Return to Work After Injury: A Sequential Prediction & Decision Problem

Erkin Ötleş[1,2], Haozhu Wang[3], Suyanpeng Zhang[2], Brian Denton[2], Jon Seymour[4], & Jenna Wiens [3]
[1]University of Michigan, Medical Scientist Training Program, [2] University of Michigan, Department of Industrial & Operations Engineering, [3] University of Michigan, Department of Electrical Engineering & Computer Science, [4]Peers Health

Background. Occupational injuries (OI) are an immense burden on the U.S. population and economy. In 2016, 4.5 million people were injured in the U.S., associated with yearly costs between $151-250 billion. Managing the recovery process of OIs is challenging, as it requires knowledge about the evolving recovery of patients be combined with knowledge about the potential benefit of therapies. Ideally, clinical decision support tools could aid in the difficult tasks of evaluating potential patient recovery and optimizing the choice of therapeutic options. To this end, we explore machine learning methods to address two problems: 1) sequential prediction of the probability of returning to work (RTW) and 2) sequential decision making regarding the selection of appropriate therapies in managing OIs.

Methods. Peers Health provided access to a dataset of OIs from the state of Ohio (2001-2011) containing over 1.2 million patient injury records. Injury records contain both demographic (age, gender, job-type) and longitudinal data (e.g., diagnoses, therapies, and medications). All individuals who did not leave work were excluded from our analysis. For the sequential prediction of RTW, we randomly assigned individuals to train, validation, and test sets. Given daily observations, we trained an RNN to predict the outcome of RTW on the next day. Discriminative performance was evaluated in terms of the area under the curve (AUC) on a bootstrapped held-out test set. We compared the performance of the proposed approach to two baselines: a) 'demographics only,' a model based on only information available at the time of injury and b) 'snapshot,' a model based on only daily info, with no longitudinal history.

For the sequential decision-making task of recommending appropriate OI management therapies, we analyzed a subset of records pertaining to lower back injuries. We investigated an offline off-policy reinforcement learning (RL) approach to match individual states to actions (i.e., therapies). Weekly states corresponded to an individual's age, gender, job type, diagnoses, treatments received, and time since injury. We clustered the state vectors and used tabular-Q learning to learn a policy mapping to any combination of potential therapeutic actions: physical therapy, chiropractic, NSAIDs, muscle relaxants, and opioids. We split the data into independent train, validation, and test sets (80/10/10%). For training, we employed a reward function based on health outcome (+10 for returning to work, and -10 for not) and used a discount factor of 0.9 to encourage sooner return. As a baseline, we compared the learned policy to a policy that always selected the "no treatment" option, a random policy, and the current policy. We estimated the value of the current policy by calculating the average cumulative reward. The other policies were evaluated using weighted importance sampling.

Results. In the sequential prediction task, 502,586 individuals had an absence from work. The RNN achieved better discriminative performance AUC=0.763 (95% CI: 0.759-0.767) compared to the two baselines demographics only 0.653 (0.650 - 0.656) and snapshot 0.596 (0.592-0.601). For the sequential decision task, 125,997 records pertained to lower back injuries. Our learned policy achieved an expected value of 8.83 (8.89 - 9.34), outperforming the expected value of the no treatment 0.09 (-0.03 - 0.86), random 0.20 (0.03 - 3.45), and current 8.57 (8.49 - 8.61) policies. During evaluation, the effective sample size was 4,632 out of 12,320 test examples. The learned policy tends to map states representing severe injuries to more treatment actions. Specifically, compared to the current policy, the learned policy tends to more frequently recommend NSAIDs and opioids.

Conclusion. Managing OIs is a challenging problem. This analysis highlights how large observational OI datasets, like the Peers Health dataset, combined with machine learning tools have the potential to generate insights into optimal treatment selection. However, before such techniques can be applied prospectively, further validation on independent datasets is needed. Moreover, validation against current practice through qualitative examination of model predictions and decisions is critical for safe deployment.