



# Next steps for Evidence Dissemination

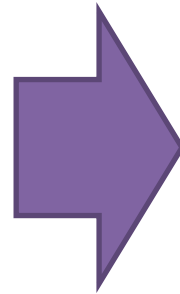
Patrick Ryan, Martijn Schuemie, Nicole Pratt



# Scaling evidence together

Current status quo:

**One Person**  
from **One Institution**  
has **One Question**  
about **One Exposure**  
and **One Outcome**  
applies **One Design**  
to **One Database**  
generating **One Result**  
disseminated in **One Publication**  
to communicate to **One Audience**

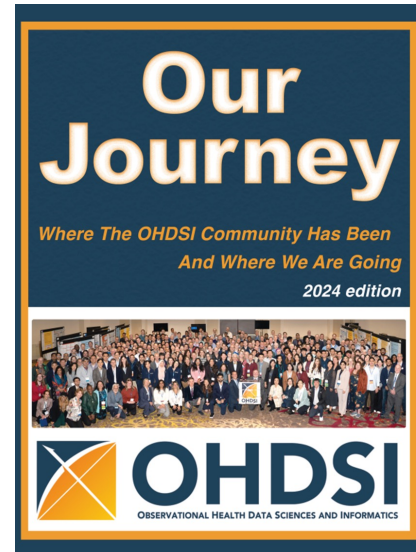


Future reality:

**One Community of 1000s of persons**  
from **Many Institutions**  
have **An Array Of Questions**  
about **All Exposures**  
and **All Outcomes**  
applies **Standardized Framework**  
**Incorporating Multiple Designs**  
to **A Network of 100s of Databases**  
Generating **Millions of Results**  
disseminated across **Multiple Channels**  
**To maximize the reach and impact across**  
**stakeholders**



# Our Journey through OHDSI



# OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

# Reflecting on the Journey So Far

Initial OHDSI vision:

“OHDSI collaborators access a network of 1 billion patients to generate evidence about all aspects of healthcare. Patients and clinicians and other decisionmakers around the world use OHDSI tools and evidence every day”

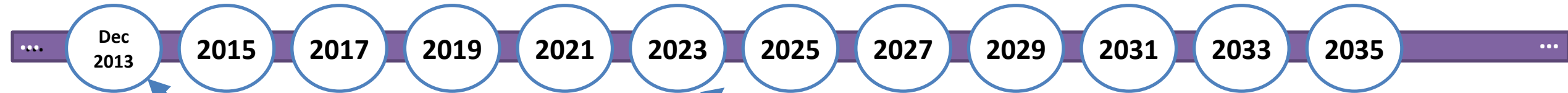


OHDSI vision revised:

“A world in which observational research produces a comprehensive understanding of health and disease.”



# What is the Journey Ahead Together?



# OHDSI

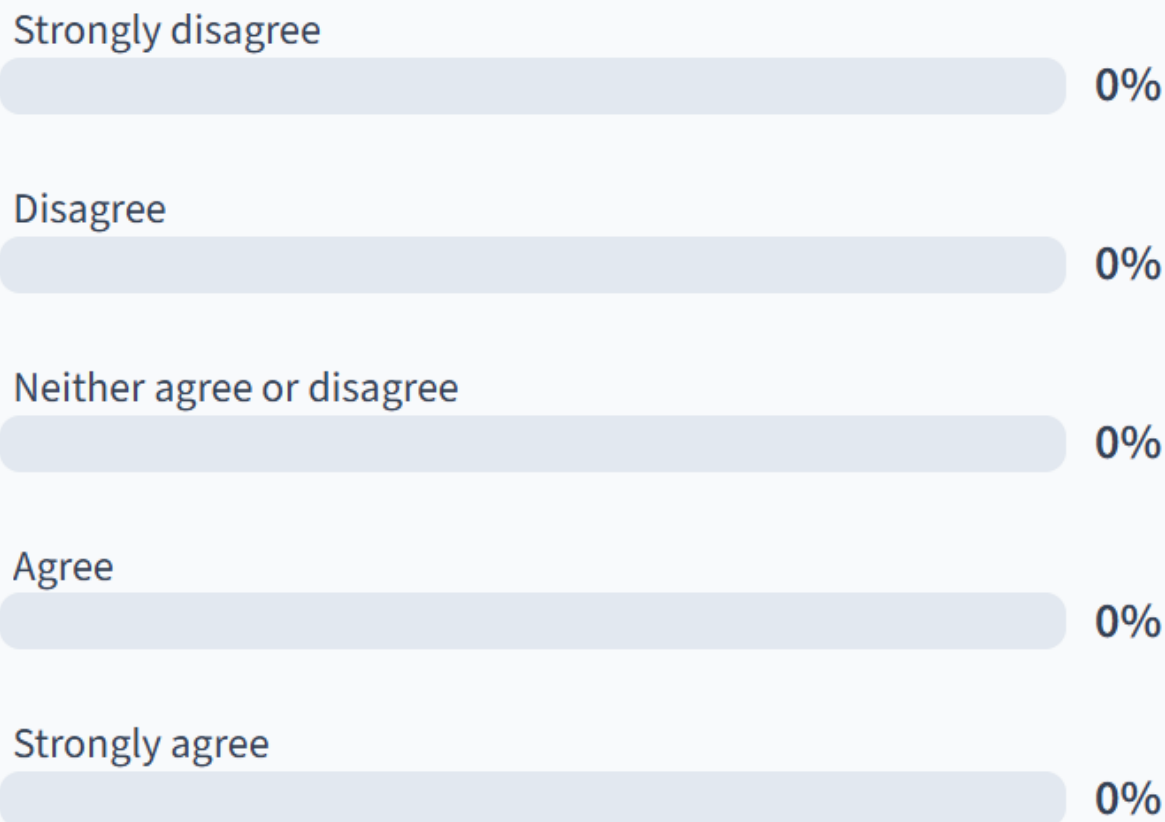
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS



## How much do you agree that by 2034, we will achieve the following prediction:



OHDSI's open science community approach to evidence generation becomes the expected behavior across stakeholders and disciplines to promote innovation, reproducibility, and collaboration.





OHDSI's open science community approach to evidence generation becomes the expected behavior across stakeholders and disciplines to promote innovation, reproducibility, and collaboration.



The OMOP Common Data Model will evolve and become recognized as the preferred international data standard for real-world evidence generation, will be seamlessly interoperable with complementary clinical data exchange standards, and will be consistently adopted across academia, industry, and government around the world.

---

---





The number of unique data sources adopting the OMOP CDM will exceed 50,000, but organizations will also use the CDM as a mechanism to partner to advance cross-organizational data linkage and participatory patient self-reporting. This will increase the completeness and longitudinal continuity of patient records, enable connections across familial generations, and improve the fitness-of-use for each integrated source across a broader set of analytic use cases.



The OHDSI Standardized Vocabularies will provide the singular resource that maps all source terminologies and unstructured medical text into a common reference ontology, with real-time updating to reflect the current state of knowledge in medicine.

---

---



Every organization collecting patient-level data during the routine course of clinical care will have established systems to standardize the data to the OMOP CDM using the latest OHDSI standardized vocabularies on a nightly basis, enabling daily reporting for disease surveillance and quality improvement.

---

---



The OHDSI community will prove that real world evidence from real world data— when adhering OHDSI’s best practices and passing all objective diagnostics— can be considered just as reliable as evidence from randomized clinical trials. Open-source systems that follow these practices will become trusted by health systems, payers, and regulators for guiding clinical care and policy decisions.



Advances in OHDSI's open-source analytic platform will decrease the time to generate reliable real-world evidence across the OHDSI distributed network; this process will be measured in minutes, not months.



The OHDSI Evidence Network will make it both commonplace and expected to see hundreds of databases, representing hundreds of millions of patients, be represented in network studies of every important public health question. This would ensure that the evidence we generate is replicable within similar populations and generalizable to patients across North and South America, Europe, Africa, Asia and Australia.



The OHDSI community will represent and support all clinical subspecialties and will become the primary source of real-world evidence to proactively fill evidence gaps needed to inform clinical guidelines around management of every disease.

---

---



The OHDSI community will design, implement, and deliver results from more than 10,000 network studies, with the majority of research questions coming directly from patients and clinicians seeking reliable evidence to address their needs at the point-of-care.

---

---





Discoveries across the OHDSI network about unrecognized effects of existing medical interventions will yield new indications that achieve regulatory approval due to the robustness of the real-world evidence produced within our community.

---

---



**OHDSI will freely disseminate its evidence through more than 100,000 scholarly publications, but it will also establish new modalities for evidence dissemination to more directly support clinical practice.**



Every disease will have a comprehensive real-world evidence summary that characterizes natural history and treatment pathways across the globe so we can understand patient heterogeneity, promote health equity, and recognize unmet medical needs.



Every medical product will have a comprehensive real-world evidence surveillance summary from OHDSI that provides characterization of the incidence of all outcomes, population-level estimation of the causally attributable risk of each outcome and comparative effectiveness with all alternative treatments, and patient-level prediction models so that individuals can accurately determine their personalized risk given their medical history.



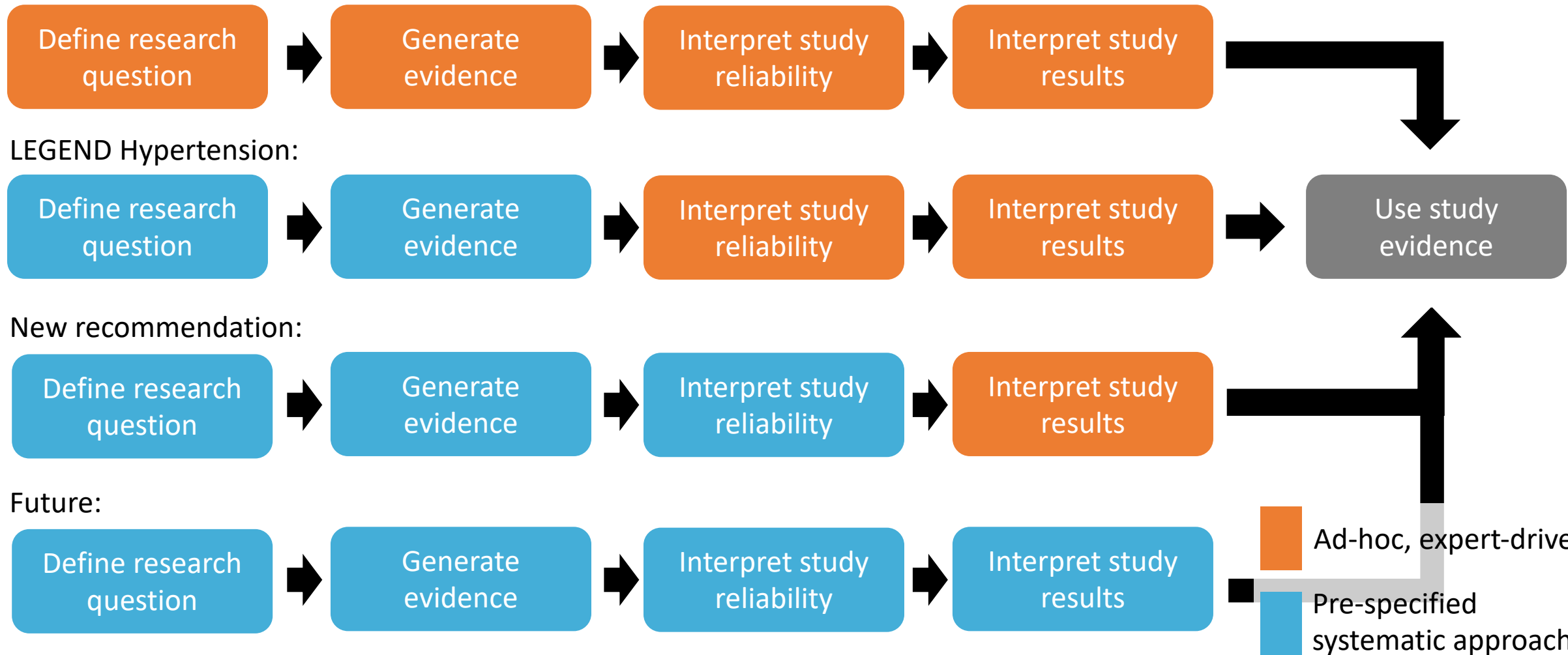
OHDSI evidence repositories will become the primary source of knowledge underpinning foundational models to promote better health decisions and better care.

---



# Pre-specification of a systematic approach

Traditional observational study:





# A rubric for interpreting studies

- Quality of study
  - Study design
    - Multiple designs to evaluate robustness (ex: comparative cohort and SCCS)
    - Multiple analyses within design (ex: PS matching vs stratification, on-treatment vs. ITT )
    - Objective diagnostics to test statistical assumptions, quantify residual error, and establish unblinding rules
  - Transparency
    - Pre-specified protocol
    - Publicly accessible analytic source code
    - Provenance of full resultset
  - Diversity of databases
    - Populations
    - Geographies
    - Data capture processes





# A rubric for interpreting studies

- Strength of evidence
    - Continuum:
      - No evidence: No databases pass diagnostics
      - Weak: One database pass diagnostics, lots of databases fail
      - Strong: lots of diverse databases pass and few fail diagnostic (including diagnostic for heterogeneity)
  - Certainty in estimate
  - Size of effect
-



# What you need to collaborate on evidence at scale

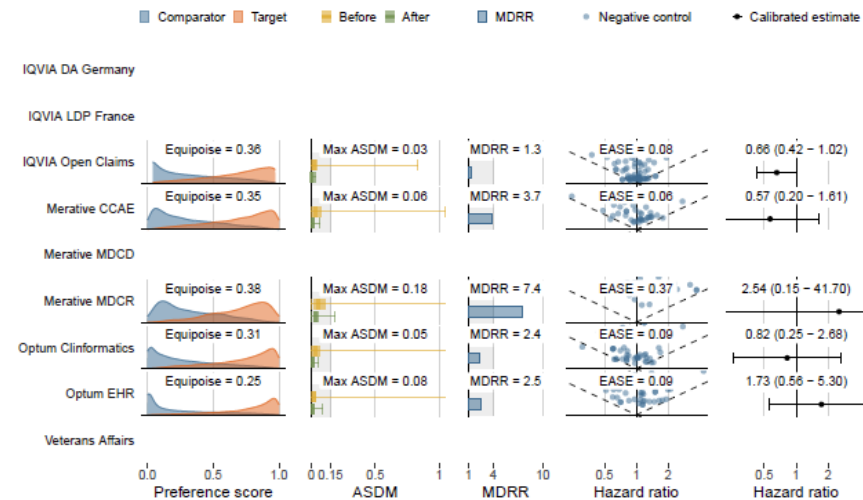
## LEGEND-T2DM Evidence Dissemination Summary

- Target (class): **Semaglutide** (GLP-1 Receptor Agonists)
- Comparator (class): **Glimepiride** (Sulfonylureas)
- Outcome: **Acute pancreatitis**

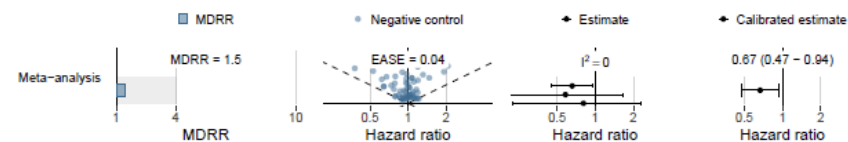
### How Often? (Incidence rates in the PS-matched target cohorts)

Data source	Persons exposed	Person-time (yrs)	Persons with outcome	IR (/1,000 PY)
IQVIA DA Germany	-	-	-	-
IQVIA LDP France	-	-	-	-
IQVIA Open Claims	99,708	52,939	60	1.13
Merative CCAE	20,240	9,388	14	1.49
Merative MDCCD	-	-	-	-
Merative MDCR	619	278	<5	<17.97
Optum Clinformatics	7,607	3,811	8	2.10
Optum EHR	6,717	2,098	7	3.34
Veterans Affairs	1,258	883	-	0.00

### How Reliable Are the Effect Estimates? (Objective diagnostics)



### What have we learned from the OHDSI Network? (Meta-analysis diagnostics and estimate)





# Database diagnostics

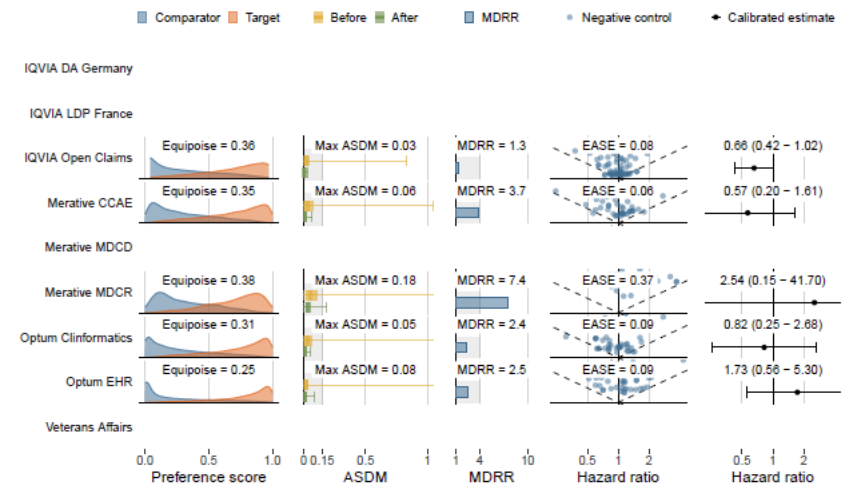
## LEGEND-T2DM Evidence Dissemination Summary

- Target (class): Semaglutide (GLP-1 Receptor Agonists)
- Comparator (class): Glimepiride (Sulfonylureas)
- Outcome: Acute pancreatitis

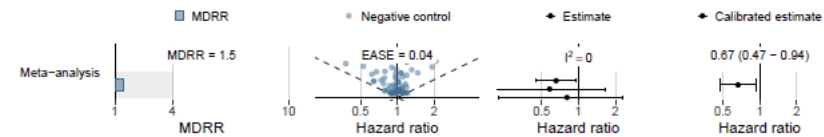
### How Often? (Incidence rates in the PS-matched target cohorts)

Data source	Persons exposed	Person-time (yrs)	Persons with outcome	IR (/1,000 PY)
IQVIA DA Germany	-	-	-	-
IQVIA LDP France	-	-	-	-
IQVIA Open Claims	99,708	52,939	60	1.13
Merative CCAE	20,240	9,388	14	1.49
Merative MDCC	-	-	-	-
Merative MDCR	619	278	<5	<17.97
Optum Clinformatics	7,607	3,811	8	2.10
Optum EHR	6,717	2,098	7	3.34
Veterans Affairs	1,258	883	-	0.00

### How Reliable Are the Effect Estimates? (Objective diagnostics)



### What have we learned from the OHDSI Network? (Meta-analysis diagnostics and estimate)





# Meta-analysis diagnostics

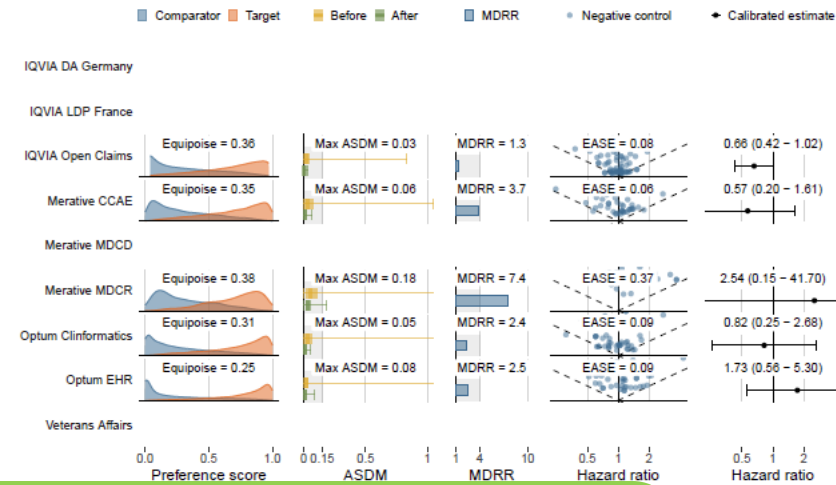
## LEGEND-T2DM Evidence Dissemination Summary

- Target (class): Semaglutide (GLP-1 Receptor Agonists)
- Comparator (class): Glimepiride (Sulfonylureas)
- Outcome: Acute pancreatitis

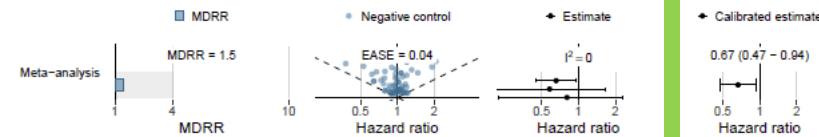
### How Often? (Incidence rates in the PS-matched target cohorts)

Data source	Persons exposed	Person-time (yrs)	Persons with outcome	IR (/1,000 PY)
IQVIA DA Germany	-	-	-	-
IQVIA LDP France	-	-	-	-
IQVIA Open Claims	99,708	52,939	60	1.13
Merative CCAE	20,240	9,388	14	1.49
Merative MDCC	-	-	-	-
Merative MDCR	619	278	<5	<17.97
Optum Clinformatics	7,607	3,811	8	2.10
Optum EHR	6,717	2,098	7	3.34
Veterans Affairs	1,258	883	-	0.00

### How Reliable Are the Effect Estimates? (Objective diagnostics)



### What have we learned from the OHDSI Network? (Meta-analysis diagnostics and estimate)





# Meta-analysis estimate

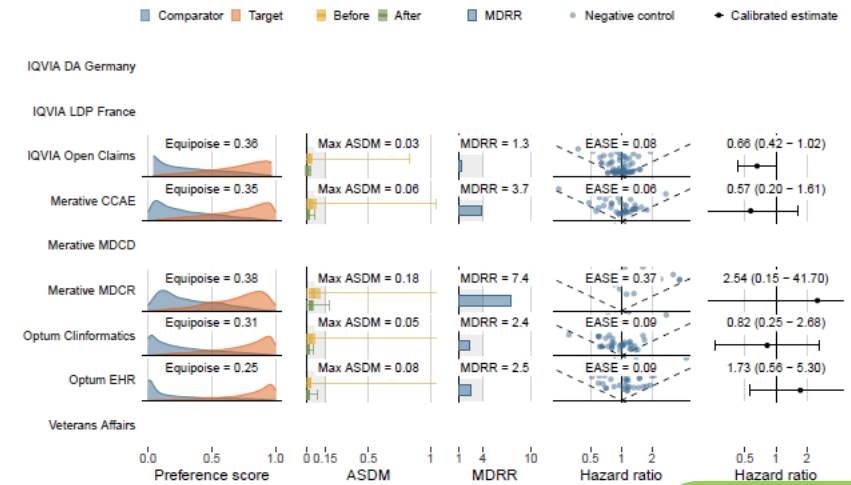
## LEGEND-T2DM Evidence Dissemination Summary

- Target (class): Semaglutide (GLP-1 Receptor Agonists)
- Comparator (class): Glimepiride (Sulfonylureas)
- Outcome: Acute pancreatitis

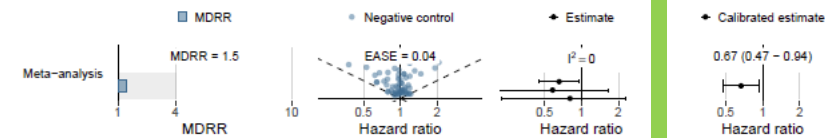
### How Often? (Incidence rates in the PS-matched target cohorts)

Data source	Persons exposed	Person-time (yrs)	Persons with outcome	IR (/1,000 PY)
IQVIA DA Germany	-	-	-	-
IQVIA LDP France	-	-	-	-
IQVIA Open Claims	99,708	52,939	60	1.13
Merative CCAE	20,240	9,388	14	1.49
Merative MDCC	-	-	-	-
Merative MDCR	619	278	<5	<17.97
Optum Clinformatics	7,607	3,811	8	2.10
Optum EHR	6,717	2,098	7	3.34
Veterans Affairs	1,258	883	-	0.00

### How Reliable Are the Effect Estimates? (Objective diagnostics)



### What have we learned from the OHDSI Network? (Meta-analysis diagnostics and estimate)





## Next steps

- Most of you have in your hand a piece of reliable evidence that could be published and make a difference
  - All of you have permission to disseminate that evidence, and explicit encouragement from the LEGEND leadership and OHDSI community to do so (just coordinate with the LEGEND team and stick to our OHDSI authorship guidelines)
  - Or even better, all of you made friends today and a dream team to move forward on the opportunity together to impact the health of millions of patients
-



# Anatomy of JACC publication

## Introduction from the Editor

JACC publishes peer-reviewed articles highlighting all aspects of cardiovascular disease, including original clinical studies, experimental investigations with clear clinical relevance, state-of-the-art papers, and viewpoints.

At JACC, we value the time you spend preparing your submissions and aim to give you a rapid decision. Toward this goal, we have minimal formatting requirements on initial submissions. If your submission progresses toward publication, we may ask for more information and some specific formatting.

We are always open to suggestions about process improvements and appreciate your support of JACC.

**Hartan M. Krumholz**  
Editor-in-Chief

## Quick Submission Guide

At JACC, **submit your paper your way**. Authors may submit their manuscript (text, figures and tables) as a single file at [www.jaccsubmit.org](http://www.jaccsubmit.org). This can be a Word or PDF file, in any format or layout, and figures and tables can be placed within the text. The only requirements for initial review are:

- Please list all author names, institutional affiliations, and relationships with industry and other entities on your title page and in the online submission system (see also [Relationships with Industry](#)).
- Please consider the word and author limit for your article type (see Article Types table below); initial submissions may exceed the word limit, but our published papers, with few exceptions, need to adhere to the requirements.

JACC now partners with [JACC: Case Reports](#) to publish their top case reports and vignettes in a one issue per week. Authors should refer to the [JACC: Case Reports instructions](#) for authors for submission and formatting.

## JACC Instructions for Authors

minimize 1. Article Types

Article Type	Description	Requirements
<b>Original Research</b>	Should relate to cardiovascular science and medicine that may include studies conducted in humans or analyses of human data that significantly advance the field.	<ul style="list-style-type: none"><li>• <b>Note: for initial submissions, please see the quick submission guide above; additional requirements requested at revision</b></li><li>• Word Count: ≤5,000, excluding references and figure legends</li><li>• Author Count: Unlimited</li><li>• Table/Figure Count: ≤6 tables and/or figures, including Central Illustration</li><li>• Reference Count: Unlimited</li><li>• Central Illustration: Optional</li><li>• Structured Abstract ≤350 words with the headings: Background, Objectives, Methods, Results, Conclusions</li><li>• Data Sharing Statement</li><li>• Follow <a href="#">EQUATOR</a> Reporting Guidelines</li></ul>

## General publication format:

- Abstract (<350 words)
- Body: (<5000 words)
  - Background
  - Methods
  - Results
    - ≤6 tables/figures
  - Conclusions
- Supporting information:
  - References
  - Supplemental Materials



# OHDSI community efforts toward supporting publications

Martijn Schuemie



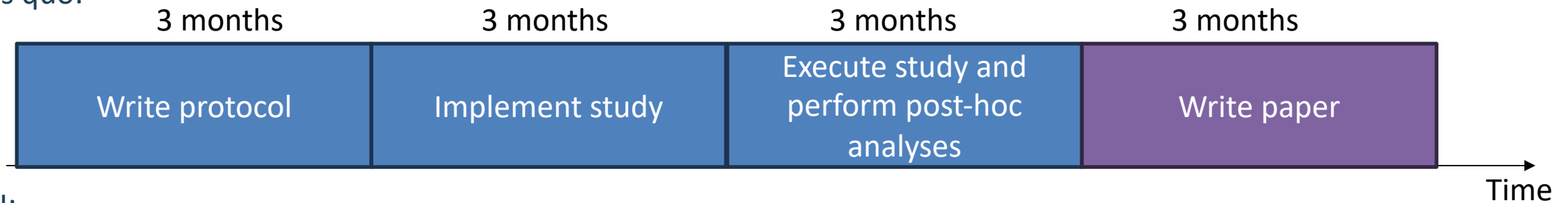


# Major improvements to observational research

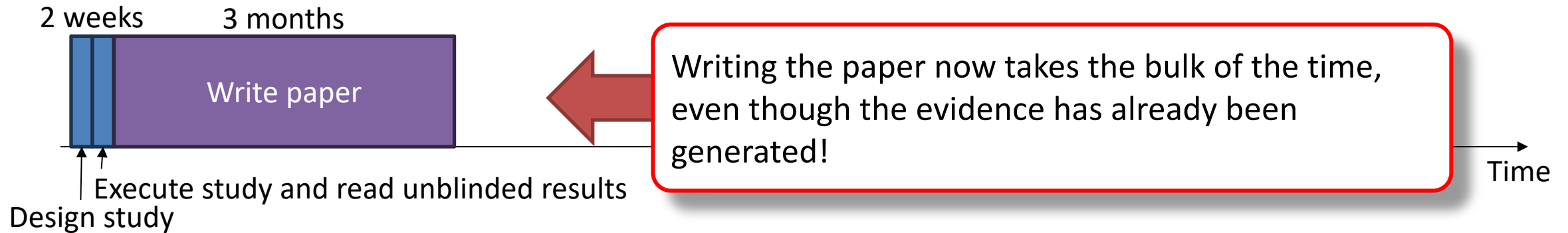
In OHDSI, we have

- Vastly **reduced time to perform observational studies** (days instead of months)
- Vastly **increased the reliability of observational studies** through use of standardized analytics, objective diagnostics, and generative evidence at scale

Status quo:



OHDSI:





# What's in a (OHDSI) scientific paper?

- Background
  - What question did we seek to answer?
  - Why is it important?
  - What is already known?
- Methods
  - What data were used?
  - How were the exposures and outcomes defined?
  - What statistical analysis was performed?
  - What objective diagnostics were used?
- Results
  - Which analyses passed diagnostics?
  - What were the results?
- Discussion
  - What have we learned?
  - How does that fit with what we already knew?

Methods and Results should be **objective descriptions** of what was done and what was observed.

When using **standardized analytics** with **standardized analysis specifications** and **standardized outputs**, this becomes a **fill-in-the-blanks** exercise that can be automated.

# LEGENDMed Central was a proof of concept

- Uses Rmarkdown to convert LEGEND Hypertension results to draft papers
- Select a target-comparator-outcome-database, and it would generate a PDF for you!
- Still alive, but a bit buggy: <https://data.ohdsi.org/LegendMedCentral>

## Acute myocardial infarction risk in new-users of ACE inhibitors versus Angiotensin receptor blockers (ARBs) for hypertension in the CCAE database

Martijn J. Schuemie<sup>1,2\*</sup>, Patrick B. Ryan<sup>1,3,4</sup>, Sang Chon<sup>1,5</sup>, Nicole Pratt<sup>1,6</sup>, David Madigan<sup>1,7</sup>, George Hlipcak<sup>1,8</sup>, and Marc A. Suchard<sup>1,2,3,4</sup>

<sup>1</sup>Observational Health Data Science and Informatics, New York, NY, USA, <sup>2</sup>Parsons Research & Development, Tyngsboro, MA, USA, <sup>3</sup>Department of Biostatistics, University of California, Los Angeles, CA, <sup>4</sup>Department of Biomedical Informatics, Columbia University, New York, NY, <sup>5</sup>Department of Biomedical Informatics, Rice University, Houston, Texas, <sup>6</sup>Department of Biostatistics, University of South Australia, Adelaide SA, Australia, <sup>7</sup>Department of Statistics, Columbia University, New York, NY, <sup>8</sup>Department of Biostatistics, University of California, Los Angeles, CA, <sup>9</sup>Department of Clinical Sciences, University of California, Los Angeles, CA.

This report was automatically compiled on November 12, 2024.

We conducted a large-scale study on the incidence of acute myocardial infarction among new users of ACE inhibitors and angiotensin receptor blockers (ARBs) from 2000 to 2018 in the CCAE database. Outcomes of interest are estimates of the hazard ratio (HR) for incident events between comparable new users under on-treatment and intent-to-treat risk models. We also generate evidence consistently by applying a systematic approach across all research questions and disseminate evidence relevant subgroup interaction with the HR. We identify 77262 ACE inhibitors and 22866 angiotensin receptor blockers (ARBs) patients for the on-treatment design, totaling 85428 and 19915 patients-years of observation, and 1598 and 468 events respectively. We control for measured confounding using propensity score trimming and stratification or matching based on an extensive propensity score model that includes all measured patient features before treatment initiation. We account for unmeasured confounding using negative and positive controls to estimate and adjust for residual systematic bias in the study design and data source, providing calibrated confidence intervals and p-values. In terms of acute myocardial infarction, ACE inhibitors had a similar risk as compared to angiotensin receptor blockers (ARBs) (HR: 1.12, 95% confidence interval (CI) 0.93 - 1.40). We conduct a large-scale study on the incidence of acute myocardial infarction among new users of ACE inhibitors and angiotensin receptor blockers (ARBs) from 2000 to 2018 in the CCAE database. Outcomes of interest are estimates of the hazard ratio (HR) for incident events between comparable new users under on-treatment and intent-to-treat risk window assumptions. Secondary analysis estimates include clinically relevant subgroup interaction with the HR. We identify 77262 ACE inhibitors and 22866 angiotensin receptor blockers (ARBs) patients for the on-treatment design, totaling 85428 and 19915 patient-years of observation, and 1598 and 468 events respectively. We control for measured confounding using propensity score trimming and stratification or matching based on an extensive propensity score model that includes all measured patient features before treatment initiation. We account for unmeasured confounding using negative and positive controls to estimate and adjust for residual systematic bias in the study design and data source, providing calibrated confidence intervals and p-values. In terms of acute myocardial infarction, ACE inhibitors had a similar risk as compared to angiotensin receptor blockers (ARBs) (HR: 1.12, 95% confidence interval (CI) 0.93 - 1.40).

Worldwide, hypertension stands as a leading cause of mortality, with an increasing prevalence over the last two decades (Ovretz et al., 2017). The 2017 American College of Cardiology (ACC) / American Heart Association (AHA) clinical practice guidelines define hypertension based on averaged blood pressure (BP) measured in a healthcare setting: systolic BP between 130 - 139 mmHg or diastolic BP between 80 - 89 mmHg characterizes stage 1 hypertension, and systolic BP > 140 mmHg or diastolic BP > 90 mmHg mark stage 2 hypertension (Whelton et al., 2018). Elevated BP contributes to approximately half of all stroke and ischemic heart disease deaths and 13% of all forms of death globally (WHO, 2009). While antihypertensive therapies carry well-established benefits in reducing BP and the risk of major cardiovascular events, the health benefits and drug safety concerns of any one class of antihypertensive drug relative to other classes as first-line therapy remain debated. Part of this debate arises from a study of large randomized controlled trials and observational studies providing head-to-head comparisons between all pairs of individual drug and drug classes. One notable counter-example is the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) (ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group, 2002) that randomized 33,357 patients to receive either atenolol (a  $\beta$ -blocker), lisinopril (an ACE inhibitor), chlorthalidone (a thiazide or thiazide-like diuretic) or lisinopril (an angiotensin converting enzyme inhibitor). However, ALLHAT employed only a single drug representative per class and did not include several other important antihypertensive classes, such as angiotensin receptor blockers and beta-blockers. Robustness et al. (2016) provide a systematic review of randomized controlled trials examining comparative benefits and harms of various antihypertensives as first-line therapy (Robustness et al., 2017). Further observational study can help refine first-line therapy recommendations in terms of both treatment effect and relative drug safety.

The large-scale Evidence Generation and Evaluation in a Network of Databases (LEGEND) project aims to generate reliable evidence on the effects of medical interventions using observational healthcare data to support clinical decision making. LEGEND fol-

lows those from the 2017 ACC/AHA guidelines and the 2018 from the Society of Cardiology (SCD) and European Society of Intension (ESI) Guidelines for the management of arterial tension (Williams et al., 2018).

### Methods

We conduct a new-user cohort study comparing new users of ACE inhibitors with new users of ARBs in the Truven Health Martini Commercial Claims and Encounters Database (CCEd) has embedded in the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) version 5 (Hippen et al., 2015; Ouchage et al., 2015; Ryan et al., 2013). The CCEd data from individuals enrolled in United States employer-sponsored insurance health plans. The data include adjudicated insurance claims (e.g. inpatient, outpatient, and outpatient therapy) as well as enrollment data from large employers and health plans who provide private healthcare coverage to employees, spouses, and dependents. Additionally, it captures laboratory for a subset of the covered lives. This administrative claims data includes a variety of fee-for-service, preferred provider organizations, and capitated health plans. The study period spans 2000-12-31, 2000-12-31 to 2018-03-31, 2018-03-31.

**Study design.** This study follows a retrospective, observational, patient-cohort design (Ryan et al., 2013). We include patients who have continuous observation in the database for at least 90 days prior to treatment initiation. We require that patients have continuous observation in the database for at least 90 days prior to treatment initiation. We exclude patients with known myocardial infarction events and less than 1 day at risk to full cohort details, include concept codes, are provided in Supporting Information. The outcome of interest is acute myocardial infarction. We begin the outcome risk window 1 day before treatment initiation and consider two design choices to define window end. First, we stop the outcome time-at-risk window at cessation of continuous drug exposure, analogous to an intent-to-treat design, and second, we end the outcome time-at-risk window when the patient is no longer observable in the database, given its intent-to-treat design. Continuous drug exposure is restricted from the available longitudinal data by considering initial prescriptions that have fewer than 30 days gap between refills.

**Risk analysis.** We conduct our cohort study using the open-source OHDSI Analytic Method R package (Schuemie et al., 2016). Large-scale analysis achieved through the Cyclotron R package (Ryan et al., 2013). We use propensity scores (PS) - estimate sample exposure probability conditional on pre-treatment baseline characteristics (Schuemie et al., 2018b) - to control for potential measured confounding and improve balance between the target (ACEs) and comparator (ARBs) groups. We additionally calibrate all HR estimates, their 95% confidence intervals (CIs) and the p-value to reject the null hypothesis of no differential effect (HR = 1). Empirical calibration serves as an important diagnostic tool to evaluate if residual systematic bias is sufficient to cast doubt on the accuracy of the unknown effect estimate.

**Results**  
**Population characteristics.** Figure 1 diagrams the inclusion of study subjects from the CCAE database under the on-treatment and intent-to-treat design. We assign these counts with cohort

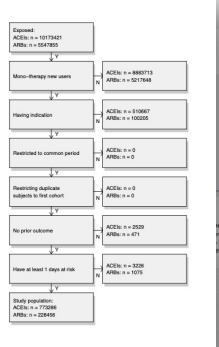


Fig. 1. Action diagram for selecting new users of ACEi and ARBs from the CCAE database.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual systematic error. In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEi and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEi to ARBs for the on-treatment and intent-to-

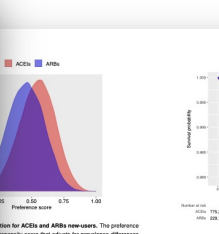


Fig. 4. Kaplan-Meier plot of acute myocardial infarction-free survival. The plot is adjusted for the propensity score stratification. The ACEi curve shows the actual observed survival. The ARBs curve applies weighting to approximate the counterfactual of what ACEi survival would look like had the ACEi cohort been exposed to ARBs instead. The shaded area denotes the 95% CI.

designs with stratification or matching. Figure 4 plots Kaplan-Meier survival curves for patients under the intent-to-treat design.

**Warning:** Using 'size' aesthetic for lines was deprecated in ggplot2. Please use 'linewidth' instead.  
**Warning:** Using 'displayed once every 8 hours' in aes() is deprecated. Use 'lifecycle::last\_lifecycle\_warnings()' to see where this warning originates.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual systematic error. In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEi and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEi to ARBs for the on-treatment and intent-to-

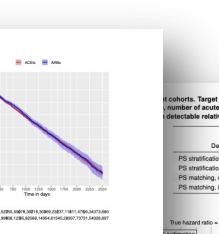


Fig. 5. Distributions for ACEi and ARBs new users. The plot shows the distributions of the hazard ratio for incident events between comparable new users under on-treatment and intent-to-treat risk window assumptions. The shaded area denotes the 95% CI.

designs with stratification or matching. Figure 4 plots Kaplan-Meier survival curves for patients under the intent-to-treat design.

**Warning:** Using 'size' aesthetic for lines was deprecated in ggplot2. Please use 'linewidth' instead.  
**Warning:** Using 'displayed once every 8 hours' in aes() is deprecated. Use 'lifecycle::last\_lifecycle\_warnings()' to see where this warning originates.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual systematic error. In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEi and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEi to ARBs for the on-treatment and intent-to-

Design	T		C		Events		IR		
	T	C	T	C	T	C	T	C	
PS stratification, on-treatment	773,266	228,456	585,422	199,149	1,596	468	2.73	2.35	1.16
PS stratification, intent-to-treat	773,266	229,170	513,163	183,043	6,275	1,855	2.96	2.72	1.08
PS matching, on-treatment	227,513	228,454	177,311	199,149	451	468	2.55	2.35	1.02
PS matching, intent-to-treat	229,230	229,168	600,743	183,043	1,881	1,855	2.89	2.72	1.10

Table 1. Target (T) cohort is ACEi new-users. Comparative (C) cohort is ARBs new-users. We report total number of patients, follow-up number of acute myocardial infarction events, and event incidence rate (IR) per 1,000 patient years (PY) in patient cohorts, as well as the desirable relative risk (RR)SH. Note that the IR does not account for any stratification or matching.

designs with stratification or matching. Figure 4 plots Kaplan-Meier survival curves for patients under the intent-to-treat design.

**Warning:** Using 'size' aesthetic for lines was deprecated in ggplot2. Please use 'linewidth' instead.  
**Warning:** Using 'displayed once every 8 hours' in aes() is deprecated. Use 'lifecycle::last\_lifecycle\_warnings()' to see where this warning originates.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual systematic error. In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEi and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEi to ARBs for the on-treatment and intent-to-

Design	T		C		Events		IR		
	T	C	T	C	T	C	T	C	
PS stratification, on-treatment	773,266	228,456	585,422	199,149	1,596	468	2.73	2.35	1.16
PS stratification, intent-to-treat	773,266	229,170	513,163	183,043	6,275	1,855	2.96	2.72	1.08
PS matching, on-treatment	227,513	228,454	177,311	199,149	451	468	2.55	2.35	1.02
PS matching, intent-to-treat	229,230	229,168	600,743	183,043	1,881	1,855	2.89	2.72	1.10

Table 1. Target (T) cohort is ACEi new-users. Comparative (C) cohort is ARBs new-users. We report total number of patients, follow-up number of acute myocardial infarction events, and event incidence rate (IR) per 1,000 patient years (PY) in patient cohorts, as well as the desirable relative risk (RR)SH. Note that the IR does not account for any stratification or matching.

designs with stratification or matching. Figure 4 plots Kaplan-Meier survival curves for patients under the intent-to-treat design.

**Warning:** Using 'size' aesthetic for lines was deprecated in ggplot2. Please use 'linewidth' instead.  
**Warning:** Using 'displayed once every 8 hours' in aes() is deprecated. Use 'lifecycle::last\_lifecycle\_warnings()' to see where this warning originates.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual systematic error. In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEi and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEi to ARBs for the on-treatment and intent-to-

Design	T		C		Events		IR		
	T	C	T	C	T	C	T	C	
PS stratification, on-treatment	773,266	228,456	585,422	199,149	1,596	468	2.73	2.35	1.16
PS stratification, intent-to-treat	773,266	229,170	513,163	183,043	6,275	1,855	2.96	2.72	1.08
PS matching, on-treatment	227,513	228,454	177,311	199,149	451	468	2.55	2.35	1.02
PS matching, intent-to-treat	229,230	229,168	600,743	183,043	1,881	1,855	2.89	2.72	1.10

Table 1. Target (T) cohort is ACEi new-users. Comparative (C) cohort is ARBs new-users. We report total number of patients, follow-up number of acute myocardial infarction events, and event incidence rate (IR) per 1,000 patient years (PY) in patient cohorts, as well as the desirable relative risk (RR)SH. Note that the IR does not account for any stratification or matching.

designs with stratification or matching. Figure 4 plots Kaplan-Meier survival curves for patients under the intent-to-treat design.

**Warning:** Using 'size' aesthetic for lines was deprecated in ggplot2. Please use 'linewidth' instead.  
**Warning:** Using 'displayed once every 8 hours' in aes() is deprecated. Use 'lifecycle::last\_lifecycle\_warnings()' to see where this warning originates.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual systematic error. In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEi and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEi to ARBs for the on-treatment and intent-to-

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Fig. 5. Distributions for ACEi and ARBs new users. The plot shows the distributions of the hazard ratio for incident events between comparable new users under on-treatment and intent-to-treat risk window assumptions. The shaded area denotes the 95% CI.

designs with stratification or matching. Figure 4 plots Kaplan-Meier survival curves for patients under the intent-to-treat design.

**Warning:** Using 'size' aesthetic for lines was deprecated in ggplot2. Please use 'linewidth' instead.  
**Warning:** Using 'displayed once every 8 hours' in aes() is deprecated. Use 'lifecycle::last\_lifecycle\_warnings()' to see where this warning originates.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual systematic error. In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEi and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEi to ARBs for the on-treatment and intent-to-

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.



# LEGENDMed Central Methods sections

- Pulled in a description of the database
- Referred to the protocol for exposure and outcome definitions
- Study design was fixed (it's a LEGEND study), so used standard text
- We did not yet use objective diagnostics at the time

OHDSI cohort definitions can be converted to human-readable text

Could modify to reflect design choices

such as those from the 2017 ACC/AHA guidelines and the 2018 European Society of Cardiology (ESC) and European Society of Hypertension (ESH) Guidelines for the management of arterial hypertension (Williams *et al.*, 2018).

## Methods

**Data source.** We conduct a new-user cohort study comparing new users of ACEIs with new users of ARBs in the Truven Health MarketScan Commercial Claims and Encounters Database (CCAE) database encoded in the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) version 5 (Hripcsak *et al.*, 2015; Overhage *et al.*, 2012; Ryan *et al.*, 2013). The CCAE represent data from individuals enrolled in United States employer-sponsored insurance health plans. The data includes adjudicated health insurance claims (e.g. inpatient, outpatient, and outpatient pharmacy) as well as enrollment data from large employers and health plans who provide private healthcare coverage to employees, their spouses, and dependents. Additionally, it captures laboratory tests for a subset of the covered lives. This administrative claims database includes a variety of fee-for-service, preferred provider organizations, and capitated health plans. The study period spans from 2000-12-31, 2000-12-31 to 2018-03-31, 2018-03-31.

**Study design.** This study follows a retrospective, observational, comparative cohort design (Ryan *et al.*, 2013). We include patients who are first time users of ACEIs or ARBs, and who have a diagnosis of hypertension on or prior to treatment initiation. We require that patients have continuous observation in the database for at least 365 days prior to treatment initiation. We exclude patients with prior acute myocardial infarction events and less than 1 day at risk. Links to full cohort details, include concept codes, are provided in the Supporting Information. The outcome of interest is acute myocardial infarction. We begin the outcome risk window 1 day after treatment initiation and consider two design choices to define the window end. First, we end the outcome time-at-risk window at first cessation of continuous drug exposure, analogous to an on-treatment design and, second, we end the outcome time-at-risk window when the patient is no longer observable in the database, analogous to an intent-to-treat design. Continuous drug exposures are constructed from the available longitudinal data by considering sequential prescriptions that have fewer than 30 days gap between prescriptions.

**Statistical analysis.** We conduct our cohort study using the open-source OHDSI CohortMethod R package (Schuemie *et al.*, 2018c), with large-scale analytics achieved through the Cyclops R package (Suchard *et al.*, 2013). We use propensity scores (PSs) – estimates of treatment exposure probability conditional on pre-treatment baseline features in the one year prior to treatment initiation – to control for potential measured confounding and improve balance between the target (ACEIs) and comparator (ARBs) cohorts (Rosenbaum and Rubin, 1983). We use an expansive PS model that includes all available patient demographics, drug, condition and procedure covariates generated through the FeatureExtraction R package (Schuemie *et al.*, 2018d) instead of a prespecified set of investigator-selected confounders. We perform PS stratification or variable-ratio matching and then estimate comparative ACEIs-vs-ARBs hazard ratios (HRs) using a Cox proportional hazards model. Detailed covariate and methods information are provided in the Supporting Information. We present PS and covariate balance metrics to assess successful confounding control, and provide HR

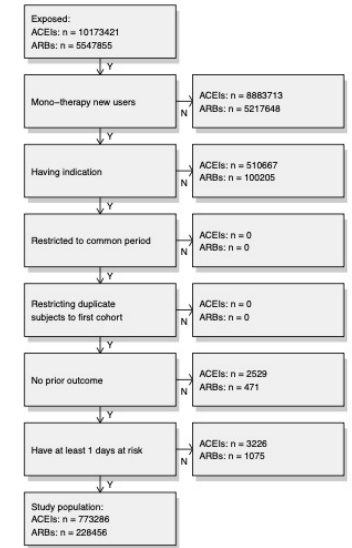


Fig. 1. Attrition diagram for selecting new-users of ACEIs and ARBs from the CCAE database.

estimates and Kaplan-Meier survival plots for the outcome of acute myocardial infarction.

Residual study bias from unmeasured and systematic sources can exist in observational studies after controlling for measured confounding (Schuemie *et al.*, 2014, 2016). To estimate such residual bias, we conduct negative control outcome experiments with 292 negative control outcomes identified through a data-rich algorithm (Voss *et al.*, 2017). We fit the negative control estimates to an empirical null distribution that characterizes the study residual bias and is an important artifact from which to assess the study design (Schuemie *et al.*, 2018a). Using the empirical null distribution and synthetic positive controls (Schuemie *et al.*, 2018b), we additionally calibrate all HR estimates, their 95% confidence intervals (CIs) and the *p*-value to reject the null hypothesis of no differential effect (HR = 1). Empirical calibration serves as an important diagnostic tool to evaluate if residual systematic error is sufficient to cast doubt on the accuracy of the unknown effect estimate.

## Results

**Population characteristics.** Figure 1 diagrams the inclusion of study subjects from the CCAE database under the on-treatment with stratification design. We augment these counts with cohort



# LEGENDMed Central Results sections

- Pulled figures and tables from the results database
  - Standard set of results artifacts, including ‘Table 1’
- Use simple logic to modify text
  - E.g. if  $\max(\text{SDM}) > 0.1$ , would call this out
- Did not use blinding when diagnostics failed

Nowadays we'd review all diagnostics and blind when appropriate

Could re-use a lot of the one-pager we printed for the closing session here!

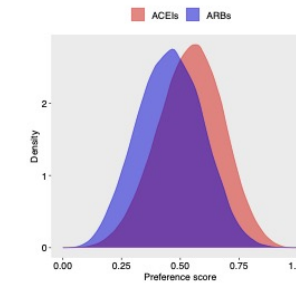


Fig. 2. Preference score distribution for ACEIs and ARBs new-users. The preference score is a transformation of the propensity score that adjusts for prevalence differences between populations. A higher overlap indicates that subjects in the two populations are more similar in terms of their predicted probability of receiving one treatment over the other.

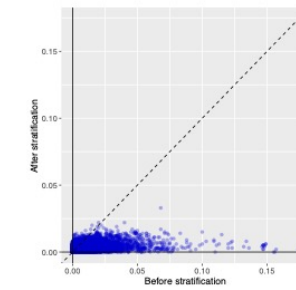


Fig. 3. Patient characteristics balance before and after stratification. As a rule-of-thumb, all values  $< 0.1$  is generally considered well-balance (Austin, 2009).

sizes we identify for the remaining designs in Table 1. This table also reports total patient follow-up time, numbers of acute myocardial infarction events these patients experience and unadjusted incidence rates. Table 2 compares base-line characteristics between patient cohorts.

**Patient characteristics balance.** Figure 2 plots the preference score distributions, re-scalings of PS estimates to adjust for differential treatment prevalences, for patients treated with ACEIs and ARBs. We assess characteristics balance achieved through PS adjustment by comparing all characteristics' standardized difference (StdDiff) between treatment group means before and after PS trimming and stratification (Table 2). Figure 3 plots StdDiff for all 5518 base-line patient features that serve as input for the PS model. Before stratification, 23 features have a StdDiff  $> 0.1$ . After stratification, the count is 0.

**Outcome assessment.** Table 3 details the time to first acute myocardial infarction or censoring distributions for patients in the ACEIs and ARBs cohorts. We report in Table 4 estimated HRs comparing ACEIs to ARBs for the on-treatment and intent-to-treat

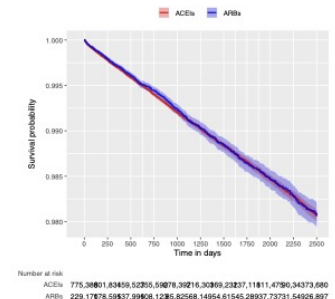


Fig. 4. Kaplan Meier plot of acute myocardial infarction-free survival. This plot is adjusted for the propensity score stratification; the ACEIs curve shows the actual observed survival. The ARBs curve applies reweighting to approximate the counterfactual of what ACEIs survival would look like had the ACEIs cohort been exposed to ARBs instead. The shaded area denotes the 95% CI.

designs with stratification or matching. Figure 4 plots Kaplan-Meier survival curves for patients under the intent-to-treat design.

```
## Warning: Using `size` aesthetic for lines was deprecated
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where
## generated.
```

**Residual systematic error.** In the absence of bias, we expect 95% of negative and positive control estimate 95% confidence intervals to include their presumed HR. In the case of negative controls, the presumed HR = 1. Figure 5 describes the negative and positive control estimates under the on-treatment with PS stratification design. Before calibration, negative and positive controls demonstrate poor coverage. After calibration, controls demonstrate acceptable coverage.

#### Conclusions

We find that ACEIs have a similar risk of acute myocardial infarction as compared to ARBs within the population that the CCAE represents.

#### Supporting Information

Here we enumerate the guiding principles of LEGEND and provide linking details on study cohorts and design.

#### LEGEND principles.

1. Evidence will be generated at large-scale.
2. Dissemination of the evidence will not depend on the estimated effects.
3. Evidence will be generated using a pre-specified analysis design.
4. Evidence will be generated by consistently applying a systematic approach across all research questions.



# What about other sections?

- Background
  - What question did we seek to answer?
  - Why is it important?
  - What is already known?
- Methods
  - What data were used?
  - How were the exposures and outcomes defined?
  - What statistical analysis was performed?
  - What objective diagnostics were used?
- Results
  - Which analyses passed diagnostics?
  - What were the results?
- Discussion
  - What have we learned?
  - How does that fit with what we already knew?

Background section is primarily a **synthesis of the context of the research question**

Large language models could nowadays help with this



# CLIO: background writer proof of concept





# Draft background section writer using generative AI

- Proof-of-concept suggests this is feasible
- Needs thorough evaluation
  - Initial results to reproduce background sections in PubMed Central show large variability in background sections
  - Will require human review
- Several methodological questions remain open
  - E.g. Should we pull in full-text articles?
- Non-trivial computational costs





# Conclusions

- Writing papers has become the bottleneck in generating and disseminating evidence
- There are opportunities for increasing efficiency
  - Methods and Results: We could create a template that is filled in with (standardized) analysis specifications and results
  - Background section: generative AI could help here, but more research is needed
- These are great opportunities for the OHDSI community!



# Join the Journey of the Evidence Translation Workgroup

Nicole Pratt



# Flashback...OHDSI APAC 2023

**Our ingredients are data**

**Our craft is science**

**Our brew is evidence**

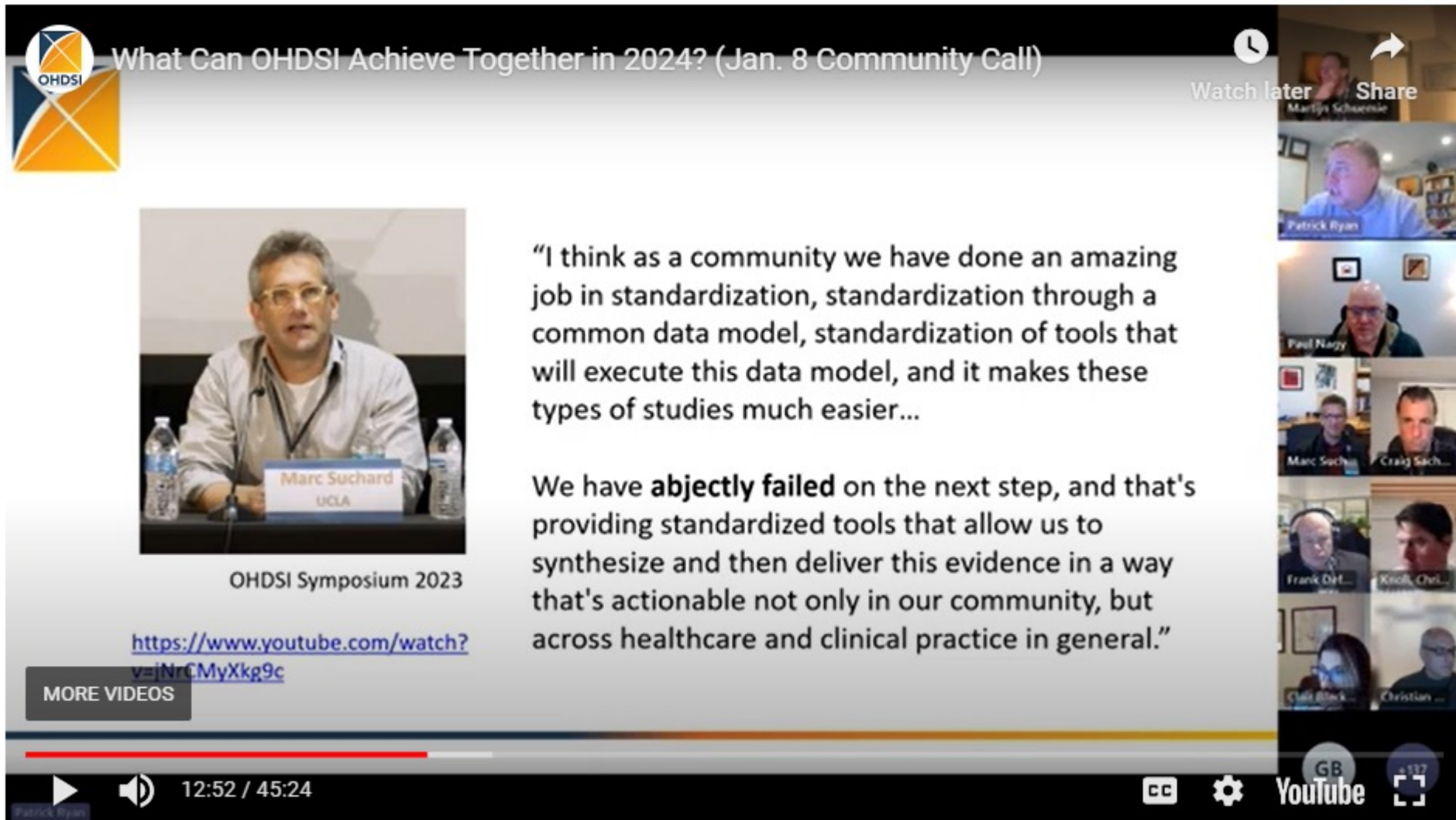
*Our duty is to share it*





OHDSI What Can OHDSI Achieve Together in 2024? (Jan. 8 Community Call)

Watch later Share



Marc Suchard  
UCLA

OHDSI Symposium 2023

<https://www.youtube.com/watch?v=jNrCMYXkg9c>

MORE VIDEOS

12:52 / 45:24

YouTube

Patrick Ryan

Martijn Schuense

Patrick Ryan

Paul Nagy

Marc Suchard

Craig Sach...

Frank Del...

Knob, Chel...

Chad Black...

Christian ...

GB +137

CC

Settings

Full Screen

“I think as a community we have done an amazing job in standardization, standardization through a common data model, standardization of tools that will execute this data model, and it makes these types of studies much easier...

We have **objectly failed** on the next step, and that's providing standardized tools that allow us to synthesize and then deliver this evidence in a way that's actionable not only in our community, but across healthcare and clinical practice in general.”

“Making evidence actionable” Patrick Ryan What can OHDSI Achieve Together in 2024 Jan 9  
<https://ohdsi.org/community-calls/>



How can OHDSI *improve the use and uptake* of the real world evidence we produce so that it is “actionable” and can be readily consumed to aid in decision making?





# We need everyone at the table to not only consume the evidence but also to **set the menu!**



Regulator

Clinician

Consumer

Researcher



How can OHDSI help *[you]* by generating evidence for the *questions that matter* to *[you]*?

What else can OHDSI do to *build trust* that the evidence we provide is reliable and can be used in decision making?

How can OHDSI *improve understanding* about real world evidence we produce so that it can be used in decision making?

**New** OHDSI Work Group:  
Evidence Translation



## Defining knowledge translation

Sharon E. Straus MD MSc, Jacqueline Tetroe MA, Ian Graham PhD

We cannot pick up a magazine or surf the Internet without facing reminders of the challenges to health care and the “sorry state” of health systems.<sup>1</sup> All health care systems are faced with the challenges of improving quality of care and reducing the risk of adverse events.<sup>2</sup> Globally, health systems fail to use evidence optimally. The result is inefficiency and a reduction in both quantity and quality of life.<sup>3,4</sup> For example, McGlynn and colleagues<sup>5</sup> found that adults in the United States received less than 55% of recommended care. Providing evidence from clinical research (e.g., through publication in journals) is necessary but not enough for the provision of optimal care.

Recognition of this issue has created interest in knowledge translation, also known as KT, which we define as the methods for closing the gaps from knowledge to practice. In this series of articles, we will provide a framework for implementing knowledge for clinicians, managers and policy-makers.

### Key points

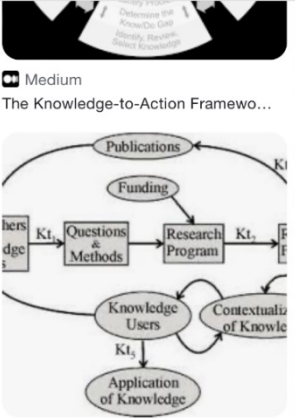
- Gaps between evidence and decision-making occur at all levels of health care, including those of patients, health care professionals and policy-makers.
- Knowledge translation involves using high-quality knowledge in processes of decision-making.
- The knowledge-to-action framework provides a model for the promotion of the application of research and the process of knowledge translation.

Knowledge creation (i.e., primary research), knowledge distillation (i.e., the creation of systematic reviews and guidelines) and knowledge dissemination (i.e., appearances in journals and presentations) are not enough on their own to ensure the use of knowledge in decision-making.

We should also clarify what knowledge translation isn't.



Q knowledge translation framework



ResearchGate CIHR Knowledge Translation (KT) M...



Medium The Knowledge-to-Action Framework...

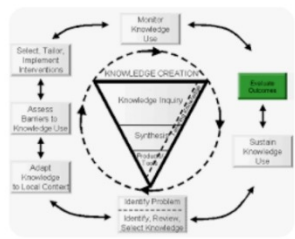


Figure 1. The knowledge-to-action framework

Q knowledge translation framework



cihr-irsc.gc.ca About us - CIHR



cihr-irsc.gc.ca Section 6.1: Methodologies to Eval...



cihr-irsc.gc.ca Knowledge-To-Action Cycle ...



cihr-irsc.gc.ca Knowledge-To-Action Cycle ...



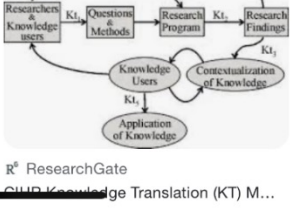
ktddr Knowledge Translation: Introduction...



cihr-irsc.gc.ca Knowledge-To-Action Cycle ...



ktddr Knowledge Translation: Introduction...



ResearchGate CIHR Knowledge Translation (KT) M...

Q knowledge translation framework

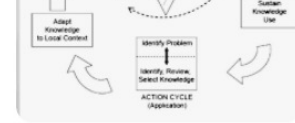


knowledge translation strategies

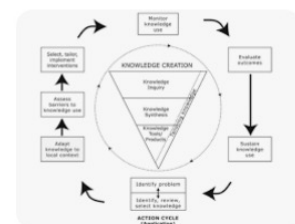
research knowledge translation

knowledge translation process

knowledge translation healthcare



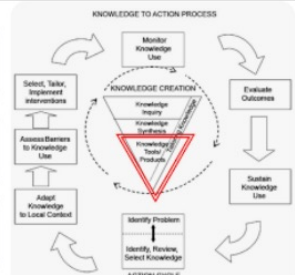
Semantic Scholar Knowledge translation: putting the ...



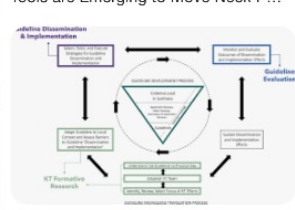
ktddr Knowledge Translation



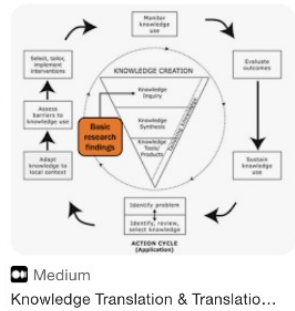
cihr-irsc.gc.ca



Benham Open Tools are Emerging to Move Neck P...



Canadian Science Publishing Canadian 24-Hour Movement Guid...

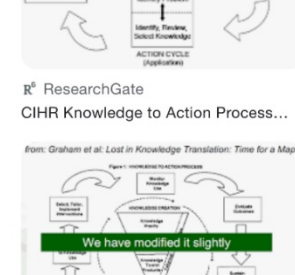


Medium Knowledge Translation & Translati...



cihr-irsc.gc.ca

Q knowledge translation framework



ResearchGate CIHR Knowledge to Action Process...

from: Graham et al. Lost in Knowledge Translation: Time for a Map?

We have modified it slightly



cihr-irsc.gc.ca Section 5.1 Knowledge disseminati...



Instituts de recherche en sant  d... Knowledge Translation Funding Pro...



cihr-irsc.gc.ca Knowledge-To-Action Cycle ...



cihr-irsc.gc.ca



Medium The Knowledge-to-Action Framework...

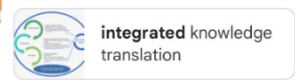
Related searches

knowledge translation examples

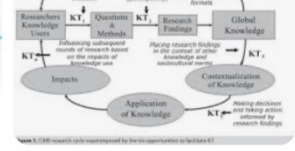
knowledge to action framework

health care knowledge translation

integrated knowledge translation



Semantic Scholar PDFJ Knowledge Translation ...

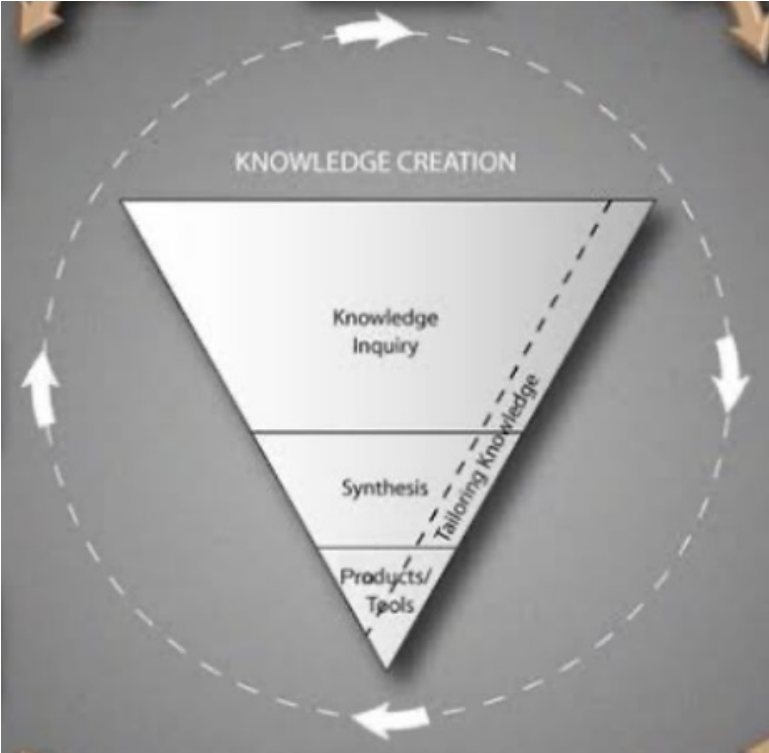


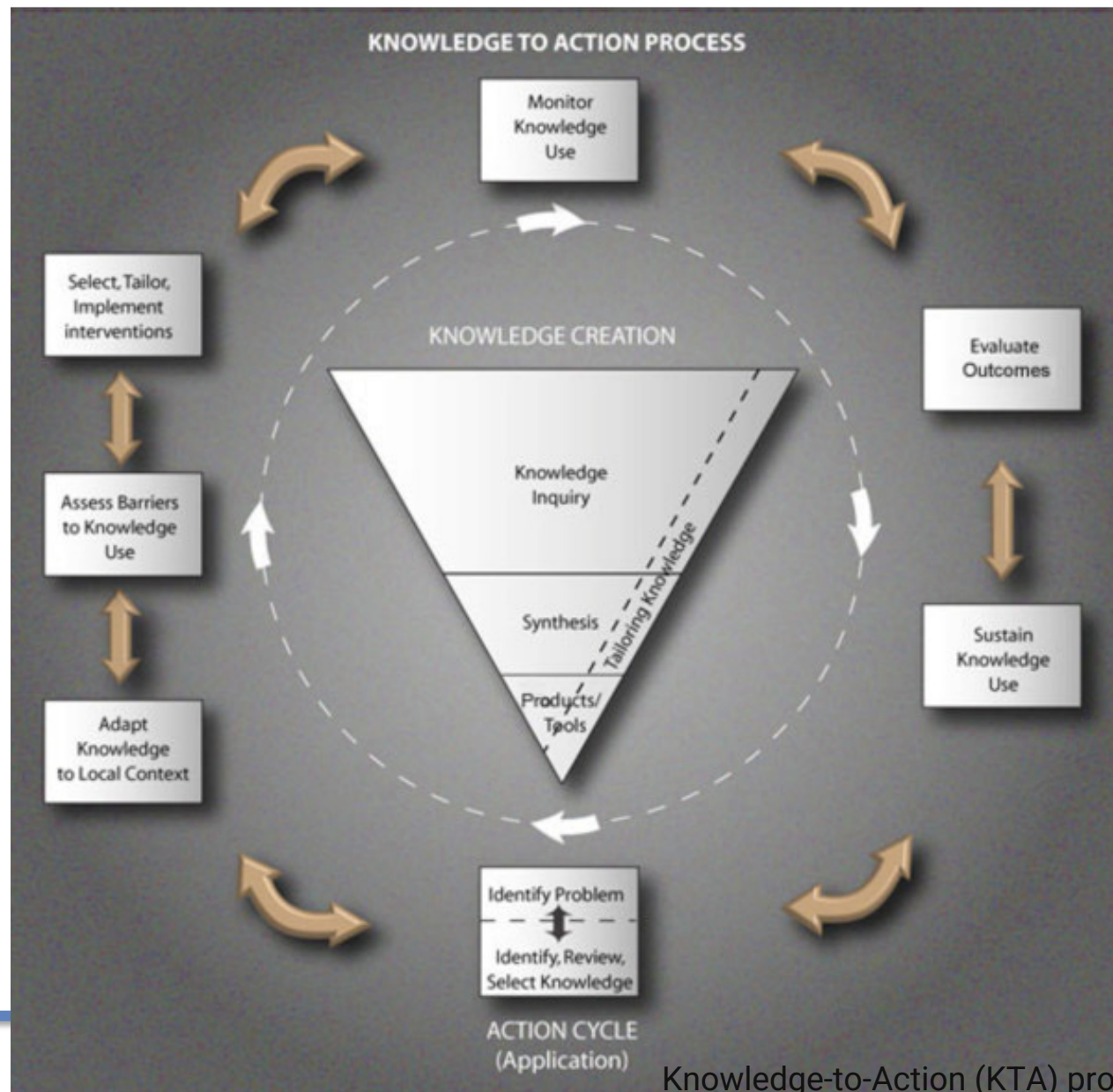
Medium Say "Knowledge Translation ...



cihr-irsc.gc.ca







Source: Cochrane Knowledge Translation Framework April 2017



Goal 1:  
Producing evidence

Prioritization and co-production

Goal 2:  
Accessible evidence

Packaging, push and support to implementation

Facilitating pull

Goal 3:  
Advocating for evidence

Exchange

Improving climate

Goal 4: Effective & sustainable

Sustainable KT processes

## Producing the Evidence

- **Prioritisation**  
Identify the questions that are important to different stakeholders
- **Co-production**  
Identify opportunities for stakeholder involvement throughout the evidence generation pipeline to ensure alignment of needs

## Making it accessible

- **Packaging, push and support to implementation**  
Identify methods and pathways for evidence dissemination
- **Facilitating pull**  
Making evidence findable accessible and developing capacity in end-users finding and using evidence

## Making evidence palatable

- **Exchange**  
Develop consumable evidence communications that are tailored to the needs of different stakeholders. Develop strategic partnerships, forums to exchange ideas

## Measuring the uptake

- **Monitoring** the uptake, reach and impact of OHDSI research into policy and/or practice

Objectives

Create an appetite

Set the table

Make it palatable

Are they eating it?

Consumer forums  
Engagement with regulators

Create user friendly findable, open-source tools "evidence libraries"

Create evidence briefs, lay summaries

Create impact stories, diffusion of evidence, audiences reached

Key Results



# Next Steps

- Join the work group: WG sign up for the community to Teams channel
  - Set up a schedule of Meetings
  - Set Objectives & Key Result (OKR)
  
  - **Purpose:** The Evidence Translation workgroup exists to promote and facilitate the dissemination and uptake of evidence generated by the OHDSI community into all aspects of health care decision-making.
-