March to Data Fitness

Clair Blacketer

Agenda

1. GDE Definition of “Fitness for use”?

2. The Role of the OHDSI Evidence Network
   - How to Join as a Data Partner

3. Evaluating “Fitness for use”?
   - Example walkthrough

Agenda

1. **GDE Definition of “Fitness for use”?**

2. The Role of the OHDSI Evidence Network
   - How to Join as a Data Partner

3. Evaluating “Fitness for use”?
   - Example walkthrough

GDE Fitness for Use

1. Interest
   - Which data partners would like be part of this effort to contribute evidence?
2. Technological Readiness
   - Data partners that indicate interest will be asked to run a Strategus test package to assess technological readiness
3. Study fitness
   - All data partners will be evaluated against all studies using data diagnostics

Agenda

1. **GDE Definition of “Fitness for use”?**

2. **The Role of the OHDSI Evidence Network**
   - **How to Join as a Data Partner**

3. Evaluating “Fitness for use”?
   - Example walkthrough

OHDSI Evidence Network

To facilitate collaborative research efforts and ensure the quality and integrity of data across the OHDSI network

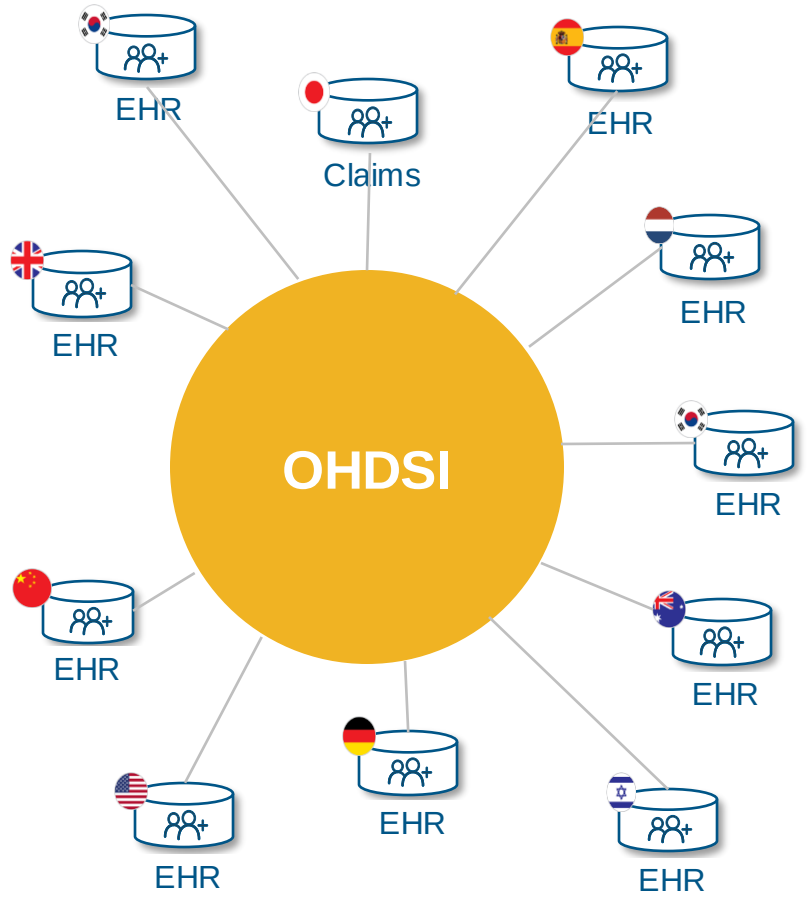
What is it?

Resource comprised of summary statistics of databases within the OHDSI network

- Held securely at the OHDSI Coordinating Center
- Used to inform network studies

Patient level data does not leave participating site

Compliance with privacy and IRB regulations



Current State of the Network

42

Data sources

23

Data Partner Organizations

https://ohdsi.github.io/EvidenceNetwork/tools/ohdsi_evidence_network.html

What you need in place

To join the network as a DPO

- Observational health data standardized to the OMOP CDM v5.3 or higher
- Data held in a relational database accessible by the organization
  - List of supported SQL environments here: https://ohdsi.github.io/SqlRender/articles/UsingSqlRender.html#translation-to-other-sql-dialects
- Approval from governance entity (i.e. IRB) to share metadata and concept counts with the OHDSI Coordinating Center (OCC)
  - Note: It is up to each individual DPO as owner or licensee of data to ensure all appropriate governance requirements are followed.
- The ability to run the DbDiagnostics R package against the data

CB

What you need to do

To join the network as a DPO

- Run the DbDiagnostics package *executeDbProfile* function to generate metadata and high-level concept counts about each data source submitted to the network
    - The aggregate information gathered by the package is listed here: https://ohdsi.github.io/DbDiagnostics/articles/SummaryStatistics.html
    - If the Achilles package was run previously and the results stored this step will take approximately 15-30 minutes, depending on the environment
    - If the Achilles package was not run previously or if the results were not stored this step will take approximately 1-8 hours, depending on the environment.
- Send the resulting information to the OCC via SFTP. Please contact evidencenetwork@ohdsi.org for the key file when you are ready

CB

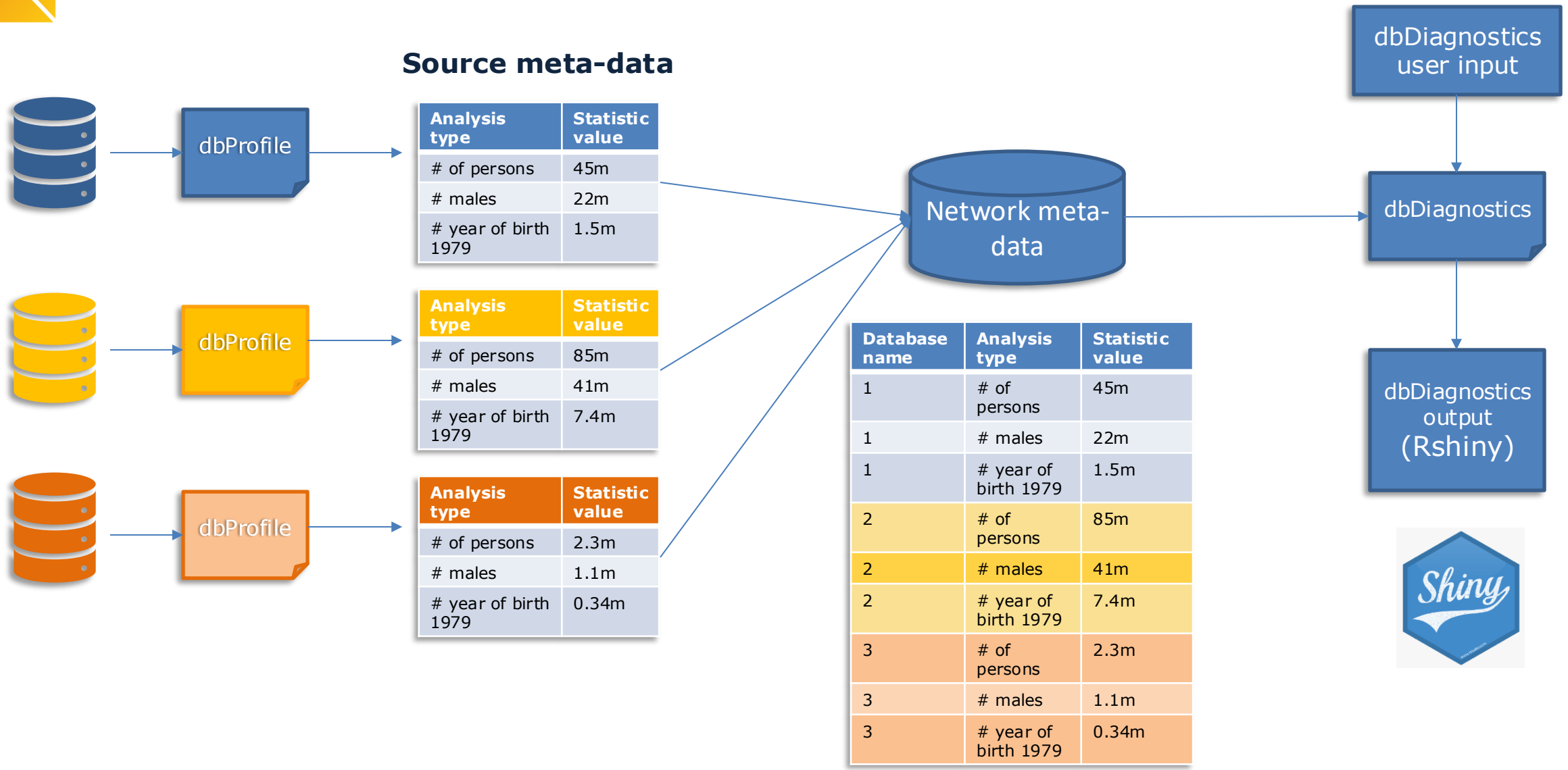
Agenda

1. GDE Definition of “Fitness for use”?

2. The Role of the OHDSI Evidence Network
   - How to Join as a Data Partner

3. Evaluating “Fitness for use”?
   - Example walkthrough

Data Diagnostics Process



Database Profile Summary Statistics

List of Summary Statistics included in executeDbProfile

The executeDbProfile function will execute and/or export the following aggregate summary statistics and DQD checks:

Persons	total
	by gender
	by year of birth
	by race
	by ethnicity
	with at least one day of observation in each month
	by observation period start month
	by number of observation periods
	by length of observation period (in 30d increments)
	with at least one visit occurrence, by visit_concept_id
	distinct patients that overlap between specific domains - including death
	with at least one concept_id, by measurement_concept_id
	with at least one concept_id, by condition_concept_id
	with at least one concept_id, by procedure_concept_id
	with at least one concept_id, by drug_concept_id
	with at least one concept_id, by device_concept_id
	with at least one concept_id, by observation_concept_id

Measurement	occurrence records, by measurement_concept_id
	occurrence records, by measurement_source_concept_id
	records with no value (numeric, string or concept)
Condition occurrence	by condition_concept_id
	by condition_source_concept_id
Drug exposure	by drug_concept_id
	by drug_source_concept_id
Procedure occurrence records	by procedure_concept_id
	by procedure_source_concept_id
Observation occurrence records	by observation_concept_id
	by observation_source_concept_id
Device exposure records	by device_concept_id
	by device_source_concept_id
Distribution of numeric values	by measurement_concept_id and unit_concept_it

Data Diagnostics User Inputs

All Available Inputs

Age range: min __ – max __

Gender: ☑ Male ☑ Female ☑ Unknown

Race: ☑ White ☑ Black ☑ Asian ☑ Unknown

Ethnicity: ☑ Hispanic ☑ Not Hispanic ☑ Unknown

Calendar time: Start Date _____ - End Date _____

Minimum longitudinal follow-up: ___ days

Target: <concept list>

Comparator: <concept list>

Indication: <concept list>

Outcome: <concept list>

Data requirements		T + C	O
Domains	Condition	x	x
	Drug	x	x
	Procedure	x	x
	Measurement	x	x
	Measurement with value	*na*	x
	Device	x	x
	Observation	x	x
	Death	x	x
Visits	Inpatient	x	x
	Emergency Room	x	x
	Outpatient	x	x

Data Diagnostics User Inputs

Ticagrelor vs. Prasugrel for Patients with Acute Coronary Syndrome undergoing Percutaneous Coronary Intervention with NACE Outcome*

Age range: min 18 – max NA

Gender: ✅ Male ✅ Female

Race: NULL (meaning there are no restrictions)

Ethnicity: NULL

Calendar time: Start Date __NULL___ - End Date __NULL__

Minimum longitudinal follow-up: 90 days

Target: ticagrelor

Comparator: prasugrel

Indication: percutaneous coronary intervention

Outcome: Net Adverse Clinical Event

Data requirements		T + C	O
Domains	Condition	x	
	Drug	x	
	Procedure	x	
	Measurement		
	Measurement with value		
	Device		
	Observation		
	Death		
Visits	Inpatient		x
	Emergency Room		x
	Outpatient		

* Study has a total of 11 outcomes

Example Study Question Inputs Across Outcomes

	NACE	All Cause Mortality	Cardiovascular Mortality	Stroke
Age restrictions	> 18	> 18	> 18	> 18
Gender(s)	male, female	male, female	male, female	male, female
Race(s)	All	All	All	All
Ethnicities(s)	All	All	All	All
Study Dates	Any	Any	Any	Any
Required Lookback	90 days	90 days	90 days	90 days
T + C Domains	condition, drug, procedure	condition, drug, procedure	condition, drug, procedure	condition, drug, procedure
O Domains	NULL	death	death	NULL
T + C Visits	NULL	NULL	NULL	NULL
O Visits	IP, ER	NULL	NULL	IP, ER
Target	Ticagrelor	Ticagrelor	Ticagrelor	Ticagrelor
Comparator	Prasugrel	Prasugrel	Prasugrel	Prasugrel
Indication	PCI	PCI	PCI	PCI
Outcome	NACE	All Cause Mortality	Cardiovascular Mortality	Stroke

Analysis variants = 11

GDE Study Participation

Diagnostics

- Study leads to provide data diagnostics settings
- Evidence Network runs diagnostics and returns results

Study Leads Contact Partners

- Study leads will receive *ohdsi.org* email addresses for data partners that meet criteria
- Study leads reach out to data partners inviting them to join

Partners Join and Run Study

- Data partners decide if they would like to join the studies
- Data partners run the studies using the strategus template provided by study leads


OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Questions?