Welcome

Land Acknowledgement

Housekeeping

Federated Analysis:
State of the Science

Réseau de recherche sur les données de santé du Canada
Health Data Research Network Canada

Join the conversation

@hdrn_rrds

#HDRNCanada
#FederatedAnalysis
#DistributedData
#MultiRegionalData

Current Approaches for Distributed Analysis

Michael Paterson
Principal Investigator, CNODES

Dr. Robert Platt
Executive Co-Lead, CNODES

CANADIAN NETWORK FOR OBSERVATIONAL
DRUG EFFECT STUDIES (CNODES)

Current Approaches for Distributed Analysis

Robert Platt and Michael Paterson

Federated Analysis: State of the Science Collective Learning Series
February 8, 2024

Outline of Session

- State of the possible in the current federated analysis environment
- Common analytic protocols vs common data models
- Types and characteristics of common data models
- CNODES' use of the Sentinel Common Data Model

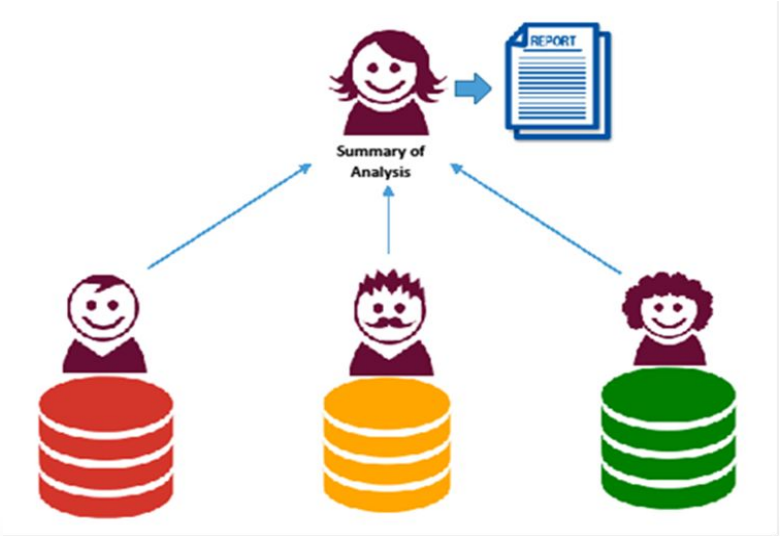
Options for federated analysis

- Separate analyses, pool results
- Common protocol-based analyses
- **Common data model-based analyses**
- Privacy-preserving fully federated analyses
  - Eg DataSHIELD

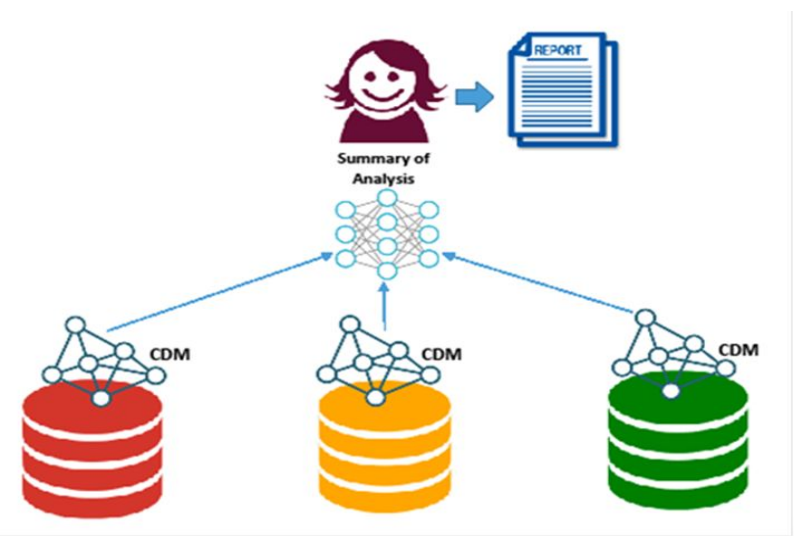
Approaches to distributed analysis

- Centres maintain possession of their own data and permit coordinated, controlled access (querying) for network purposes

Common Protocol



Common Data



What is a Common Data Model (CDM)?

- Tool to standardize data into a **common format**
  - Structure
    - Basic dataset formatting
  - Format
    - Value sets, variable formats
  - Meanings
    - Encounters etc.
- **May** include analytic code/analytics layer

Advantages and Limitations of Common Protocol Approach

- **Advantages**
  - Analytical flexibility: design, data sources, statistical methods
  - Local capacity building
    - Data repository development
    - Analytical and methods expertise
- **Limitations**
  - Timeliness
  - Risk for inefficiency, error

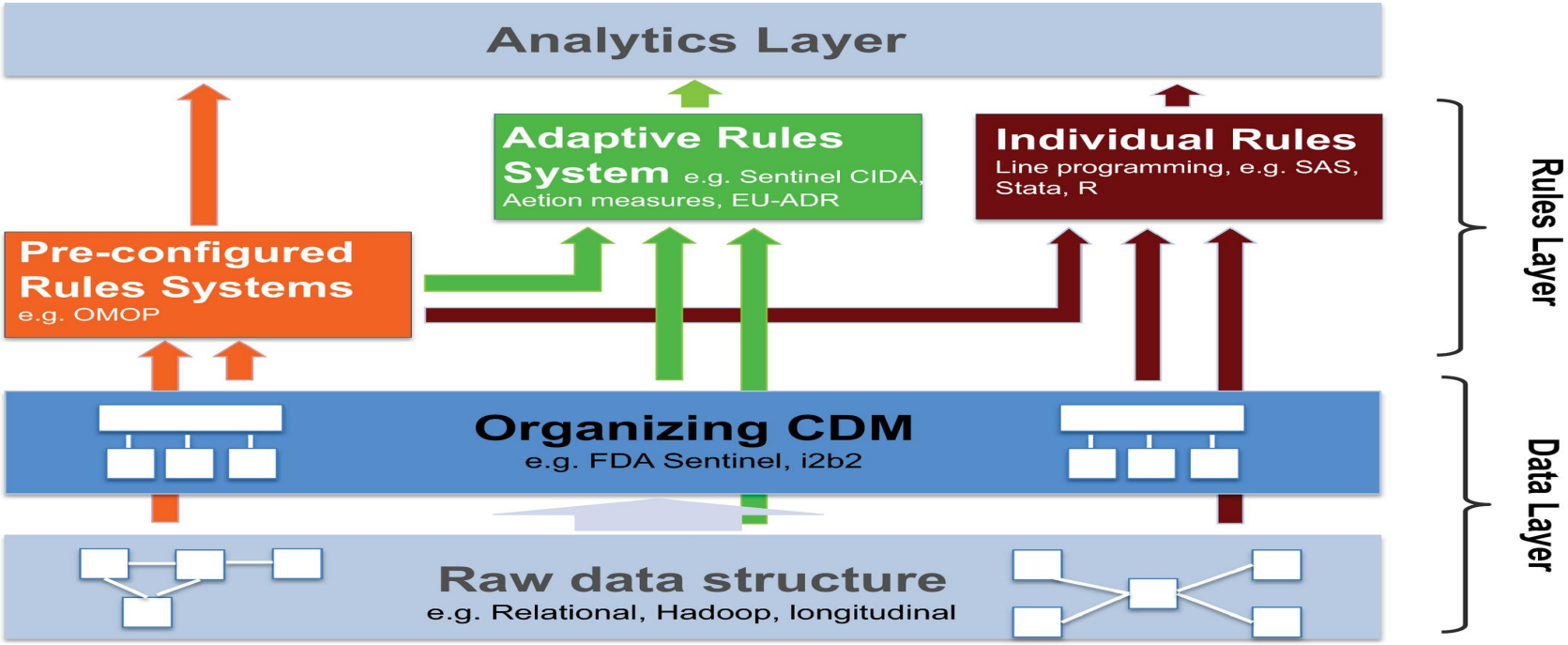
Why CDM?

- Standardized code
  - Reliable, tested against data formats
  - Reproducible
  - Rapid – no need for modification
  - Guardrails – prevents inappropriate analyses
  - At extreme, standardized/automated reports
- Multi-database studies
  - Run **the same code** across multiple sites

Types of CDMs

- Organizing CDM (e.g., Sentinel)
  - Reorganize data files into a standard data structure
  - No modifications to raw data **content**
  - Project-specific coding **(“adaptive rules system”)**
- Mapping CDM (e.g., OMOP)
  - Map raw data to concepts/constructs
  - Mapping usually comprehensive and not study-specific
  - **“Preconfigured rules system”**

Types of CDMs



Clin Pharma and Therapeutics, Volume: 107, Issue: 4, Pages: 827-833, First published: 22 July 2019, DOI: (10.1002/cpt.1577)

Examples of CDMs

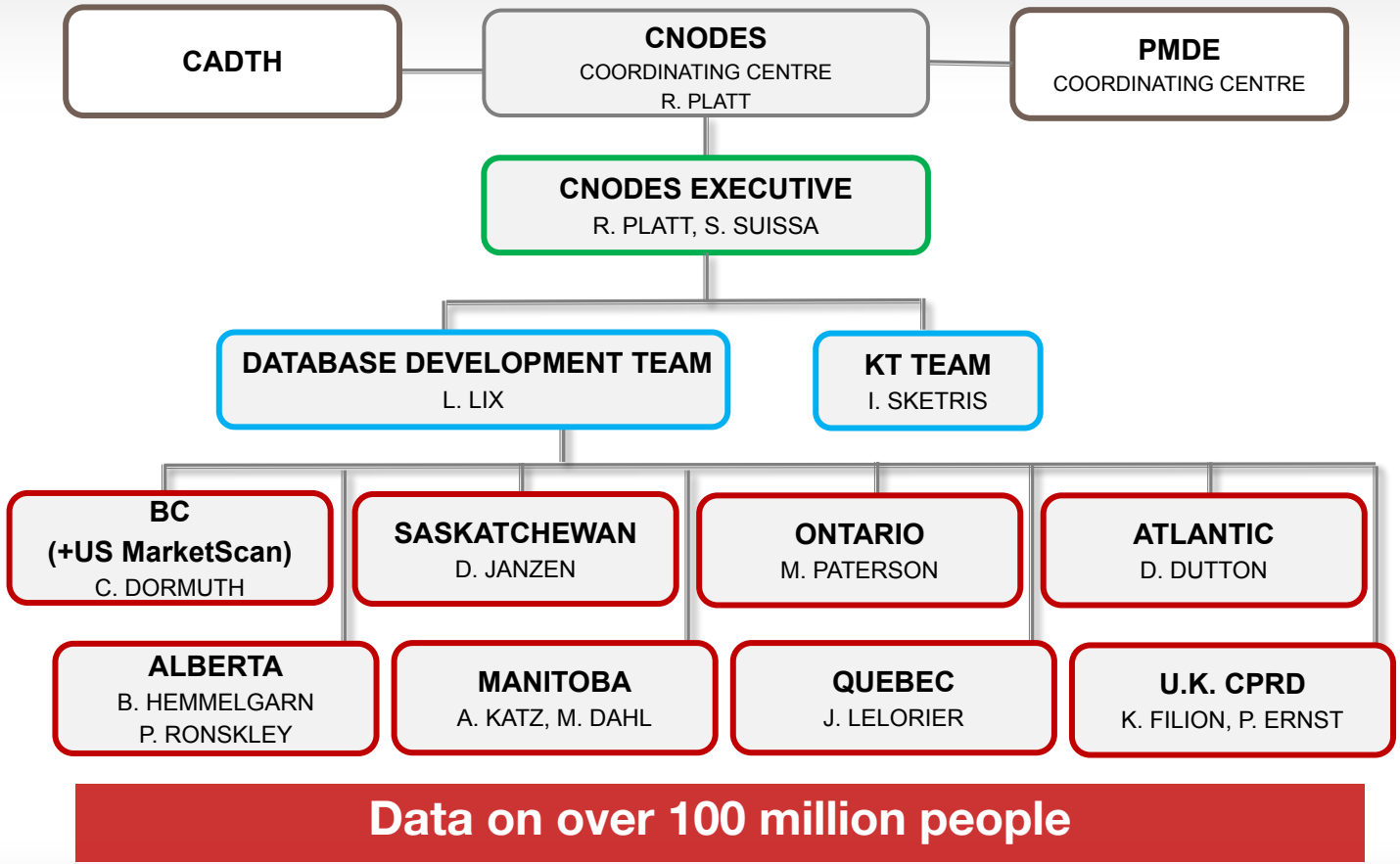
- Sentinel – upcoming presentation
- Area-specific CDMs, e.g., ConcePTION for pregnancy/linked mother-baby
- Open research CDMs (PCORnet)
- Commercial CDMs (Aetion, Panalgo, TriNetX)
- OMOP/OHDSI
  - DARWIN-EU
  - OHDSI – open-source community of researchers

CDM Analytics

- CDMs usually come with a set of analytic packages
  - Sentinel, OMOP – open source SAS, R code respectively
  - Aetion and other proprietary platforms (TriNetX, Panalgo, others)
- Increases speed, reliability, reproducibility
  - Pre-specified reports
- Limits flexibility
  - Only what's pre-written

CNODES at a glance

Canadian Network for Observational Drug Effect Studies (CNODES) uses ***population-based administrative healthcare data*** to provide ***timely responses*** to queries for Canadian public stakeholders regarding drug safety and effectiveness



Data on over 100 million people

CNODES
CANADIAN NETWORK FOR OBSERVATIONAL DRUG EFFECT STUDIES

Data Sources

- Linked provincial administrative health data

CORE	SUPPLEMENTARY
• health insurance registries • prescription drug claims • physician service claims • hospital discharge abstracts • emergency department records • vital statistics	• cancer registries • pregnancy registries • laboratory test results • health surveys • CPRD: EMR-based risk factor data (e.g., smoking status, alcohol use, BMI, blood pressure, lipids, etc.)

FDA Sentinel CDM: Main Components

1) **Standardized, regularly refreshed common data tables**
   - Enrollment, Demographic, Dispensing, Encounter, Diagnosis, Procedure, Death, Mother-Infant Linkage
2) **Query Tool: reusable, flexible modular programs with standardized outputs (stratified by age, sex, year, site)***
   - Numbers newly exposed to specific drugs over a specified time period, overall and nested within specific diagnoses
   - Outcome events (ED visits, hospitalizations, procedures) following a new drug exposure
   - Numbers exposed to multiple drugs concomitantly, persistence, switching, etc

*CIDA: Cohort Identification and Descriptive Analysis (www.sentinelinitiative.org/sentinel/surveillance-tools/routine-querying-tools/routine-querying-system)

Sentinel CDM Data Tables

Administrative Data

Enrollment	Demographic	Dispensing	Encounter	Diagnosis	Procedure
Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID
Enrollment Start & End Dates	Birth Date	Dispensing Date	Service Date(s)	Service Date(s)	Service Date(s)
Drug Coverage	Sex	National Drug Code (NDC)	Encounter ID	Encounter ID	Encounter ID
Medical Coverage	Zip Code	Days Supply	Encounter Type and Provider	Encounter Type and Provider	Encounter Type and Provider
Medical Record Availability	Etc.	Amount Dispensed	Facility	Diagnosis Code & Type	Procedure Code & Type
			Etc.	Principal Discharge Diagnosis	Etc.

Clinical Data

Lab Result	Vital Signs
Patient ID	Patient ID
Result & Specimen Collection Dates	Measurement Date & Time
Test Type, Immediacy & Location	Height & Weight
Logical Observation Identifiers Names and Codes (LOINC®)	Diastolic & Systolic BP
Etc.	Tobacco Use & Type
	Etc.

Registry Data

Death	Cause of Death	State Vaccine
Patient ID	Patient ID	Patient ID
Death Date	Cause of Death	Vaccination Date
Source	Source	Admission Date
Confidence	Confidence	Vaccine Code & Type
Etc.	Etc.	Provider
		Etc.

Inpatient Data

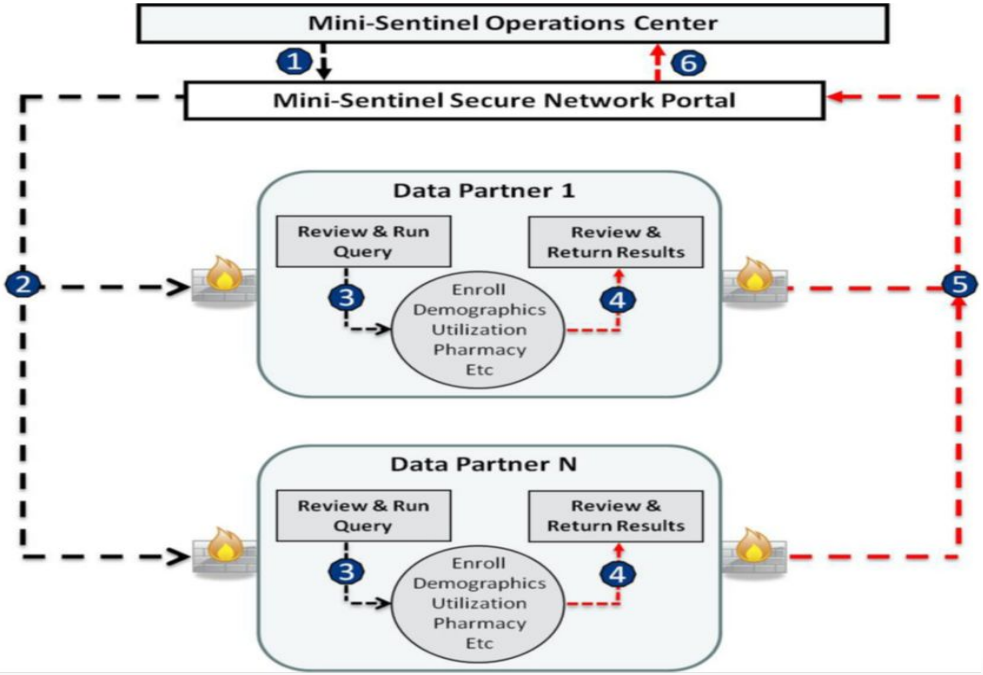
Inpatient Pharmacy	Inpatient Transfusion
Patient ID	Patient ID
Administration Date & Time	Administration Start & End Date & Time
Encounter ID	Encounter ID
National Drug Code (NDC)	Transfusion Administration ID
Route	Transfusion Product Code
Dose	Blood Type
Etc.	Etc.

Mother-Infant Linkage Data

Mother-Infant Linkage
Mother ID
Mother Birth Date
Encounter ID & Type
Admission & Discharge Date
Child ID
Child Birth Date
Mother-Infant Match Method
Etc.

Sentinel Query Process

- Standardized SAS programs are distributed to partner sites, run against the CDM tables, and results are securely returned for aggregation



1. User submits query
2. Data partners retrieve code
3. Partners review and run code against local data
4. Results generated
5. Partners review and return output via secure network
6. Results aggregated at Operations Center and returned to requestor

Examples of FDA Sentinel Analyses/Programs

1) CONCOMITANT DRUG USE

- Concomitant Use of Opioids and Cytochrome P450 (CYP) Inhibitors and Inducers (posted 02/01/2018)
- https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Report_CDER_mpl1r_wp055_nsdp_v01.pdf

2) PATIENT CHARACTERISTICS AND DRUG USE

- Characteristics of Gout Patients and Use of Urate-Lowering Therapies (posted 03/22/2019)
- https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Sentinel_Final_Report_cder_mpl1r_wp123_Report_1.pdf

3) DRUG USE AND CRUDE OUTCOME EVENT RATES

- Dabigatran, Warfarin and GI Bleed, Intracerebral Hemorrhage (posted 02/11/2013)
- https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_Modular-Program-Report_MSY3_MPR41_Dabigatran-Warfarin-GIH-ICH_Part-1_1.pdf

Examples of FDA Sentinel Analyses, cont'd

4) PROPENSITY SCORE-MATCHED COHORT STUDY

- Use of Dabigatran, Warfarin and Gastrointestinal Hemorrhage (posted 08/17/2018)
- https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/MS_Brief_Report2_to16_cap_mpl2r_wp005_nsdp_v02.pdf

5) SELF-CONTROLLED RISK INTERVAL ANALYSIS

- Seizure following Ranolazine Use: A Self-Controlled Risk Interval Analysis (posted 01/03/2019)
- https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Sentinel_Report_Ranexa_and_Seizures_L2.pdf

CNODES CDM Pilot Project

- CDM tables constructed and quality assured
  - Phase 1: SK, MB, ON, NS
  - Phase 2: BC, AB, US MarketScan
- Pilot studies
  - Use of New Molecular Entities approved by Health Canada in 2015
  - Gastrointestinal or intracerebral hemorrhage following new use of oral anticoagulants among patients with atrial fibrillation
  - Rhabdomyolysis and acute kidney injury following new use of statins

CDM Pilot Project Timelines in Phase 1

Site	(1) Data Custodial/REB Approval Time (Mean Days)	(2) Study Conduct	(3) Permission to Share Small Cells Approval Time (Mean Days)	(4) Results Pooling	Total
SK	1) 10, 2) 32, 3) 28 = 23 d	~14 d	42 d (typically, 7-10 d)	~3 d	~80 d
MB	1) 40, 2) 54, 3) 54 = 50 d		**Not permissible (serves as pooling site)**		
ON	1) 3, 2) 10, 3) 21 = 10 d		Permissible on query approval		
NS	Commitment to 5 d		Blanket agreement for pilot project		

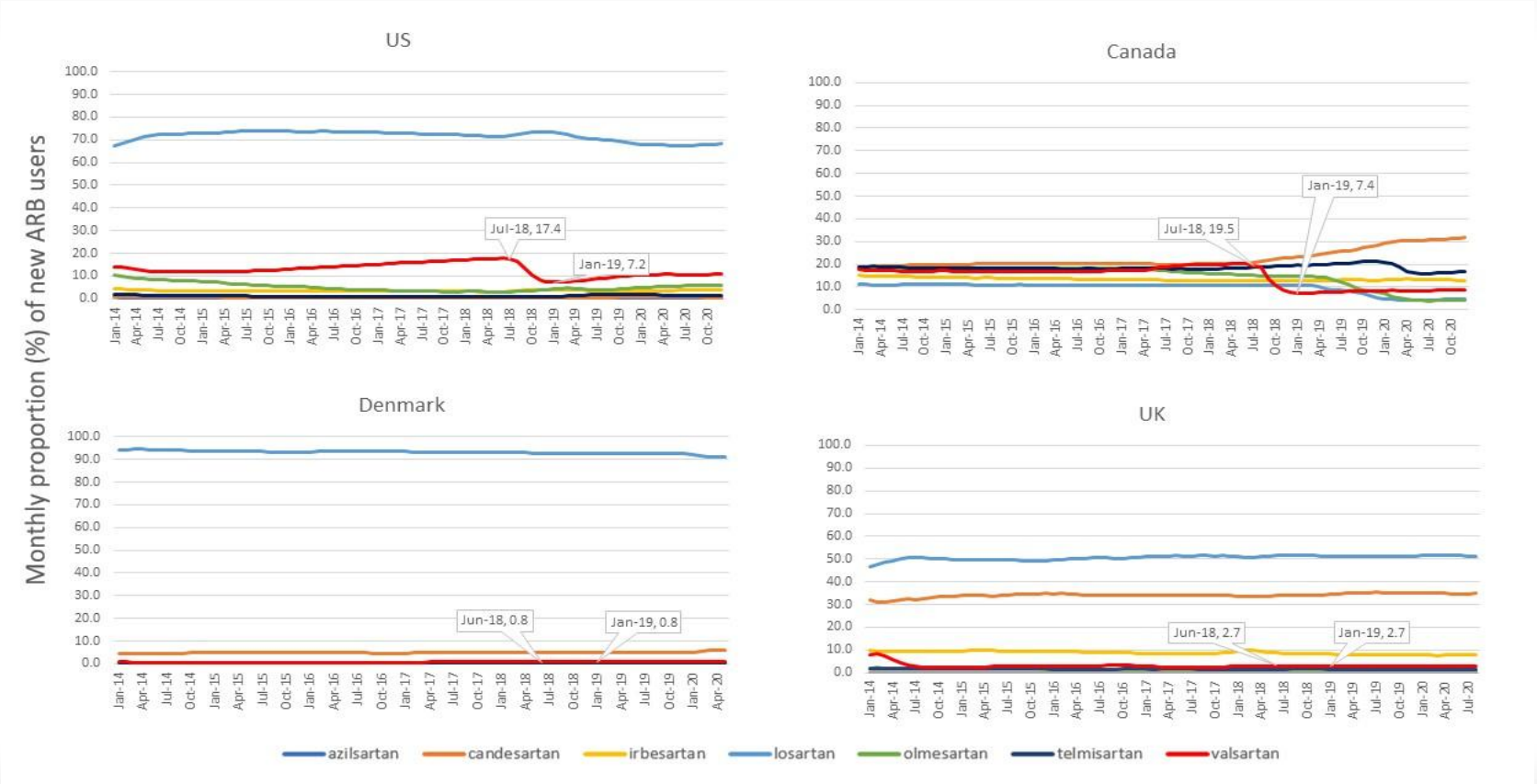
Lessons from CDM Pilot

- Need to continue to work to shorten ethical and data custodial approval times
  - Test of project timeliness based upon a limited number of simple queries that did not include time for consultations with query submitters
- Unable to secure access to Quebec data
- Data missing from CDM tables
  - Laboratory test results, vaccines, drugs dispensed in hospital, infused chemotherapies
    - Have since added a lab table with COVID-19 PCR lab tests

Health Canada Projects Completed since CDM Pilot

Query title	Objective
Utilization of **angiotensin II receptor blockers (ARBs) following nitrosamine recalls** - in collaboration with FDA, EU teams	• To describe and compare differences in ARB utilization before and after nitrosamine drug safety communications (DSCs)
Utilization of **corticosteroids in COVID-19 patients** - in collaboration with FDA; comparable studies undertaken by FDA and EMA	• To describe utilization of systemic corticosteroids for the treatment of COVID-19 among outpatients
Natural history of **coagulopathy in COVID-19 patients** - in collaboration with FDA and EMA	• To determine the incidence of arterial and venous thromboembolic complications and their consequences among patients infected with COVID-19

Utilization of ARBs following nitrosamine recalls



CANADIAN NETWORK FOR OBSERVATIONAL DRUG EFFECT STUDIES

Systemic Corticosteroid Treatment in Patients with COVID-19

- UK-based RECOVERY trial found reduced mortality with dexamethasone treatment in ventilated patients or those with supplemental oxygen
- Observational studies have reported mixed outcomes, especially in mild cases, leading to recommendations for non-use of systemic corticosteroids
- **Aim:** To describe the utilization of systemic corticosteroids for treatment of COVID-19 in outpatients within 90 days of diagnosis

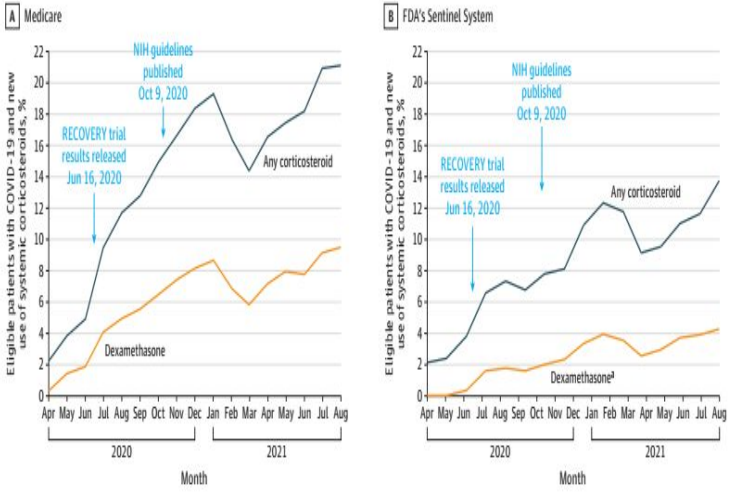
US Sentinel

- Medicare: 2.2% in Apr 2020 to 21.1% in Aug 2021
- Overall: 2.2% in Apr 2020 to 13.8% in Jul 2021

Canada (BC, MB, ON)

- < 1% in Apr 2020 to 5% in Jan 2021; range: 9% in ON, < 2% BC, MB

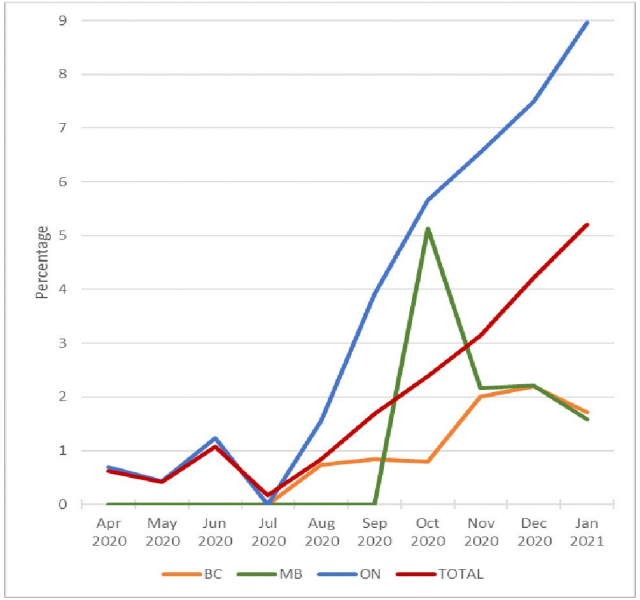
Figure. Proportion of Patients With COVID-19 Initiating Systemic Corticosteroids Within 14 Days of Diagnosis



FDA indicates Food and Drug Administration; NIH, National Institutes of Health; RECOVERY, Randomised Evaluation of COVID-19 Therapy.

[a] The name of the corticosteroid was only available for pharmacy dispensings.

Figure 1. Monthly percentage of COVID-19 outpatients dispensed systemic corticosteroids, overall and by province, April 2020 - January 2021

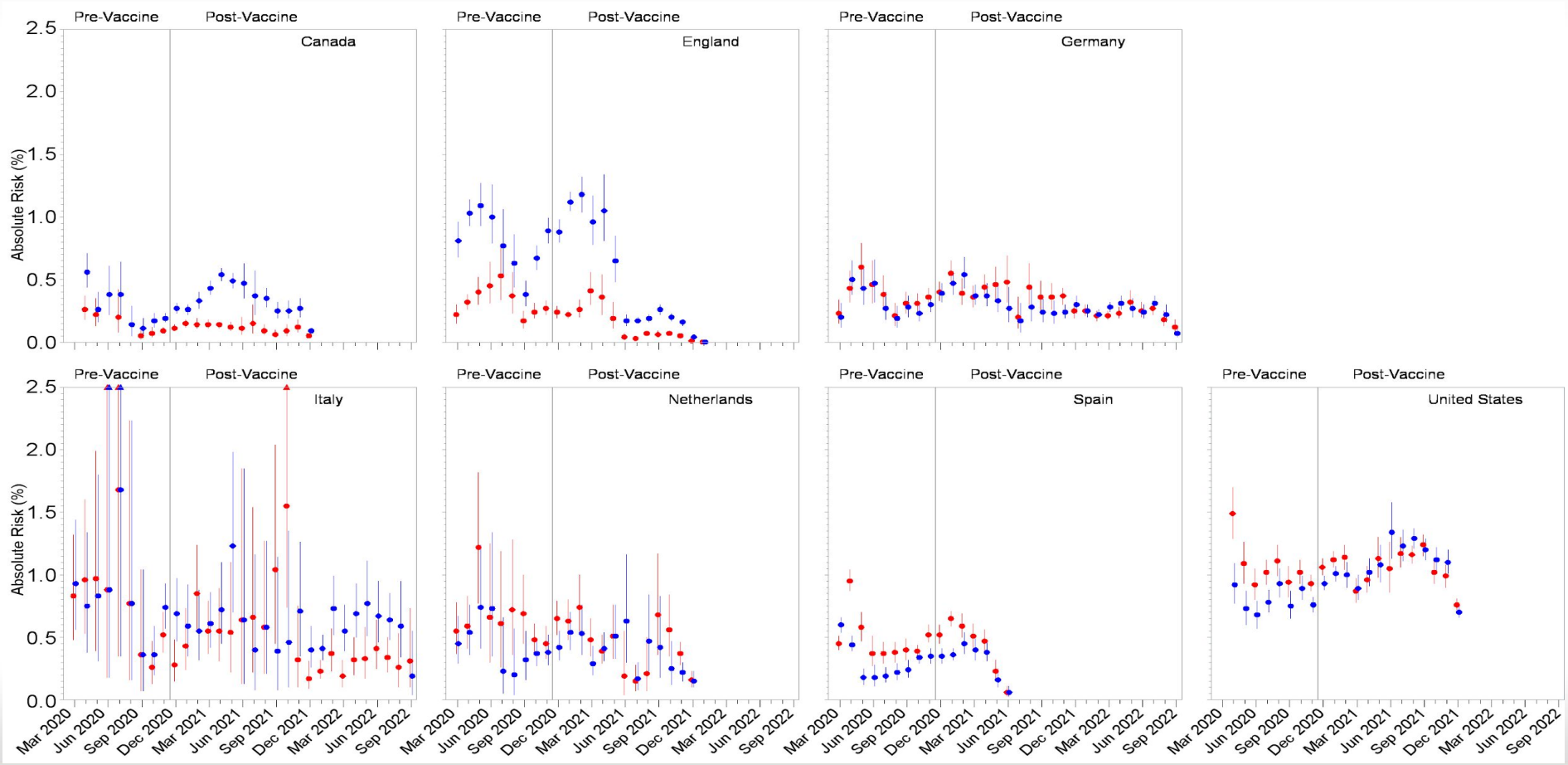


CNODES
CANADIAN NETWORK FOR OBSERVATIONAL DRUG EFFECT STUDIES

COVID-19 and risk of thrombotic events

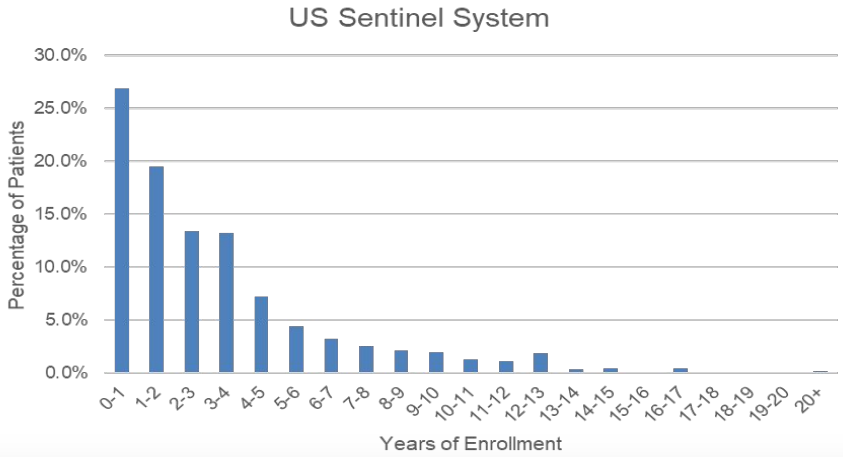
- **Aim:** Assess risk of arterial (myocardial infarction and stroke) and venous (pulmonary embolism, deep venous thrombosis) thrombotic events within 90 days of COVID-19 diagnosis in the ambulatory setting and the hospital setting

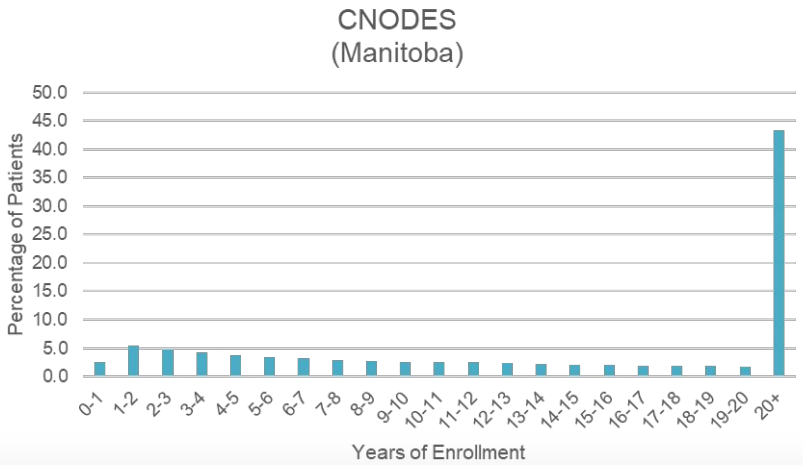
COVID-19 and risk of thrombotic events



Advantages of International Collaboration: US more populous, but Canada has greater follow-up time

	US	Canada
Total Patient IDs with Medical and Drug Coverage	342M	30M
Total Patients Currently Accruing with Medical and Drug Coverage	113M	12.4M





CNODES
CANADIAN NETWORK FOR OBSERVATIONAL DRUG EFFECT STUDIES

Latest Canada-US CDM Demonstration Projects

Project title	Objective
Safety Monitoring Following Ozempic Use in Patients with Diabetes	• To determine whether there are potential safety signals among adults with diabetes who are newly treated with Ozempic relative to similar patients newly treated with sitagliptin
Utilization of Antidiabetic Drugs During Pregnancy	• To describe the utilization of antidiabetic drugs during pregnancy

Looking Forward

- Further improve efficiency through ongoing negotiations for shorter ethics and data custodial approval times
- Continue to explore opportunities to acquire QC data
- Expand data sources and analytic tools available to CDM queries
- Demonstrate experience with, value of the new data/tools through new CoLab demonstration studies (eg, TreeScan)

Thank you! Questions?

Visit us at www.cnodes.ca

Federated Analysis:
State of the Science

Réseau de recherche sur les données de santé du Canada
Health Data Research Network Canada

Q&A

Got feedback?

Scan QR code to fill out our feedback form

Federated Analysis

State of the Science Collective Learning Series

Panel Discussion:

Current Approaches for Distributed Analysis

Dr. Judith Maro

James Weaver

Michael Paterson

Thursday, March 14

10:00 a.m. PT | 1:00 p.m. ET

Thank you!

For events & updates, follow HDRN Canada:

hdrnrrds

@hdrn_rrds

@hdrncanada

hdrn.ca