Abstract: Annually, the United States of America spends more than $3.0 trillion in healthcare, and the cost has a trend to grow faster in recent years. Although the cost is high, several patients fail to get timely and effective medical treatments. Therefore, there is a crucial need to predict the recovery time and the death time of patients to achieve timely and effective healthcare services. Albeit the last few years have witnessed an explosive increase of healthcare data in terms of volume, variety and veracity, it is insufficient to build a robust prediction model in various scenarios due to time, geographical and domain inherent constraints. Survival analysis can overcome time constraints to build a robust model at an early stage, where there are only a few fully observed patients but numerous partially observed patients. Transfer learning models can be used to overcome geographical constraints to help small hospitals in rural areas, where they do not have enough medical records, to build robust prediction models using the healthcare information exchange. Recently, precision medicine has become a new nationwide effort, which aims at providing personalized prevention methods and treatment strategies with the integrating of genomic and traditional clinical information. Some of the fundamental challenges that arise in this domain is the presence of longitudinal information of the patients, insufficient number of samples, and preserving patient privacy. This talk will describe several ways to overcome these challenges and discusses several new exciting research directions.

Bio: Dr. Yan Li is a Postdoc fellow in the Department of Computational Medicine and Bioinformatics at University of Michigan, Ann Arbor. He received his Ph.D. and M.S. from Wayne State University and B.S. from Xidian University. His primary research interests are Data Mining and Machine Learning with applications to Healthcare Analytics and Bioinformatics. His research works have been published in leading conferences and journals including SIGKDD, ICDM, WSDM, SDM, CIKM, DMKD, and Information Science.