Panel Discussion

Ionut Bebu
The Biostatistics Center
Department of Biostatistics and Bioinformatics
The George Washington University
ibebu@bsc.gwu.edu

A Wonderful Tradition

- The Annual UPenn Conference in CLinical Trials is on our calendars as a must-attend meeting for a good reason: wonderful talks on extremely important (both methodological and clinical) topics.
- This year is no different:
    - Eight very exciting presentations
    - Excellent discussions
    - A wonderful dinner
- We are all trying to replicate this conference in one way or another, but...
- Many thanks: Susan Ellenberg and Mary Putt, and the entire organizing committee.

Multistate Models and Life History Processes

- Professors Therneau and Cook advocate for the use of multistate models in describing (perhaps complex) life history processes.
- One can approach it from (at least) two angles: efficiency, and clinical relevance.
- The most important part of a study are the participants: there is no study without participants...
- Efficient use of all the relevant data from each participant is crucial. Completely agree with going beyond a single yes/no, or a single p-value.
- This is both an efficiency argument as well as an ethical argument.

Multistate Models and Life History Processes

- A simple yes/no (or even a single outcome) may not even address the real scientific questions.

- Individuals with T1D have higher risk of micro- and macrovascular complications, and mortality compared to non-diabetetic individuals.

- The DCCT study (1983-1993) investigated the glycemic hypothesis: controling blood glucose levels close to non-diabetic levels prevents incidence of complications, and reduces risk of progression to more severe disease among those with early disease.

- The DCCT study was only powered to prove the effect for 3-step progression of retinopathy on the ETDRS scale. Intensive therapy aimed at controling glycemia became standard of care for T1D.

- The real clinical question was about advanced complications!

Multistate Models and Life History Processes

- With longer follow-up, the EDIC observational study (1994–present) has shown the beneficial effect of glycemic control with respect to retinopathy, nephropathy, neuropathy, CVD, and mortality.

- It will depend of course on the specific disease and outcomes, but the typically limitted length of a study dictates what is feasible. This was the case for the DCCT/EDIC study.

- The multistate models are perhaps particularly helpful for understanding potentially complex life history processes with longer follow-up.

- This has the potential for very important medical decisions and public health recommendations.

Multistate Models and Life History Processes: Example
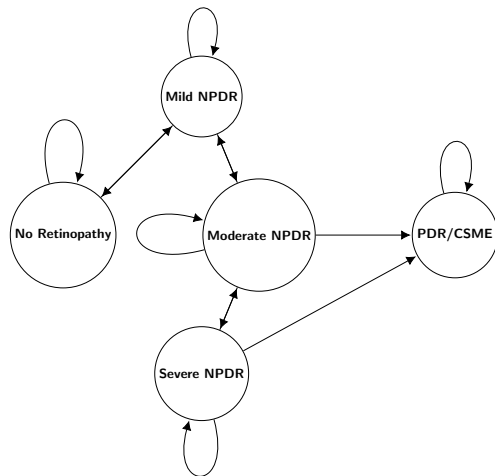
Multistate Models and Life History Processes: Example

- Time to next visit (years) as a function of the current state such that the risk of reaching PDR/CSME is approximately 5%.

Current State	Time to Next Visit	Probability
No retinopathy	5.250	0.049
Mild NPDR	3.583	0.048
Moderate NPDR	0.333	0.045
Severe NPDR	0.083	0.057

- The annual screening schedule leads to an expected undetected time of 0.606 years and an average of 18.37 visits for $L = 20$.
- Over the same time horizon, a practical screening schedule of $(4, 3, 0.5, 0.25)$ leads to an expected undetected time of 0.415 years and an average of 7.65 visits.
- The $(4, 3, 0.5, 0.25)$ schedule dominates the $(1, 1, 1, 1)$ schedule both in terms of effectiveness with an expected 0.19 years lower average undetected time, and costs with an expected 10.7 fewer number of visits over up to 20 years of follow-up.

Non- and Semiparametric Analysis of Composite Time-to-Event Endpoints

- Dr. Mao provides an excellent overview of the state-of-the-art results on generalized pairwise comparisons, such as win ratio of proportion in favor of treatment.
- The idea: Finkelstein & Schoenfeld (1999) and Pocock et al. (2012).
- Both potentially more efficient (uses more events), and more clinically relevant (accounts for the severity of the individual outcomes).
- Scott Evans and collaborators are advocating for a similar approach desirability of outcome ranking (DOOR).
- Similar to a standard TTFE analysis, for fixed marginal cumulative incidence functions, the GPC parameters further depend on the correlations between the two events. Therefore, the risk of each individual outcome may be equal in the the two groups (i.e., the same burden of disease), but the win ratio may be different from 1.
- Fey et al (2018): causal estimands associated with WMW tests.

Component-Wise Censored Composite Endpoints

- Dr. Eaton is to be commended for bringing to light one of (several) difficulties when dealing with multiple outcomes. Namely that not all individual outcomes may be assessed based on the same screening schedule.

- Retinopathy: composite PDR/CSME (say assessed every two years) and/or laser surgery (in continuous time).

- The question is how to account in the standard models/tests for the difference in assessment.

- John Lachin showed that the power of a test using a discrete or grouped time-to-event data approaches the power of a test using continuous-time event data as the frequency of evaluations increases.

- In some applications the screening schedule is likely to depend on social-economic factors (e.g., access to care). The (naive) standard approach is then likely to yield biased the results for factors associated with the screening (e.g., race/ethnicity).

References

- Nathan D *et al.* Frequency of evidence-based screening for retinopathy in type 1 diabetes, *NEJM*, 376(16):1507–1516, 2017.

- Bebu I, Lachin JM. Optimal screening schedules for disease progression with application to diabetic retinopathy, *Biostatistics*, 19(1):1–13, 2018.

- Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med*, 18:1341-–1354, 1999.

- Pocock SJ *et al.* (2012). The win ratio: a new approach to the analysis of composite outcomes in clinical trials based on clinical priorities. *Eur Heart J*, 33:176–182.

- Evans SR et al. Desirability of outcome ranking (DOOR) and response adjusted fordDuration of antibiotic risk (RADAR), *CID*, 61:800–806.

- Bebu I, Lachin JM. Properties of composite time-to-first-event versus joint marginal analyses of multiple outcomes, *Stat Med*, 37(27):3918–3930, 2018.

- Lachin JM. Power of the Mantel–Haenszel and othertests for discrete or grouped time-to-event data under a chained binomial model, *Stat Med*, 32:220–229, 2013.

- Fey MP *et al.* Causal estimands and confidence intervals associated with Wilcoxon-Mann-Whitney tests in randomized experiments, *Stat Med*, 37:2923-2937, 2018.