**OHDSI Clinical Trials Working Group** March 1, 2019

What is the origin of the phrase “Rabbit, rabbit”.

*All members can edit the text.*

Roll Call

* Sonia Araujo (IQVIA) – UK
* Chris Roeder (Univ of CO) - Colorado
* Maxim Moinat (the Hyve) – The Netherlands
* Kees van Bochove (the Hyve) – The Netherlands
* Greg Klebanov (Odysseus)
* Rupa Makadia (Janssen)
* Andrew Williams (Tufts)
* Josh Ransom (Medidata/SHYFT)
* Shawn D – (Gates Foundation)
* Someone named Ayesha Kang seemed to dial in but was not speaking

**New members**: first time at this meeting: **Josh Ransom**, Medidata/SHYFT Analytics, uses OHDSI (was author on a poster), improving clinical trial improvement; Chris Roeder (past and anonymized clinical trials in OHDSI), may not be using OHDSI optimally; **Maxim and Kees**, The Hyve (Utrecht); have seen interest, involved in some things already (Pioneer Project), how do you represent clinical trials in a good way in OMOP; **Sonia Araujo**, in the UK, (Pioneer Project), ex-Veeva, deep experience in clinical trials, sees possibilities in OHDSI and RWE re clin trials.

From email Shawn sent earlier about the diversity in interest shown in meeting 1 from members, how to organize that?

1. There is one **cohort** of people who are interested in use cases, understanding all of them, ranking them, understanding them by user org or user type, benefit.
2. There is one **cohort** of people interested in sort of data-centric approach: what data exists out there, what is missing, how can RWD data be used in trials; novel uses of RWD in this area, statistics, *methods*, creating de novo or synthetic data
3. There is one **cohort** of people who are interested in the OHDSI tools and CDM—getting data in, what else is needed, needed features

Open dialogue about what group could do or what topic groups are there

AW: the list is bigger perhaps than this.

JR: wrote a white paper on use cases (area 1). **Trial design** is one batch of use cases; can you optimize a design; can you find a site; patient identification; this uses RW data sets up front. Also, there is the use of the CDM for **storing the data** has been collected, either registry or trial itself, and using OHDSI tools for data mgmt. of trial or reg. This is already being done for registries, and maybe pilots for trials. All of US trial will use CDM. Then using **OHDSI tools** for **analyzing**, once it is stored there, could be aggregated or **data set of data sets** and then analyze across trials. OHDSI Network, could do an OHDSI Network of de-identified trial data sets and then federate a study. (JR will send white paper).

SA: maybe 2 ways to think about this: clinical and observational. Can interact. How can observational augment trials.

GK: cross-study analysis is a thing. Pooling clinical and observational.

Kees: One challenge in cross-study analysis is how to represent interventional data (which is different! From the DNA of CDM).

SD: Should one put interventional into CDM directly?

KvB A big question, and no clear answer within the community exists. Patrick Ryan has thought about this. One view is a data model has a scope, and thus some things should be out of scope, and maybe transaction-esque data should be out of scope. Clarifying, intervention to KvB doing the intervention introduces bias (1 vs another), the nature and purpose of a trial, and that doesn’t mix well vs just observing flat data in the world. In want to do inter justice, you have to change CDM from right now.

AW: clinical trial mgmt. system is transactional, excellent solutions already used, and fit for that purpose. Seems not a fruitful direction. Use case of accumulating trial data is quite different from using it as ODS (so to speak). Might be more than a table or two, even radical change, might be needed to leverage the advantages of OMOP for trials. As is, will or may miss a lot. Vocabularies may be the most formidable challenge in that regard. It may be a valuable piece of work, even if daunting. Within NIH adherence to FAIR may provide additional push to this, since a priority.

RM: on the 3 areas, they are broad, but what is coming out in discussion main be main sub points. At \_\_\_ firm, they do pool trial data and then use it for feasibility, to estimate how many patients could be recruited. Can we think about making a reality to use observational data to identify standard of care (baseline or control) in trial. Rupa’s use cases are more about using existing RWE data (not the trial generated data). Can RWE be used to obviate a clinical trial or something in that vector.

CR: focus on using OMOP as a point to harmonize studies for downstream. DWH. Once trials are completed. FAIR is interesting to U Colo. Could E.H.R. data that goes into CDM be used as an in silico clinical trial rather than a new trial.

KvB: maybe FHIR is better with E.H.R. and then do in code your trial-esque analytics, don’t need to put it into CDM.

AW: Epic and Cerner will be tackling how to reduce the burden of executing a trial on-site at a clinic that has their sw. Some areas already have feature/function in those softwares. Might be walled off, data-wise.

SD: let me look into that.

KvB: analytics may not be the right word for all, there is a ‘doing trials better’. Analytics vs trial optimization.

KvB One question we could answer – why do trials in OMOP?

FAIR principle 1.3 is of interest (SD: what is this?)

A number of reasons—why are we talking about this—we should ask this

Potential output for the group (three areas below differ in **consumability**)

* Convene expert presentations – recorded, highly consumable ex post facto to an external audience.
* Produce group deliverables? Made to be consumable by the outside (there is some interest in this) (some deliverable may be to OHDSI for new vocabularies for example; others may be for the world, not tech in nature) (another potential deliverable is, per SA, what are we doing already vs what could we do better)
* Answer questions just for this group (not specifically a deliverable) / support—(tailored to be consumed internally as dialogue occurs), someone would need to listen to a whole recording to ferret out the answer because it is not ‘produced’ (could we help specifically on the *cusp* of what is next for an existing situation)

Presentations This Year Committed

* Odysseus (Greg) (maybe with Josh)
* The Hyve (Kees and Maxim)
* What Janssen Does (not confidential info) Today (Rupa)
* Andrew open to presenting once the group settles into topics of interest

We will try to have some meetings specific to a single topic. Please submit your topics that could be a whole meeting in themselves. For example:

**Example: CDM as ODS as a single meeting topic, where is the gray area**. Intervention data in there is different than using CDM as a clinical trial mgmt. system. Let’s do this one after Hyve and Odysseus presentation. Distinguishing between late phase studies where an intervention is already in the vocab, vs early studies where an intervention is pre any dictionary and this needs custom codes.

**Example: Pooling trial data sets and the multi-trial set as a ‘database’ or ‘data lake’ that can be analyzed**.

**Example: How to optimize a trial with RWE data you have in hand**.