J Proteomics Bioinform 2018, Volume 11
DOI: 10.4172/0974-276X-C1-113

Developing a novel computational method for uncovering temporal correlation among chronic diseases using longitudinal medical records

Qun Feng Dong
Loyola University Chicago, USA

The availability of large-scale electronic medical records provides an unprecedented opportunity to investigate potential temporal correlations among different diseases using computational approaches. In this study, we present a novel computational method, implemented in R, which automatically builds retrospective matched cohorts from longitudinal electronic medical records to examine whether the occurrences of any diseases show significant temporal correlations using both cox proportional hazards regression and random forest survival analysis. In the method, time is correctly modeled as a continuous variable, accurately accounting for the temporal space between the onsets of different diseases. In addition, our method is flexible for incorporating relevant confounding factors such as age, gender and other demographic and medical information in the analysis. The output of our method is a disease correlation network, which is displayed using our web-based visualization tool, implemented in JavaScript, for users to interactively explore the correlations among diseases of interest based on statistical significance of the correlations and graph-theory-based network topology. We have successfully applied the method to a longitudinal electronic medical record dataset at Loyola with 10,832,319 distinct encounters detailing 92 diseases from 425,122 patients. Many uncovered disease correlations are strongly supported by in-depth literature reviews. Our computational package is freely available for download.

qdong@luc.edu