Answers to Questions Asked During the Webinar

Provided by: Walter Dempsey, Ph.D.
University of Michigan
Methods: Mind the Gap
Tuesday, November 5, 2019

1. **Slide 27, association example - are these sample-wide effects. or single-subject effects?**

The estimates are sample-wide. One can account for patient heterogeneity via frailty models; however, the covariate effects will still be sample-wide. For recurrent events, one can account for patient heterogeneity via mixed effect models and get subject-specific effects.

2. **Did you try the time-dependent AUC to compare the prediction accuracy?**

We focused on the Brier score in this paper but did check time-dependent AUC and saw improved performance. That said, we were less interested in building a classifier and more interested in making probabilistic statements about event risk which is why we focused on Brier score.

3. **I am interested in learning more from you about how you evaluate the impacts of using mHealth intervention on the participants by taking into account the complicated process of behavioral or health outcomes changes. For example, the patients' behaviors might have been changed because the interventions have been provided when we build in tailored and interactive mechanisms in the mobile App. So, after they started to use the mobile, the timing and content of the intervention are probably all different for the participants and the changes in outcomes are measured in different time points when they do occur.**

In recent work, we have designed the micro-randomized trial to assess time-varying effects of treatment and how they are moderated by contextual information which itself may be impacted by prior treatment. There are many challenges in such work such as those described in the question. We estimate effects called *causal excursions*. Our primary analyses aim at estimating the impact of receiving an intervention versus not receiving an intervention on some proximal outcome of interest. The effect is conditional on some moderators of interest, and marginal over prior treatment and all other variables in the observed history. The primary goal is to assess if the intervention has any effect on the outcomes of interest. If we can show this to be true, then we try and build intervention policies that send the messages at important times and contexts. That is, we see these trials as exploratory rather than confirmatory. They are to assess whether an intervention component should be included in a final intervention package. We then run a sequence of trials to better decide when and in which contexts to provide this component to the patient.

4. **What are the challenges in statistical modelling given that if we are interested in combining providers' and patients' data to evaluate the intervention effects? For example, providers use the tool to do referrals or provide consultation and the patients receive that information and make their decisions?**

In our current mHealth studies, we do not take into account information from providers. However, in the future, this will be an issue. As stated previously, we see the interventions as one component of an overall intervention package which may include referrals or consultation. In such cases, the intervention policy for a particular component will account for this additional information. In particular, the timing and context in which the component will be triggered may change based on this information.

5. **As Brier score is not a universally agreed metric for prediction accuracy. Also how is the model's calibration performance? How many random effects did you used in the model? And what's the computation burden like?**

The model appears well calibrated; however, we note the small sample size. We would like to assess calibration on a much larger sample size in the near future. No random effects were included in this model. The CT-HMM can be extended in this direction; however, we wanted to keep the model as simple as possible given small sample size and computation burden. For the current model, we introduced a simple EM algorithm that converges quickly. Joint modeling in general entails higher computational burden compared to simple regression methods. However, they are necessary for unbiased estimation of time-varying covariate effects on survival and the other questions discussed in the talk.

6. **I wonder if multiple imputation is the right way for EMA data since we are trying to "reconstruct" people's behaviors and mental experiences in real-life settings. Also, how about missing data in sensors, do we also need to deal with it before modeling?**

In the Sense2Stop example, we are imputing the latent smoking times. We are less interested in reconstructing what the person's answer would have been, and more interested in figuring out when they smoked given various pieces of information (i.e., random EMA, event-contingent EMA, end-of-day EMA, and sensor output). Currently we do not impute missing sensor data. Instead, we assume sensor data satisfies a missing-at-random assumption (i.e., the sensors become detached for issues unrelated to if the user is smoking or not). The sensor information is then used to help pinpoint smoking times if and when the sensors are giving high quality data.

7. **For all the models you talked about, the outcome measures could be either continuous, dichotomous, ordinal, or discrete choices?**

The longitudinal data can be continuous, dichotomous, ordinal, or discrete. Joint modeling is readily equipped to handle such measures by using generalized linear models to account for these various measurement types.

8. **Could you interpret that as Odds Ratio?**

The odds ratio and hazard ratio are distinct. The odds ratio concerns the relative change in odds of an event occurring based on covariates, while the risk ratio concerns the instantaneous risk of an event changing based on covariates. The two quantities are not mathematically equivalent. In the Cox regression example, the hazard ratio of treatment versus no treatment was exp(0.2129) = 1.237. This means the instantaneous risk of the event is 1.237 times higher under treatment. To make this clear, if the baseline risk were constant (K), then the probability of an event in the first unit time is 1 - exp(-K) under no treatment and 1 – exp(-1.237 K). For K=1, the probability is 0.63 versus 0.71. The odds ratio is 1.42 which is not the same as the risk ratio of 1.237. If we set K = 2, the odds ratio increases to 1.70, but the risk ratio stays the same.

9. Were these covariates adjusted or not?

The covariates were not adjusted. The positive and negative affect were based on factor models applied to EMA questions. The factor models were used to build risk models that were interpretable and go from a higher to lower dimensional covariate space (i.e., the model includes 2 covariates rather than many highly collinear covariates).