

厚生労働科学研究費補助金  
(政策科学総合研究事業(臨床研究等 ICT 基盤構築・人工知能実装研究事業))  
分担研究報告書

Precision medicine の確立に資する統合医療データベースの利活用に関する研究

研究分担者 中島 直樹 九州大学病院メディカル・インフォメーションセンター 教授  
研究分担者 野原 康伸 熊本大学大学院先端科学研究部 特任准教授

研究要旨

病院診療データを対象とした機械学習手法によるデータ解析を行った。医療データには非線形な要素が多いことなどから、既存の手法を上回る予測精度が得られた。また、機械学習の解釈手法の改善を行い、より妥当な解釈を行うことが可能となった。機械学習手法による解析は、統合医療データベースに対しても、妥当性の高い結果が期待できる。

A. 研究目的

病院診療データを対象とした機械学習手法によるデータ解析を行うことを通じて、統合医療データベースに機械学習手法を適用するにあたっての利点や問題点を洗い出す。

B. 研究方法

済生会熊本病院の脳梗塞患者データを用いて、患者の予後スコアおよび肺炎併発の有無について、機械学習手法による予測モデルを構築し、既存モデルとの性能比較をおこなった。また、なぜそのような結果が出たかを Shapley Additive Explanation(SHAP)等の手法を用いて解釈を試みた。

(倫理面への配慮)

病院データの利用に当たっては、病院倫理審査委員会による許諾を得ている。

C. 研究結果

脳梗塞患者の予後予測について、Gradient Boosting Decision Tree や Random Forest とい

った機械学習アルゴリズムを用いることで、線形回帰モデルによる予測および IScore、ASTRAL といった臨床で広く用いられる予後スコアを上回る予測精度が得られた[論文 1-1]。また、肺炎併発の有無をアウトカムとする予測モデルを構築し、そのモデルを解釈した結果、入院時の意識レベルが重要であることを見出した。そこで入院時の意識レベルに基づいたリスク評価を行い、口腔ケアの実施等の業務改善を実施した。その結果、肺炎発症を有意に減少させることができた[論文 1-3]。

機械学習モデルの解釈手法である SHAP について、変数重要度の計算方法の改善、および、変数グループ化という新たな解析手法を提案した。線形モデルで用いられている変数重要度との完全に整合性が取れるようになり、予測モデルの再構築を伴うことなく、複数の説明変数がアウトカムに与える影響を調べることが可能となった[論文 2-2]。

## D. 考察

医療データは、非線形な要素が多いため、これまで医療で広く用いられてきた線形モデルと比べて高い予測精度を実現できたものと考えられる。また、機械学習手法は、なぜそのような結果が得られたか分からずブラックボックス的と問題点を指摘されることが多いが、線形モデルとの対比を通じた適切な解釈手法を開発したことで、臨床家の直観ともより合う結果が期待できる。

## E. 結論

病院診療データに対して機械学習手法を適用することで、既存の手法を上回る予測精度が得られた。また、機械学習モデルの解釈手法の改善法を提案した。機械学習手法による解析は、統合医療データベースに対しても、妥当性の高い結果が期待できる。

## F. 研究発表

### 1. 論文発表

1. Koutarou Matsumoto, Yasunobu Nohara, Hidehisa Soejima, Toshiro Yonehara, Naoki Nakashima, Masahiro Kamouchi, “Stroke Prognostic Scores and Data-Driven Prediction of Clinical Outcomes After Acute Ischemic Stroke”, *Stroke*, Vol.51, No.5, pp. 1477 - 1483, May 2020
2. Keisuke Ido, Ryota Kurogi, Ai Kurogi, Kunihiro Nishimura, Koichi Arimura, Ataru Nishimura, Nice Ren, Akiko Kada, Ryu Matsuo, Daisuke Onozuka, Akihito Hagihara, So Takagishi, Keitaro Yamagami, Misa Takegami, Yasunobu Nohara, Naoki Nakashima, Masahiro Kamouchi, Isao Date, Takanari Kitazono, Koji Iihara, J-ASPECT Study Collaborators, “Effect

of treatment modality and cerebral vasospasm agent on patient outcomes after aneurysmal subarachnoid hemorrhage in the elderly aged 75 years and older”, *PloS one*, Vol.15, No.4, e0230953, Apr. 2020

3. Koutarou Matsumoto, Yasunobu Nohara, Yoshifumi Wakata, Takanori Yamashita, Yukio Kozuma, Rui Sugeta, Miki Yamakawa, Fumiko Yamauchi, Eri Miyashita, Tatsuya Takezaki, Shigeo Yamashiro, Toru Nishi, Jiro Machida, Hidehisa Soejima, Masahiro Kamouchi, Naoki Nakashima. Impact of a learning health system on acute care and medical complications after intracerebral hemorrhage. *Learn Health Sys.* 2020; e10223. <https://doi.org/10.1002/lrh2.10223>, Mar. 2020

4. Tasuku Okui, Jinsang Park, Naoki Nakashima, An Interrupted Time Series Analysis Method for Healthcare Data Using the INGARCH Model: An Application to Psychotropic Drug Prescription Data in Japan, *Asian Pacific Journal of Health Economics and Policy*, 2:2-12, Feb.2020.

5. Jinsang Park, Tasuku Okui, Hiroko Furuhashi, Shoji Tokunaga, Naoki Nakashima, Evaluation of polypharmacy in Japan using the national health insurance claims database in 2015-2017, *European Journal of Public Health*, 29(4):351, Nov.2019.

### 2. 学会発表

1. Nice Ren, Ataru Nishimura, Kunihiro Nishimura, Koichi Arimura, Ai Kurogi, Ryu Matsuo, Ryota Kurogi, Naoki Nakashima, Yasunobu Nohara, Takanari Kitazono, Koji

lihara, J-ASPECT study group (2020). Abstract TP298: Measuring Quality of Care for Acute Ischemic Stroke Treated with Acute Reperfusion Therapy in Japan: A Nationwide Quality Improvement Initiative of the J-aspect Study. Stroke, 51(Suppl\_1), ATP298-ATP298.

2. Yasunobu Nohara, Koutarou Matsumoto, Hidehisa Soejima, Naoki Nakashima, "Explanation of Machine Learning Models Using Improved Shapley Additive Explanation", Proc. of the 10th ACM Conference on Bioinformatics, Computational Biology and Health Informatics (ACM BCB 2019), pp. 546, Sep. 2019

3. Rafiqul Islam, Naoki Nakashima, Yasunobu Nohara, Ashir Ahmed, Nazneen Sultana, Md Jiaur Rahman, "PORTABLE HEALTH CLINIC: An Advanced Tele-Healthcare System for Unreached Communities", Proc. of the 17th World Congress on Medical and Health Informatics (MedInfo2019), Vol. 264 of Studies in Health Technology and Informatics, pp.616-619, Aug. 2019

4. Rieko Izukura, Yoshifumi Wakata, Chinatsu Nojiri, Yasunobu Nohara, Takanori Yamashita, Atsushi Takada, Jinsang Park, Naoki Nakashima, "The development of an electronic phenotyping algorithm for identifying rhabdomyolysis patients in the MID-NET database" , Proc. of the 17th World Congress on Medical and Health Informatics (MedInfo2019), Vol. 264 of Studies in Health Technology and Informatics, pp.1498-1499, Aug. 2019

5. Jinsang Park, Takanori Yamashita,

Atsushi Takada, Chinatsu Nojiri, Rieko Izukura, Yasunobu Nohara, Taeko Hotta, Dongchon Kang, Naoki Nakashima , "Effect of Governance Functionality for Data Standardization Management of the Medical Information Database Network Project", Proc. of the 17th World Congress on Medical and Health Informatics (MedInfo2019), Vol. 264 of Studies in Health Technology and Informatics, pp.1562-1563, Aug. 2019

(発表誌名巻号・頁・発行年等も記入)