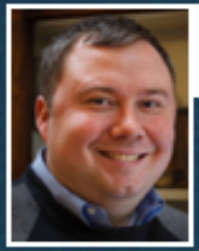
OHDSI 2020 Plenary Session:

Large-Scale Network Phenotype Development, Evaluation and Characterization



Patrick Ryan, PhD

A Framework for Phenotype Development and Evaluation

Anna Ostropolets, MD

Concept Prevalence — a OHDSI Network Study
Design Diagnostics — PHOEBE

Gowtham Rao, MD, PhD

Data Diagnostics
OHDSI Phenotype Library

This session will be shown twice (1 am ET, 8 am ET) during the OHDSI Global Symposium so collaborators around the world have an opportunity to see these exciting presentations.

Anthony Sena

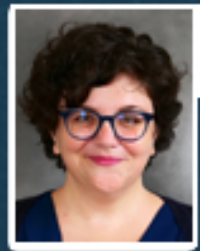
A Framework for Large-Scale Characterization

Talita Duarte-Salles, PhD

CHARYBDIS — Large-Scale Characterization of COVID-19 Disease Natural History

Dani Prieto-Alhambra, MD, PhD

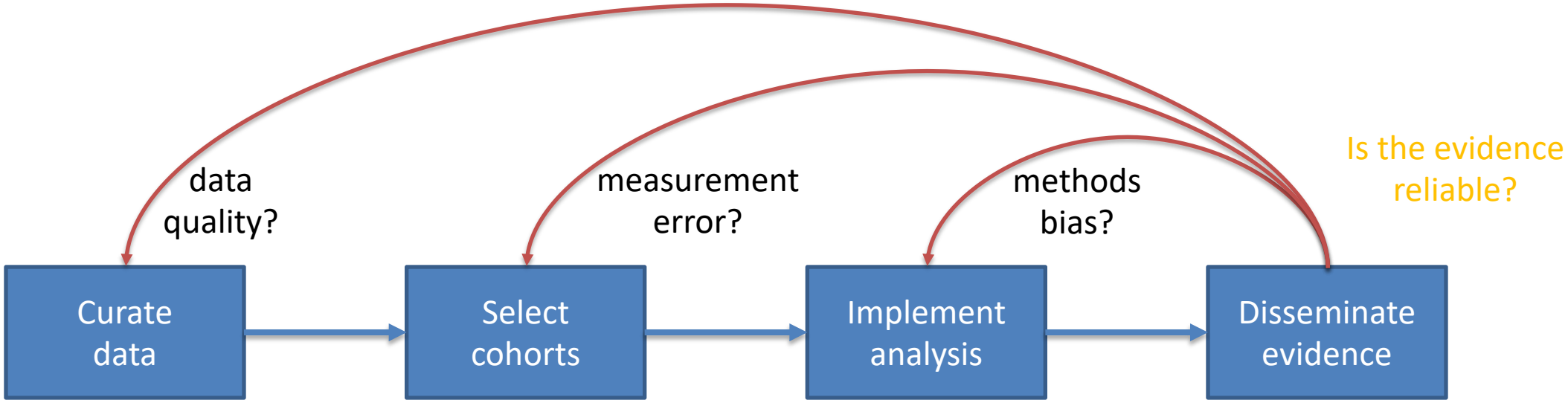
SCYLLA — Large-Scale Characterization of COVID-19 Treatment Utilization



Noémie Elhadad, PhD

HERA — Large-Scale Characterization of Health Equity

Current status quo in observational research



JAMA®

Opinion

VIEWPOINT

Weighing the Benefits and Risks of Proliferating Observational Treatment Assessments

Observational Cacophony, Randomized Harmony

Robert M. Califf, MD
Verily Life Sciences (Alphabet), South San Francisco, California.

Adrian F. Hernandez, MD, MHS
Duke Clinical Research Institute, Durham, North Carolina; and Division of Cardiology, Department of Medicine, Duke University School of Medicine, Durham, North Carolina.

Martin Landray, MBChB
Nuffield Department of Population Health, University of Oxford, Headington, Oxford, United Kingdom.

Amid the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic, substantial effort is being directed toward mining databases and publishing case series and reports that may provide insights into the epidemiology and clinical management of coronavirus disease 2019 (COVID-19). However, there is growing concern about whether attempts to infer causation about the benefits and risks of potential therapeutics from nonrandomized studies are providing insights that improve clinical knowledge and accelerate the search for needed answers, or whether these reports just add noise, confusion, and false confidence. Most of these studies include a caveat indicating that "randomized clinical trials are needed." But disclaimers aside, does this approach help make the case for well-designed randomized clinical trials (RCTs) and accelerate their delivery?[1] Or do observational studies reduce the likelihood of a properly designed trial being performed, thereby delaying the discovery of reliable truth?

The growth of structured registries and organization of claims and electronic health record data have directly involved in discourse about treatments they assert are effective. The natural desire of all elements of society to find effective therapies can obscure the difference between a proven fact and an exaggerated guess. Nefarious motives are not necessary for these problems to occur.

The role of regulators in this context is crucial. In the United States, the 21st Century Cures Act and user fee agreements require industry, academia, and regulators to advance the use of data and evidence from clinical settings.[3] This legislation directed the US Food and Drug Administration (FDA) and the National Institutes of Health (NIH) to work with the clinical research ecosystem to develop robust methods for generating such evidence and clear guidance for applying it. Historically, the FDA has insisted on high-quality evidence as a condition for granting marketing approval for drugs and devices, and for specific marketing claims.

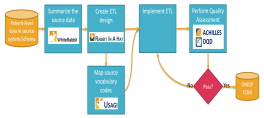
Considerable progress has been made in defining appropriate methods for improving the quality of observational treatment comparisons. Both NIH- and FDA-funded work fosters transparency by publishing study

Alternative title (and motivation for OHDSI):
Scientific Cacophony,
Harmony achieved through ***collaboration***, not randomization

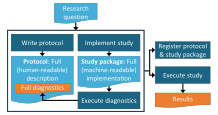
OHDSI's approach to observational research

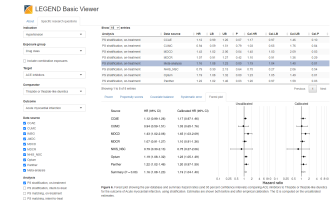
Curate data **network** → Select cohorts **within the network** → Implement analysis **across the network** → Disseminate evidence **from the network**

Standardized processes:

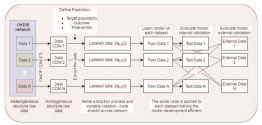


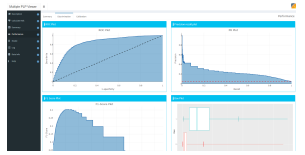
Estimation





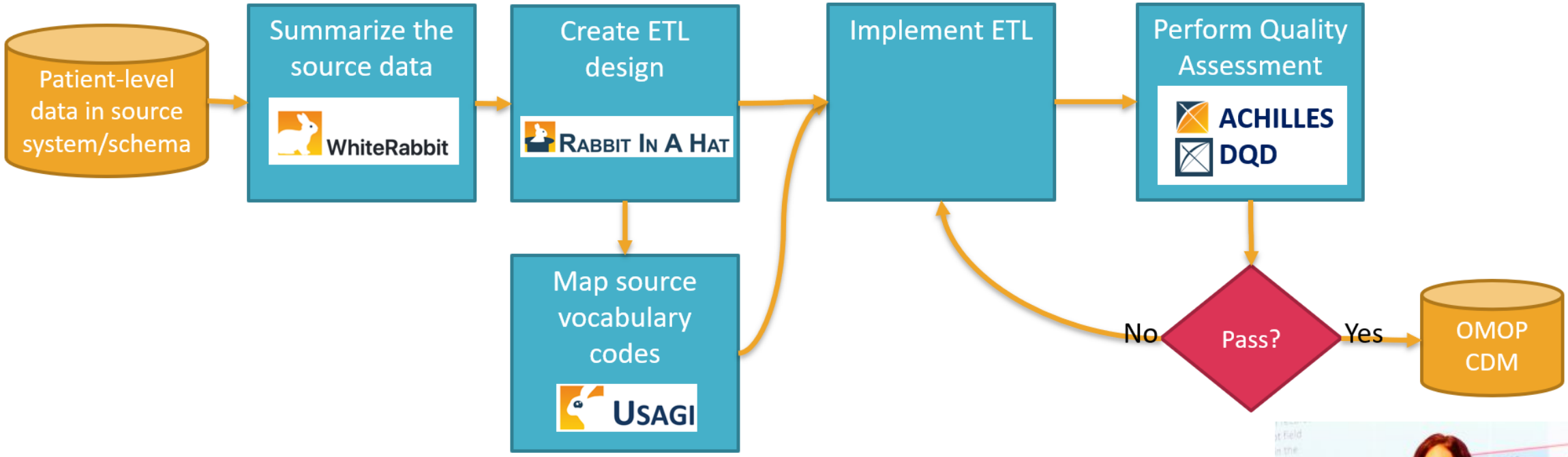
Prediction





Characterization

Standardized process for data curation



Standardized process for analysis implementation

Research question

Write protocol

Protocol: Full (human-readable) description

Full diagnostics

Implement study

Study package: Full (machine-readable) implementation

Execute diagnostics

Register protocol & study package

Execute study

Results

Standardized process for prediction model development and validation

Journal of the American Medical Informatics Association, 0(0), 2018, 1
doi: 10.1093/jamia/ocy032
Research and Applications

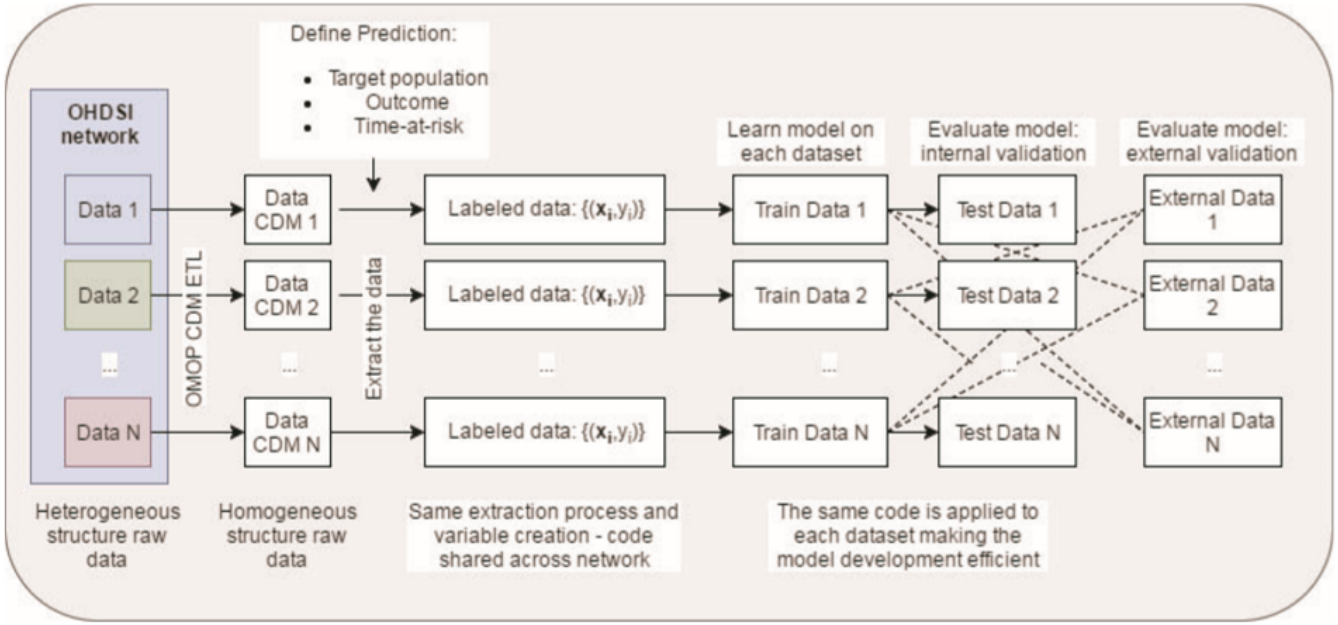
Research and Applications

Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data

Jenna M Reps,[1] Martijn J Schuemie,[1] Marc A Suchard,[2] Patrick B Ryan,[1] and Peter R Rijnbeek[3]

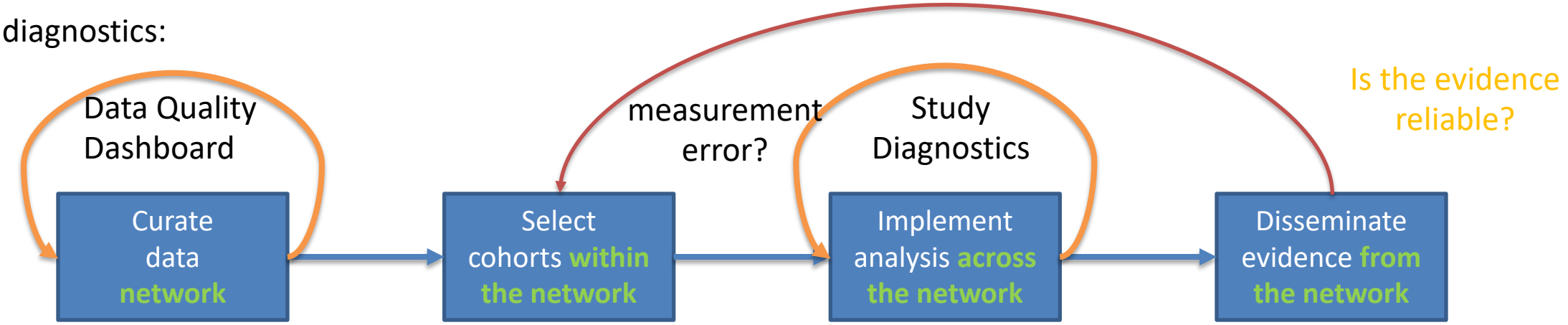
[1]Janssen Research and Development, Raritan, NJ, USA, [2]Department of Biomathematics, UCLA School of Medicine, CA, USA, and [3]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam,The Netherlands

Corresponding Author: Dr Jenna M Reps, Janssen Research and Development, Raritan, New Jersey, USA; jreps@its.jnj.com

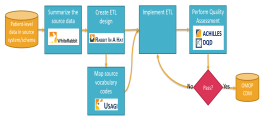


OHDSI's approach to observational research

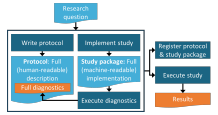
Empirical diagnostics:



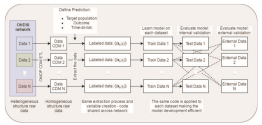
Standardized processes:

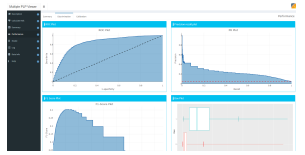


Estimation



Prediction





Characterization

Journal of the American Medical Informatics Association, 0(0), 2017, 1–6
doi: 10.1093/jamia/ocx110
Perspective

Perspective

High-fidelity phenotyping: richness and freedom from bias

George Hripcsak[1] and David J Albers[1]

[1]Department of Biomedical Informatics, Columbia University Medical Center, New York, NY, USA

- A phenotype is a specification of an observable, potentially changing state of an organism (as distinguished from the genotype, derived from genetic makeup).
- The term phenotype can be applied to patient characteristics inferred from electronic health record (EHR) data.
- The goal is to draw conclusions about a target concept based on raw EHR data, claims data, or other clinically relevant data.
- Phenotype algorithms – ie, algorithms that identify or characterize phenotypes – may be generated by domain experts and knowledge engineers, or through diverse forms of machine learning to generate novel representations of data.

OHDSI's definition of 'cohort'

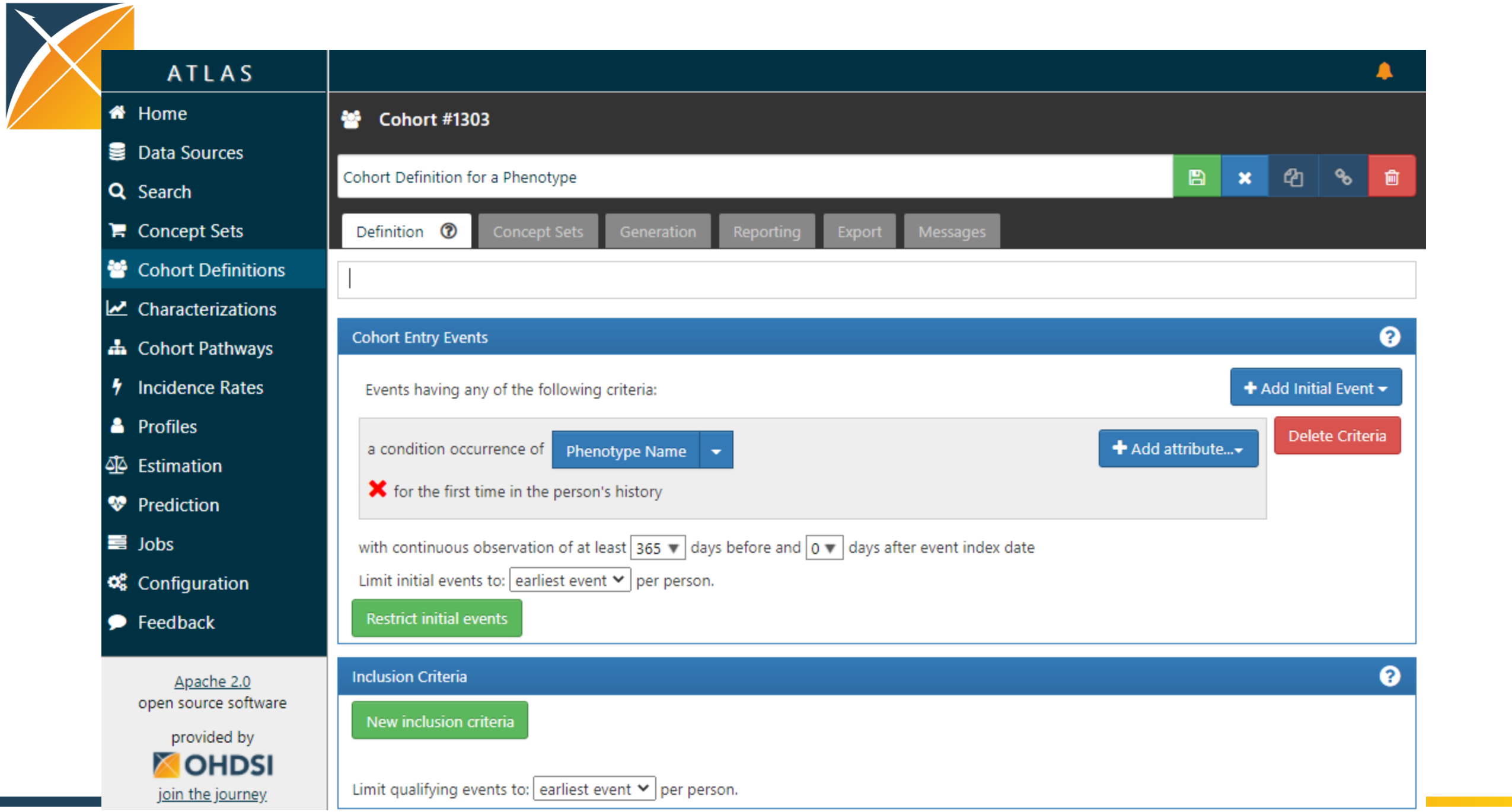
Cohort = a set of persons who satisfy one or more inclusion criteria for a duration of time

Cohort era = a continuous period during which a person has satisfied a cohort's inclusion criteria

Cohort definition = the specification for how to identify a cohort

A codeset is NOT a cohort…

…logic for how to use the codes in criteria is required


ATLAS
Home
Data Sources
Search
Concept Sets
Cohort Definitions
Characterizations
Cohort Pathways
Incidence Rates
Profiles
Estimation
Prediction
Jobs
Configuration
Feedback
Apache 2.0
open source software
provided by
OHDSI
join the journey
Cohort #1303
Cohort Definition for a Phenotype
Definition
Concept Sets
Generation
Reporting
Export
Messages
Cohort Entry Events
Events having any of the following criteria:
+ Add Initial Event
a condition occurrence of
Phenotype Name
+ Add attribute...
Delete Criteria
for the first time in the person's history
with continuous observation of at least 365 days before and 0 days after event index date
Limit initial events to: earliest event per person.
Restrict initial events
Inclusion Criteria
New inclusion criteria
Limit qualifying events to: earliest event per person.

The phenotype of Type 2 Diabetes Mellitus

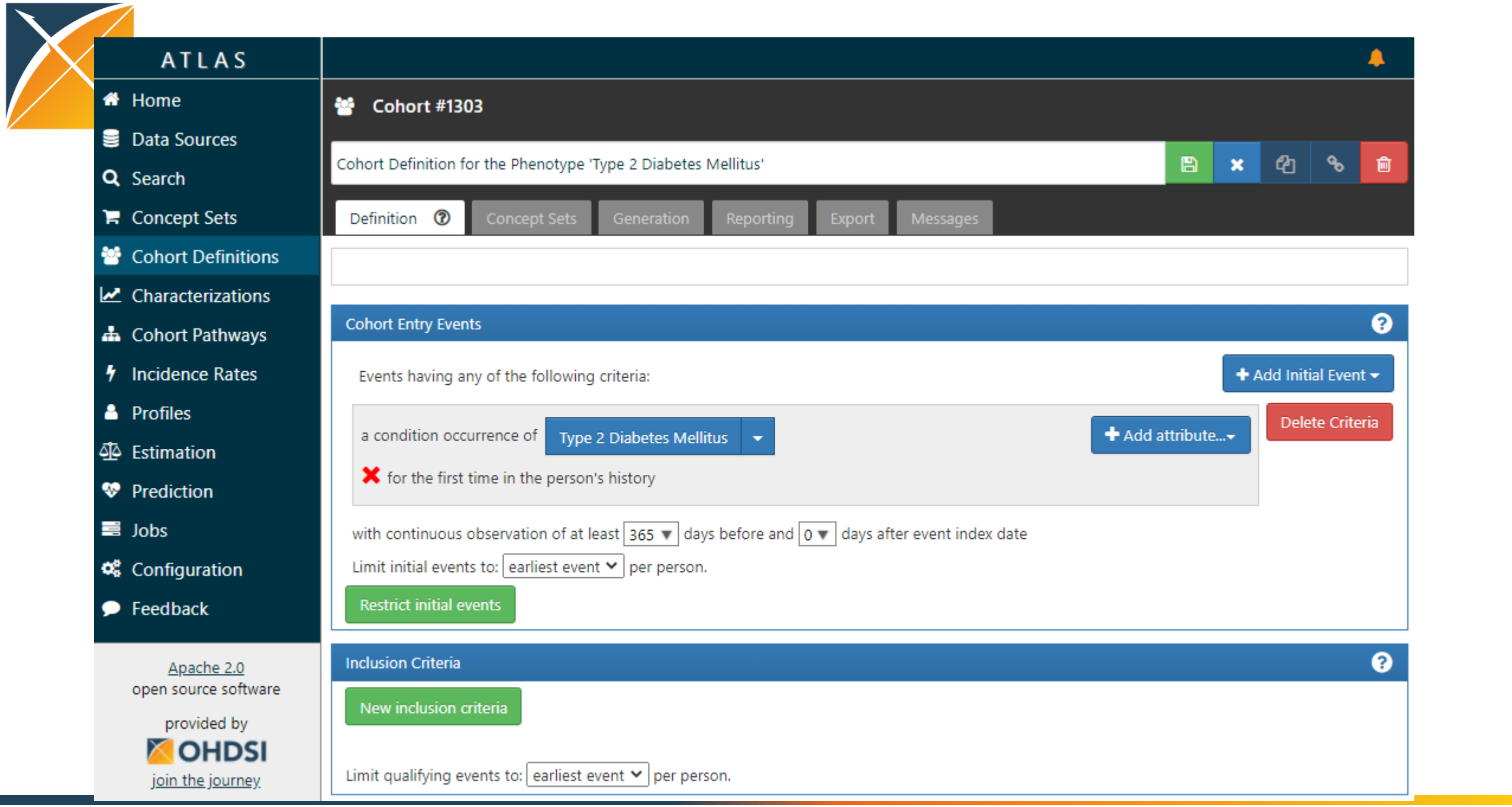
Overview: Type 2 DM accounts for 90% of all DM, and the generic use of Diabetes Mellitus almost always refers to Type 2. It is hyperglycemia and related complications usually due to progressive loss of insulin secretion from the pancreatic beta cells with background of insulin resistance. It is a state of 'relative' insulin deficiency - where insulin even when present may be less effective due to resistance.

Presentation: The classic symptoms of hyperglycemia (including polyuria, polydipsia, nocturia, blurred vision, and weight loss), but these are most observed in retrospect with most common mode of diagnosis is screening in asymptomatic patients or seeking emergency care for hyperosmolar hyperglycemic state without ketoacidosis.

Assessment: Oral glucose tolerance test and HbA1c, fasting blood glucose, lipid, liver function tests

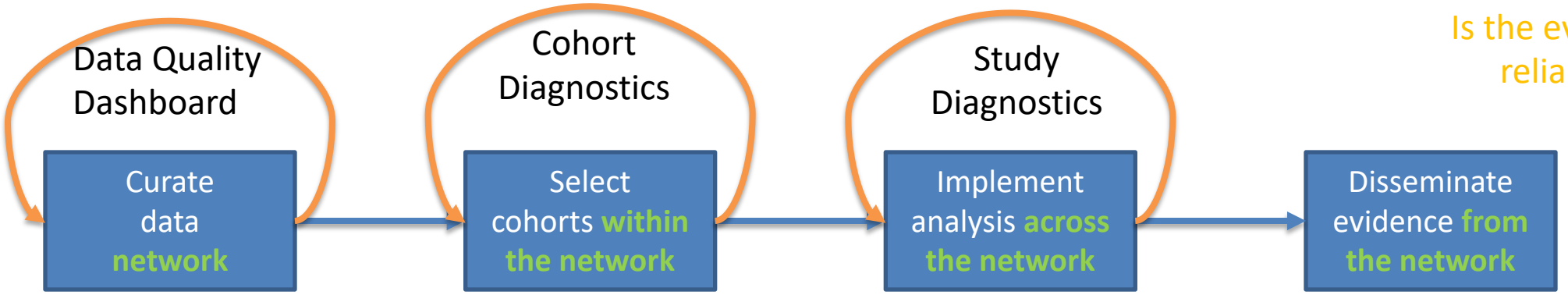
Plan: Initial management is to achieve normoglycemic state with lifestyle modification - weight, diet. Metformin is commonly recommended first line, but treatment choices vary with presence of other risk factors. Progression may lead to dependence on insulin.

Prognosis: Life-long disease that is amenable and if not well managed may lead complications.


ATLAS
Home
Data Sources
Search
Concept Sets
Cohort Definitions
Characterizations
Cohort Pathways
Incidence Rates
Profiles
Estimation
Prediction
Jobs
Configuration
Feedback
Apache 2.0
open source software
provided by
OHDSI
join the journey
Cohort #1303
Cohort Definition for the Phenotype 'Type 2 Diabetes Mellitus'
Definition
Concept Sets
Generation
Reporting
Export
Messages
Cohort Entry Events
Events having any of the following criteria:
+ Add Initial Event
a condition occurrence of
Type 2 Diabetes Mellitus
+ Add attribute...
Delete Criteria
for the first time in the person's history
with continuous observation of at least 365 days before and 0 days after event index date
Limit initial events to: earliest event per person.
Restrict initial events
Inclusion Criteria
New inclusion criteria
Limit qualifying events to: earliest event per person.

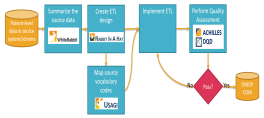
Today's session:
OHDSI innovations to further strengthen the reliability of the observational research workflow and evidence

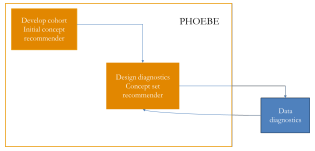
Empirical diagnostics:



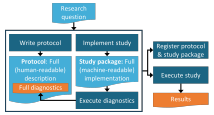
Is the evidence reliable?

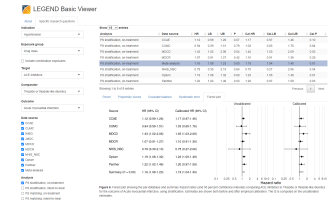
Standardized processes:



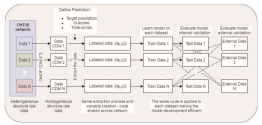


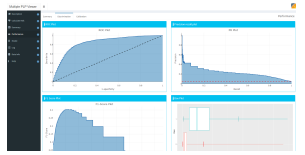
Estimation





Prediction





Characterization