

## ポスター4

### 医療データ解析・テキストマイニング

2020年11月20日(金) 11:20 ~ 12:20 D会場(コンgressセンター4階・43~44会議室)

#### [3-D-2-04] 地域包括ケアシステムを見据えた脳梗塞患者の転院先予測モデルの開発

\*水迫 友和<sup>1,2</sup>、馬場 祥吾<sup>1</sup>、岩穴口 孝<sup>3</sup>、宇都 由美子<sup>3</sup>（1. 鹿児島大学大学院医歯学総合研究科医療システム情報学, 2. 鹿児島大学病院看護部, 3. 鹿児島大学病院医療情報部）

\*Tomokazu Mizusako<sup>1,2</sup>, Shogo Baba<sup>1</sup>, Takashi Iwaanakuchi<sup>3</sup>, Yumiko Uto<sup>3</sup>（1. 鹿児島大学大学院医歯学総合研究科医療システム情報学, 2. 鹿児島大学病院看護部, 3. 鹿児島大学病院医療情報部）

キーワード : Cerebral infarction, The regional comprehensive care system, Machine learning

【背景】脳梗塞患者は突然発症し、急性期病院にて加療を受けるが、その多くは麻痺や高次脳機能障害を来し転院を余儀なくされる。我々の先行研究では、脳梗塞患者の約半数が回復期リハビリテーション病棟を有する施設へ転院しており、そのまた約半数がDPC入院期間IIIでの転院となっていた。心身機能の回復に於いて、早期のリハビリテーションの開始が重要であることは明らかであるが、急性期病院での入院期間が延長することにより、専門的リハビリテーションを受ける機会の喪失が生じる可能性がある。我々は、脳梗塞患者の入院期間が長期に及ぶ要因を分析した結果、その要因として「退院決定の遅れ」「転院先施設との調整」が影響することが示唆された。また、転院先施設として急性期治療も実施可能な医療提供体制が整備されている施設への偏りがみられ、転院先施設の選択において、医療提供能力も要因の一つとして示唆された。

【目的】本研究では地域包括ケアシステムの構築を踏まえ、患者・家族の思いや医療者の判断等において、どのような後方受け入れ施設の選定が行われているかを明らかにし、機械学習を用いて早期の転院先決定支援を目指した予測モデルを開発する。入院早期から適切な転院先の候補を予測することで、患者・家族の選択の幅を広げることができ、さらに転院先施設との調整が早期より取り組まれることにより、入院期間の短縮が期待される。【方法】特定機能病院 A病院に入院した脳梗塞患者のデータを病院情報システムより抽出し、機械学習を用いた転院先予測モデルの作成と評価を行う。【結果・考察】入院早期に転院先を予測する転院先予測モデルの精度は一定の水準を満たすことが出来たが、地域の医療リソースの影響を受けていることも確認できた。この予測モデルを早期転院支援のツールとして活用することは、入院期間が長期に及ぶ転院調整待ちの回避につながることを期待できる。

# 地域包括ケアシステムを見据えた脳梗塞患者の転院先予測モデルの検討

水迫 友和<sup>\*1</sup>、馬場 祥吾<sup>\*1</sup>、岩穴口 孝<sup>\*2</sup>、宇都 由美子<sup>\*2</sup>

<sup>\*1</sup> 鹿児島大学大学院医歯学総合研究科医療システム情報学、<sup>\*2</sup> 鹿児島大学病院医療情報部

## A transfer destination prediction model for cerebral infarction patients in view of the regional comprehensive care system

Tomokazu Mizusako<sup>\*1</sup>, Shogo Baba<sup>\*1</sup>, Takashi Iwaanakuchi<sup>\*2</sup>, Yumiko Uto<sup>\*2</sup>

<sup>\*1</sup>Kagoshima University Graduate school of Medical and Dental Sciences, Medical Information Science,

<sup>\*2</sup> Kagoshima University Hospital, Medical Information

[Background and purpose] Many patients with cerebral infarction experience sudden onset of symptoms and are treated at their local acute care hospital. The patients may then be forced to transfer to another hospital due to various factors. The purpose of this study was to develop and evaluate a transfer destination prediction model using machine learning, based on factors that prolong the length of hospital stays of patients with cerebral infarction admitted to acute care hospitals.

[Methods] Data on cerebral infarction patients were extracted from the hospital information system and a transfer destination prediction model was created and evaluated using machine learning.

[Results and Discussion] The results of analysis demonstrated that the accuracy of the model for predicting the outcome destination at an early stage of hospitalization satisfied a high standard. However, the prediction model that classified the transfer destinations by the medical care provision system was not sufficiently accurate. Therefore, it is necessary to take the analysis of medical resources into consideration when constructing a model to predict transfer destinations.

**Keywords:** Cerebral infarction, The regional comprehensive care system, Machine learning

### 1. 背景

我が国において脳血管疾患は、悪性新生物、心疾患に次いで、死因の第3位である。また、脳血管疾患は発症後に後遺症や合併症を併発する頻度が高く、患者は治療や療養の長期化を余儀なくされ、要介護となる主要原因となっている。中でも脳梗塞は脳血管疾患の約6割を占めている<sup>1)</sup>。

発症後救急搬送された脳梗塞患者の多くは、急性期病院で加療や早期のリハビリテーションを受け、早期の在宅復帰や地域への回復期リハビリテーション病棟への転院を目指している。また、少子高齢化が進んでいる我が国においては、地域包括ケアシステムを構築し、住み慣れた地域でのリハビリテーションや治療を継続することが求められている。我々の先行研究<sup>2)</sup>において、脳梗塞患者の入院期間が長期に及ぶ要因として「退院決定の遅れ」、「転院先施設との調整」が影響することを明らかにした。また、転院先施設として急性期治療も実施可能な医療提供体制が整備されている施設への偏りがみられ、転院先施設の選択において、医療提供能力も要因の一つである可能性が示唆された。

そこで、本研究では、脳梗塞患者の入院期間が長期に及ぶ要因である「退院決定の遅れ」に着目し、入院早期から得られるデータをもとに機械学習のアルゴリズムを用いた転院先予測モデルの開発を検討し評価を行った。

### 2. 目的

機械学習のアルゴリズムを用いて、脳梗塞患者の地域包

括ケアシステムを見据えた転院先決定支援予測モデルを開発し、入院早期からの退院支援につなげる。

### 3. 方法

#### 3.1 本研究で用いたプログラミング言語

本研究では、プログラミング言語 Python と Python のオープンライブラリとして機械学習の分野で多くの研究で用いられている scikit-learn を用いて試行を行った。Python を使用する利点は Numpy や pandas, matplotlib といった数値計算やグラフ描画といった分析・結果の可視化を行うための多彩なライブラリが備わっていることである。また scikit-learn には、分類や回帰、クラスタリング等の機能が実装されているだけでなく、様々な評価尺度やクロスバリデーション、パラメータのグリッドサーチ等の機能も含まれている。

#### 3.2 期間

2018年4月1日~2019年10月31日

#### 3.3 対象

上記期間に A 特定機能病院に入院した診断群分類 (DPC) の疾患コード「010060」の患者 85 例のうち、死亡退院患者 3 例を除外した 82 例。

### 3.4 データセット概要

解析に使用するデータは、先行研究にて明らかにした脳梗塞患者の入院期間に影響を与える因子として「看護必要度」、「退院支援カンファレンスの有無」、「入院前の要介護・要支援認定の有無」、「転帰先」とした。ただし、「看護必要度」に関しては、疾患の特性上、救急搬送されるケースが多く、入院2日目のデータを取り扱った。「退院支援カンファレンスの有無」、「入院前の要介護・要支援認定の有無」に関しては、2値のダミー変数に変換しデータを取り扱った。転帰先については同様に、「自宅退院」か「自宅退院以外」として2値のダミー変数に変換した。自宅退院以外の患者で回復期リハビリテーション病棟へ転院した患者を抽出し、転院先である回復期リハビリテーション病棟の医療提供体制(脳外科又は脳神経内科を有する)についても調査し、急性期治療が可能な回復期リハビリテーション病棟か、それ以外かで2値のダミー変数に変換した。

### 3.5 予測モデルの開発及び評価

#### 1) 機械学習アルゴリズムの概要

本研究は、機械学習の1つであるSVM(Support Vector Machine)を用いて解析を行う。様々な識別方法の中からどれを選択するかは重要であるが、本研究では、十分なデータを準備出来なかったこと、2クラス分類問題の代表的手法であり、未知データに対して高い予測精度を持つ分類器(関数)が構築できることが報告されている<sup>2)</sup>ことからSVMを選択した。SVMは教師あり学習を用いる識別手法の一つであり、線形SVMと非線形SVMに分類される。本研究では、線形SVMを用いた。また、分類における代表的な関数として「線形カーネル」を用いる。

予測モデル $h_{\theta}(X)$ はカーネル関数 $K(X^{(i)}, X)$ を用いて、次の式(1)に一般化できる。予測モデルはパラメータと特徴量の訓練データ $m$ 個の線形和になる<sup>3)</sup>。

$$h_{\theta}(X) = \sum_{i=1}^m a^{(i)} t^{(i)} K(X^{(i)}, X) + \theta_0 \quad (1)$$

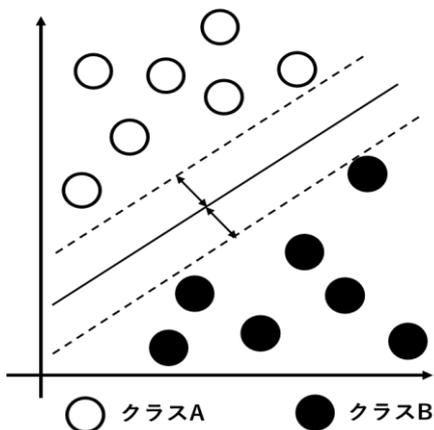


図1 SVMの識別面

#### 2) 予測モデル開発のステップ

本研究においては、得られたデータから2クラス分類にて解析を行うため、ステップ1として脳梗塞患者の転帰先が自宅であるのか、それとも自宅以外なのかを分類する。次にステップ2として、転帰先が自宅以外の患者の中から、

転帰先が回復期リハビリテーション病棟なのか、それ以外なのか、ステップ3として転院先の回復期リハビリテーション病棟が脳外科または脳神経内科を有する施設であるのかそうでないかの分類を行った。

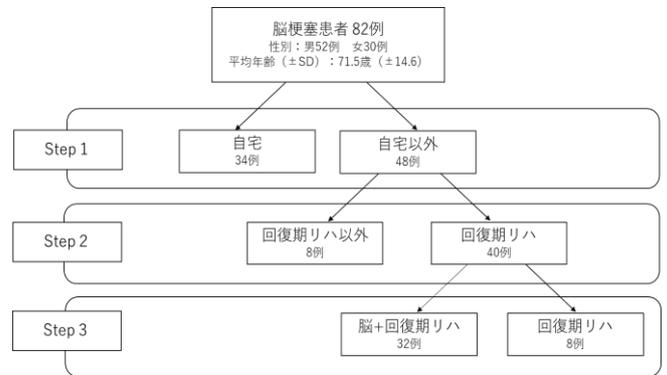


図2 予測モデル開発ステップの概要

### 3.6 倫理的配慮

本研究は鹿児島大学大学院医歯学総合研究科疫学研究等倫理委員会の承認(受付番号:190252 疫)を受けて実施した。得られた情報を取り扱う際は、個人を特定できるデータは除外し、匿名加工情報等として扱い、データ分析を行った。

### 4. 結果

SVMによるモデルの開発及び評価を行った。特徴量を「看護必要度」、「退院支援カンファレンスの有無」、「入院前の要介護・要支援認定の有無」とし、モデルに組み入れた。モデルの分類性能の評価には、Accuracy、F値、AUCを用いた。モデルを構築するためのトレーニングデータとモデルを検証するためのテストデータにランダムに振り分け、トレーニングデータで構築された推定モデルをもとにして、テストデータへ適応したときの評価を行った。ステップ1及びステップ2・3におけるモデルのAccuracy、F値、AUCの結果を以下に示す。ステップ1においては、Accuracy:0.88、F値:0.87、AUC:0.89、ステップ2においてはAccuracy:0.81、F値:0.45、AUC:0.5、Accuracy:0.85、F値:0.46、AUC:0.41であった(表1)。また、各ステップのROC曲線を図3-5に示す。

ステップ1では、高い精度を示す結果を得たが、ステップ2・3では、十分な結果を示すことが出来なかった。

表1 各モデルの評価結果

	Accuracy	F値	AUC
STEP1	0.88	0.87	0.89
STEP2	0.81	0.45	0.5
STEP3	0.85	0.46	0.41

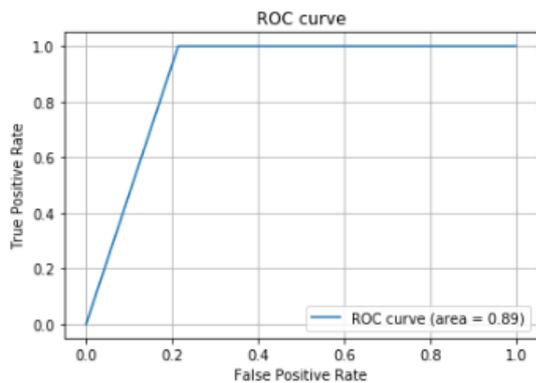


図3 ステップ1のROC曲線

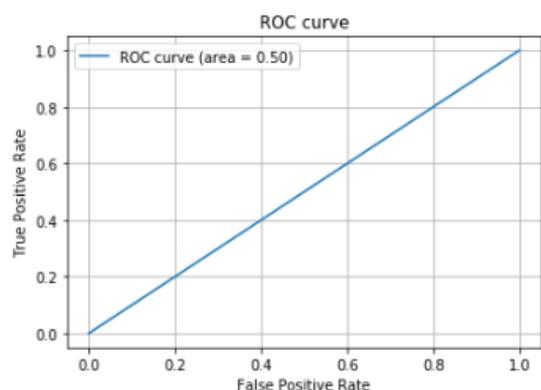


図4 ステップ2のROC曲線

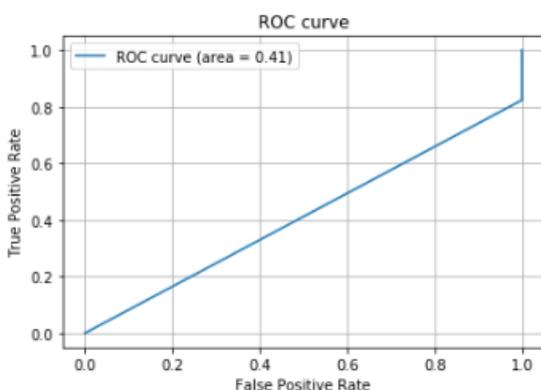


図5 ステップ3のROC曲線

## 5. 考察

本研究は、我々の先行研究をもとに、脳梗塞患者の急性期における入院期間が長期化する要因を特徴量とした。また、その特徴量は入院早期から得られるデータである。その特徴量を用いて機械学習の一つである SVM のアルゴリズムを用いて転院先の予測モデルの検討を行った。

表1に示した通り、転帰先が自宅かそれ以外であるかの予測においては高い精度を得ることが出来た。入院期間が延長する要因として、治療完了後の不要な転院調整待ちがその一つとして挙げられる<sup>4)</sup>。また、脳梗塞患者が急性期病院から自宅へ退院する場合、家屋環境や介護保険サービスなどの準備や調整に時間を要する。すなわち、入院早期から得られるデータから、入院早期に転帰先が予測出来れば、入院

早期からの退院調整、転院調整が期待できる。

一方、転帰先が回復期リハビリテーション病棟か、それ以外かの分類については、期待する結果は得られなかった。その要因は、テストデータにおける正解率である Accuracy は 0.81 と一定の水準を満たしているものの、F 値は 0.45、AUC は 0.5 と低いことから、データの偏りの影響が最も考えられる。脳梗塞患者の転帰先の選択において、疾患の特性上、回復期リハビリテーション病棟を選択することが多く、その理由には、機能障害による ADL 低下によるものが大きい。また、回復期リハビリテーション病棟ではなく、それ以外の医療施設を選択されるのは、重度な機能障害による回復が望めないとの判断があったからと考えられる。今回は、患者の重症度を示す看護必要度を特徴量の一つとしてモデルを構築しているが、転帰先が回復期リハビリテーション病棟か、それ以外かの選択においては、重症度以外の特徴量があることが示唆され、今後の検討が必要である。

患者が回復期リハビリテーション病棟へ転院する場合の転院先施設の選択では、医療提供能力も影響すると考え、脳外科または脳神経内科を有する施設であるかそうでないかの分類を行ったが、期待する結果は得られなかった。今回、転院先施設として、急性期治療も実施可能な医療提供体制が整備されている施設への偏りがみられているが、回復期リハビリテーション病棟の転院先施設の選択においても、看護必要度が示す患者の重症度や、転院先の医療提供体制以外の転院先選択要件があることが示唆された。また、本研究の限界としては、1 施設から抽出したデータを扱っている点、患者の居住地が鹿児島保健医療圏に集中していた点があげられる。同一医療圏内における転院が多く、その中で回復期リハビリテーション病棟への転院が選択されていることから、医療圏内における回復期リハビリテーション病棟の医療提供体制の偏りも影響していることが考えられる。今後は地域包括ケアシステムの構築を踏まえ、患者・家族の思いや医療者の判断等にも視点を広げ、どのように後方受け入れ施設の選定が行われているのかを検討していく必要がある。また、医療提供体制以外の転院先施設のデータを分析し検討する必要がある。

## 6. 結論

今後も、地域包括ケアシステムを踏まえた脳梗塞患者の転院先予測モデルの開発および予測精度を高めることを目指したい。本研究による予測精度をさらに向上させることは、入院早期からの転院先の候補を予測することにつながり、患者・家族の選択の幅を広げることができる。さらに、転院先施設との調整が早期から取り込まれることにより、急性期病院における入院期間を短縮でき、患者状態に応じた医療やケアの提供が受けられる適切な支援につながると考える。

## 参考文献

- 1)厚生労働省政策統括官付参事官付保健統計室. 平成 29 年患者調査 (傷病分類編) : 厚生労働省, 2017. [ <https://www.mhlw.go.jp/toukei/saikin/hw/kanja/10syoubu/01/h29syobyu.pdf> (cited 2020-Aug-25)].
- 2)水迫友和,馬場祥吾,岩穴口孝,宇都由美子. 脳梗塞患者の入院期間が長期に及ぶ要因分析. 第 21 回日本医療情報学会看護学術大会論文集, 2020:91-92.
- 3)竹内一郎,鳥山昌幸. サポートベクトルマシン. 講談社. 2019:1-29.
- 4)毛利拓也,北川廣野,澤田千代子,谷一徳. scikit-learn データ分析

- 実装ハンドブック. 秀和システム. 2019:183-201.
- 5)林谷昌洋,大野友嗣,久保雅洋,森口真由美. 電子カルテデータを用いた退院方向性予測による早期退院の実証. 医療情報学会・人工知能学会 AIM 合同研究会資料 SIG-AIMED-004-13,2017. [[https://jsai.ixsq.nii.ac.jp/ej/?action=repository\\_uri&item\\_id=8962&file\\_id=1&file\\_no=1](https://jsai.ixsq.nii.ac.jp/ej/?action=repository_uri&item_id=8962&file_id=1&file_no=1) (cited 2020-Aug-25)].
  - 6)Julian Avila,Trent Hauck.scikit-learn 活用レシピ 80+.インプレス.2019:239-250.
  - 7)國枝洋太,三木敬嗣,今井智也,新田収,星野晴彦. 急性期脳梗塞患者における自宅退院の可否に関連する因子の検討. 理学療法学 第 42 巻第 7 号,2015:562-568.
  - 8)三浦真美,山崎雄斗,矢板沙来美,安納崇之,小林信行. 脳梗塞入院患者の在院日数に関わる要因の実態調査. 日農医誌 67 巻 4 号,2018:492-499.
  - 9)八木麻衣子,川口明子,吉岡亭ら. 急性期病院の脳梗塞患者における退院先に関連する因子の検討-自宅退院群と回復期病院群における検討. 理学療法学 第 39 巻第 1 号,2012:7-13.