# Minutes of the Population-Level Estimation Workgroup

March 30, 2017

Present: George Hripcsack, Jamie Weaver, Rachel Melamed, Rui Duan, Yong Chen, Yuxi Tian, Alejandro Schuler , Yuriy Khoma

Martijn gave a status update on the Method Evaluation Task Force, and the discussion that have taken place on the forums and in the Eastern Hemisphere meeting.

Martijn noted that there is some tension between the goals of

1. finding the overall best method
2. finding the best method for a specific clincial question

Alejandro noted that this ambiguity can be found in epi in general. When methods are first proposed, they are presented as being good in general, but for a specific clinical question researchers will have strong preferences for applying one particular method and not another.

Yuxi commented that we could learn from computer science, where there are several benchmarks available for specific tasks, and any new technology needs to be tested on these benchmarks. Yuxi is suprised no such benchmark exists for population-level estimation methods.

Alejandro indicated there can be value in finding the overall best method. Some methods will work well in general, and some won’t. This can help (to some extent) inform selecting the appropriate method when facing a specific clinical question.

Martijn presents on the two approaches for synthesizing positive controls that are currently on the table: injection on top of real negative controls, and simulating all outcomes. He lists pros and cons of both approaches.

Alejandro notes that although we may feel good that negative controls have a null effect on average, there may be patient-specific effects that average out.

George wonders whether simply dropping covariates to simulate unmeasured confounding will lead to realistic evaluations. ‘Hard’ confounding might be mostly present in the unmeasured bit, and can’t be simulated by making measured confounding unmeasured. George suggests a head-to-head comparison of real negative controls to simulated negative controls. Alejandro disagrees with George on his point of ‘hard’ confounding being mostly unmeasued, and mentions age as a counter example which is measured and often introduces very complicated confounding.

Martijn mentions that a simple version of injection on top of real negative controls is already implemented and seems to work as tested in a small pilot study.

Alejandro mentions that his framework of simulating all outcomes still requires an implementation for time-to-event outcomes.

A decision was made to work with injection on real controls for now. Later (for comparative effectiveness studies) we’ll compare this approach to Alejandro’s framework.