

「新しい社会システムデザインに向けた情報基盤技術の創出」
2016 年度採択研究者

2018 年度
実績報告書

河添 悦昌

東京大学大学院医学系研究科
特任准教授

医療ビッグデータからの病態進行のシミュレーションによる先制医療に向けた研究開発

§ 1. 研究成果の概要

【研究の背景】

- 過去 3 回のブームを経て、医療への AI 応用が本格化している。現在の主要な AI 技術である Convolutional Neural Network は画像に対する識別性能が高いことから、医用画像や動画を対象とした AI は撮影モダリティの数だけ開発され、実用になるまでの期間も短い。その一方で、投薬情報や検査結果といった、画像以外の電子カルテデータを AI 技術により活用することで、臨床的イベントを予測するモデルの開発も盛んとなってきている。ここで、電子カルテデータは longitudinal なデータを多く含み、これは 1 つの対象について経過を追い複数の時点で観察結果を記録したものであるが、このことが予測モデルの開発を困難としている。例えば、血液や尿などの検体検査結果は典型的な longitudinal データであり、極めて多くの項目からなる多変数の時系列データであるが、全項目が同時に測定されず、また測定間隔が不規則であることにより時系列データとして扱う際に多くの欠損を生じ、解析の上での問題となる。また、検体検査結果以外にも、投薬オーダや保険登録病名などの情報を利用する上でも同様の問題が生じる。

【研究の狙い】

- 本研究は複数の医療データを時系列データとして利用する手法を開発する。これまで、不等間隔で測定される血液検査の欠損値の補間に畳み込みオートエンコーダを使った手法を適用し、線形補間に比べ精度の向上を認めること、ならびに、糖尿病検査項目の系列から、欠損を考慮したオートエンコーダを用いて低次元特徴量を取り出し、この特徴量を元に予後分類を行う Neural Network モデルの効果を報告してきた。

【本年度の報告】

- 本年度は、臨床的イベントとして深部静脈血栓症(Deep Vein Thrombosis:DVT)の発生有無を予測する LSTM モデルの性能を評価した。また、Self-Supervised Learning を適用することで、欠損値を復元する過程で得られる、より低次元の表現が予測精度の向上に寄与するかを検討

した。DVT 診断のゴールドスタンダードは、超音波検査で血管内の血栓を確認することであるが、人的リソースの問題から網羅的に行なえず、また、DVT 予測スコア(Wells スコア)が知られているが、それには熟練した医療者が判断する必要がある項目が含まれるため簡便に実施できない。そこで、日々発生する電子的診療データから計算機により予測可能となれば、より網羅的なスクリーニングにつながることを期待される。

§ 2. 研究実施体制

- ① 研究者:河添 悦昌 (東京大学大学院医学系研究科 特任准教授)
- ② 研究項目:
 - 研究計画の策定
 - データの抽出
 - 解析プログラムの作成
 - 実験の評価
 - 論文の執筆