University of California, Irvine
Statistics Seminar

Assessment of Surrogacy Using Potential Outcomes and Bayesian Modeling

Jeremy M G Taylor
Pharmacia Research Professor of Biostatistics
University of Michigan

4 p.m., Thursday, February 10, 2022

In randomized clinical trials, an outcome variable (S) can be measured before the true outcome of interest (T) and may provide early information regarding the treatment (Z) effect on T. The available data from a single trial consists of T, S, Z and possibly baseline variables (X). The question of interest is whether S can be used as a surrogate outcome to assess the effect of Z on T. A basic metric for assessing surrogacy can be based on estimation of the quantity $m(s)=E[T(1)-T(0)|S(1)-S(0)=s]$, where $S(0)$, $S(1)$, $T(0)$ and $T(1)$ are the potential outcomes. As a starting point we propose a Bayesian estimation strategy when the joint distribution of potential surrogate and outcome measures is multivariate normal, and from this inference about the quantity $m(s)$ can be derived. As the model is not fully identifiable from the data, we propose some reasonable prior distributions and assumptions that can be placed on weakly identified parameters to aid in estimation. Examples of the type of assumptions that could be plausible in the surrogacy setting are $|corr(S(0),S(1))|> |corr(S(0),T(1))|$, or that $S(0)$ and $T(1)$ are conditionally independent given $S(1)$, $T(0)$ and X. This basic structure can be extended to handle different data types. One example is for an ordered categorical surrogate and a censored event time true endpoint for which we use Gaussian copulas. Another case is in which T is longitudinal, in which case random effects are included in the model. When both S and T are time to event outcomes, we formulate their joint distribution using illness-death models. The specific example I will use in this talk is for a gene therapy trial in Duchenne muscular dystrophy, for which T is longitudinal and from domain knowledge we know that $S(0)=0$.