|  |
| --- |
|  |
| **Janssen Research & Development\*** |
|  |
| **Study Protocol for Retrospective Observational Studies Using Secondary Data** |
|  |
| **Paracetamol versus Ibuprofen and the risk of Myocardial Infarction, Stroke, Gastrointestinal Bleeding and Renal Dysfunction** |
|  |
| **Protocol** PCSCVM000322 |

\* Janssen Research & Development (Janssen R&D) is a global organization that operates through different legal entities in various countries. Therefore, the legal entity acting as the sponsor for studies of Janssen R&D may vary. The term "sponsor" is used throughout the protocol to represent these various legal entities.

**Status:** Final

**Date:** 21 August 2018

**Prepared by:** Janssen Research & Development, LLC; Janssen Research & Development,

**EDMS number:**

**Confidentiality Statement**

The information in this document contains trade secrets and commercial information that are privileged or confidential and may not be disclosed unless such disclosure is required by applicable law or regulations. In any event, persons to whom the information is disclosed must be informed that the information is *privileged* or *confidential* and may not be further disclosed by them. These restrictions on disclosure will apply equally to *all* future information supplied to you that is indicated as *privileged or confidential*.

# TABLE OF CONTENTS

1. TABLE OF CONTENTS 2

2. LIST OF ABBREVIATIONS 3

3. RESPONSIBLE PARTIES 3

3.1. Investigator(s) and Authors 3

3.2. Sponsor 3

4. ABSTRACT 3

5. AMENDMENTS AND UPDATES 3

6. RATIONALE AND BACKGROUND 3

7. STUDY Objectives 4

7.1. Primary Objective(s) 4

7.2. Secondary Objective(s) 4

8. Research METHODs 4

8.1. Study Design and Setting 4

8.2. Describe Data Source(s) 4

8.3. Study Population(s) 4

8.3.1. [Treatment Group] [Cases] 5

8.3.2. [Comparator Group] [Controls] 5

8.4. Outcome(s) of Interest 5

8.5. Exposure(s) of Interest 7

8.6. Other Variables of Interest (Demographic Characteristics, Confounders, Effect Modifiers) 7

9. SAMPLE SIZE AND STUDY POWER 9

10. DATA ANALYSIS PLAN 9

10.1. Calculation of Time-at-Risk 9

10.2. Patient Characteristics Summary 10

10.3. Model Specification 10

10.4. Evidence Evaluation 10

11. STRENGTHS AND LIMITATIONS OF THE RESEARCH METHODS 11

12. PROTECTION OF HUMAN SUBJECTS 11

13. MANAGEMENT AND REPORTING OF ADVERSE EVENTS AND ADVERSE REACTIONS 11

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS 11

15. references 11

16. list of tables & figures 12

ANNEX. LIST OF STAND-ALONE DOCUMENTS 13

# LIST OF ABBREVIATIONS

|  |  |
| --- | --- |
| abbreviation | description of abbreviated term |
| CPRD | Clinical Practice Research Database |
|  |  |

# RESPONSIBLE PARTIES

## Investigator(s) and Authors

Rachel Weinstein, Joel Swerdel, Jesse Berlin, Patrick Ryan, Martijn Schuemie, Daniel Fife

## Sponsor

Janssen Research & Development, LLC

# ABSTRACT

Purpose: In a recent study, we observed evidence of channeling bias, or confounding by contraindication, for the prescription of paracetamol versus ibuprofen. We developed and tested an approach to adjusting for the channeling bias in the prescription of paracetamol versus ibuprofen with outcomes that were known not to be causally associated with adverse outcomes of either medicine. We were able to do this since we knew that the exposure and outcome should not have been associated. The purpose of the current study is to use the new method to assess whether adverse outcomes identified in the published literature, such as myocardial infarction and stroke, continue to be associations with paracetamol versus ibuprofen after adjustment.

Methods: In a cohort of new users first prescribed paracetamol alone or ibuprofen alone, use large scale propensity score matching to assess the relative risk of myocardial infarction (MI), stroke, gastrointestinal (GI) bleeding or renal disease after paracetamol prescription versus ibuprofen. Negative controls will be used to assess potential bias and to calibrate p-values. In addition, positive controls will be generated by simulation of relative risks of size 1.5, 2.0 and 4.0 in the negative controls. Positive controls will be used to refine confidence bound estimates of the relative risks of interest (MI, stroke GI bleed, renal disease).

Analysis: Two propensity score models will be developed (for each outcome). One based on those typical in the published literature using hand-picked covariates. The other will be a partially automated model including all possible covariates in the database—including 1,000+ variables. One-to-one matching will be implemented in on-treatment Cox models of first paracetamol prescription versus ibuprofen for the outcomes of interest and negative controls. Visualizations of hazard rates versus standard errors will be the primary output. Such graphics include traditional and calibrated regions for assessing significant adverse event associations.

# AMENDMENTS AND UPDATES

| Number | Date | Section of study protocol | Amendment or update | Reason |
| --- | --- | --- | --- | --- |
| 1 | July 11, 2018 | 7.8 Other Variables of Interest | Clarify the time windows for condition and drug groups. | There was no mention of the time frame for condition and drug groups |
| 2 | July 11, 2018 | 15 List of Tables and Figures | Added tables/figures corresponding to calibrated and uncalibrated, negative and positive control outcomes | Table/Figures for positive controls, uncalibrated and calibrated outcomes had been omitted. |
| 3 | August 21, 2018 | 9.1 Calculation of Time-at-Risk | Two intent-to-treat analyses were added as sensitivity analyses | Based on the diagnostic plots of the negative control outcomes only to understand the robustness of results of the primary analysis. |

# RATIONALE AND BACKGROUND

In a prior study, we presented evidence of channeling in the recommendation of paracetamol versus ibuprofen. We also used a variety of models to show that most, if not all, previous publications have inadequately adjusted for this source of bias and we proposed a more effective way to adjust for confounding. It is now reasonable to conduct a new study to assess the association of paracetamol versus ibuprofen use and the severe adverse event outcomes of gastrointestinal bleed, myocardial infarction, stroke and renal insufficiency.

Numerous cohort and case-control studies over the past three decades have found an association of major adverse events, such as renal disease, myocardial infarction, stroke and gastrointestinal (GI) bleeding, with paracetamol use compared to use of ibuprofen and other over-the-counter (OTC) analgesics (Roberts at al 2015, Lipworth 2003 et al, De Vrie et al 2010, Chan et al 2006). Most of these studies mention channeling as a possible source of bias influencing the results, but do not attempt to measure the impact of such bias. In our prior study (Weinstein et al 2017), we found evidence of channeling in the prescription of paracetamol versus ibuprofen and then used two propensity score models for matching on negative control outcomes to assess the extent of channeling. (These are conditions for which we have high confidence there should be no association with the exposure(s) of interest.) That study suggested a strategy for reducing the impact of this form of bias. In the current study, we will apply those methods to examine the association of paracetamol versus ibuprofen for those adverse events reviewed in the prior studies.

# STUDY Objectives

## Primary Objective(s)

The primary objective is to assess whether paracetamol, compared to ibuprofen, is associated with an increased risk of myocardial infarction, stroke, GI bleeding and renal disease.

## Secondary Objective(s)

A secondary objective is to assess the extent of residual bias in the estimation of the effect as specified in the primary objective using negative and positive control hypotheses. Any observed residual bias will be incorporated in an empirically calibrated p-value and confidence interval. Research METHODs

## Study Design and Setting

This is a retrospective cohort study. We will include adults in UK general practices with a prescription for single-ingredient paracetamol or ibuprofen in 2005-2014.

## Describe Data Source(s)

We will use the Clinical Practice Research Datalink (CPRD), which was used in one of the prior studies, as well as in the most recent channeling bias paper. The CPRD is a governmental, not-for-profit research service, jointly funded by the NHS National Institute for Health Research (NIHR) and the Medicines and Healthcare Products Regulatory Agency (MHRA), a part of the Department of Health, United Kingdom (UK). CPRD consists of data collected from a sample of UK primary care physicians for patients of all ages. This includes conditions, observations, measurements, and procedures that the general practitioner is made aware of in additional to any prescriptions by the general practitioner. In addition to primary care, there are also linked secondary care records for a small number of people.

## Study Population(s)

**Inclusion criteria**

Subjects are included in the study if they received a first prescription for either single-ingredient paracetamol or single-ingredient ibuprofen in 2005-2014. The focus is on single ingredient drugs so the indication is not complicated by other ingredients in a combination drug. The date of this prescription is the index date. Subjects are included if they are age eighteen or older on the index date and enrolled for the two years prior to and at least one day after the index date. The 2-year observation period is intended to provide adequate information about covariates. To reduce the risk of observing prevalent prescription use, we require 12 months of continuous observation without prescriptions of paracetamol or ibuprofen prior to index date. We will classify analgesic use into two cohorts: 1) “Paracetamol only”—patients with new, single-ingredient paracetamol exposure without concomitant ibuprofen, and 2) “Ibuprofen only”— patients with new, single-ingredient ibuprofen exposure without concomitant paracetamol.

**Exclusions**

Patients who receive prescriptions for both paracetamol and ibuprofen on their index date will be excluded from the study. Both the paracetamol and the ibuprofen cohort members should be free of prior prescription of both paracetamol or ibuprofen in the 12 months prior to their index dates. In addition, subjects with a prescription for other NSAIDs or aspirin in the 12 months prior to or on the index date are excluded. In addition, subjects with a prescription for any paracetamol- or ibuprofen- containing combination products in the 12 months prior to and including the index date will be excluded from the study. The rationale for this exclusion is that these prescriptions could be driven primarily by the other ingredient, which distracts from the purposes of the study.

### Treatment Group

The treatment cohort is “Paracetamol only”, that is patients with new paracetamol exposure alone.

### Comparator Group

The comparator group is “Ibuprofen only”, that is, patients with single-ingredient ibuprofen alone.

## Outcome(s) of Interest

The four primary outcomes of interest are incident GI bleeding, myocardial infarction, stroke and renal insufficiency. Variables for presence or absence of a given diagnosis will be developed based on diagnoses in the database following the index date, regardless of the amount of time subjects are in the database following the index date. Review of all available time prior to the outcome is the constraint in determining whether an event was incident or not. We require all patients to have at least 1 day in the database after the index date. Disease codes will be developed based on disease vocabularies available within the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).

**Negative Control Outcomes**

Negative control outcomes are those determined a priori to have no association with the exposure of interest. The intention in identifying negative control outcomes, is to identify exposure-outcome pairs that do not have a causal association. Therefore, we identify outcomes as follows: Person counts of all potential drug-condition pairs are reviewed in observational data; this person count data helps determine which pairs are even probable for use in calibration. Given the list of potential drug-condition pairs, the concepts in the pairs must meet the following requirements to be considered as negative controls: (1) that there is no Medline abstract where the MeSH terms suggest a negative association between the drug and the condition (Winnenburg et al 2015), (2) that there is no mention of the drug-condition pair on a US Product Label in the “Adverse Drug Reactions” or “Postmarketing” section (Duke et al 2013), (3) there are no US spontaneous reports suggesting that the pair is in an adverse event relationship (Banda et al 2016; Evans et al 2001), (4) that the OMOP Vocabulary does not suggest that the drug is indicated for the condition, (5) that the concepts are usable (i.e. not too broad, not suggestive of an adverse event relationship, not pregnancy related), and (6) the exact concept itself is utilized in patient level data (i.e. concepts that are not usually used within the data are usually indicative a broad concept that has a child that is more specific).  The remaining concepts are “optimized”, meaning parent concepts remove children as defined by the OMOP Vocabulary (e.g. if both “Non-Hodgkin’s Lymphoma” and “B-Cell Lymphoma” we selected, child concept “B-Cell Lymphoma would be removed for its parent “Non-Hodgkin’s Lymphoma”).  Once potential negative control candidates are selected, manual clinical review to exclude any pairs that may still be in a causal relationship or similar to the study outcome should be performed to select the top 50 or so concepts by patient exposure.

Models which adequately control for confounding factors should produce hazard ratio (HR) estimates of the null value (1.0) for these negative controls outcomes. These models will allow for the examination of the extent of bias in the database, study design, and analysis to the degree that they produce significant RR estimates different from 1.0. In addition, an empirical distribution of the HR under the null is developed from these negative control outcomes. We will fit the series of cox regression models using 39 controls belowusing criteria mentioned.

The 39 negative control outcomes we will be using from the prior study are as follows:

|  |  |
| --- | --- |
| 1 | Achilles tendinitis |
| 2 | Atrophic vaginitis |
| 3 | Breath smells unpleasant |
| 4 | Bronchiectasis |
| 5 | Disorders of initiating and maintaining sleep |
| 6 | Ear problem |
| 7 | Erythema nodosum |
| 8 | Falls |
| 9 | Foot-drop |
| 10 | Ganglion and cyst of synovium, tendon and bursa |
| 11 | Hemangioma |
| 12 | Hydrocele |
| 13 | Hyperthyroidism |
| 14 | Impaired glucose tolerance |
| 15 | Impingement syndrome of shoulder region |
| 16 | Impotence |
| 17 | Incontinence of feces |
| 18 | Interpersonal relationship finding |
| 19 | Irregular periods |
| 20 | Irritability and anger |
| 21 | Joint stiffness |
| 22 | Loss of sense of smell |
| 23 | Mixed hyperlipidemia |
| 24 | Osteitis deformans |
| 25 | Panic attack |
| 26 | Perforation of tympanic membrane |
| 27 | Pes planus |
| 28 | Polymyalgia rheumatica |
| 29 | Premature menopause |
| 30 | Prolapse of female genital organs |
| 31 | Pure hypercholesterolemia |
| 32 | Respiratory symptom |
| 33 | Restless legs |
| 34 | Restlessness and agitation |
| 35 | Rosacea |
| 36 | Simple goiter |
| 37 | Skin sensation disturbance |
| 38 | Snapping thumb syndrome |
| 39 | Urinary symptoms |

**Positive control outcomes**

In addition to negative control outcomes, we will also include synthetic positive control outcomes. These are outcomes based on the negative controls described above, but where the true effect size is artificially increased to a desired effect size by injection of additional, simulated outcomes. To preserve confounding, these additional outcomes are sampled from predicted probabilities generated using a fitted predictive model. For each negative control outcome, three positive control outcomes will be generated with true relative risk is 1.5, 2, and 4. Using both negative and positive controls, we will fit a systematic error model and perform confidence interval calibration (Scheumie et al *forthcoming*).

## Exposure(s) of Interest

Not applicable. See section 7.5 above.

## Other Variables of Interest (Demographic Characteristics, Confounders, Effect Modifiers)

The other variables of interest are those used in the propensity score models. Two propensity score models will be fit with different sets of variables. One will be referred to as the “publication variables” and are intended to be equivalent to the list of variables seen commonly in publications on adverse effects of paracetamol and/or ibuprofen. The propensity score model based on the publication variables list is intended to replicate the analyses seen in prior publications on this topic.

*Publication variables* are as follows:

Obese, morbidly obese, smoker, alcohol abuser , upper GI events , osteoarthritis, rheumatoid arthritis, ischemic heart disease, heart failure, hypertension, cerebrovascular disease, diabetes mellitus , hyperthyroidism , stroke or transient ischemic attack , cancer [excluding non-melanoma skin cancer], inflammatory bowel, autoimmune disease, depression, drug abuse, anticoagulants, oral glucocorticoids, diuretics, cardiac glycosides, statins, angiotensin receptor blockers, hypnotics and anxiolytics, antipsychotics, antibacterials, aminosalicylates, antidepressant, aspirin, oral corticosteroids, proton-pump inhibitors, histamine-2 receptor antagonists, and hyperlipidemia.

In the second propensity score model, a much larger set of baseline covariates will be defined, which we will refer to as the ‘full set of covariates available’ or ‘large-scale propensity score covariates to characterize patient demographics, prior conditions, drugs, procedures and health service utilization patterns:

Age in 5-year increments

Sex

Race

Year of index date

Month of index date

Charlson Comorbidity Index

Diabetes Complications Severity Index (DCSI) score

CHADS2 score

Conditions

Presence/absence of a condition in a 365-day window prior to or on the index date

Presence/absence of a condition in a 30-day window prior to or on the index date

Presence/absence of a condition diagnosed in an inpatient setting in a 180-day window prior to or on the index date

Presence/absence of an aggregation of episodes of care over time for a condition (‘condition era’) any time prior to or on the index date

Presence/absence of a condition era that overlaps the index date

Condition group (based on the SNOMED hierarchy of conditions)

Presence/absence of an aggregation of episodes of care over time for a condition group era any time prior to or on the index date

Presence/absence of a condition group era that overlaps the index date

Drugs

Presence/absence of length of time of exposure to a drug product (‘drug era’) in a 365-day window prior to or on the index date

Presence/absence of a drug era in a 30-day window prior to or on the index date

Presence/absence of a drug era that overlaps the index date

Presence/absence of a drug era any time prior to or on the index date

Presence/absence of length of time of exposure to a drug group era in a 365-day window prior to or on the index date

Drug groups are based on the Anatomical Therapeutic Chemical hierarchy

Presence/absence of a drug group era in a 30-day window prior to or on the index date

Presence/absence of a drug group era that overlaps the index date

Presence/absence of a drug group era any time prior to or on the index date

Procedures, observations and measurements

Presence/absence of a procedure in a 365-day window prior to or on the index date

Presence/absence of a procedure in a 30-day window prior to or on the index date

Procedures classes (based on the SNOMED hierarchy of procedures)

Presence/absence of an observation in a 365-day window prior to or on the index date

Presence/absence of an observation in a 30-day window prior to or on the index date

# SAMPLE SIZE AND STUDY POWER

For the cohort study design, the table below presents the minimum detectable risk (Armstrong 1987) for given number of events in a cohort study. The distributions of exposures and events in the table come from those observed in the prior study [1]. The interpretation of the numbers in the table, in the context of sample size and power is that, for example, if the number of observed MIs in the paracetamol cohort is 1,148 and in the ibuprofen cohort is 288, then the minimum detectable relative risk would be 1.07, with 80% power. A hypothesized minimum number of events is also included in order to obtain the RR when a much smaller than expected number of events is observed.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Minimum detectable risk, alpha=0.05, power=0.80 | | | |  |
| Prior paper | N paracetamol | N ibuprofen | Total Events | Minimum Detectable Relative risk |
| MI | 1148 | 288 | 1436 | 1.067 |
| Renal Disease | 5382 | 1887 | 7269 | 1.041 |
| Stroke | 1791 | 494 | 2285 | 1.096 |
| GI Bleeding | 2201 | 1480 | 3681 | 1.092 |
| Minimum | 125 | 125 | 250 | 1.452 |

Note also that A Cox regression of the log hazard ratio on a covariate with a standard deviation of 0.1000 based on a sample of 143130 observations (from the prior study) achieves 87% power at a 0.05000 significance level to detect a regression coefficient equal to 0.4055 (=ln(1.5)). The sample size was adjusted for an anticipated event rate of 0.0400.

For the current study period, (1/1/2005 – 12/31/2014) we have the following numbers of subjects:

For acetaminophen - 349,738 subjects

For ibuprofen - 512,556 subjects

# DATA ANALYSIS PLAN

In a prior publication, we developed a model for adjustment of channeling bias in the prescription of paracetamol and ibuprofen using large scale propensity score matching. In this study, we will re-examine the association between paracetamol and the incidence of severe adverse events using that same approach, although the final outcome models will be Cox models, rather than logistic regression model. The Cox models are more appropriate in this context which is more focused on safety than on the methods of controlling for confounding. Outcomes that will be assessed are acute myocardial infarction, stroke, GI bleed, and renal insufficiency.

This will be an on-treatment analysis approach and will focus on the first treatment in keeping with the framework of physician-prescribed channeling. Therefore, subjects who get a prescription for paracetamol alone and, at a later time in the study year, get a prescription for ibuprofen are classified as paracetamol, however they are censored at the time if the ibuprofen prescription; and vice versa.

## Calculation of Time-at-Risk

Because the underlying hypothesis is that differential prescribing is the source of association, exposure is classified by the drug prescribed on the index date. Time-at-risk will be from the index date (the date of treatment initiation) to the end of treatment allowing for a 30-day gap between consecutive prescriptions. Drug eras end at the end of the last prescription, according to the days’ supply.

Two sensitivity analyses will be implemented to extend the treatment window and contribute to our understanding of the robustness of the results. They will be intent-to-treat (ITT) analyses with 1) a 90-day window and 2) a 1-year window. For these ITT analyses, exposure is assigned based on the drug at index and only censored if the patient leaves the practice or dies before the end of the window.

## Patient Characteristics Summary

Descriptive analyses will be comprised of covariate balance of age group, gender, and selected variables before and after matching for the publication variables model and the all possible covariate models.

## Model Specification

Cohort members’ status with respect to gastrointestinal bleeding, myocardial infarction, stroke, and kidney disease will be assessed based on the presence or absence of a diagnosis following the index date for the respective analyses.The look-back period of two years is to account for comorbid conditions and outcome events which may have occurred in the relatively distant past, but still influence prescribing behaviors. Two Cox models will be employed with propensity score matching to evaluate the relative risk of exposure to paracetamol versus ibuprofen and the degree to which prescription bias or residual confounding may influence the observed association with paracetamol versus ibuprofen prescription. Each outcome model will exclude those with a prior history of the condition and, in addition, each outcome is treated independently. The Cox models will estimate the effect of first exposure (paracetamol vs. ibuprofen) on the risk of each outcome while on treatment, without further adjustment.

1) Publication variable propensity score matched outcome model:  We will perform a multivariate logistic regression to estimate a propensity score that predicts treatment assignment (paracetamol vs. ibuprofen) using baseline covariates for sex, age in 5-yr increments, and a set of 36 binary variables defined below, as identified in the publication by De Vries et al.  The propensity score (PS) will be used to perform 1-to-1 matching (using standardized caliper of 0.25\*PS standardized deviation).   The matched sets will then be used within a univariate cox proportional hazard regression, which will estimate the effect of exposure (paracetamol vs. ibuprofen) on the incidence of each outcome, without further adjustment. The publication variables were listed above in Section 8.6.

The publication variables are based on the De Vries et. al (2010) paper and are similar to those seen in other papers on this topic, but are meant to represent a selected set of potential confounders as distinguished from an approach where all possible covariates are entered into the propensity score model. Although the exact codes included in the definition of each variable were not published, the authors indicated that the above conditions and exposures were controlled at baseline. The definitions of the custom covariates are in Annex 1.

2) Large-scale propensity score matched model:  We will perform a regularized logistic regression using L1 (LASSO) shrinkage to estimate a propensity score that predicts treatment assignment (paracetamol vs. ibuprofen) using a large array of baseline covariates to characterize patient demographics, prior conditions, drugs, procedures and health service utilization patterns, as defined above.  The propensity score will be used to perform 1-to-1 matching (using standardized caliper of 0.25\*PS standardized deviation).   The matched sets will then be used within a univariate cox regression model, which will estimate the effect of first exposure (paracetamol vs. ibuprofen) on the incidence of each outcome within 1 year post-exposure start, without further adjustment.

Specifically, because we have transformed CPRD into the OMOP common data model, v5, we will apply standardized covariates within the open-source OHDSI application “CohortMethod” (https://github.com/ohdsi/cohortmethod), which include SNOMED-coded concepts and higher-level MedDRA classifications for conditions, drugs coded at the RxNorm ingredient and ATC class levels, and procedures hierarchically represented in SNOMED. The optimal regularization hyper-parameters will be estimated using 10-fold cross-validation. Baseline variables will be evaluated based on data available prior to the cohort index date. To avoid over-fitting models and to accommodate a large number of predictors, the RLRM will be fit using a cyclic coordinate descending (CCD) method with L1 penalty (i.e., least absolute shrinkage and selection operator (LASSO)) (Tibshirani, 1996).

## Evidence Evaluation

Preference scores will be calculated for the probability of exposure to a first prescription for paracetamol versus ibuprofen, corrected for the overall prevalence of the two drugs, to mimic the situation in which the two drugs are equally prevalent. Preference score plots as well as covariate balance scatter plots of the 2 exposure cohorts will be generated.

Visual characterization of bias and confounding will be shown for each of the 4 primary outcomes and 2 propensity score models via scatter plots of HRs versus standard errors shaded for traditional and calibrated significance regions.

# STRENGTHS AND LIMITATIONS OF THE RESEARCH METHODS

This is an observational study and as such is subject to the inherent bias because we were not able to randomize treatment, for example. In this data source, the drug data is based on prescriptions written, not dispensed, so we do not know if the drugs were taken and how. For those prescriptions that were repeated, the likelihood that the medication was ingested increases. Exposure attribution based on a single prescription may be seen as a limitation, however we would argue that any other measures such as dose or drug era would be similarly confounded with underlying health.

# PROTECTION OF HUMAN SUBJECTS

Approval for CPRD use will be obtained from the Independent Scientific Advisory Committee (ISAC).

# MANAGEMENT AND REPORTING OF ADVERSE EVENTS AND ADVERSE REACTIONS

This study uses coded data that already exist in an electronic database. In this type of database, it is not possible to link (i.e., identify a potential causal association between) a particular product and medical event for any individual. Thus, the *minimum criteria for reporting an adverse event (i.e., identifiable patient, identifiable reporter, a suspect product, and event) are not available* and adverse events are not reportable as individual AE reports. The study results will be assessed for medically important results.

# PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

We will register with the OHDSI website after the protocol has been finalized and approved.

# references

Armstrong B. A simple estimator of minimum detectable relative risk, sample size or power in cohort studies. *American Journal of Epidemiology* 1987;126(2):356-358.

Banda JM, Evans L, Vanguri RS, Tatonetti NP, Ryan PB, Shah NH. A curated and standardized adverse drug event resource to accelerate drug safety research. *Sci Data*. 2016 May 10;3:160026. doi: 10.1038/sdata.2016.26. PubMed PMID: 27193236; PubMed Central PMCID: PMC4872271.

Chan AT, Manson JE, Albert CM, et al. Nonsteroidal anti-inflammatory drugs, paracetamol, and the risk of cardiovascular events. *Circulation* 2006;113:1578–87.

Duke J, Friedlin J, Li X. Consistency in the safety labeling of bioequivalent medications. Pharmacoepidemiol Drug Saf. 2013 Mar;22(3):294-301. doi: 10.1002/pds.3351. Epub 2012 Oct 8. PubMed PMID: 23042584.

Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf*. 2001 Oct-Nov;10(6):483-6. PubMed PMID: 11828828.

Lipworth L, Friis S, Mellemkjaer L, et al. A population-based cohort study of mortality among adults prescribed paracetamol in Denmark. *J Clin Epidemiol* 2003;56:796–801.

Roberts E, Delgado Nunes V, Buckner S, Latchem S, Constanti M, Miller P, Doherty M, Zhang W, Birrell F, Porcheret M, Dziedzic K, Bernstein I, Wise E, Conaghan PG. Paracetamol: not as safe as we thought? A systematic literature review of observational studies. *Ann Rheum Dis* 2015;0:1-8.

Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety*. Drug Saf*. 2013;36(Suppl 1):S33–47. doi:10.1007/s40264-013-0097-8.

Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proceedings of the National Academy of Sciences*. 2018. 201708282; DOI:10.1073/pnas.1708282114

Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc Series B Stat Methodol* 1996;58: 267-288

de Vries F, Setakis E, van Staa TP, et al. Concomitant use of ibuprofen and paracetamol and the risk of major clinical safety outcomes. *Br J Clin Pharmacol* 2010;70:429–38.

Weinstein RB, Ryan P, Berlin JA, Matcho A, Schuemie M, Swerdel J, Patel K, Fife D. Channeling in the use of nonprescription paracetamol and ibuprofen in an electronic medical records database: evidence and implications.  *Drug Safety* Published online: 05 August 2017.

Winnenburg R, Sorbello A, Ripple A, Harpaz R, Tonning J, Szarfman A, Francis H, Bodenreider O. Leveraging MEDLINE indexing for pharmacovigilance - Inherent limitations and mitigation strategies. J *Biomed Inform*. 2015 Oct;57:425-35. doi: 10.1016/j.jbi.2015.08.022. Epub 2015 Sep 2. PubMed PMID: 26342964; PubMed Central PMCID: PMC4775467.

# list of tables & figures

Table 1. Covariate balance before and after matching for selected covariates for description

Table 2 or Figure 7: Hazard ratio estimates and confidence intervals for the 4 outcomes of interest before and after calibration, publication variables propensity score model and large scale propensity score model

Figure 1. Propensity score curves for paracetamol and ibuprofen cohorts

Figures 2a) Covariate balance under matched publication variable model. b) covariate balance of all covariates matched on publications variable model. c) covariate balance of all covariates matched on all covariates model.

Figures 3.1-3.16) For each outcome, plot all negative and positive control **uncalibrated** estimates RR versus standard error for **publications variable** propensity score model

Figures 4.1-4.16) For each outcome, plot all negative and positive control **uncalibrated** estimates RR versus standard error with calibrated p-value regions for the **large-scale propensity score** model

Figures 5.1-5.16) For each outcome, plot all negative and positive control **calibrated** estimates RR versus standard error for **publications variable** propensity score model

Figures 6.1-6.16) For each outcome, plot all negative and positive control **calibrated** estimates RR versus standard error with calibrated p-value regions for the **large-scale propensity score** model

1. ANNEX. LIST OF STAND-ALONE DOCUMENTS

|  |  |  |
| --- | --- | --- |
| **Document number** | **Date** | **Title** |
| 1 |  | List of variable definitions and code sets for variables included in the protocol |
| 2 |  |  |
| ... |  |  |

**SAMPLE Variable Code List** [an excel file with source codes can be maintained as a separate file]

|  |  |
| --- | --- |
| **Variable/Condition/Procedure** | **Link to concept set created in Atlas** |
| Acetaminophen Single Ingredient | https://epi.jnj.com/atlas/#/conceptset/6533/details |
| Acetaminophen All | https://epi.jnj.com/atlas/#/conceptset/6534/details |
| Ibuprofen Single Ingredient | https://epi.jnj.com/atlas/#/conceptset/5818/details |
| Ibuprofen All | https://epi.jnj.com/atlas/#/conceptset/6535/details |
| NSAIDs plus Aspirin | https://epi.jnj.com/atlas/#/conceptset/6536/details |
| MI | https://epi.jnj.com/atlas/#/conceptset/6241/details |
| Stroke | https://epi.jnj.com/atlas/#/conceptset/6345/details |
| GI Bleed | https://epi.jnj.com/atlas/#/conceptset/6243/details |
| Acute Renal Failure | https://epi.jnj.com/atlas/#/conceptset/6244/details |
| Negative Controls | https://epi.jnj.com/atlas/#/conceptset/6056/details |

