OHDSI's Journey for 2017: Where we should go together

Patrick Ryan, PhD
Janssen Research and Development
Columbia University Medical Center
17 January 2017

What's our journey in 2017?

The Journey of Odysseus

OHDSI's areas of focus:
A look forward at 2017…

Methodological research → Open-source analytics development → Clinical applications

Observational data management

Clinical characterization

Population-level estimation

Patient-level prediction

- **Generate and disseminate more clinical evidence**
- Maintain and evolve open community data and vocabulary standards
- Develop and improve tools to enable large-scale analysis
- Establish and promote community best practices
- Strengthen and expand collaborations across OHDSI research network
- Advance scholarship in observational data science through publication, presentations, and education
- **Generate and disseminate more clinical evidence**

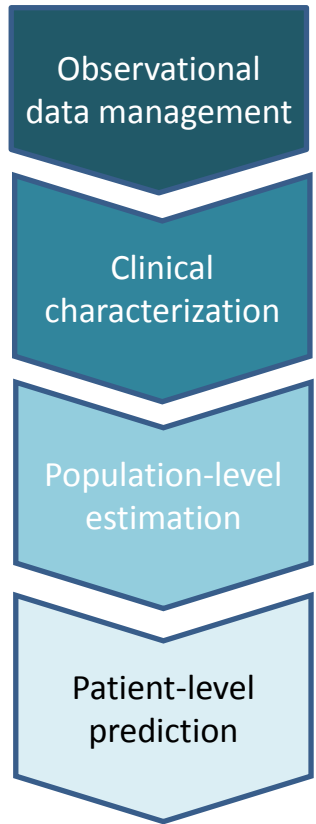# Building the LHC of observational research?

Some ambitious OHDSI community goals for 2017

- Large-scale clinical characterization
  - For all medical interventions, what is the population-level incidence of experiencing all possible outcomes of interest?

- Large-scale patient-level prediction
  - For all medical interventions, what is the personalized prediction of experiencing all possible outcomes of interest?

- Large-scale population-level effect estimation
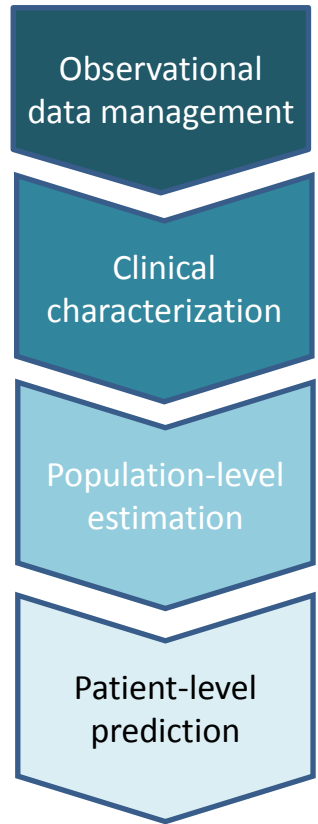- Large-scale data management

Open questions to achieve large-scale goals

Methodological research → Open-source analytics development → Clinical applications

Observational data management

Clinical characterization

Population-level estimation

Patient-level prediction

- How do we define cohorts for 'all medical interventions' and 'all outcomes of interest' that can be applied across a network of observational datasets?
- How do we evaluate and compare large-scale phenotyping approaches: naïve, heuristic, probabilistic?
- What model(s) and how much data are required to achieve the greatest predictive accuracy?
- How do we appropriately measure performance to ensure the evidence we disseminate is reliable?
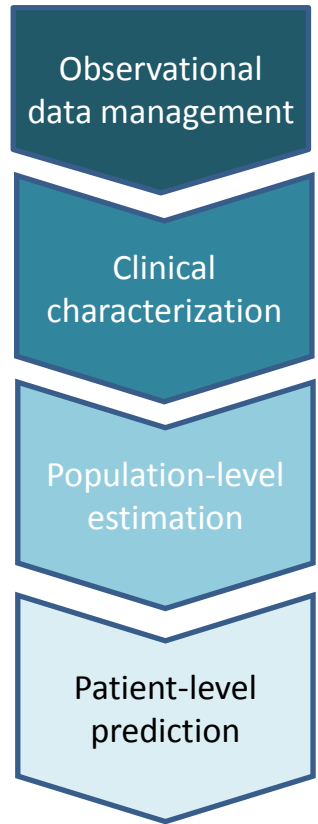
Open questions to achieve large-scale goals

| Methodological research | Open-source analytics development | Clinical applications |

Observational data management

Clinical characterization

Population-level estimation

Patient-level prediction

- How do we efficiently scale incidence rate calculations for all exposures and all outcomes, across various strata and time-at-risk?
- What approach is computational feasible for data partners to implement predictive model learning?
- How can we transport and apply learned models to other datasets?
- What type of central evidence repository do we need to build to ensure all parties are comfortable with sharing results?

Open questions to achieve large-scale goals

| Methodological research | Open-source analytics development | Clinical applications |

Observational data management

Clinical characterization

Population-level estimation

Patient-level prediction

- How do we disseminate the evidence that is generated?
- What API framework would be useful to allow evidence to be automatically integrated into existing clinical decision support systems?
- What type of user interface design would allow patients to ask the question: "What'll happen to me?" and effectively navigate to evidence that can meaningfully inform their decision making?

OHDSI F2F Agenda:
Day 1 – March 17

Time	Description
8:30 – 9:00am	Registration
9:00 – 9:30am	Welcome Session
9:30 – 10:30am	Working Group Breakout Session – Part I •Common data model and vocabulary •Population-level estimation •Hadoop
10:30 – 11:00am	Break
11:00 – 12:00pm	Working Group Breakout Session – Part II •Architecture •Patient-level prediction •LAERTES knowledge base •Natural language processing
12:00 – 12:30pm	Reconvene the community •Summary of key points from each work group
12:30 – 1:30pm	Lunch
1:30 – 3:00pm	Target problems to tackle •Patrick Ryan: Presentation of target problems •Community discussion to frame each problem
3:00 – End of day	Hack-a-thon – Three tracks •Phenotyping and cohort building •Large scale statistical programming •Design session: UI experience and information dissemination

OHDSI F2F Agenda: Day 2 – March 18

Time	Description
8:00 – 12:30pm	Continue hack-a-thon activities across three tracks: •Phenotyping and cohort building •Large scale statistical programming •Design session: UI experience and information dissemination
12:30 – 1:30pm	Lunch
1:30 – 3:00pm	Reconvene the community •Review outcomes from each hack-a-thon track
3:00 – 3:30pm	Break
3:30 – 5:30pm	Open community discussion •Next steps for following through on group projects •Other priorities for collaborative projects •Other ways to engage the community and make contributions
5:30pm	Wrap-up

OHDSI F2F Logistics

- Meeting: March 17-18 in Atlanta, GA (Georgia Institute of Technology)

- OHDSI F2F is not like OHDSI Symposium
  - No lectures → no passive participation
  - Everyone who comes is expected to roll up their sleeves and make contributions to our shared goals

- Meeting space limited, so you must register to participate

- Registration: http://www.ohdsi.org/events/2017-ohdsi-collaborator-face-to-face/

Join the journey

- Discussion / questions / comments

ryan@ohdsi.org