Phenotype Development and Evaluation Workgroup

OHDSI Community Call Update
October 5th 2021
Who are we? What are we? Why are we?

“Cohort Definitions”

We meet twice a month
- Second Friday
- Fourth Friday
@1pm EST
Mission: We improve the quality and reliability of the evidence generated from observational data, by advancing the science of phenotype development and evaluation.

Areas of focus:

1. Content development and dissemination
2. Best practice development and evaluation
3. Science of phenotyping
4. Tool design and innovation
Content development and dissemination

“We build better cohort definitions and make it available for everyone.”

- Develop phenotypes
  - Define the clinical idea
  - Rule based and probabilistic cohort definitions

- Evaluate performance characteristics
  - Generalizability
  - Impact on measurement bias
  - Misclassification error of persons, index date etc.
  - Agreement vs disagreement between alternatives

- Make available
  - Repository of definition
  - Catalogued, easily retrievable
  - Used/reused in OHDSI studies
Best practice development and evaluation

“We enable a community to build better cohort definitions”

To develop, document and use best practices for improving the reproducibility and quality of cohort definitions – and disseminate these best practices through teaching, education and training.

<table>
<thead>
<tr>
<th>Develop best practice</th>
<th>Document</th>
<th>Disseminate</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Process mapping</td>
<td>- Document the process</td>
<td>- Teaching</td>
</tr>
<tr>
<td>- What went right</td>
<td>- Use the process</td>
<td>- Training - workshop</td>
</tr>
<tr>
<td>- What went wrong</td>
<td>- Reproducible and efficient</td>
<td>- Education</td>
</tr>
<tr>
<td>- Gotchya, mistakes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Science of phenotyping

“We evaluate if we are building better cohort definitions and innovate on methods to improve this science.”

Conduct methods research on performance of cohort definitions, develop measurement systems that provide absolute and comparative performance of a cohort definition.

- Sensitivity
- Specificity
- Positive predictive value
- Measurement bias
- Misclassification bias

Computational feasibility – complex logic on very large dataset may not be possible
Portability of definitions that apply to rich datasets vs not so granular datasets (including missing data).
Applicability and acceptance of variations (and resultant errors) based on the nature of source data (open vs closed, EHR vs claim)
Tool design and innovation

“We facilitate systematic use of best practices by adapting, innovating and promoting the use of OHDSI standard tools.”

Enable the community to adopt best practices facilitated by OHDSI developed tools such as CirceR, CAPR, Aphrodite, Phevaluator, Cohort Diagnostics, Phenotype Library.

Not tool developers → HADES

We are designers to help adopt best practices
How is work done

• Co-ordinate using two formal monthly meetings
  – We report out to other OHDSI workgroups that are leading studies

• We plan synchronously
  – We use theme of the month approach
  – Get it done - ‘Get it to completion’

• We execute asynchronously
  – Asynchronous communication
  – Asynchronous projects or efforts (sub-group/leads)
Theme of the Month – get it done

How confident are we that persons with a diagnosis code for Covid-19 actually have Covid disease?

Attendance
Gowtham Rao
Williams, Andrew E
Suleiman, Lina
Roggenkamp, Tracie
Weaver, James [JANUS]
Evan Patrick Mitty
Makadia, Rupa [JRDFS]
Shahabi, Azza [JRDFS]
Kenny, Ryan
Bruce Bray
Chip Shaw
Patrick Ryan
Vojtech Huzar
Wilkins, Kenneth [NIH/NIDDK] [E]
Natarajan, Karthik
OHDSI Phenotype development and evaluation workgroup - Fourth Friday meeting - 1pm to 2pm EST


Researchers performing analysis on observational data, assume that coders and health care providers have adhered and applied the guidelines while identifying the diagnosis being reported. Per guidelines, a confirmed COVID-19 diagnosis should be assigned an ICD-10-CM code of U07.1, with a few exceptions for uncertain diagnosis at the time of discharge (section II, H). Disease confirmation maybe with a positive test result for COVID-19 or a provider’s documentation that the individual has COVID-19. In the absence of positive test, provider documentation of uncertain diagnosis (e.g. suspected, possible etc.) are not considered appropriate for use of U07.1.

What is unknown is the adherence to such guidelines i.e. the potential for and quantification of misclassification bias? We also do not know if over the course of the pandemic, has the classification/misclassification changed over time. What is the correlation between positive test result
Universal deliverables (upcoming)

- OHDSI Phenotype Library
Universal deliverables (upcoming)

- Phenotype library study
Phenotype Development and Evaluation Workgroup

Come join us

If you workgroup/study needs cohort definitions – come talk to us