

# Design and Analysis of Clinical Trials, and Analysis of Correlated Data

Li-An Lin\*

Department of Biostatistics, University of Texas Health Science Center, USA

## Design and Analysis of Clinical Trials

Today, human disorders are thought to be heterogeneous with regard to their molecular and genomic properties. Recent advances in biotechnology have resulted in a shift toward molecularly targeted agent. These new therapeutics works on specific genetic pathways which are only dysregulated in some smaller subset of patients with the disease. New and general statistical methods are needed so that genetics or other biomarkers can be used to assist the analysis of clinical data for personalized medicine. The challenge raised by personalized medicine is how to find the predictive biomarkers that can identify patients who are likely or unlikely to benefit from a specific treatment within large number of biomarkers to be considered and how to incorporate biomarkers information in randomized clinical trial while considering issues of both efficiency and medical ethics. Several statistical methods for evaluating and comparing biomarkers for patient treatment selection have been developed. Currently available methods are far from sufficient to deal with the complications of patient treatment selection in clinical trial practice. A good design of clinical trials can use less number of total patients, send more patient to the better treatment arms, and expose less patients to highly toxic treatments. We will work in these areas to help advance clinical trial design.

Once predictive biomarkers identified, phase III trial is needed to adaptively modify the target patient population that serves as the basis for comparing the new treatment to the control. In such designs, enrollment would initially open to a broad patient population with the option to restrict future enrollment, following an interim analysis, to only those biomarker subgroups that appeared to be benefiting from the experimental therapy. However, the benefits of such design or method could be limited if it takes very long to obtain the primary endpoint measurements at the interim. Considerable interest has been drawn toward the short-term endpoint. Several prognostic events and prognostic measurements can be used to predict the primary endpoint. However, it's statistical challenge to combine all type of prognostic measurements to adaptive enrichment trial. We consider the delayed outcomes as missing and propose mitigating the delay effect by imputing from the short-term endpoint. Such designs can both increase the efficiency of the clinical trial and protect patients from exposure to treatments with serious toxicities to which they may have little likelihood of benefit.

## Analysis of Correlated Data

In clinical and epidemiological studies, the observations of interesting are rarely independent. For time-to-event data, the events may occur many times for a given subject, such as such as repeated lung infections in people with cystic fibrosis, recurrent shunt failures in children with hydrocephalus, and recurrent strokes in older adults. An important feature of recurrent events is that the event times are correlated, including within-subject correlation, event-specific dependence, between event-type correlation, etc. Thus, correlation structure need to be studied and adjusted when conduct statistical modeling. We consider two types of correlation among recurrent event times which is subject-specific heterogeneity and event-specific dependence. Subject-specific heterogeneity represents unmeasured variables which induce the within-subject correlation among event times. Alternatively, correlation may be induced by recurrent event processes where the incidence of events may elevate or reduce the risk of future recurrent events. We will consider both types of correlation and quantify how the incidences of events change the risk of future events and the long-term risk factor effects that could vary as event incidence increases.

Moreover, the independent variables may not linearly related to the response variable. For example, age is one of the largest risk factors for cardiovascular diseases, and its functional form for risk of cardiovascular diseases may not be linear. Multiple studies have reported that the risk of cardiovascular diseases increases greatly for older adults. In the nonparametric framework, the shape of the regression function is determined by the data, whereas in the parametric framework the shape is determined by the model and constrained by the linearity assumption. The typical application for drug discovery and clinical trial study is PK/PD model, dose-finding study, time-dependent treatment effect, GWAS study, predictive model, functional data analysis, and so on. In the classical approach, the choice of the smoothing parameters via cross validation is difficult if the number of smooth functions in the model is large as then the computational effort to compute an optimal solution becomes intractable whereas the simultaneous estimation of functions and smoothing parameters in the Bayesian approach works quite well.

**\*Corresponding author:** Li-An Lin, Department of Biostatistics, University of Texas Health Science Center, 7900 Cambridge St, Apt 29-1D Houston, TX 77054, USA, Tel: +281-760-6209; E-mail: [Li.An.Lin@uth.tmc.edu](mailto:Li.An.Lin@uth.tmc.edu)

**Received** January 09, 2016; **Accepted** January 11, 2016; **Published** January 18, 2016

**Citation:** Lin LA (2016) Design and Analysis of Clinical Trials, and Analysis of Correlated Data. J Biom Biostat 7: 280. doi:10.4172/2155-6180.1000280

**Copyright:** © 2016 Lin LA. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.