Probabilistic Time Series Modeling of Tumor Growth in Patient-Derived Xenografts

[1]Rohan Subramanian and [1]Wesley Tansey
[1]Computational Oncology, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center

Pre-clinical drug screens often fail to predict clinical success due to limited generalizability of cell line models. Patient-derived xenografts (PDXs), where human tumor tissue is engrafted into mice, enable more accurate *in vivo* drug screening. Mathematical modeling of tumor growth in PDXs allows prediction of drug response and tumor growth dynamics. Previous studies have used ordinary differential equations to simulate natural tumor growth (Voulgarelis 2022) or deep learning to predict drug responses separately (Mourragui 2021). We propose an autoregressive Hidden Markov Model (HMM) that simultaneously predicts tumor growth and drug response. The model has tumor volumes as emissions and learns discrete latent states. Tumor volume on a given day depends on the previous day and the latent state, and the state transition matrix is determined by learned embeddings for the PDX, drug, current and next state. We applied our model to the Novartis PDX Encyclopedia containing 4699 curves from 62 drugs and 277 PDXs across six cancer types (Gao 2015). Our model successfully learns PDX embeddings whose latent structure corresponds to the cancer type. It identifies five latent states with different growth regimes and predicts interpretable latent state sequences given observed tumor volumes. State durations are significantly associated with response categories (mRECIST), linking latent states to clinically relevant outcomes. Simulation of growth curves by sampling for specific PDX-drug combinations predicts different regime-switching dynamics corresponding to expected response. Our findings highlight the utility of HMMs in modeling tumor growth and identifying growth regimes related to disease states. Joint learning of time series dynamics and PDX-drug interactions improves prediction of PDX curves. The flexibility of this framework supports extensions such as modeling survival probabilities and integration of mutational and transcriptomic data.