

# **Robust Imputation-Based Q-Learning for Optimizing Dynamic Treatment Regimes with Right-Censored Data**

Lingyun Lyu, Yu Cheng, and **Abdus S. Wahed**

## **Abstract:**

Q-learning has been one of the most commonly used methods for optimizing dynamic treatment regimes (DTRs) in multi-stage decision-making. Right-censored survival outcome poses a significant challenge to Q-Learning due to its reliance on parametric models for counterfactual estimation which are subject to misspecification and sensitive to missing covariates. Limited works reported for estimating optimal DTRs for survival outcomes with Q-learning mostly used parametric models. In this talk, we propose an imputation-based Q-learning (IQ-learning) where flexible nonparametric or semiparametric models are employed to estimate optimal treatment rules for each stage, and then weighted hot-deck MI and direct-draw MI are used to predict optimal potential survival times. Missing data are handled using inverse probability weighting and MI, and the non-random treatment assignment among the observed is accounted for using a propensity-score approach. We investigate the performance of IQ-learning via extensive simulations and show that it is more robust to model misspecification than existing Q-Learning methods, imputes only plausible potential survival times contrary to parametric models, and provides more flexibility in terms of baseline hazard shape. We demonstrate IQ-Learning by developing an optimal DTR for leukemia treatment based on a randomized trial with observational follow-up.