Leaf & Atlas Tutorial: Self-Service Query Tools for MSDW2

Fabrício Kury, MD
Clinical Data Strategist
Scientific Computing
Icahn School of Medicine at Mount Sinai

November 3rd, 2021
Objectives

Learn:

1. What is MDSW2, what are clinical query tools.

2. What do Leaf and Atlas offer; how to request access.
   i. What is the OMOP CDM and the OHDSI community.

3. How to build an example query in each tool.

4. When and how to request a custom dataset.
Objectives

Learn:

1. **What is MDSW2, what are clinical query tools.**

2. **What do Leaf and Atlas offer; how to request access.**
   i. What is the OMOP CDM and the OHDSI community.

3. **How to build an example query in each tool.**

4. **When and how to request a custom dataset.**
Mount Sinai Data Warehouse (2)

https://labs.icahn.mssm.edu/msdw/
Inclusion Criteria:
• Type 2 DM for over six months
• Medical history of glaucoma
• A1c > 7.5% but < 10 despite treatment with 2 drugs
• 18 years of age or older

Exclusion Criteria:
• Pregnancy as determined by a serum β HCG

Examples of eligibility criteria taken from:
Efficacy Evaluation of Different Medication Combination in Type 2 Diabetes Treatment - Tabular View - ClinicalTrials.gov
Objectives

Learn:

1. What is MDSW2, what are clinical query tools.

2. What do Leaf and Atlas offer; how to request access.
   i. What is the OMOP CDM and the OHDSI community.

3. How to build an example query in each tool.

4. When and how to request a custom dataset.
Clinical Query Tools (CQTs)

<table>
<thead>
<tr>
<th></th>
<th>Leaf</th>
<th>Atlas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development</td>
<td>Nic Dobbins, Univ. of Washington, plus collabs., including at MSSM</td>
<td>OHDSI community: <a href="http://www.ohdsi.org">www.ohdsi.org</a></td>
</tr>
<tr>
<td>License</td>
<td>Free and Open-Source Software (FOSS)</td>
<td></td>
</tr>
<tr>
<td>Tradeoff</td>
<td>Easier, quicker, less powerful</td>
<td>Harder, laborious, more powerful</td>
</tr>
<tr>
<td>Data available</td>
<td>De-identified only</td>
<td>De-identified or PHI (with IRB)</td>
</tr>
</tbody>
</table>
| Capabilities | • Simple Boolean logic  
• Predefined stats and viz.  
• Can download lists of patients (privacy-preserving IDs) | • Sophisticated logic  
• Customized stats and viz.  
• Save your work and reuse parts  
• Run entire statistical analyses  
• **No data downloads.** |
What is PHI? What is De-identification?

“PHI (Protected Health Information) is information (demographic, financial, social, clinical) relating to an individual’s past, present, or future health history, treatment, or payment for health care services that is held or transmitted by a CE or its BA that identifies the individual or for which there is a reason to believe it can be used to identify the individual.”

De-identification is the process by which PHI is rendered not individually identifiable. The HIPAA Privacy Rule establishes two methods to de-identify PHI:
https://labs.icahn.mssm.edu/msdw/ -> “Request Access” (under Leaf)

Scientific Computing

Cohort Selection Tools

Welcome! You can raise a Cohort Selection Tools request from the options provided.

What do you need help with?

[Search]

Query Tool Access Request Form
Unified request form for data access

Report a bug
Tell us the problems you're experiencing.

Other questions
Don't see what you're looking for? Select this option and we'll help you out.
Objectives

Learn:

1. What is MDSW2, what are clinical query tools.

2. What do Leaf and Atlas offer; how to request access.
   i. What is the OMOP CDM and the OHDSI community.

3. How to build an example query in each tool.

4. When and how to request a custom dataset.
live Leaf demo

https://leaf.mssm.edu/
Leaf Application Status and Roadmap

Last Updated: October 27, 2021

Status

Leaf is currently in beta release at Mount Sinai and is available for user testing. Features and functionality are occasionally updated. Leaf currently supports only the following query domains:

- Conditions (diagnoses) using ICD-10-CM
- Procedures using CPT
- Demographics (such as age, ethnicity, gender, race, and vital status)
- Vitals
- Visit Location

Leaf beta contains the following known bugs:

- Multiple important clinical domains are unsupported
- Queries that run longer than 1 minute are terminated
- Age is often missing from the Patient List view
Map of OHDSI collaborators as of August, 2019
May I introduce you to the **OMOP CDM**?
May I introduce you to the *OMOP CDM*?
Characterizing treatment pathways at scale using the OHDSI network

George Hripcsak\textsuperscript{a,b,c,d}, Patrick B. Ryan\textsuperscript{e,f}, Jon D. Duke\textsuperscript{g,h}, Nigam H. Shah\textsuperscript{a,i}, Rae Woong Park\textsuperscript{j,k}, Vojtech Huser\textsuperscript{a,c,d}, Marc A. Suchard\textsuperscript{l,m,n}, Martijn J. Schuemie\textsuperscript{d}, Frank J. DeFalco\textsuperscript{d}, Adler Perotte\textsuperscript{e,f}, Juan M. Banda\textsuperscript{d}, Christian G. Reich\textsuperscript{d}, Lisa M. Schilling\textsuperscript{a,c,d}, Michael E. Matheny\textsuperscript{a,c,d}, Daniella Meeker\textsuperscript{a,c,d}, Nicole Pratt\textsuperscript{a,c,d}, and David Madigan\textsuperscript{a,c,d}

\textsuperscript{a}Department of Biomedical Informatics, Columbia University Medical Center, New York, NY 10032; \textsuperscript{b}Medical Informatics Services, New-York Presbyterian Hospital, New York, NY 10032; \textsuperscript{c}Observational Health Data Sciences and Informatics, New York, NY 10032; \textsuperscript{d}Epidemiology Analytics, Janssen Research and Development, Titusville, NJ 08532; \textsuperscript{e}Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN 46202; \textsuperscript{f}Center for Biomedical Informatics Research, Stanford University, CA 94305; \textsuperscript{g}Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea, 463-898; \textsuperscript{h}Lester Hill National Center for Biomedical Communications (National Library of Medicine), National Institutes of Health, Bethesda, MD 20894; \textsuperscript{i}Department of Biostatistics, University of California, Los Angeles, CA 90095; \textsuperscript{j}Department of Biomedical Informatics, University of California, Los Angeles, CA 90095; \textsuperscript{k}Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90089; \textsuperscript{l}Department of Pediatrics, University of Southern California, Los Angeles, CA 90089; \textsuperscript{m}Division of Health Sciences, University of South Australia, Adelaide, SA, Australia 5001; and \textsuperscript{n}Department of Statistics, Columbia University, New York, NY 10027

Edited by Richard M. Shiffman, Indiana University, Bloomington, IN, and approved April 5, 2016 (received for review June 14, 2015)

Observational research promises to complement experimental research by providing large, diverse populations that would be infeasible for an experiment. Observational research can test its own clinical hypotheses, and observational studies also can contribute to the design of experiments and inform the generalizability of experimental research. Understanding the diversity of populations and the variance in care is one component. In this study, the Observational Health Data Sciences and Informatics (OHDSI) collaboration created an international data network with 11 data sources from four countries, including electronic health records and administrative claims data on 250 million patients. All data were mapped to common data standards, patient privacy was maintained by using a distributed model, and results were aggregated centrally. Treatment pathways were elucidated for and depression. The pathways toward more consistent treatment locations, but significant benefits were observed in the favored a single treatment, even if it was not indicated. The extent of hypertension or depression patients and all followed a treatment pathway. Aside from factors such as academic medical centers, health records data and administrative results. Large-scale international observational research | data mine

Health Data Sciences and Informatics is an international data network with data on 250 million patients. All data are mapped to common data standards, patient privacy was maintained by using a distributed model, and results were aggregated centrally. Treatment pathways were elucidated for and depression. The pathways toward more consistent treatment locations, but significant benefits were observed in the favored a single treatment, even if it was not indicated. The extent of hypertension or depression patients and all followed a treatment pathway. Aside from factors such as academic medical centers, health records data and administrative results. Large-scale international observational research | data mine

A learning health system is a feedback loop of medical decision-making. It is a system to test the effects of medical interventions. A learning health system is a feedback loop of medical decision-making. It is a system to test the effects of medical interventions. A learning health system is a feedback loop of medical decision-making. It is a system to test the effects of medical interventions.

Without sufficiently broad databases available in the first stage, randomized trials are designed without explicit knowledge of the disease status and treatment practice. Literature reviews are restricted to the population choices of previous investigations, and pilot studies usually are limited in scope. By exploiting the ClinicalTrials.gov national trial registry (9) and electronic health records, researchers have demonstrated the discrepancy between targeted populations and populations available for study (10), raising the concern that designs may not be optimal. Designs cannot be based simply on current treatment recommendations: Local stakeholders (patient, family, physician, and consultant) and global stakeholders (industry, regulators, academics, and the media) interact in complex ways across media.
Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)

…brings the data to the code, instead of the reverse.
OHDSI Network Studies

Network Study Workflow
https://atlas.msdw.mountsinai.org/
live Atlas demo

https://atlas.msdw.mountsinai.org/
ATLAS Application Status and Roadmap

Last Updated: October 27, 2021

Status

ATLAS application is currently in beta release at Mount Sinai and is available for user testing. Features and functionality are occasionally updated.
ATLAS functionality is currently limited to include:

- De-identified data source only
- OMOP standard concept IDs are contained within only the following clinical domains: conditions, procedures, visits, vital measurements
- All other domains can be queried using concepts in Epic vocabularies
- Only the following left-hand sidebar functions are offered: Data Sources, Search, Concept Sets, Cohort Definitions

ATLAS beta contains the following limitations:

- Data containing PHI, including datamarts, are not yet available for query
- Most ATLAS data uses OMOP standard concept IDs, but some live data uses non-standard vocabularies

https://labs.icahn.mssm.edu/msdw/atlas-application-status/
Objectives

Learn:

1. What is MDSW2, what are clinical query tools.

2. What do Leaf and Atlas offer; how to request access.
   i. What is the OMOP CDM and the OHDSI community.

3. How to build an example query in each tool.

4. When and how to request a custom dataset.
Need something unavailable through the self-service tools?

Some examples:
- Chained contingencies
  - treated with drug X after discharge from readmission within 30 days of hospitalization due to cirrhotic liver $(C \rightarrow B \rightarrow A)$
- Mathematical formulas
  - ASCVD Risk score above 10%
  - Child-Pugh score between 7-9
- Arbitrarily layered Boolean logic
  - $(A \land (B \lor C) \land (D \lor \neg(E \lor F)) \lor G)…$
- Judgment calls
  - In good health except for T2DM…
- …can I just have the raw data, please? Thank you.

Augmenting datasets with data from Clarity or Caboodle is possible.
custom dataset request form demo

https://scicomp.mssm.edu/jira/servicedesk/customer/portal/4
Thank you!

More help, tutorials & videos: https://labs.icahn.mssm.edu/msdw/

Fabrício Kury, MD
fabricio.kury@mssm.edu
Clinical Data Strategist
Scientific Computing
Icahn School of Medicine at Mount Sinai
November 4th, 2021