Balancing Efficacy and Toxicity by Constrained Q-learning in Multi-Decision Stage

Chin-Fu Hsiao[1] and Chian Chen[1]

[1]*Institute of Population Health Sciences, National Health Research Institutes, Taiwan*

Abstract

A dynamic treatment regime (DTR) is a set of sequential decision rules, each corresponding to a key decision point in a disease or disorder process, where each rule takes as input patient information and returns the treatment option he/she should receive. Most of current work focuses on estimating optimal treatment strategy targeting on one side of clinical decision making – the clinical efficacy, but they ignore the potential risk due to the "optimal" and potentially aggressive treatments. The important issue of controlling risk while maximizing benefit has long been recognized in the clinical community where safety concerns for medications often arise, since the most efficacious medication for a patient may also lead to a higher risk. We developed the statistical framework for balancing efficacy and toxicity in the DTR settings. In the DTR settings, information gain improves decision making at subsequent decision points. The proposed method found treatment policies which aims to maximize efficacy under the constrained toxicity through combing the statistical framework in the DTR settings and a variety of reinforcement learning techniques.