OHDSI Phenotype Phebruary: lessons learned

2022 OHDSI Symposium
14 October 2022 Bethesda MD

Azza Shoaibi, Joel Swerdel, Allan Wu, Gowtham Rao, Adam Black, Evan Minty, Asieh Golozar, Rupa Makadia, Jill Hardin, Erica Voss, Tiffany J. Callahan, Juan Banda, Anna Ostropolets, Claudia Pulgarin, Marcela Rivera, David Vizcaya, Patrick Ryan
Phenotype development and evaluation is yet to become a completed chapter in the book of OHDSI

- Phenotypes are the foundational elements in almost every real-world analysis.
- The reliability of the generated evidence depends on the validity of the phenotypes.
- Yet, the science of phenotype development and evaluation is relatively immature.
- We have built best practices and end to end process, tools and packages for characterization, estimation and prediction.
- But for phenotyping- “addressed in the limitation section.”
"Phenotype Phedbruary": I realized that becoming a master of karate was not about learning 4,000 moves but about doing just a handful of moves 4,000 times.—Chet Holmes

- We collectively started a discussion on 28 phenotypes over 28 days
- Followed 5 step process:
  1. Join the conversation
     - Discussions will be held on forums.ohdsi.org
     - Each day will be a new thread
     - Explore the definitions and review the results provided
     - Reply with your thoughts, reflections, insights and question
  2. Evaluate the cohort definitions in your data
     - Execute cohort definitions and CohortDiagnostics in your CDM
     - Share Insights you learn from your data on the forums
     - Share results to compile across the network on data.ohdsi.org
15 phenotypes were developed, evaluated and discussed and we learned few things

**Thematic learnings:** clinical description, phenotype development and phenotype evaluation. The themes identified belonged to 5 different types of lessons: tips, strategies, debatable topics, challenges, and opportunities.
Lessons learned: Phenotype development

Tips/strategies

Model the clinical idea not the analytical use case

Code selection is a clinical choice. The material consequences should be empirically investigated

Notions like "primary position" need to be standardized and evaluated

Differentiate between situations where a data is not "fit-for use" and situations where logic is not "fit" for the data

Opportunities

Systematically assess multiple look back periods and recommend one.

Combine Probabilistic -based approach (APHRODITE) with a rule-based approach (Atlas)

Develop a PubMed search strategy to find published/evaluated phenotypes

Do we customize phenotypes to specific data sources/analytical use?
Lessons learned: Phenotype evaluation

**Tips/strategies**

- Evaluate all types of measurement error
- Use patient profile to get a sense of validity. Identify disqualifying patterns.
- Explore in CD: temporal stability, expected trends, patients composition, index event misclassification
- Estimate measurement error & quantify trade-offs by PheValuator or APHRODITE

**Challenges**

- Subjective
- Time-consuming and complex
- Lack an approach to evaluate exit criteria & washout periods for reoccurring events

What objective diagnostic criteria can we apply to determine fitness-for-use?
The choices are NOT:
1: Code list
2: Code list with chart reviews

Phenotyping is complex, multidimensional and requires exchange of knowledge, learnings and insights across collaborators from different background and expertise.

Large scale characterization (e.g. CD), Diagnostic predictive models (e.g., PheValuator) and structured review of patient’s profile are potentially effective and novel strategies for phenotype evaluation.

We are getting closer to a standardized process. But further collaboration is needed to formalize a scalable and reproducible processes and establish empirically-driven objective diagnostics.