

Protocol

# High-Throughput Computing to Detect Harmful Drug-Drug Interactions in Older Adults: Protocol for a Population-Based Cohort Study

Neda Rostamzadeh<sup>1\*</sup>, PhD; Rishabh Sharma<sup>2,3\*</sup>, PharmD, MSc; Sheikh S Abdullah<sup>1,2,3,4</sup>, PhD; Eric McArthur<sup>2,3</sup>, MSc; Niaz Chalabianloo<sup>1,3,5</sup>, PhD; Jessica M Sontrop<sup>2,3,6</sup>, PhD; Matthew A Weir<sup>2,3,6,7</sup>, MD, MSc; Kamran Sedig<sup>1</sup>, PhD; Amit X Garg<sup>2,3,6,7</sup>, MD, PhD; Flory T Muanda<sup>2,3,5,6</sup>, MD, PhD

<sup>1</sup>Department of Computer Science, Western University, London, ON, Canada

<sup>2</sup>London Health Sciences Centre Research Institute, London, ON, Canada

<sup>3</sup>ICES Ontario, London, ON, Canada

<sup>4</sup>Department of Computer Science, MacEwan University, Edmonton, AB, Canada

<sup>5</sup>Department of Physiology and Pharmacology, Western University, London, ON, Canada

<sup>6</sup>Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada

<sup>7</sup>Division of Nephrology, Department of Medicine, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada

\* these authors contributed equally

**Corresponding Author:**

Flory T Muanda, MD, PhD

Department of Physiology and Pharmacology

Western University

Medical Sciences Building, 1151 Richmond St, Room 287

London, ON, N6A 5C1

Canada

Phone: 1 519 661 2111 ext 84885

Email: [fmuatdat@uwo.ca](mailto:fmuatdat@uwo.ca)

## Abstract

**Background:** Drug-drug interactions (DDIs) are a major concern, especially for older adults taking multiple medications. Although Health Canada and the US Food and Drug Administration (FDA) use population-based studies to identify adverse drug events, detecting harmful DDIs is challenging due to the millions of potential drug combinations. Traditional pharmacoepidemiologic studies are slow and inefficient, often missing important harmful DDIs.

**Objective:** This protocol outlines a novel approach to efficiently identify harmful DDIs using administrative health care data.

**Methods:** Using high-throughput computing, we will conduct multiple population-based, new-user cohort studies using Ontario's linked administrative health care data. The cohorts will be selected from the population of Ontario residents aged 66 years and older who filled at least one oral outpatient drug prescription from 2002 to 2023. In each cohort, the exposed group will comprise individuals who are regular users of one drug (drug A) who start a new prescription for a second drug (drug B); the referent group will comprise regular users of drug A not taking drug B. We will evaluate 74 acute outcomes within 30 days of cohort entry, including hospitalizations, emergency department visits, and mortality. Propensity score methods will balance exposed and referent groups on more than 400 baseline health characteristics. Modified Poisson and binomial regression models will estimate risk ratios (RRs) and risk differences (RDs). To ensure findings are both statistically and clinically meaningful, we will apply prespecified thresholds for effect sizes (eg, lower bounds of 95% CIs  $\geq 1.33$  for RRs and  $\geq 0.1\%$  for RDs) and control the false discovery rate at 5% using the Benjamini-Hochberg procedure to address multiplicity. Subgroup and sensitivity analyses, including negative control outcomes and E-values, will assess robustness.

**Results:** In a preliminary analysis, we identified approximately 3.8 million older adults who filled prescriptions for over 500 unique medications during the study period (2002-2023), and therefore, approximately 200,000 potential drug combinations will be available for study. The initial drug pair cohorts had a median of 583 new users per cohort (IQR 237-2130); the median overlap in drug pair prescriptions was 57 days (IQR 30-90). The protocol was finalized on August 30, 2025, and outlines the analysis of data from 2002 to 2023. The analysis is scheduled to be completed by fall 2026, with results interpreted in 2027. The final manuscript submission is planned for December 2028.

**Conclusions:** This study aims to identify credible signals of harmful DDIs in older adults in routine care. This study will use an innovative approach that leverages data from provincial administrative health care databases and integrates high-throughput computing and rigorous pharmacoepidemiologic methods to generate robust real-world evidence that can inform safer prescribing practices and regulatory decision-making.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/77224

(*JMIR Res Protoc* 2025;14:e77224) doi: [10.2196/77224](https://doi.org/10.2196/77224)

## KEYWORDS

older adults; high-throughput computing; drug-drug interactions; machine learning; drug safety

## Introduction

Drug-drug interactions (DDIs) occur when two or more medications interact in ways that alter their effectiveness or safety [1,2]. The drug being affected is termed the *object drug*, while the drug causing the interaction is known as the *precipitant drug* [3,4]. The precipitant drug can influence the object drug through pharmacokinetic (PK) mechanisms (affecting absorption, distribution, metabolism, or excretion) or pharmacodynamic (PD) mechanisms (altering the drug's effects at its target site) [3]. An interaction may increase the effect of an object drug, leading to toxicity. Alternatively, an interaction may lessen the effect of an object drug, leading to reduced efficacy or treatment failure [5].

DDIs pose a significant risk for older adults due to polypharmacy and age-related changes in drug metabolism and elimination [6-8]. Approximately one in six hospitalizations among older adults is attributed to adverse drug events [9]. Additionally, older adults are often excluded from clinical trials, limiting our understanding of how DDIs affect this population. Drugs with a narrow therapeutic index are particularly concerning, as even minor interactions can have serious clinical consequences [10,11].

Although premarket clinical trials evaluate PK DDIs in healthy volunteers, many clinically significant DDIs are only identified after drug approval through case reports, small observational studies, and large-scale pharmacoepidemiologic research [12,13]. Traditional DDI detection methods often analyze one drug pair at a time, which can be slow and inefficient. These limitations are a concern worldwide: international efforts using electronic health records, claims data, or national prescription registries have made progress, but such approaches may not scale fast enough to keep pace with the thousands of new drug combinations entering the market each year, highlighting the need for rapid, large-scale, and systematic surveillance methods.

High-throughput computing and automated analytical pipelines have been applied in other contexts to rapidly screen for potential DDIs. For example, researchers in the United States and Europe have used these approaches with large electronic health care databases to identify clinically relevant DDI signals [14,15]. Bykov et al [14] validated a case-crossover-based method by successfully flagging known interactions in large US claims databases. Similarly, Leonard et al [15] demonstrated the ability of high-throughput pharmacoepidemiologic screening using a self-controlled case series design to confirm known DDI risks and identify novel potential interactions in real-world data.

Although Bykov et al [14] and Leonard et al [15] have demonstrated the potential of high-throughput screening for DDIs in US health care databases using self-controlled designs, our study complements and extends these approaches by applying high-throughput computing to Ontario's population-based, linked administrative health care data within a universal health care system using a population-based, new-user cohort design [16]. This will allow us to systematically evaluate over 200,000 medication combinations in older adults, integrating PK, PD, and epidemiologic evidence for robust, real-world signal detection at a scale and scope not previously achieved. By applying these methods in a Canadian context, we aim to identify potentially harmful DDIs that may warrant further validation, address key knowledge gaps, and generate findings relevant to both national and international health care systems. To the best of our knowledge, this represents the first application of such an approach in Canada, leveraging the strengths of a universal health care setting to further enhance the generalizability and clinical relevance of detected DDI signals. The results of this work have the potential to support safer prescribing practices, generate signals that could inform clinical decision-making once validated, and guide health care policy aimed at reducing adverse drug events in older adults.

This protocol outlines our plan to detect harmful DDIs more efficiently in routine care using high-throughput computing and automation. To do this, we will simultaneously conduct more than 200,000 population-based, new-user drug safety studies using 21 years of data from Ontario's administrative health care databases (2002-2023). These databases contain encrypted data on all health care visits and hospitalizations of Ontario residents—all of whom have universal access to hospital care and physician services through a government-funded single-payer system, with Ontarians aged 65 years and older also receiving universal outpatient prescription drug coverage.

Following the application of inclusion and exclusion criteria, we will analyze a refined set of medication cohorts, each containing an exposed group (individuals prescribed a certain pair of medications) and a referent group (individuals prescribed a single medication). The exposed and referent groups will be statistically balanced on more than 400 baseline health characteristics. We will assess 74 clinically relevant outcomes and conduct extensive sensitivity and bias analyses to minimize false discoveries. Harmful DDIs that meet prespecified criteria will be further examined in additional and exploratory analyses. This study's objective is to identify credible signals of harmful DDIs in older adults in routine care.

## Methods

### Study Design and Data Sources

Using high-throughput computing and automation, we will conduct over 200,000 population-based, new-user drug safety cohort studies using linked administrative health care databases in Ontario, Canada. The study period will span from July 1,

2002, to March 1, 2023. All study data will be obtained from Ontario’s linked administrative health care databases housed at the Institute for Clinical Evaluative Sciences (ICES) [17]. These databases contain encrypted, person-level data for Ontario residents, who have universal access to hospital care and physician services through a government-funded, single-payer system. A description of the databases is provided in [Table 1](#).

**Table 1.** Description of study databases.

Database	Description
<b>National databases/registries</b>	
Canadian Institute for Health Information’s Discharge Abstract Database (CIHI-DAD)	Contains diagnostic and procedural information for all hospitalizations, collected by each province
Canadian Organ Replacement Register (CORR)	National registry that includes information on vital organ transplantation and dialysis activities
National Ambulatory Care Reporting System (NACRS)	Contains information on hospital and community-based ambulatory care visits
<b>Ontario databases</b>	
Ontario Mental Health Reporting System (OMHRS)	Contains adult inpatient mental health information
Registered Persons Database (RPDB)	Contains information on patient demographics, including sex, birth dates, and death dates
ICES <sup>a</sup> Physician Database (IPDB)	Contains physician-related information, such as birth dates, sex, education, and specializations
Ontario Drug Database	Contains claims for prescription drugs received under the ODB <sup>b</sup> program
Ontario Health Insurance Plan (OHIP)	Contains diagnostic information and health claims for inpatient and outpatient services
Ontario Laboratories Information System (OLIS)	Contains laboratory test orders and results from hospitals, community labs, and public health labs

<sup>a</sup>ICES: Institute for Clinical Evaluative Sciences.

<sup>b</sup>ODB: Ontario Drug Benefit.

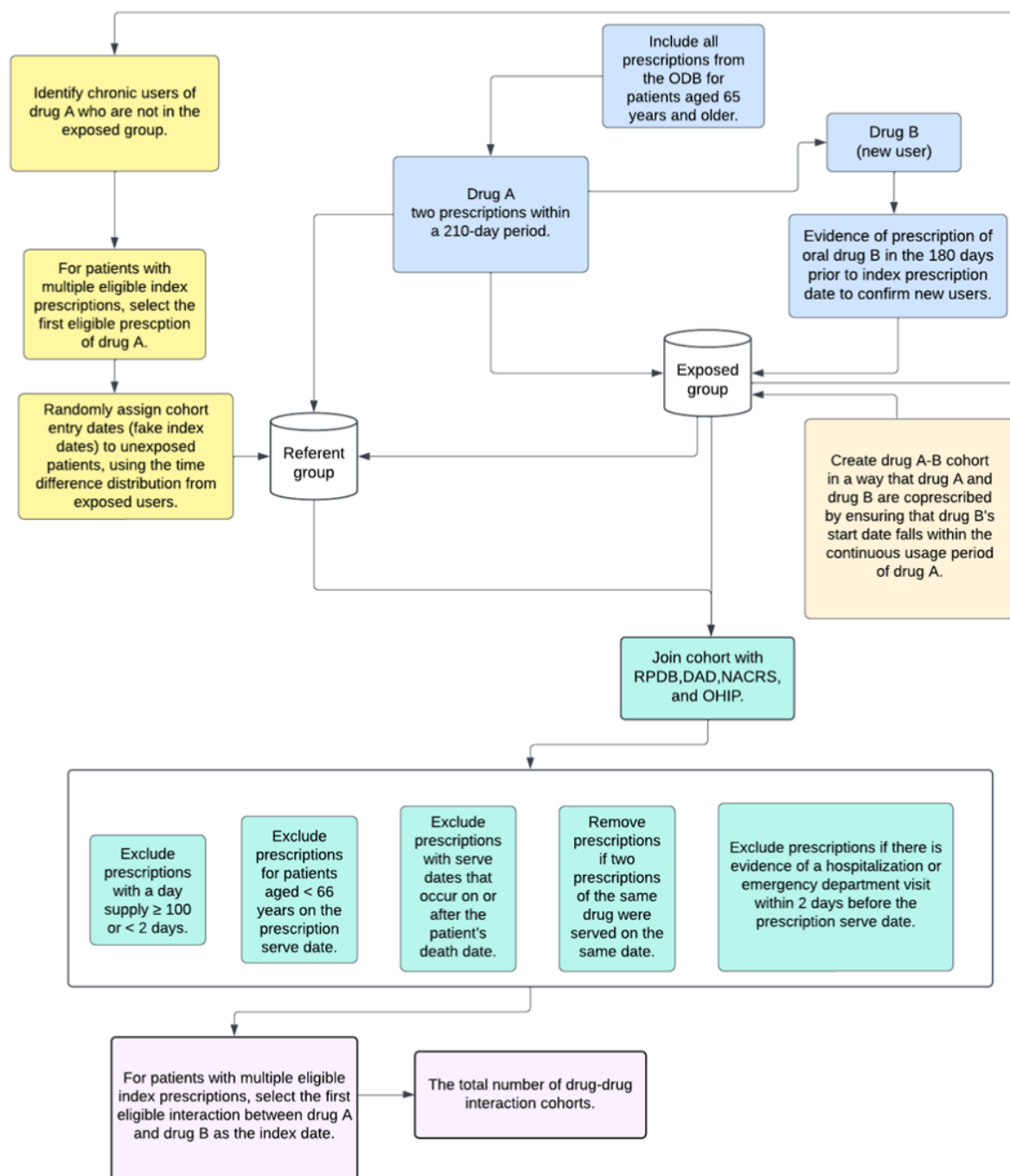
Residents aged 65 years and older receive universal outpatient prescription drug coverage through the Ontario Drug Benefit (ODB) program (while residents of all ages only receive universal prescription drug coverage for medications dispensed in emergency departments or during hospital admissions). Consequently, our study will leverage comprehensive outpatient prescription data for residents aged 65 years and older. We have previously used these databases to examine adverse drug events and health outcomes [16,18-21].

These databases are expected to be complete for all study variables, except for prescriber specialty (10%-20% missing), rural residence, and neighborhood income quintile (less than

2% missing). The only source of loss to follow-up is emigration from Ontario, which occurs at an estimated rate of 0.5% per year [22].

The study conduct and reporting will adhere to recommended guidelines for observational pharmacoepidemiology studies that use routinely collected health data ([Multimedia Appendix 1](#)) [23,24]. More details on the databases, study variables, and the codes used to define baseline comorbidities are provided in [Multimedia Appendix 2](#). The integration of data from different sources is summarized in [Multimedia Appendix 3](#). The process of building the cohorts is summarized in [Figure 1](#), and the analytic steps are summarized in [Multimedia Appendix 4](#).

**Figure 1.** Study flow diagram. DAD: Discharge Abstract Database; NACRS: National Ambulatory Care Reporting System; ODB: Ontario Drug Benefit; OHIP: Ontario Health Insurance Plan; RPDB: Registered Persons Database.



## Source Population

The source population will comprise adults aged 66 years and older who filled at least one oral outpatient drug prescription with a supply duration of at least 3 days, dispensed between July 1, 2002, and March 1, 2023 (codes for nonoral versions are provided in [Multimedia Appendix 5](#)). The age criterion will ensure that individuals in this population have at least 1 year of prior ODB prescription drug coverage.

## Cohort Builds

The study cohorts will be constructed using the process described later. Each cohort will be selected from the source population and contain an exposed group (individuals prescribed a pair of medications: drugs A and B) and a referent group (individuals prescribed drug A), as defined next.

### Regular Users of Drug A

We will first select individuals with at least two prescriptions for the same medication (drug A) within a 210-day period, with

a maximum allowable gap equivalent to 150% of the days' supply of the previous prescription.

### ***New Users of Drug B***

From the sample of regular users of drug A, we will select individuals who start a prescription for a new medication (drug B) that is taken concurrently with drug A, meaning drug B's start date is on the same date as the most recent drug A fill date or falls within drug A's supply duration.

### ***Cohort Entry Dates for Exposed and Referent Groups***

The exposed group will comprise regular users of drug A and new users of drug B, as described earlier. The cohort entry date for individuals in the exposed group will be the start date of drug B. The referent group will comprise regular users of drug A who are not coprescribed drug B. The cohort entry date for individuals in the referent group will be assigned based on the distribution of the time differences between the most recent prescription for drug A and the initiation of drug B for individuals in the exposed group, ensuring similar observation periods between groups [25].

### ***Individual-Level Exclusion Criteria Applied to the Exposed and Referent Groups***

We will exclude individuals who meet any of the following criteria:

- Individuals prescribed drug B within 180 days before the cohort entry date (to ensure those in the exposed group are new users of drug B and that those in the referent group were not previously exposed).
- Individuals prescribed another medication in the same class as drugs A and B on the cohort entry date. For example, if

drug A or B is sertraline, which is a serotonin reuptake inhibitor (SSRI), individuals prescribed a nonsertraline SSRI on the cohort entry date will be excluded.

- Individuals discharged from the hospital or emergency department within 2 days before the cohort entry date (to ensure that prescriptions are new outpatient prescriptions).
- Individuals whose day supplies for drug A or B are  $\geq 100$  days or  $< 2$  days.
- Individuals who are nonpermanent Ontario residents, are aged  $< 66$  or  $> 105$  years on the cohort entry date, or have a date of death on or before the cohort entry date (data cleaning exclusions).

An individual can only enter a cohort once; if an individual meets the eligibility criteria multiple times, the first eligibility date will serve as the cohort entry date.

### ***Cohort-Level Exclusion Criteria***

To ensure meaningful results, only cohorts with a minimum sample size of 1000 patients, with at least 200 individuals in the exposed group and 200 individuals in the referent group, will be considered for further analysis.

### ***Outcomes***

We will examine 74 outcomes within 30 days of cohort entry (Textbox 1). The outcomes were chosen using a prespecified analytic approach and considering the clinical significance of the outcomes and the observed rates in the study region [16]. Many of the algorithms used to identify these outcomes have been validated in previous studies and demonstrate high sensitivity and specificity [26-30]. Shorter and longer follow-up times for outcomes will be examined in additional analysis (see the *Additional Analysis to Characterize the Signals* section).

**Textbox 1.** Final set of 74 clinically relevant outcomes (ordered alphabetically).

**Clinically relevant individual outcomes (n=56):**

- Abdominal aortic aneurysm repair/aortic bypass
- Achilles tendon rupture
- All-cause mortality (prespecified)
- Arthroscopy
- Asthma
- Atrial fibrillation/flutter or other arrhythmia (pacemaker insertion, palpitations, tachycardia unspecified, atrioventricular block, supraventricular tachycardia, or other conduction disorders)
- Bowel obstruction
- Bradycardia
- Bronchoscopy
- Carotid endarterectomy
- Carotid ultrasound
- Cataract, unspecified
- Chest X-ray
- Cholecystitis, unspecified
- Clostridium difficile colitis
- Colposcopy
- Computed tomography: abdomen
- Computed tomography: head
- Computed tomography: thorax
- Cutaneous drug reaction
- Cystoscopy
- Depression
- Disorders of calcium metabolism
- Emergency department visit for any reason (prespecified)
- Gout
- Guillain-Barré syndrome
- Head trauma
- Hearing loss
- Hospital admission for any reason (prespecified)
- Hyperkalemia
- Hypoglycemia
- Hypokalemia
- Hyponatremia
- Intensive care unit admission
- Kidney stones
- Laparoscopy
- Liver disease toxicity
- Lower limb amputation
- Major abdominal surgery
- Major cardiovascular surgery
- Osteoarthritis



- Pancreatitis
- Parkinson's disease
- Physician visit for any reason (prespecified)
- Pneumonia
- Renal biopsy
- Retinal detachment
- Rhabdomyolysis
- Rheumatoid arthritis
- Self-harm
- Sepsis
- Surgery (for any reason other than major abdominal surgery, major cardiovascular surgery, lower limb amputation and fracture, and major ortho procedure)
- Severe respiratory depression
- Urinary incontinence
- Urinary retention
- Venous thromboembolism

**Clinically relevant composite outcomes (n=18):**

- Acute dialysis or acute kidney injury
- Aplastic anemia (unspecified), neutropenia, or other agranulocytosis, or thrombocytopenia (unspecified)
- Cardiac arrest or ventricular arrhythmia
- Cardiac catheterization or coronary revascularization
- Colonoscopy or esophagogastroduodenoscopy
- Diabetic ketoacidosis, hyperglycemia, or hyperosmolar nonketotic coma
- Delirium or other encephalopathy
- Delirium, fall, or other encephalopathy
- Echocardiography or Holter monitoring
- Electroencephalography or seizures
- Emergency department visit or hospital admission for any reason (prespecified)
- Fall, fracture or major orthopedic procedure, or fracture
- Fall or hypotension
- Fall or syncope
- Heart failure or myocardial infarction (most-responsible diagnosis), or ischemic stroke (most-responsible diagnosis)
- Hemarthrosis or hemorrhage (not elsewhere classified), intracerebral hemorrhage (unspecified), lower gastrointestinal bleed, other nontraumatic intracranial hemorrhage, subarachnoid hemorrhage (unspecified), or upper gastrointestinal bleeding
- Hypotension or syncope
- Urinary tract infection or urine culture

## Statistical Analysis Plan

All analyses will be programmed and automated using Python, SAS, the JavaScript library D3, and R (R Foundation for Statistical Computing) and will be conducted on the Health Artificial Intelligence Data Analytic Platform, a high-performance computing environment for analyzing ICES data. The open source JavaScript library D3 will be used to develop interactive visualizations.

## *Analysis to Balance the Exposed and Referent Groups in Each Cohort*

Propensity scores will be calculated using multivariable logistic regression, incorporating more than 400 baseline characteristics (the complete list and coding definitions are provided in [Multimedia Appendix 2](#)). Characteristics include calendar year, prior use of drug A, demographics (eg, age, sex, neighborhood income quintile, urban vs rural residence, geographic region [31], and long-term care home residency), comorbidities (eg,

asthma, diabetes, coronary artery disease) assessed over the 5 years before cohort entry, health care use in the year prior (eg, hospitalizations, emergency room visits, specialist consultations), and prescription drug dispensing within 120 days before cohort entry. The 120-day lookback period accounts for potential delays in prescription refills, ensuring comprehensive capture of medication dispensing.

The propensity scores will estimate the probability of a patient being in the exposed group versus the referent group, given their baseline characteristics. To enhance the comparability of the exposed and referent groups, we will exclude (1) individuals in the referent group with propensity scores lower than the lowest 1% of scores in the exposed group and (2) individuals in the exposed group with propensity scores higher than the highest 99% of scores in the referent group. This will also ensure that we do not violate the positivity assumption in causal inference, which requires that there should be a nonnull probability across each stratum of baseline characteristics within both the exposed and referent groups [32].

Inverse probability of treatment weighting (IPTW) using average treatment-effect-in-the-treated (ATT) weights will then be used to balance the exposed and referent groups on baseline characteristics as follows: Individuals in the referent group will be weighted using ATT weights, defined as their odds of being in the exposed group ( $\text{propensity score}/[1 - \text{propensity score}]$ ). Individuals in the exposed group (the medication pair group) will receive weights of 1. This method will produce a weighted pseudosample of individuals in the referent group with a distribution of measured characteristics similar to that of the exposed group [33,34]. Between-group differences in baseline characteristics in the weighted samples will be examined using standardized differences (SDs) [35].

Cohorts will be excluded from further analysis if fewer than 95% of the baseline characteristics are balanced between the exposed and referent groups (ie, the SDs must be  $<10\%$  for  $\geq 95\%$  of the characteristics).

### Regression Analyses

A visual summary of the regression and additional exploratory analyses is provided in [Multimedia Appendix 4](#). In each cohort, outcomes will be eligible for analysis if there are at least six outcome events in both the exposed and referent groups. Reporting fewer than six events could lead to potential patient reidentification and substantial estimation errors.

For each eligible cohort, we will use a modified Poisson regression model [36] to estimate the weighted risk ratio (RR) and its 95% CI for each outcome in the exposed group compared to the referent group. Additionally, a binomial regression model with an identity link function will be used to estimate the weighted risk difference (RD) and its 95% CI. If all 74 outcomes for a particular cohort are eligible, up to 74 regression models will be run.

### Criteria for Identifying Potential Signals of Harmful DDIs

Associations between a drug pair and an adverse outcome that meet the following criteria will be considered potential signals of harm and analyzed further:

- Thresholds for RRs and RDs. The lower bound of the 95% CI for the RR must be  $\geq 1.33$ , and for the RD, it must be  $\geq 0.1\%$ . These thresholds were chosen to ensure that only clinically significant increases in risk (of sufficient magnitude to be of concern) are considered. This analysis will focus solely on potential harms from drug combinations rather than protective benefits.
- Statistical significance. The  $P$  values from the modified Poisson regression model (for RRs) and the binomial regression model (for RDs) must be statistically significant after applying the Benjamini-Hochberg correction. This correction accounts for multiple comparisons across all outcomes and drug pair cohorts. The false discovery rate will be set at 5%, and the highest  $P$  value below the Benjamini-Hochberg critical value will be considered statistically significant, along with all  $P$  values below it. This correction will minimize the risk of false-positive discoveries (type I errors).

Drug pair cohorts that do not generate significant signals will not undergo further analysis.

### Sensitivity and Bias Analyses to Guard Against Spurious Discoveries

Signals of potentially harmful DDIs will undergo the following sensitivity and bias analyses to ensure that identified signals are not artifacts of statistical or methodological bias but are credible associations worthy of further investigation:

- Survival analysis. To assess whether the initial signal persists when analyzed using a time-to-event approach, we will conduct a Cox proportional hazards regression with a 30-day follow-up period, censoring on death when all-cause mortality is not the outcome of interest. We will test the proportional hazards assumption by including an interaction term between the new-user status and follow-up time. This analysis will help determine whether the signal is robust across different modeling techniques.
- Negative-control exposure analyses. To detect potential biases from unobserved confounders, we will perform negative-control exposure analyses. A negative-control exposure is a variable that shares the same sources of bias as the drug of interest but cannot plausibly cause the outcome. We will rerun the regression analyses for the initial signals using a control exposure period starting 90 days before cohort entry [37]. This analysis will help identify any confounding factors that may have influenced the initial signal. Notably, the outcome of all-cause mortality cannot be examined with this method, as all patients lived to their cohort entry date. A positive association (ie, RRs and RDs significantly different from null, with  $P < .05$ ) in this analysis would suggest potential confounding, whereas the absence of an association would bolster the credibility of the original signal.



- E-value analysis. We will conduct E-value analyses to quantify the minimum strength of association an unmeasured confounder would need to have with both the drug combination and the outcome to negate the observed signal [38]. Larger E-values indicate that substantial unmeasured confounding would be required to explain away the signal, whereas smaller E-values suggest that minimal unmeasured confounding could do so. For an initial signal to be considered robust, the E-value must exceed 2.
- Propensity score matching. To further validate the robustness of the signal, we will use 1:1 propensity score matching to balance the exposed and referent groups on baseline characteristics [39]. This method estimates the average treatment effect in the treated and mitigates the influence of extreme weights, which can occur with IPTW. Consistency between the results of this analysis and the initial findings will strengthen the validity of the signal.
- Bootstrap analysis. We will use a multiplier bootstrap procedure to compute RRs and RDs across 200 random samples, each of the same size as the original cohort. This method helps correct SEs and CIs, ensuring they have the correct coverage rates [36,40]. For the signal to be considered robust, at least 90% of the bootstrap samples must yield results consistent with the original findings, with  $P < .05$  and CIs for RRs and RDs in the same direction as initially observed.

### ***Additional Analysis to Characterize the Signals***

Signals of potentially harmful DDIs that are supported by the aforementioned sensitivity and bias analyses will be further explored and characterized as follows:

- Longer and shorter follow-up periods. To better understand the effects of prolonged medication exposure, we will extend the follow-up period beyond 30 days. We will rerun the regression analyses to examine the RRs and RDs for outcomes occurring within 60 and 90 days after cohort entry. We will also assess shorter follow-up periods by examining outcomes that occur within 7 and 14 days of cohort entry.
- Number needed to harm (NNH). To further quantify the absolute risk, we will calculate the NNH, which is derived from the RD ( $NNH = 1/RD$ ). This metric indicates how many patients would need to take the medication combination to result in harm to one additional patient who otherwise would not have been harmed. A lower NNH indicates a greater potential for harm [41].
- Population attributable fraction (PAF). We will calculate the PAF using the formula  $Pe(\text{relative risk} - 1) / [1 + Pe(\text{relative risk} - 1)]$ , where  $Pe$  represents the proportion of the population exposed to the drug combination [42]. A PAF close to 1 suggests that eliminating the drug pair exposure could significantly reduce the number of adverse outcomes in the population, while a PAF close to 0 implies that reducing exposure would have minimal impact at the population level. Separate PAFs will be calculated for each outcome, and the results will be combined as a composite outcome before calculating the PAF for each cohort.

### ***Bidirectional Exposure Analysis to Assess Signal Robustness***

We will specifically highlight instances where the addition of drug B to drug A, and, conversely, drug A to drug B, yield concordant results for the outcome of interest. Observing consistent effects in both directions, particularly for chronic medications, will strengthen confidence in the validity of the signal and suggest that findings are less likely to be explained by residual confounding or treatment selection bias.

### ***Exploring the Data Using Interactive Visualization***

To facilitate data exploration, we will develop an interactive visualization tool. The tool will enable users to visually explore the outputs from our analyses, providing an overview of the data with the ability to access, filter, and modify the displayed information [43,44]. Users will be able to apply filters based on RRs or other metrics to identify associations that meet specific criteria. The tool will support manual exploration, enabling users to interactively query and sort the data.

### ***Overall Framework to Identify, Characterize, and Validate Signals of Harmful DDIs***

Signals identified in the high-throughput computing analysis will be rigorously evaluated for clinical relevance by examining evidence from PK, PD, and epidemiological studies, as well as by using clinical validation tools to prioritize and confirm the most clinically significant interactions as follows (Figure 2):

- Step 1: signal identification. Signals of potentially harmful DDIs identified in the high-throughput analysis must (1) meet predefined thresholds for RRs and RDs, (2) remain statistically significant after the Benjamini-Hochberg correction for multiple comparisons, and (3) be supported in sensitivity and bias analyses.
- Step 2: PK and PD assessment. PK and PD evidence will be evaluated to assess the biological plausibility of the interaction. For PK assessment, published studies on enzyme inhibition, drug metabolism, and clearance will be reviewed. For example, sulfamethoxazole inhibits CYP2C9, which metabolizes warfarin, leading to increased warfarin levels and a higher risk of bleeding [11,45]. For PD assessment, synergistic and antagonistic interactions will be examined. For example, the concurrent use of two antiplatelet agents can increase the risk of bleeding.
- Step 3: cross-validation with drug interaction checkers and product monographs. Signals will be cross-validated using DDI checkers, such as Lexicomp, Micromedex, and Drug Interaction Facts, along with regulatory product monographs (eg, from Health Canada and the US Food and Drug Administration [FDA]). Interaction severity will be classified as major, moderate, or minor based on standardized clinical guidelines. Recommended interventions for major interactions include (1) avoiding coadministration, (2) adjusting doses, and (3) monitoring relevant biomarkers (eg, the International Normalized Ratio [INR] for warfarin interactions).
- Step 4: prioritization of high-quality epidemiological evidence. Evidence from epidemiological studies will be examined and prioritized as randomized controlled trials

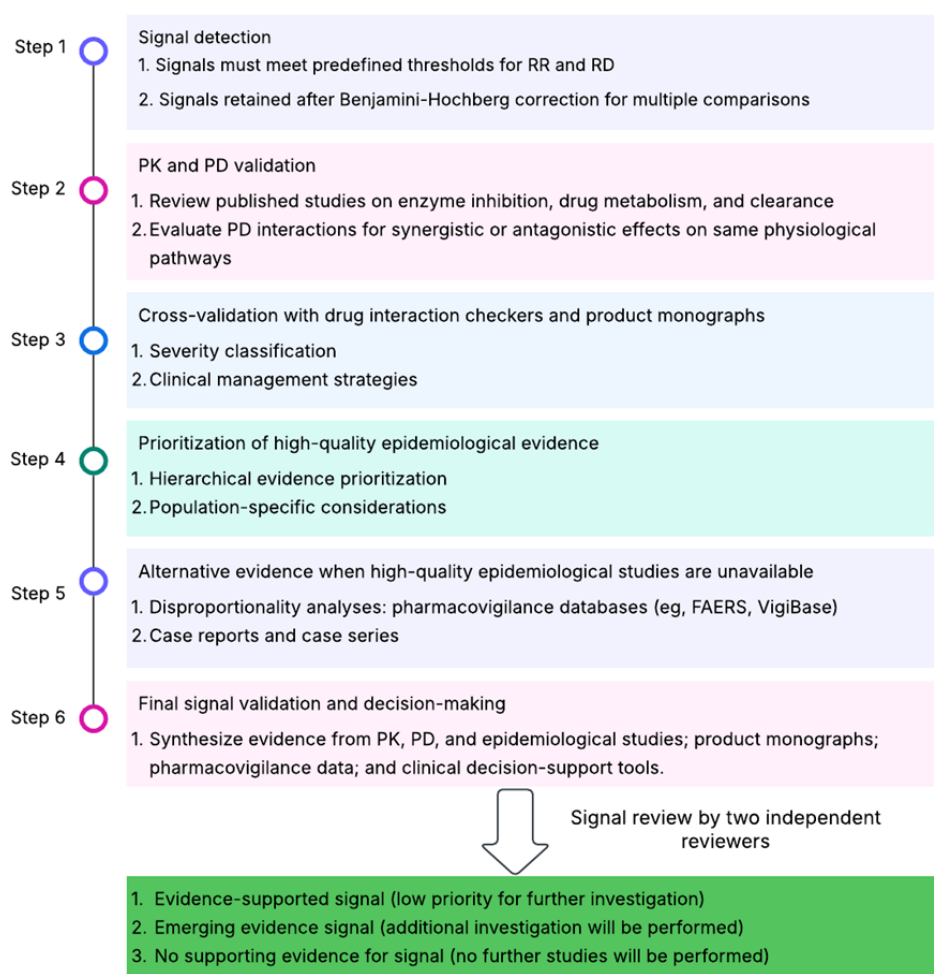
(RCTs), high-quality cohort studies, and case-control studies. We will also prioritize evidence from studies of vulnerable populations (eg, older adults, patients with chronic kidney disease, and polypharmacy cohorts).

- Step 5: alternative evidence when high-quality epidemiological studies are unavailable. Complementary signal detection methods (eg, disproportionality analyses of pharmacovigilance databases, such as FAERS and VigiBase), case reports, and case series will be reviewed to assess the interaction's potential clinical significance when high-quality epidemiological evidence is unavailable.

For disproportionality analyses, pharmacovigilance databases (eg, FAERS, VigiBase) will be analyzed to determine whether drug coadministration is disproportionately associated with adverse events compared to the administration of a single medication. When large-scale epidemiological data are lacking, case reports and case series may provide supporting evidence.

- Step 6: final signal validation and decision-making. Evidence from PK, PD, and epidemiological studies, pharmacovigilance data, and clinical decision-support tools will be synthesized to assess the credibility of the signal.

**Figure 2.** Framework to identify, characterize, and validate signals of harmful DDIs. DDI: drug-drug interaction; PD: pharmacodynamic; PK: pharmacokinetic; RD: risk difference; RR: risk ratio.



Each signal will be independently reviewed by two reviewers using the aforementioned criteria. The reviewers will classify each signal based on the strength and consistency of supporting data as an evidence-supported signal, an emerging evidence signal, or a signal with no supporting evidence. To minimize subjectivity, we will calculate an interrater agreement measure (eg, Cohen  $\kappa$ ) to assess the level of concordance between the reviewers and determine whether they reach the same conclusion [46,47]. Any discrepancies will be resolved through discussion and, if needed, adjudicated by a third reviewer.

The criteria for classifying a signal as an evidence-supported signal (regulatory and clinical action may be warranted), an emerging evidence signal (additional investigation will be performed), or a signal having no supporting evidence (no further studies will be performed) are summarized in Table 2. Briefly, a signal will be considered evidence supported if multiple lines of evidence support a causal effect, it will require further study if the existing evidence is conflicting, and it will be dismissed in the absence of supportive evidence.

**Table 2.** Summary table for criteria for signal confirmation, further study, or dismissal.

Category and criteria	Key indicators	Next steps
<b>Evidence-supported signal<sup>a</sup> (low priority for further investigation)</b>		
Multiple independent sources align	<ul style="list-style-type: none"><li>• PK<sup>b</sup>/PD<sup>c</sup> evidence supports a plausible interaction.</li><li>• High-quality epidemiological data (RCTs<sup>d</sup>, cohort studies) confirm increased risk.</li><li>• Product monographs or clinical decision-support tools (eg, Lexicomp) classify it as a major interaction.</li></ul>	Regulatory and clinical action may be warranted, such as labeling changes or guideline updates.
<b>Emerging evidence signal<sup>a</sup> (additional investigation to be performed)</b>		
Conflicting evidence	<ul style="list-style-type: none"><li>• PK/PD data suggest a risk, but epidemiological studies do not confirm it.</li><li>• Epidemiological findings are inconsistent or based on low-quality studies (eg, small sample size, case reports only).</li><li>• Product monographs or clinical decision support tools give mixed or uncertain classifications.</li></ul>	A pharmacoepidemiological study should be conducted.
Limited population-specific data	<ul style="list-style-type: none"><li>• High-risk populations (patients with chronic kidney disease, older adults, polypharmacy patients) are underrepresented in studies.</li><li>• Current evidence is based on general population data but lacks subgroup analyses.</li></ul>	A pharmacoepidemiological study should be conducted.
Emerging signal in pharmacovigilance databases	<ul style="list-style-type: none"><li>• Disproportionality analysis (eg, FAERS, Vigibase) suggests a safety concern but lacks validation in clinical studies.</li><li>• Case reports describe severe but rare adverse events needing further confirmation.</li></ul>	A pharmacoepidemiological study should be conducted.
<b>No supporting evidence for signal<sup>a</sup> (no further studies to be performed)</b>		
No clear biological or clinical risk	<ul style="list-style-type: none"><li>• There is a lack of PK/PD plausibility (eg, drugs do not share metabolic pathways or receptor targets).</li><li>• There is no statistically significant association in large epidemiological studies.</li><li>• There are no case reports/case series.</li><li>• Product monographs or clinical decision-support tools classify it as a low-risk (minor) interaction or report no interaction.</li></ul>	No further action is needed.

<sup>a</sup>Signals retained after Benjamini-Hochberg correction for multiple comparisons.

<sup>b</sup>PK: pharmacokinetic.

<sup>c</sup>PD: pharmacodynamic.

<sup>d</sup>RCT: randomized controlled trial.

**Exploring Harmful DDIs Using Machine Learning**

We will explore an alternative strategy for further investigating harmful DDIs using a machine learning algorithm. This complementary approach will integrate key components from our six-step validation framework, including empirical evidence from PK and PD data; findings from epidemiological studies; and clinical validation tools, such as drug interaction checkers (Lexicomp, Micromedex) and product monographs (eg, from

Health Canada and the US FDA). We will combine these diverse sources of evidence to train the model, aiming to identify patterns that can help prioritize high-risk DDIs for further investigation. We will also use natural language processing (NLP) algorithms to extract important insights from clinical monographs, published research papers, and other medical literature. This step will help us better understand the context around each signal and provide a richer foundation for analysis. Although still exploratory, this method seeks to refine our ability

to distinguish clinically significant interactions from spurious associations, ultimately improving how we prioritize potential risks in clinical practice.

We will first use known, previously validated DDIs identified through our six-step validation process to train the model. All sources of evidence collected—including high-throughput computing, PK and PD data, clinical validation tools (eg, drug interaction checkers and product monographs), epidemiological studies, and pharmacovigilance databases—along with insights from the six-step validation framework, will form a comprehensive training dataset for our machine learning model, enabling it to generate risk scores and rank potential interactions based on predicted clinical significance. To ensure robustness, the model will undergo k-fold cross-validation and bootstrap sampling. We will apply interpretability techniques, such as Shapley additive explanations (SHAP), to identify the most influential factors driving each prediction. By automating signal ranking and prioritization, this machine learning approach will explore new opportunities to enhance the identification of clinically meaningful DDIs, supporting safer prescribing and improved postmarket drug surveillance.

## Ethical Considerations

This research is exempt from Research Ethics Board (REB) review because the data to be used are authorized under Section 45 of Ontario's Personal Health Information Protection Act (PHIPA), which allows for the use of personal health information for health system evaluation and planning without requiring REB review [48]. The data will be analyzed at ICES [17], an independent, not-for-profit research institute authorized under PHIPA to collect and use personal health information without consent for these authorized purposes. Individual consent from Ontario residents will not be required for this study, as ICES's legal status under PHIPA authorizes the secondary use of provincial health care records.

All analyses will be performed at ICES using linked, coded (deidentified) data. Researchers will not have access to any identifiable personal information, and strict privacy and security safeguards will be in place to protect confidentiality. No compensation will be provided, as this study will use existing health administrative data and will not involve direct participant recruitment or interaction. This study will not include images, videos, or any identifiable information from individual participants. On completion, the results of this population-based study will be submitted to a peer-reviewed biomedical journal for publication and presented at relevant conferences.

## Results

In a preliminary analysis, we identified approximately 3.8 million older adults who filled prescriptions for over 500 unique medications during the study period. Each drug can serve as drug A or drug B in various combinations, resulting in approximately 200,000 potential drug pairs for analysis. The initial drug pair cohorts had a median of 583 new users per cohort (IQR 237-2130), with a median overlap duration of 57 days (IQR 30-90). For example, in a chronic-chronic drug pair cohort where levothyroxine was drug A and rosuvastatin was

drug B, the median overlap duration was 135 days among 128,762 patients, with 73.1% (n=94,036) females and a median age of 66 years. In a chronic-acute drug pair cohort (eg, levothyroxine as drug A and amoxicillin trihydrate as drug B), the median overlap was 10 days among 197,057 patients, with 74.8% (n=147,399) females and a median age of 67 years. In an acute-chronic drug pair cohort (eg, amoxicillin trihydrate as drug A and rosuvastatin as drug B), the median overlap was 10 days among 2702 patients, with 51.6% (n=1394) females and a median age of 66 years. Lastly, in an acute-acute drug pair cohort (eg, amoxicillin trihydrate as drug A and ciprofloxacin as drug B), the median overlap was 10 days among 4220 patients, with 59.4% (n=2507) females and a median age of 70 years. The protocol, finalized on August 30, 2025, outlines the analysis of data from 2002 to 2023. The high-throughput computing analysis is planned for completion by fall 2026, and findings from the analysis will be assessed in 2027 for clinical relevance by integrating PK, PD, and epidemiological evidence, supported by clinical validation tools. The final results will be prepared for dissemination, with manuscript submission planned for December 2028.

## Discussion

### Summary

This study aims to identify harmful DDIs by analyzing Ontario health care databases using high-throughput computing. We will process data from over 3.8 million residents aged 66 years and older, evaluating more than 200,000 medication pairs and examining associations with 74 acute outcomes, including hospitalizations, emergency department visits, and mortality. High-throughput computing will enable us to identify and test thousands of medication combinations. We will compare individuals prescribed medication pairs with those on single medications, using propensity score methods to balance comparison groups on over 400 baseline health characteristics. We will conduct extensive sensitivity and bias analyses to ensure that potential signals are not artifacts of statistical or methodological bias. Additionally, machine learning algorithms will be used to complement the high-throughput computing approach.

High-throughput computing and automated analytical pipelines have proven effective for detecting potential DDIs in large health care datasets in other countries, such as the United States. Previous studies by Bykov et al [14] and Leonard et al [15] have demonstrated that these methods can efficiently confirm known interactions and identify novel signals. However, these approaches have largely been limited to single-object drugs, specific outcomes, or nonrepresentative datasets (derived from a specific subset of the population, such as those with a certain type of insurance, rather than the entire population), which may restrict generalizability. Building on these advances, our study will use high-throughput computing to analyze Ontario's population-based, linked administrative databases, which capture comprehensive prescription and outcome data across the province's universal health care system. Using a population-based, new-user cohort design, we will systematically evaluate nearly 200,000 medication pairs among



3.8 million older adults—an effort that would be impractical with traditional methods. This approach will not only fill a critical gap by providing large-scale, real-world evidence in a Canadian context but also enhance generalizability, generate signals that could inform clinical decision-making once validated, and strengthen postmarketing pharmacovigilance efforts for older adults. To the best of our knowledge, this represents the first application of such an approach in Canada, leveraging the strengths of a universal health care setting to further enhance the generalizability and clinical relevance of detected DDI signals.

The cohort study is designed to mimic typical clinical scenarios, for example, an individual who has been taking digoxin for 7 months and is then newly prescribed sulfamethoxazole-trimethoprim to treat a urinary tract infection. In DDI studies, such scenarios are categorized as follows:

- Precipitant triggered: Individuals initiate the object drug first, followed by the precipitant, with follow-up starting on the day the precipitant is initiated.
- Object triggered: Individuals initiate the precipitant first, followed by the object drug, with follow-up starting on the day the object is initiated.

Our approach will be hypothesis free since it does not assume beforehand whether drug A or drug B will function as the object or the precipitant. Analyzing initiation orders separately is highly recommended in traditional DDI studies—if feasible—as it addresses different clinical questions, better controls confounding, and increases the validity of the results [2].

### Strengths of the Study

A key strength of this study is the use of Ontario's administrative health care databases. These databases include comprehensive, encrypted data on health care visits, hospitalizations, lab results, and prescription records for residents aged 65 years and older [17]. This extensive data coverage enables the capture of a wide range of adverse outcomes, providing a detailed picture of the potential risks associated with specific drug interactions [16].

The application of high-throughput computing and automation will significantly enhance the efficiency and scalability of our analyses [49]. This approach will allow us to simultaneously evaluate numerous drug interactions, which would be impractical using traditional methods. By automating the creation of cohorts and statistical analysis, we can quickly and accurately identify potential DDIs, facilitating more rapid updates to clinical guidelines and regulatory policies [4,50].

### Challenges and Limitations

Despite using advanced statistical techniques to control for confounding, residual confounding may still be present [51]. Factors that are not captured in the administrative data or that are measured with error could bias our results [52]. To address this, we will conduct multiple sensitivity analyses and use

negative control outcomes to assess the robustness of our findings [37].

Our study focuses on acute (30-day) outcomes, which will not capture the effect of long-term exposure [53]; however, we will conduct additional analyses to assess longer follow-up times.

The study is designed to identify harmful DDIs with a moderate-to-high risk in magnitude in cohorts of at least 1000 individuals, and therefore, it will have limited statistical power to detect small risk increases.

Our study does not include data on nonprescription (over-the-counter) drugs, drugs prescribed in hospitals, or comprehensive outpatient prescription drug use in children or adults aged under 65 years. These exclusions may limit the generalizability of our findings to a broader population. However, the focus on older adults is justified, given their higher risk of polypharmacy and adverse drug interactions.

Given the large number of comparisons being made, there is a risk of spurious findings due to multiple testing [16]. To mitigate this, we will apply the Benjamini-Hochberg procedure to control the false discovery rate and use E-values to assess the robustness of our findings against unmeasured confounding [31].

### Implications for Clinical Practice and Policy

This study has the potential to significantly improve drug safety, particularly in older adults, by identifying harmful DDIs more efficiently. If replicated in other jurisdictions, the findings could provide clinicians with new insights into the risks associated with specific drug combinations, leading to more informed prescribing practices and enhanced patient monitoring. The evidence generated could also inform regulatory agencies, such as Health Canada and the US FDA, helping refine product label warnings and support regulatory decisions. Ultimately, the findings could contribute to safer medication use on a population level, improving medication safety practices worldwide. Moving forward, we will prioritize and verify high-throughput signals across Ontario and beyond, ensuring the continued refinement and application of this approach to achieve more precise and informed medication safety practices.

### Conclusion

This protocol outlines an innovative approach to detecting harmful DDIs in older adults using high-throughput computing and automation. By leveraging Ontario health care data and using advanced analytical techniques, we aim to enhance the efficiency and comprehensiveness of postmarket drug surveillance. The results of this study, if replicated in other jurisdictions, have the potential to significantly improve medication safety, inform clinical practice, and guide regulatory policies. Ultimately, this research aims to reduce the burden of adverse drug events in older adults and lay the foundation for more precise, data-driven medication safety practices worldwide.

### Acknowledgments

This study was funded by the Western Interdisciplinary Development Initiatives (IDI; grant award number 55760). The study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). The study was completed at the ICES Western site,



where core funding is provided by the Academic Medical Organization of Southwestern Ontario, the Schulich School of Medicine and Dentistry, Western University, and the Lawson Health Research Institute. This document used data adapted from the Statistics Canada Postal CodeOM Conversion File, which is based on data licensed from the Canada Post Corporation, and data adapted from the Ontario Ministry of Health Postal Code Conversion File, which contains data copied under license from the Canada Post Corporation and Statistics Canada. We thank IQVIA Solutions Canada Inc for use of their Drug Information File. Parts of this material are based on data and information compiled and provided by the Canadian Institutes of Health Information (CIHI) and by the Ontario MOH. The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred. The study sponsors played no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The conclusions, opinions, and statements expressed herein are those of the authors and do not necessarily reflect those of the funding or data sources; no endorsement is intended or should be inferred.

---

## Authors' Contributions

FTM, SSA, NR, AXG, JS, EM, KS, and MW obtained the funding and developed the study idea. FTM, SSA, AXG, EM, and RS designed the study. FTM, SSA, AXG, EM, NC, and RS contributed to the study design, methodology, and analysis plan. NR and RS drafted the initial manuscript, and all authors critiqued the protocol manuscript. All authors have approved the attached manuscript for publication and are accountable for all aspects of the work.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Checklist of recommendations for reporting of observational studies using the RECORD (Reporting of Studies Conducted Using Observational Routinely Collected Health Data) statement.

[\[DOCX File , 21 KB-Multimedia Appendix 1\]](#)

---

## Multimedia Appendix 2

Study variables and the codes used to ascertain baseline comorbidities (sorted alphabetically).

[\[DOCX File , 35 KB-Multimedia Appendix 2\]](#)

---

## Multimedia Appendix 3

Process of integrating data from different sources.

[\[PNG File , 708 KB-Multimedia Appendix 3\]](#)

---

## Multimedia Appendix 4

Overview of the plan to perform the primary analysis and additional exploratory analysis.

[\[PNG File , 179 KB-Multimedia Appendix 4\]](#)

---

## Multimedia Appendix 5

Codes for nonoral drugs.

[\[DOCX File , 14 KB-Multimedia Appendix 5\]](#)

---

## References

1. Rodrigues AD. Drug-Drug Interactions. Boca Raton, FL. CRC Press; 2019.
2. Chen C, Pham Nguyen TP, Hughes JE, Hennessy S, Leonard CE, Miano TA, et al. Evaluation of drug–drug interactions in pharmacoepidemiologic research. *Pharmacoepidemiol Drug Saf*. Jan 13, 2025;34(1):e70088. [[FREE Full text](#)] [doi: [10.1002/pds.70088](https://doi.org/10.1002/pds.70088)] [Medline: [39805810](https://pubmed.ncbi.nlm.nih.gov/39805810/)]
3. Aronson JK, Grahame-Smith DG. Clinical pharmacology. Adverse drug interactions. *Br Med J (Clin Res Ed)*. Jan 24, 1981;282(6260):288-291. [[FREE Full text](#)] [doi: [10.1136/bmj.282.6260.288](https://doi.org/10.1136/bmj.282.6260.288)] [Medline: [6779990](https://pubmed.ncbi.nlm.nih.gov/6779990/)]
4. Hennessy S, Leonard C, Gagne J, Flory J, Han X, Brensinger C, et al. Pharmacoepidemiologic methods for studying the health effects of drug-drug interactions. *Clin Pharmacol Ther*. Jan 23, 2016;99(1):92-100. [[FREE Full text](#)] [doi: [10.1002/cpt.277](https://doi.org/10.1002/cpt.277)] [Medline: [26479278](https://pubmed.ncbi.nlm.nih.gov/26479278/)]

5. Becker ML, Kallewaard M, Caspers PW, Visser LE, Leufkens HG, Stricker BH. Hospitalisations and emergency department visits due to drug-drug interactions: a literature review. *Pharmacoepidemiol Drug Saf.* Jun 11, 2007;16(6):641-651. [FREE Full text] [doi: [10.1002/pds.1351](https://doi.org/10.1002/pds.1351)] [Medline: [17154346](https://pubmed.ncbi.nlm.nih.gov/17154346/)]
6. Pretorius RW, Gataric G, Swedlund SK, Miller JR. Reducing the risk of adverse drug events in older adults. *Am Fam Physician.* Mar 01, 2013;87(5):331-336. [FREE Full text] [Medline: [23547549](https://pubmed.ncbi.nlm.nih.gov/23547549/)]
7. Davies EA, O'Mahony MS. Adverse drug reactions in special populations - the elderly. *Br J Clin Pharmacol.* Oct 22, 2015;80(4):796-807. [FREE Full text] [doi: [10.1111/bcp.12596](https://doi.org/10.1111/bcp.12596)] [Medline: [25619317](https://pubmed.ncbi.nlm.nih.gov/25619317/)]
8. Raut A, Pawar A, Diwan A, Naruka G, Sonar C. Clinically significant drug-drug interactions and their association with polypharmacy in elderly patients. *J Adv Sci Res.* May 10, 2013;4(2):6-12. [FREE Full text]
9. Chan M, Nicklason F, Vial JH. Adverse drug events as a cause of hospital admission in the elderly. *Intern Med J.* May 20, 2001;31(4):199-205. [FREE Full text] [doi: [10.1046/j.1445-5994.2001.00044.x](https://doi.org/10.1046/j.1445-5994.2001.00044.x)] [Medline: [11456032](https://pubmed.ncbi.nlm.nih.gov/11456032/)]
10. Blix HS, Viktil KK, Moger TA, Reikvam A. Drugs with narrow therapeutic index as indicators in the risk management of hospitalised patients. *Pharm Pract (Granada).* Jan 2010;8(1):50-55. [FREE Full text] [doi: [10.4321/s1886-36552010000100006](https://doi.org/10.4321/s1886-36552010000100006)] [Medline: [25152793](https://pubmed.ncbi.nlm.nih.gov/25152793/)]
11. Hines LE, Murphy JE. Potentially harmful drug-drug interactions in the elderly: a review. *Am J Geriatr Pharmacother.* Dec 2011;9(6):364-377. [FREE Full text] [doi: [10.1016/j.amjopharm.2011.10.004](https://doi.org/10.1016/j.amjopharm.2011.10.004)] [Medline: [22078863](https://pubmed.ncbi.nlm.nih.gov/22078863/)]
12. Downing NS, Shah ND, Aminawung JA, Pease AM, Zeitoun J, Krumholz HM, et al. Postmarket safety events among novel therapeutics approved by the US Food and Drug Administration between 2001 and 2010. *JAMA.* May 09, 2017;317(18):1854-1863. [FREE Full text] [doi: [10.1001/jama.2017.5150](https://doi.org/10.1001/jama.2017.5150)] [Medline: [28492899](https://pubmed.ncbi.nlm.nih.gov/28492899/)]
13. Wallach JD, Egilman AC, Dhruva SS, McCarthy ME, Miller JE, Woloshin S, et al. Postmarket studies required by the US Food and Drug Administration for new drugs and biologics approved between 2009 and 2012: cross sectional analysis. *BMJ.* May 24, 2018;361:k2031. [FREE Full text] [doi: [10.1136/bmj.k2031](https://doi.org/10.1136/bmj.k2031)] [Medline: [29794072](https://pubmed.ncbi.nlm.nih.gov/29794072/)]
14. Bykov K, Schneeweiss S, Glynn RJ, Mittleman MA, Gagne JJ. A case - crossover-based screening approach to identifying clinically relevant drug-drug interactions in electronic healthcare data. *Clin Pharmacol Ther.* Jul 18, 2019;106(1):238-244. [FREE Full text] [doi: [10.1002/cpt.1376](https://doi.org/10.1002/cpt.1376)] [Medline: [30663781](https://pubmed.ncbi.nlm.nih.gov/30663781/)]
15. Leonard CE, Zhou M, Brensing CM, Bilker WB, Soprano SE, Pham Nguyen TP, et al. Clopidogrel drug interactions and serious bleeding: generating real - world evidence via automated high - throughput pharmacoepidemiologic screening. *Clin Pharmacol Ther.* Nov 04, 2019;106(5):1067-1075. [FREE Full text] [doi: [10.1002/cpt.1507](https://doi.org/10.1002/cpt.1507)] [Medline: [31106397](https://pubmed.ncbi.nlm.nih.gov/31106397/)]
16. Abdullah SS, Rostamzadeh N, Muanda FT, McArthur E, Weir MA, Sontrop JM, et al. High-throughput computing to automate population-based studies to detect the 30-day risk of adverse outcomes after new outpatient medication use in older adults with chronic kidney disease: a clinical research protocol. *Can J Kidney Health Dis.* Jan 06, 2024;11:20543581231221891. [FREE Full text] [doi: [10.1177/20543581231221891](https://doi.org/10.1177/20543581231221891)] [Medline: [38186562](https://pubmed.ncbi.nlm.nih.gov/38186562/)]
17. ICES data. ICES. URL: <https://www.ices.on.ca/Data-and-Privacy/ICESdata> [accessed 2024-12-16]
18. Muanda FT, Weir MA, Bathini L, Blake PG, Chauvin K, Dixon SN, et al. Association of baclofen with encephalopathy in patients with chronic kidney disease. *JAMA.* Nov 26, 2019;322(20):1987-1995. [FREE Full text] [doi: [10.1001/jama.2019.17725](https://doi.org/10.1001/jama.2019.17725)] [Medline: [31705755](https://pubmed.ncbi.nlm.nih.gov/31705755/)]
19. Hwang YJ, Dixon SN, Reiss JP, Wald R, Parikh CR, Gandhi S, et al. Atypical antipsychotic drugs and the risk for acute kidney injury and other adverse outcomes in older adults: a population-based cohort study. *Ann Intern Med.* Aug 19, 2014;161(4):242-248. [FREE Full text] [doi: [10.7326/m13-2796](https://doi.org/10.7326/m13-2796)] [Medline: [25133360](https://pubmed.ncbi.nlm.nih.gov/25133360/)]
20. Fleet JL, Dixon SN, Kuwornu PJ, Dev VK, Montero-Odasso M, Burneo J, et al. Gabapentin dose and the 30-day risk of altered mental status in older adults: a retrospective population-based study. *PLoS One.* Mar 14, 2018;13(3):e0193134. [FREE Full text] [doi: [10.1371/journal.pone.0193134](https://doi.org/10.1371/journal.pone.0193134)] [Medline: [29538407](https://pubmed.ncbi.nlm.nih.gov/29538407/)]
21. Muanda FT, Blake PG, Weir MA, Ahmadi F, McArthur E, Sontrop JM, et al. Low-dose methotrexate and serious adverse events among older adults with chronic kidney disease. *JAMA Netw Open.* Nov 01, 2023;6(11):e2345132. [FREE Full text] [doi: [10.1001/jamanetworkopen.2023.45132](https://doi.org/10.1001/jamanetworkopen.2023.45132)] [Medline: [38010652](https://pubmed.ncbi.nlm.nih.gov/38010652/)]
22. Migration: interprovincial, 2011/2012. Statistics Canada. URL: <https://www150.statcan.gc.ca/n1/pub/91-209-x/2014001/article/14012-eng.htm> [accessed 2024-12-16]
23. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, et al. RECORD Working Committee. The reporting of studies conducted using observational routinely-collected health data (RECORD) statement. *PLoS Med.* Oct 6, 2015;12(10):e1001885. [FREE Full text] [doi: [10.1371/journal.pmed.1001885](https://doi.org/10.1371/journal.pmed.1001885)] [Medline: [26440803](https://pubmed.ncbi.nlm.nih.gov/26440803/)]
24. Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *PLoS Medicine.* Oct 16, 2007;4(10):1628-1654. [FREE Full text] [doi: [10.1371/journal.pmed.0040297](https://doi.org/10.1371/journal.pmed.0040297)] [Medline: [17938389](https://pubmed.ncbi.nlm.nih.gov/17938389/)]
25. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf.* Mar 2007;16(3):241-249. [FREE Full text] [doi: [10.1002/pds.1357](https://doi.org/10.1002/pds.1357)] [Medline: [17252614](https://pubmed.ncbi.nlm.nih.gov/17252614/)]
26. Jean S, Candas B, Belzile E, Morin S, Bessette L, Dodin S, et al. Algorithms can be used to identify fragility fracture cases in physician-claims databases. *Osteoporos Int.* Feb 19, 2012;23(2):483-501. [FREE Full text] [doi: [10.1007/s00198-011-1559-4](https://doi.org/10.1007/s00198-011-1559-4)] [Medline: [21336492](https://pubmed.ncbi.nlm.nih.gov/21336492/)]

27. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J*. Jan 2009;16(6):183-188. [FREE Full text] [doi: [10.1155/2009/963098](https://doi.org/10.1155/2009/963098)] [Medline: [20011725](https://pubmed.ncbi.nlm.nih.gov/20011725/)]
28. Tu K, Nieuwlaat R, Cheng SY, Wing L, Ivers N, Atzema CL, et al. Identifying patients with atrial fibrillation in administrative data. *Can J Cardiol*. Dec 2016;32(12):1561-1565. [FREE Full text] [doi: [10.1016/j.cjca.2016.06.006](https://doi.org/10.1016/j.cjca.2016.06.006)] [Medline: [27742459](https://pubmed.ncbi.nlm.nih.gov/27742459/)]
29. Hodge MC, Dixon S, Garg AX, Clemens KK. Validation of an international statistical classification of diseases and related health problems 10th revision coding algorithm for hospital encounters with hypoglycemia. *Can J Diabetes*. Jun 2017;41(3):322-328. [FREE Full text] [doi: [10.1016/j.cjcd.2016.11.003](https://doi.org/10.1016/j.cjcd.2016.11.003)] [Medline: [28268189](https://pubmed.ncbi.nlm.nih.gov/28268189/)]
30. Tu K, Mitiku T, Guo H, Lee D, Tu J. Myocardial infarction and the validation of physician billing and hospitalization data using electronic medical records. *Chronic Dis Inj Can*. Sep 2010;30(4):141-146. [FREE Full text] [doi: [10.24095/hpcdp.30.4.06](https://doi.org/10.24095/hpcdp.30.4.06)] [Medline: [20946715](https://pubmed.ncbi.nlm.nih.gov/20946715/)]
31. Figure 1 Ontario local health integration networks (LHINs) and sub-LHINs. Government of Canada, Statistics Canada. URL: <https://www150.statcan.gc.ca/n1/pub/82-003-x/2015003/article/14144/c-g/fig1-eng.htm> [accessed 2023-04-16]
32. Stürmer T, Wyss R, Glynn RJ, Brookhart MA. Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. *J Intern Med*. Jun 13, 2014;275(6):570-580. [FREE Full text] [doi: [10.1111/joim.12197](https://doi.org/10.1111/joim.12197)] [Medline: [24520806](https://pubmed.ncbi.nlm.nih.gov/24520806/)]
33. Brookhart MA, Wyss R, Layton JB, Stürmer T. Propensity score methods for confounding control in nonexperimental research. *Circ Cardiovasc Qual Outcomes*. Sep 2013;6(5):604-611. [FREE Full text] [doi: [10.1161/circoutcomes.113.000359](https://doi.org/10.1161/circoutcomes.113.000359)] [Medline: [24021692](https://pubmed.ncbi.nlm.nih.gov/24021692/)]
34. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. May 2011;46(3):399-424. [FREE Full text] [doi: [10.1080/00273171.2011.568786](https://doi.org/10.1080/00273171.2011.568786)] [Medline: [21818162](https://pubmed.ncbi.nlm.nih.gov/21818162/)]
35. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med*. Dec 10, 2015;34(28):3661-3679. [FREE Full text] [doi: [10.1002/sim.6607](https://doi.org/10.1002/sim.6607)] [Medline: [26238958](https://pubmed.ncbi.nlm.nih.gov/26238958/)]
36. Zou G. A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. Apr 01, 2004;159(7):702-706. [FREE Full text] [doi: [10.1093/aje/kwh090](https://doi.org/10.1093/aje/kwh090)] [Medline: [15033648](https://pubmed.ncbi.nlm.nih.gov/15033648/)]
37. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. May 2010;21(3):383-388. [FREE Full text] [doi: [10.1097/EDE.0b013e3181d61eeb](https://doi.org/10.1097/EDE.0b013e3181d61eeb)] [Medline: [20335814](https://pubmed.ncbi.nlm.nih.gov/20335814/)]
38. Haneuse S, VanderWeele TJ, Arterburn D. Using the E-value to assess the potential effect of unmeasured confounding in observational studies. *JAMA*. Feb 12, 2019;321(6):602-603. [FREE Full text] [doi: [10.1001/jama.2018.21554](https://doi.org/10.1001/jama.2018.21554)] [Medline: [30676631](https://pubmed.ncbi.nlm.nih.gov/30676631/)]
39. Austin P. Statistical criteria for selecting the optimal number of untreated subjects matched to each treated subject when using many-to-one matching on the propensity score. *Am J Epidemiol*. Nov 01, 2010;172(9):1092-1097. [FREE Full text] [doi: [10.1093/aje/kwq224](https://doi.org/10.1093/aje/kwq224)] [Medline: [20802241](https://pubmed.ncbi.nlm.nih.gov/20802241/)]
40. Austin PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Stat Med*. Dec 30, 2016;35(30):5642-5655. [FREE Full text] [doi: [10.1002/sim.7084](https://doi.org/10.1002/sim.7084)] [Medline: [27549016](https://pubmed.ncbi.nlm.nih.gov/27549016/)]
41. Number needed to harm. Wikipedia. URL: [https://en.wikipedia.org/w/index.php?title=Number\\_needed\\_to\\_harm&oldid=1105549469#cite\\_note-02-1](https://en.wikipedia.org/w/index.php?title=Number_needed_to_harm&oldid=1105549469#cite_note-02-1) [accessed 2024-12-16]
42. Attributable fraction for the population. Wikipedia. URL: [https://en.wikipedia.org/w/index.php?title=Attributable\\_fraction\\_for\\_the\\_population&oldid=1091375478](https://en.wikipedia.org/w/index.php?title=Attributable_fraction_for_the_population&oldid=1091375478) [accessed 2024-12-16]
43. Abdullah SS, Rostamzadeh N, Sedig K, Garg AX, McArthur E. Multiple regression analysis and frequent itemset mining of electronic medical records: a visual analytics approach using VISA\_M3R3. *Data*. Mar 29, 2020;5(2):33. [FREE Full text] [doi: [10.3390/data5020033](https://doi.org/10.3390/data5020033)]
44. Keim DA, Mansmann F, Thomas J. Visual analytics: how much visualization and how much analytics? *SIGKDD Explor Newsl*. May 27, 2010;11(2):5-8. [FREE Full text] [doi: [10.1145/1809400.1809403](https://doi.org/10.1145/1809400.1809403)]
45. Hale SF, Lesar TS. Interaction of vitamin K antagonists and trimethoprim-sulfamethoxazole: ignore at your patient's risk. *Drug Metabol Drug Interact*. Nov 14, 2014;29(1):53-60. [FREE Full text] [doi: [10.1515/dmdi-2013-0049](https://doi.org/10.1515/dmdi-2013-0049)] [Medline: [24231121](https://pubmed.ncbi.nlm.nih.gov/24231121/)]
46. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res Social Adm Pharm*. May 2013;9(3):330-338. [FREE Full text] [doi: [10.1016/j.sapharm.2012.04.004](https://doi.org/10.1016/j.sapharm.2012.04.004)] [Medline: [22695215](https://pubmed.ncbi.nlm.nih.gov/22695215/)]
47. Rau G, Shih Y. Evaluation of Cohen's kappa and other measures of inter-rater agreement for genre analysis and other nominal data. *J Engl Acad Purp*. Sep 2021;53:101026. [FREE Full text] [doi: [10.1016/j.jeap.2021.101026](https://doi.org/10.1016/j.jeap.2021.101026)]
48. ICES. URL: <https://www.ices.on.ca/data-repository-requirements/> [accessed 2024-12-16]
49. Gbedawo WV, Dzikunu A, Nyamadi M. Scalability and efficiency of deep learning models on high-performance computing clusters: bibliometric analysis. *Afr J Appl Res*. Dec 28, 2024;10(2):283-305. [FREE Full text] [doi: [10.26437/ajar.v10i2.807](https://doi.org/10.26437/ajar.v10i2.807)]

50. Hughes JE, Waldron C, Bennett KE, Cahir C. Prevalence of drug-drug interactions in older community-dwelling individuals: a systematic review and meta-analysis. *Drugs Aging*. Feb 24, 2023;40(2):117-134. [FREE Full text] [doi: [10.1007/s40266-022-01001-5](https://doi.org/10.1007/s40266-022-01001-5)] [Medline: [36692678](https://pubmed.ncbi.nlm.nih.gov/36692678/)]
51. Assimon MM. Confounding in observational studies evaluating the safety and effectiveness of medical treatments. *Kidney360*. Jul 29, 2021;2(7):1156-1159. [FREE Full text] [doi: [10.34067/kid.0007022020](https://doi.org/10.34067/kid.0007022020)]
52. Smith M, Lix L, Azimaee M, Enns J, Orr J, Hong S, et al. Assessing the quality of administrative data for research: a framework from the Manitoba Centre for Health Policy. *J Am Med Inform Assoc*. Mar 01, 2018;25(3):224-229. [FREE Full text] [doi: [10.1093/jamia/ocx078](https://doi.org/10.1093/jamia/ocx078)] [Medline: [29025002](https://pubmed.ncbi.nlm.nih.gov/29025002/)]
53. Snowdon DA, Hau R, Leggat SG, Taylor NF. Does clinical supervision of health professionals improve patient safety? A systematic review and meta-analysis. *Int J Qual Health Care*. Sep 09, 2016;28(4):447-455. [FREE Full text] [doi: [10.1093/intqhc/mzw059](https://doi.org/10.1093/intqhc/mzw059)] [Medline: [27283436](https://pubmed.ncbi.nlm.nih.gov/27283436/)]

## Abbreviations

**ATT:** average treatment-effect-in-the-treated  
**DDI:** drug-drug interaction  
**FDA:** Food and Drug Administration  
**ICES:** Institute for Clinical Evaluative Sciences  
**IPTW:** inverse probability of treatment weighting  
**NNH:** number needed to harm  
**ODB:** Ontario Drug Benefit  
**PAF:** population attributable fraction  
**PD:** pharmacodynamic  
**PHIPA:** Personal Health Information Protection Act  
**PK:** pharmacokinetic  
**RCT:** randomized controlled trial  
**RD:** risk difference  
**REB:** Research Ethics Board  
**RR:** risk ratio  
**SD:** standardized difference  
**SSRI:** serotonin reuptake inhibitor

*Edited by A Schwartz; The proposal for this study was peer-reviewed by the Western University Interdisciplinary Development Initiatives Review Committee; Submitted 09.May.2025; accepted 15.Sep.2025; published 10.Oct.2025.*

### *Please cite as:*

Rostamzadeh N, Sharma R, Abdullah SS, McArthur E, Chalabianloo N, Sontrop JM, Weir MA, Sedig K, Garg AX, Muanda FT  
*High-Throughput Computing to Detect Harmful Drug-Drug Interactions in Older Adults: Protocol for a Population-Based Cohort Study*

*JMIR Res Protoc* 2025;14:e77224

URL: <https://www.researchprotocols.org/2025/1/e77224>

doi: [10.2196/77224](https://doi.org/10.2196/77224)

PMID:

©Neda Rostamzadeh, Rishabh Sharma, Sheikh S Abdullah, Eric McArthur, Niaz Chalabianloo, Jessica M Sontrop, Matthew A Weir, Kamran Sedig, Amit X Garg, Flory T Muanda. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 10.Oct.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.