Identification and Assessment of Longitudinal Biomarkers Using Frailty Models in Survival Analysis

葛豐壽
國家衛生研究所博士後研究

Abstract

A biomarker is a measurement which can be used as a predictor or sometimes even a surrogate for a biological endpoint that directly measures a patient's disease or survival status. Biomarkers are often measured over time and so are referred to as longitudinal biomarkers. Biomarkers are of public health interest because they can provide early detection of life threatening or fatal diseases.

In this dissertation, we introduce a method employing a frailty model to identify longitudinal biomarkers or surrogates for a time to event outcome. Our method is an extension of earlier work by Tsiatis, et al., 1995, Wulfson and Tsiatis, 1997, Song et al., 2004 where it was assumed that the event times have the same baseline hazard. To do this, we develop a joint likelihood function which combines the likelihood functions of the longitudinal biomarkers and the survival times. Henderson, et al. (2000, 2002) put random effects in the longitudinal component or in the Cox model to determine if a longitudinal biomarker is associated with time to an event. In our method, we extend Henderson, et al and allow random effects to be present in both the longitudinal biomarker and underlying survival function. The random effect in the biomarker is introduced via an explicit term while the random effect in the underlying survival function is introduced by the inclusion of frailty parameters into the model. We use simulations to explore how the number of individuals, the number of time points per individual and the functional form of the random effects from the longitudinal biomarkers influence the power to detect the association of the longitudinal biomarker and the survival time. We also explore effect of missingness on how a biomarker predicts a time to event outcome. We conclude that for a given sample size, the biomarker effectiveness for relatively small numbers of subjects and large numbers of observed time points is better than for relatively large numbers of subjects and small numbers of observed time points. We also conclude that when the missing data mechanism is missing at random (MAR), our method works reasonably well. However, when the missing data mechanism is non-ignorable, our method doesn't perform well in determining whether or not potential biomarkers are good predictors of a time to event outcome. Finally, we apply our method to liver cirrhosis data and conclude that prothrombin is a good predictor of time to liver cirrhosis and thus, can be used as a potential surrogate for liver failure