# Minutes of OHDSI workgroup in Population Level Estimation

July 2, 2019

Present: Jon Duke, Clair Blacketer, David Madigan, Anthony Sena, Martijn Schuemie, Greg Klebanov

# \*\*\* Progress report \*\*\*

Chapters:

I OHDSI Community

1 Mission, vision, values: George has taken this on. Will not be ready for July 9

2 Collaborators: Patrick is owning this

3 Open Science: Kees is working on this

4 Where to begin: Kristin has this

II Uniform Data Representation

5 The Common Data Model: Draft complete

6 Standardized Vocabularies: Mui and Christian are working on this.

7 Extract Transform Load: Clair was working on EHDEN course that can be used to put something together. Will focus on WR and RiaH first.

III Data Analytics

8 Data Analytics Use Cases: Draft is finished

9 OHDSI Analytics Tools: Frank will contribute part on ATLAS

10 SQL and R: Draft is complete

11 Building the building blocks: cohorts: Kristin is leading this.

12 Characterization: Anthony Sena is working on it. Will send draft to Dani, David and Martijn

13 Population-level estimation: Draft is finished.

14 Patient Level Prediction: Jenna has done massive update. Martijn has provided some feedback that Jenna is processing.

IV Evidence Quality

15 Evidence Quality: Jon has created initial draft. Martijn has provided feedback, Jon is rewriting.

16 Data Quality: Unclear. Martijn will contact Andrew and Vojtech

17 Clinical Validity: Chan and Joel are leading this.

18 Software Validity: Martijn is leading.

19 Method Validity: Martijn is working on this, will be done on time.

V OHDSI Studies

20 Study steps: Sara Dempster is working on this.

21 OHDSI Network Research: Kristin and Greg are working on this.

Martijn to ping Chan and Joel on clinical validity

# \*\*\* Summaries \*\*\*

Clair: Key take-aways instead of “summary”?

# \*\*\* BroadSea? \*\*\*

Clair: I’m using OHDSI-in-a-box for Moodle.

Anthony: Integrated solutions should be “Deployment options”

OHDSI-in-a-box is used for tutorials. Ping James if used for actual analytics. <https://aws.amazon.com/blogs/big-data/creating-data-science-environments-on-aws-for-health-analysis-using-ohdsi/>

BroadSea: used at George tech for actual analytics. Docker instance

OHDSI in a box is more complete. BroadSea is ‘cloud neutral’

# \*\*\* Data quality chapter \*\*\*

Anthony: database level characteristics go in Characterization chapter.

Themis: part of CDM or ETL, not Data Quality. Clair will integrate

Vojtech already has very elaborate paper on Achilles Heel: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5226382/>

Martijn: will write bit on unit tests for Data Quality chapter

Martijn and David will coordinate with Vojtech and Andrew.

Will revisit Data Quality next week.

Martijn mentions Frank has created a tiny simulated dataset for Eunomia that can be used to create exercises. Martijn will try to have some example exercises to discuss next meeting.

Next meeting: July 9.