The Journey of OHDSI: Where can we go together?

Patrick Ryan, PhD
Janssen Research and Development
Columbia University Medical Center
“No one could reasonably maintain that there has, up to the present, been a shortage of printed material on the adverse reactions to drugs; the problem, and the justification of this Annual, is just the reverse...

Leo Meyler’s first ‘Side Effects of Drugs’, a valiant effort to tame the rising flood, was a readable little book of 128 pages; the latest edition of the encyclopaedic work which it has become runs more than 1100 pages of small print.

At this rate, one would by the end of the century need a treatise of ten volumes to supply the physician with even a summary view of what is known, or supposed to be known, about adverse drug reactions. “
“But the exponential growth of ‘Meyler’ is merely symptomatic of the underlying problem. The number of published papers on drugs and their wanted or unwanted effects has become staggering and virtually indigestible.

Truth, ever evasive, has now become embedded in a vast haystack of repetitions, assertions, denials, arguments, and irrelevancies. Are there any answers to such a problem?”
The journey to real-world evidence

Patient-level data in source system/schema

Reliable evidence
The journey to real-world evidence

Different types of observational data:
- **Populations**
  - Pediatric vs. elderly
  - Socioeconomic disparities
- **Care setting**
  - Inpatient vs. outpatient
  - Primary vs. secondary care
- **Data capture process**
  - Administrative claims
  - Electronic health records
  - Clinical registries
- **Health system**
  - Insured vs. uninsured
  - Country policies
The journey to real-world evidence

Types of evidence desired:
- **Cohort identification**
  - Clinical trial feasibility and recruitment
- **Clinical characterization**
  - Treatment utilization
  - Disease natural history
  - Quality improvement
- **Population-level effect estimation**
  - Safety surveillance
  - Comparative effectiveness
- **Patient-level prediction**
  - Precision medicine
  - Disease interception

Patient-level data in source system/schema

Reliable evidence
### Desired attributes for reliable evidence

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeatable</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Reproducible</td>
<td>Identical</td>
<td>Different</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Replicable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Similar</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Generalizable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Different</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Robust</td>
<td>Identical</td>
<td>Same or different</td>
<td>Same or different</td>
<td>Different</td>
<td>Similar</td>
</tr>
<tr>
<td>Calibrated</td>
<td>Similar (controls)</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Statistically consistent</td>
</tr>
</tbody>
</table>
Minimum requirements to achieve reproducibility

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproducible</td>
<td>Identical</td>
<td>Different</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
</tbody>
</table>

- Complete documented specification that fully describes all data manipulations and statistical procedures
- Original source data, no staged intermediaries
- Full analysis code that executes end-to-end (from source to results) without manual intervention

Patient-level data in source system/schema

Reliable evidence
How a common data model + common analytics can support reproducibility

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproducible</td>
<td>Identical</td>
<td>Different</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
</tbody>
</table>

- Use of common data model splits the journey into two segments: 1) data standardization, 2) analysis execution
- ETL specification and source code can be developed and evaluated separately from analysis design
- CDM creates opportunity for re-use of data step and analysis step

Patient-level data in source system/schema

Patient-level data in CDM

Reliable evidence
Challenges to achieve replication

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Similar</td>
<td>Identical</td>
<td>Similar</td>
</tr>
</tbody>
</table>

• If analysis procedure is not identical across sources, how do you determine if any differences observed are due to data vs. analysis?
How a common data model + common analytics can support replication

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Similar</td>
<td>Identical</td>
<td>Similar</td>
</tr>
</tbody>
</table>

Source 1
...
Source i
...
Source n

Source 1 CDM

Source i CDM

Source n CDM

Similar evidence
Reliable evidence
Similar evidence
How a common data model + common analytics can support robustness

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robust</td>
<td>Identical</td>
<td>Same or different</td>
<td>Same or different</td>
<td>Different</td>
<td>Similar</td>
</tr>
</tbody>
</table>

- Sensitivity analyses can be systematically conducted with parameterized analysis procedures using a common input.
How a common data model + common analytics can support calibration

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibrated</td>
<td>Similar (controls)</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Statistically consistent</td>
</tr>
</tbody>
</table>

- With a defined reproducible process, you can measure a system’s performance and learn how to properly interpret the system’s outputs.

Source data

Patient-level data in CDM

One-time

Repeated
Why reliable evidence requires a community effort

<table>
<thead>
<tr>
<th>Desired attribute</th>
<th>Question</th>
<th>Researcher</th>
<th>Data</th>
<th>Analysis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeatable</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Replicable</td>
<td>Identical</td>
<td>Different</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
</tr>
<tr>
<td>Generalizable</td>
<td>Identical</td>
<td>Same or different</td>
<td>Similar</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Robust</td>
<td>Identical</td>
<td>Same or different</td>
<td>Different</td>
<td>Identical</td>
<td>Similar</td>
</tr>
<tr>
<td>Calibrated</td>
<td>Similar (controls)</td>
<td>Identical</td>
<td>Identical</td>
<td>Identical</td>
<td>Similar</td>
</tr>
</tbody>
</table>

- Community of researchers with important public health questions
- Data network using community standards
- Analyses sharing community open-source tools
- Application of community best practices for evaluation
- Evidence sharing across community
Complementary evidence to inform the patient journey

Clinical characterization: What happened to them?

Patient-level prediction: What will happen to me?

Population-level effect estimation: What are the causal effects?

Observation

Inference

Causal inference
How should patients with major depressive disorder be treated?

Treating Major Depressive Disorder
A Quick Reference Guide

Pharmacotherapy

- The effectiveness of antidepressant medications is generally comparable between and within classes of medications, including selective serotonin reuptake inhibitors (SSRIs), serotonin norepinephrine reuptake inhibitors (SNRIs), bupropion, tricyclic antidepressants (TCAs), and monoamine oxidase inhibitors (MAOIs). Therefore, choose a medication largely based on the following:

  - Patient preference
  - Nature of prior response to medication
  - Safety, tolerability, and anticipated side effects
  - Co-occurring psychiatric or general medical conditions
  - Pharmacological properties of the medication (e.g., half-life, actions on cytochrome P450 enzymes, other drug interactions; consult the full guideline or a current drug database)
  - Cost

- For most patients, a SSRI, a SNRI, mirtazapine, or bupropion is optimal.

- In general, the use of MAOIs should be restricted to patients who do not respond to other treatments.

Based on Practice Guideline for the Treatment of Patients With Major Depressive Disorder, Third Edition, originally published in October 2010. A guideline watch, summarizing significant developments in the scientific literature since publication of this guideline, may be available.
How are patients with major depressive disorder ACTUALLY treated?

Hripcsak et al, PNAS, 2016
How are patients with major depressive disorder ACTUALLY treated?

• Substantial variation in treatment practice across data sources, health systems, geographies, and over time

• Consistent heterogeneity in treatment choice as no source showed one preferred first-line treatment

• 11% of depressed patients followed a treatment pathway that was shared with no one else in any of the databases

Hripcsak et al, PNAS, 2016
What questions does this answer?
What question does it prompt to ask?

Which treatment did patients choose after diagnosis?

Which patients chose which treatments?

How many patients experienced the outcome after treatment?

What is the probability I will experience the outcome?

Does one treatment cause the outcome more than an alternative?

Does treatment cause outcome?

Hripcsak et al, PNAS, 2016
What do the treatment pathways look like across Europe?
What do the treatment pathways look like across the world?
Questions we can answer with reliable evidence

- **Clinical characterization**: What happened to them?
  - What treatment did they choose after diagnosis?
  - Which patients chose which treatments?
  - How many patients experienced the outcome after treatment?

- **Patient-level prediction**: What will happen to me?
  - What is the probability that I will develop the disease?
  - What is the probability that I will experience the outcome?

- **Population-level effect estimation**: What are the causal effects?
  - Does treatment cause outcome?
  - Does one treatment cause the outcome more than an alternative?
Journey toward reliable evidence

Evidence Generation
• How to produce evidence from the data?

Evidence Evaluation
• How do we know the evidence is reliable?

Evidence Dissemination
• How do we share evidence to inform decision making?
Large-scale Evidence Generation and Evaluation in a Network of Databases

17,718 estimates
87.0% of CIs include 1

Literature
What do the population-level effect estimates look like across Europe?
What do the population-level effect estimates look like across the world?
Building the LHC of observational research?
OHDSI community

We’re all in this journey together...
Coffee Break