

## Analyzing Large Scale Recurrent Event Data with a Divide-And-Conquer Approach

**Authors :** Jerry Q. Cheng

**Abstract :** Currently, in analyzing large-scale recurrent event data, there are many challenges such as memory limitations, unscalable computing time, etc. In this research, a divide-and-conquer method is proposed using parametric frailty models. Specifically, the data is randomly divided into many subsets, and the maximum likelihood estimator from each individual data set is obtained. Then a weighted method is proposed to combine these individual estimators as the final estimator. It is shown that this divide-and-conquer estimator is asymptotically equivalent to the estimator based on the full data. Simulation studies are conducted to demonstrate the performance of this proposed method. This approach is applied to a large real dataset of repeated heart failure hospitalizations.

**Keywords :** big data analytics, divide-and-conquer, recurrent event data, statistical computing

**Conference Title :** ICCSS 2019 : International Conference on Computational and Statistical Sciences

**Conference Location :** London, United Kingdom

**Conference Dates :** April 24-25, 2019