalovirus, parv		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Bacterial infection		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Organ transplant		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Autoimmune diseases (e.g. Lupus, type 1 diabetes, rheumatoid arthritis, psoriatic arthritis,		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Chemotherapy		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Cancer		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Acute b-lymphoblastic leukemia		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Medicines for gout (e.g. Allopurinol)		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Sulfa antibiotics (e.g. Bactrim, septra),		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Sertraline		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
“oxicam” anti-inflammatory drugs (e.g. Meloxicam, piroxicam)		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Alcohol use disorders		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Liver disease		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Gallstones		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Cystic fibrosis		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Kawasaki disease		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Reye’s syndrome		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Hemolytic uremic syndrome (HUS)		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Thrombotic thrombocytopenic purpura (TTP)		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Hyperparathyroidism		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Migraine		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Bipolar disorder		Include	Binary, 1 if diagnosis on a claim, otherwise 0	[-183, -1]	Any	n/a	Exposure, comparator
Empirically Defined Covariates	**Settings**	**Confounder in primary analysis?**	**Categories/Specification**	**Lookback window**	**Care Settings[1]**	**Primary Dx**	**Measured for:**
high dimensional propensity score (hdPS) - exposure based selection	**Dimensions:** IP/OP ICD9 diagnoses ICD9/HCPCS/CPT procedures FDB generic names **Top_N** = 200 **k** = 50 **Ranking_method** = exp_assoc **Outcome_zero_cell_corr** = 1 **Default settings for all other macro parameters**	Include	Binary and categories of frequency	[-183, -1]	Any	n/a	Exposure, comparator

Outcome		Type of entry	Washout window	Care Settings[1]	Primary Dx	Measured for:
Event/Outcome	MLCCS tree (incidence defined at level 3 nodes)	Incident	[-183, 0]	IP, ED	Yes	Exposure, comparator
Follow up						
Begins	Day 1					
Ends	**Check all that apply**	**Specify**				
Day X *(specify day)*	X	183				
End of study period	X	30-Sep-15				
Outcome						
Death	X					
Disenrollment	X					
End of exposure *(specify grace and/or stockpiling algorithm)*						
Add/switch exposure *(specify)*						
Other *(specify)*						

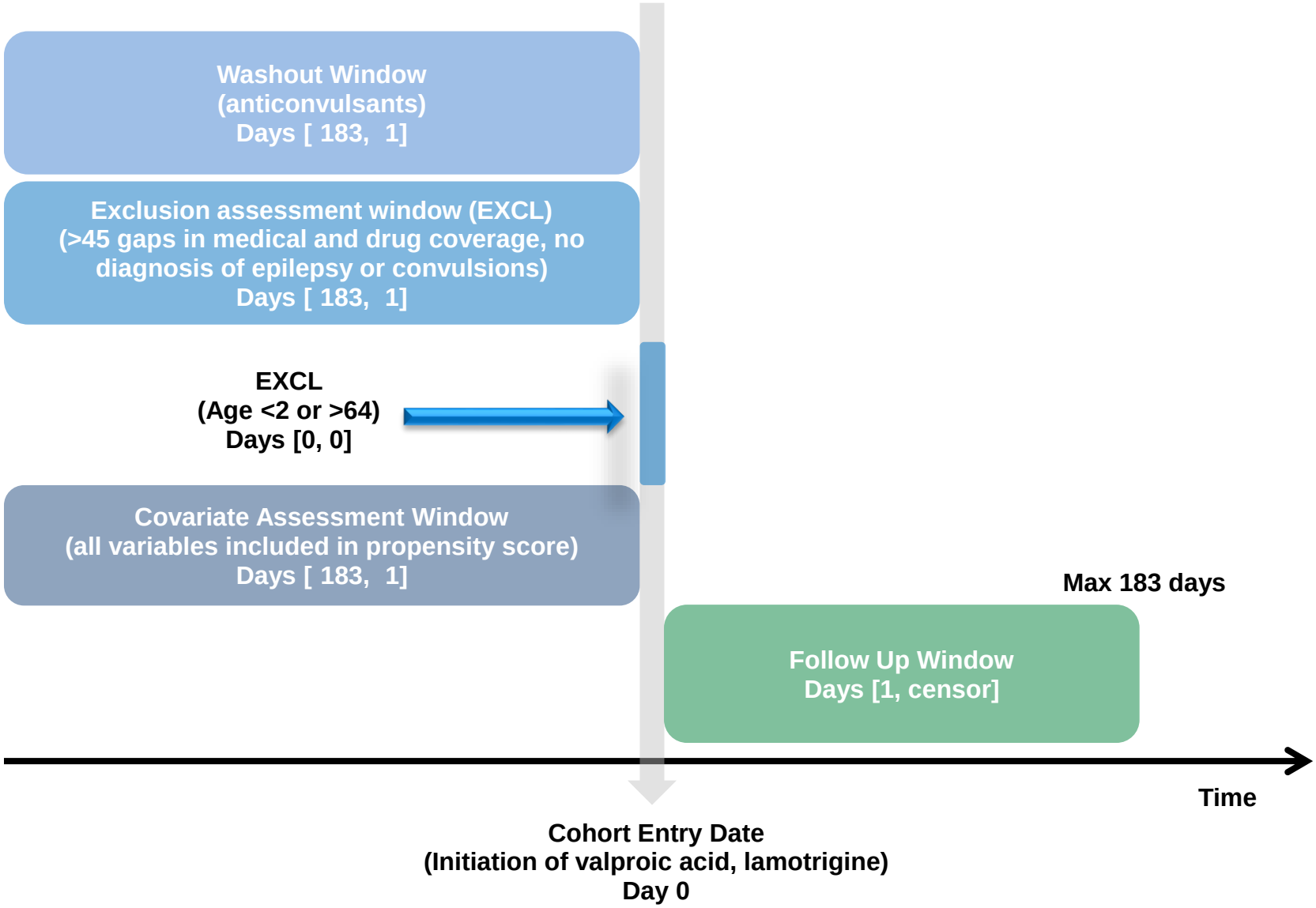
Code algorithms for cohort entry date are provided in Appendix A, study entry criteria in Appendix B, covariates in Appendix C and D, outcomes in Appendix E (a * in code algorithm indicates use of a wildcard)

All temporal windows anchored on study population entry date (Day 0) unless otherwise specified.

( ) represent open intervals that do not include the end points

[ ] represent closed intervals that do include the end points

[1] Please enter all that apply. Valid entries: IP = inpatient, OP = outpatient, ED = emergency department, any, other, n/a = not applicable. See Appendix F for details on how care setting is defined

DESIGN DIAGRAM

Example 4 Valproic acid versus lamotrigine for epilepsy