

ポスター | 医療データ解析

## ポスター6

### 医療データ解析・医療支援

2021年11月20日(土) 09:00 ~ 10:00 P会場 (イベントホール)

#### [3-P-1-02] 機械学習を用いた新型コロナウイルス感染症重症化予測モデルの検討

\*植松 直哉<sup>1</sup>、真辺 篤<sup>1</sup>、藤村 義明<sup>1</sup>、高橋 則之<sup>1</sup>、尾崎 勝彦<sup>1</sup>、堀内 滋人<sup>2</sup>、日比野 真<sup>2</sup>、亀井 徹正<sup>2</sup>（1. 徳洲会インフォメーションシステム株式会社, 2. 医療法人徳洲会 湘南藤沢徳洲会病院）

\*Naoya Uematsu<sup>1</sup>, Atsushi Manabe<sup>1</sup>, Yoshiaki Fujimura<sup>1</sup>, Noriyuki Takahashi<sup>1</sup>, Katsuhiko Ozaki<sup>1</sup>, Shigeto Horiuchi<sup>2</sup>, Makoto Hibino<sup>2</sup>, Tetsumasa Kamei<sup>2</sup>（1. 徳洲会インフォメーションシステム株式会社, 2. 医療法人徳洲会 湘南藤沢徳洲会病院）

キーワード：Machine Learning, COVID-19, NEWS Score

背景・目的：新型コロナウイルス感染症によるパンデミックが世界中で進行し、重症化感染者数が病床を上回ることにより医療逼迫が続く、自宅療養患者の死亡も報告されている。そこで我々は、早期に重症化を予測することを目的に、状況に応じた2つの重症化予測モデル（24時間以内・24~48時間以内）を検討した。現在徳洲会グループでは、英国早期警告スコア(NEWS)を用いて重症化を予測している。NEWSは、一時点のバイタル情報（呼吸数、脈拍数、収縮期血圧、体温、SpO2、意識レベル、酸素投与あり・なし）をスコアに変換し、重症化予測を行う手法である。NEWSは患者の特性とバイタル情報の変動が考慮されず、早期の重症化予測が困難であった。そこで我々は、予測精度を改善させる目的で、機械学習を用いた機械学習早期警告システム(ML-EWS(COVID-19))の開発を行った。

方法：対象データ：徳洲会グループが保有する徳洲会メディカルデータベース（TMD）から、2020年7月から現在までの2,000名を超える新型コロナ陽性患者情報及びその患者のNEWSで利用されているバイタル情報を取得した。

解析手法：本研究では、重症化発生以前のバイタル情報を3回分利用し、ランダム・フォレスト法を用いた機械学習モデル(ML-EWS(COVID-19))を構築した。また各重症化予測モデルに対して、バイタル情報の各項目の重要度分析も行った。

結果・考察：受信者動作特性分析では、NEWSより ML-EWS(COVID-19)が良い予測精度を示した。24時間以内と24~48時間以内の重症化予測モデルによる重要度分析の詳細については、発表会当日に報告する。今後の課題としては、病名情報等を活用しながら、予測精度の改善を行う予定である。

# 機械学習を用いた新型コロナウイルス感染症重症化予測モデルの検討

植松直哉<sup>\*1</sup>, 真辺篤<sup>\*1</sup>, 藤村義明<sup>\*1</sup>, 高橋則之<sup>\*1</sup>, 尾崎勝彦<sup>\*1</sup>,  
堀内滋人<sup>\*2</sup>, 日比野真<sup>\*2</sup>, 亀井徹正<sup>\*2</sup>

<sup>\*1</sup> 徳洲会インフォメーションシステム株式会社,

<sup>\*2</sup> 医療法人徳洲会 湘南藤沢徳洲会病院

## Examination of a machine learning based severity prediction of patients with COVID-19

Naoya Uematsu<sup>\*1</sup>, Atsushi Manabe<sup>\*1</sup>, Yoshiaki Fujimura<sup>\*1</sup>, Noriyuki Takahashi<sup>\*1</sup>, Katsuhiko Ozaki<sup>\*1</sup>,  
Shigeto Horiuchi<sup>\*2</sup>, Makoto Hibino<sup>\*2</sup>, Tetsumasa Kamei<sup>\*2</sup>

<sup>\*1</sup> Tokushukai Information System Co., Ltd.,

<sup>\*2</sup> Tokushukai Medical Corporation, Shonan Fujisawa Tokushukai Hospital

The pandemic of COVID-19 has been increasing all over the world, and we have some reports that patients who could not be hospitalized have died in hotels. To predict the deterioration of patients with COVID-19 at an early stage, we proposed the following two prediction models, in 12-36 hours and in 24-48 hours. Recently, hospitals in Tokushukai group have predicted the deterioration by using National Early Warning Score (NEWS). NEWS is a scoring method that converts a set of vital signs which are respiratory rate, heart rate, systolic blood pressure, body temperature, SpO<sub>2</sub>, consciousness, and any supplemental oxygen, to a score at a point in time. It has been difficult to predict the deterioration at the early stage because NEWS does not consider the fluctuation of the vital signs information and the characteristic features of the patients.

To improve the prediction accuracy, we propose an ML-EWS (COVID-19) by using a random forest algorithm. For numerical experiments, we used the information of 3,586 patients admitted to 47 hospitals in Tokushukai group between July 2020 and February 2021. By using Receiver Operating Characteristic (ROC) curve, ML-EWS (COVID-19) performed better than NEWS. Moreover, the algorithm analyzed the feature importance for each vital sign, and the model predicting the deterioration in 12-36 hours showed that SpO<sub>2</sub> was highly important, and the other model in 24-48 hours showed that SpO<sub>2</sub>, body temperature, and respiratory rate were relatively important. As future work, we will improve the prediction accuracy by using other information such as chronic conditions.

**Keywords:** Machine Learning, COVID-19, National Early Warning Score

### 1. 背景と目的

新型コロナウイルス感染症によるパンデミックが世界中で進行し、重症化感染者数が病床を上回ることにより医療逼迫が続き、ホテル療養患者や自宅療養患者の死亡も報告されている。そこで我々は、早期に重症化を予測することを目的に、状況に応じた2つの重症化予測モデル(重症化の12-36時間以内・24-48時間以内)を検討した。本研究では、重症化を高流量酸素療法(高流量鼻カニューラ、非侵襲的陽圧換気療法(NPPV)、侵襲的人工呼吸療法、Extracorporeal Membrane Oxygenation (ECMO))を適用した場合と定義した。

現在徳洲会グループでは、英国標準の早期警告スコア National Early Warning Score (NEWS) を用いて新型コロナウイルス陽性患者の重症化予測<sup>1)</sup>を行っている。NEWSは、予期せぬ心停止や死亡等を防ぐ目的に開発され、一時点の Vital Sign (VS)情報(呼吸数、脈拍、収縮期血圧、体温、SpO<sub>2</sub>、意識レベル、酸素投与あり・なし)をスコアに変換して重症化予測を行う手法である<sup>2)</sup>。算出された NEWS スコアが高ければ高いほど、重症化するリスクが高くなる。NEWS は、ある一時点の VS 情報のみで予測が可能であり有用性が高いが、患者の特性と過去の VS 情報が考慮されず、早期の重症化予測が困難であることが課題であった。心停止予測分野では、予測精度を改善させるため、使用状況等に応じてスコア等を変更させた Modified EWS (MEWS)<sup>3)</sup>や深層学習を用いた Deep-learning-based EWS (DEWS)<sup>4)</sup>などの開発研究が盛んに行われている。

本研究では、早期の新型コロナウイルス陽性患者の重症化予測精度を改善する目的で、3回分の VS 情報を用いた時系

列データを利用し、ランダムフォレスト法を用いた機械学習 (Machine Learning) モデル ML-EWS (COVID-19)を提案し、その有効性に関して議論する。(以下 ML-EWS とする。)

### 2. 従来手法 (NEWS)

NEWS は、一時点の VS 情報を表1に従ってスコアに変換する手法である。例えば、患者の一時点での VS 情報:呼吸数:22, SpO<sub>2</sub>:95, 酸素使用:なし, 体温:36.5, 収縮期血圧:105, 脈拍:70, 意識レベル:P は、NEWS スコア 7 (= 2 + 1 + 0 + 0 + 1 + 0 + 3) になる。NEWS スコアが取得可能な値は、0 から 20 であり、一般的に、1-4 点が低リスク、5-6 が中リスク、7 以上が高リスクとして知られている<sup>5)</sup>。

### 3. 提案手法 (ML-EWS)

本研究では、機械学習のアルゴリズムの一つであるランダムフォレスト法<sup>6)</sup>を利用し、ML-EWS を構築した。ランダムフォレスト法は、決定木による弱学習器を複数組み合わせることで汎化能力を向上させる教師あり機械学習(アンサンブル学習アルゴリズム)である。線形関係だけでなく非線形関係も分析できることから、分類・回帰の用途でも利用されている。説明変数が多い場合でも効率的に働くことから、本研究では、説明変数を VS 情報、目的変数を重症化するか否かと設定し利用した。さらに、VS 情報の時系列変動を考慮する目的で、複数の VS 情報を利用した。また、ML-EWS では、分類に寄与した各説明変数の寄与率を算出する重要度分析も可能である。ある説明変数の重要度が 0 に近い場合、分類に無関係な説明変数であると考えることができる。その場合、変数選択 (feature selection)

表 1: National Early Warning Score (意識レベルの A, V, P, U: A: Alert 覚醒, V: 声かけに反応 P: 痛みに反応 U: 無反応)

|                  | 3      | 2      | 1         | 0         | 1         | 2       | 3          |
|------------------|--------|--------|-----------|-----------|-----------|---------|------------|
| 呼吸数              | ≤ 8    |        | 9-11      | 12-20     |           | 21-24   | ≥ 25       |
| SpO <sub>2</sub> | ≤ 91   | 92-93  | 94-95     | ≥ 96      |           |         |            |
| 酸素使用             |        | あり     |           | なし        |           |         |            |
| 体温               | ≤ 35.0 |        | 35.1-36.0 | 36.1-38.0 | 38.1-39.0 | ≥ 39.1  |            |
| 収縮期血圧            | ≤ 90   | 91-100 | 101-110   | 111-219   |           |         | ≥ 220      |
| 脈拍               | ≤ 40   |        | 41-50     | 51-90     | 91-110    | 111-130 | ≥ 131      |
| 意識レベル            |        |        |           | A         |           |         | V, P または U |

を行い、重要度の高い説明変数のみを利用し、再度予測モデルを構築し、予測精度の改善を行うことも可能となる。

#### 4. データ

徳洲会グループが保有する徳洲会メディカルデータベース(TMD)から、徳洲会グループ 47 施設で 2020 年 7 月から 2021 年 2 月までに新型コロナウイルス感染症で入院していた 3,586 件を利用した。そのデータ内容は、重症化 346 件(24 施設)、平均年齢 66.5 (±19.5)、女性割合 44.1%であった。各患者に対して、重症化時刻は、高流量鼻カニューラ、非侵襲的陽圧換気療法(NPPV)、侵襲的人工呼吸療法、ECMO の内、最初に行われた治療の開始時刻とした。本研究では、重症化時刻を取得できた患者を重症化患者と定義し、重症化患者以外を非重症化患者と定義した。

TMD は、各患者の VS 情報を連続値ではなく、各患者のある計測時刻での VS 情報(+拡張期血圧)を保持している。本研究の重症化予測では、データ数と重症化の定義の関係により、意識レベルと酸素使用の項目を利用せず、拡張期血圧を追加して予測を行った。

##### 4.1 誤入力・欠損データ

本研究では正常値の設定を以下の通りに行った。呼吸数: 3-60 breaths/min, 体温: 30-45 度, 収縮期血圧: 30-300 mmHg, 拡張期血圧: 30-300mmHg, 脈拍: 10-300 beats/min とした。今回、SpO<sub>2</sub> は NEWS に変更したスコアを利用したので 1-3 とした。正常値の範囲外の場合、異常値とみなし欠損値と設定した。

前記の通り、VS 情報は、各患者に対して連続値ではなく 1 日数回計測されているが、病院によっては毎回全ての 6 項目の VS 情報を計測しない場合がある。各項目で欠損値かつ前回値がある場合、前回値で欠損値を補完する手続きを時系列の古い情報から繰り返し行った。

#### 5. 数値実験

##### 5.1 教師データと設定

12-36 時間以内モデルの重症化データは、各患者の重症化時刻から遡ること 12 時間と 36 時間の間に計測された連続 3 回の VS 情報を利用した。24-48 時間以内も同様の手続きで重症化データを作成した。非重症化データは、非重症化患者のコロナ陽性確認日から 10 日間に計測された連続 3 回分の VS 情報(時系列の最初と最後の VS 情報の時間差が 3 時間以上 24 時間以内である)を利用した。

比較実験において NEWS と ML-EWS に利用する学習および検証データ数を同数にするために、NEWS は ML-EWS で作成した連続 3 回の VS 情報の最新時刻の VS 情報のみを利用した。

表 2: 教師データ数(括弧内: データ作成に利用した患者数)

|         | 12-36 時間以内     | 24-48 時間以内     |
|---------|----------------|----------------|
| 重症化データ  | 307 (96)       | 287 (85)       |
| 非重症化データ | 42,456 (1,702) | 42,456 (1,702) |

表 3: AUROC 値

|        | 12-36 時間以内  | 24-48 時間以内  |
|--------|-------------|-------------|
| NEWS   | 0.79        | 0.77        |
| ML-EWS | <b>0.88</b> | <b>0.89</b> |

表 4: 感度(括弧内: 特異度) (Youden Index を利用)

|        | 12-36 時間以内         | 24-48 時間以内         |
|--------|--------------------|--------------------|
| NEWS   | 0.73 (0.72)        | 0.71 (0.71)        |
| ML-EWS | <b>0.83 (0.80)</b> | <b>0.86 (0.80)</b> |

連続 3 つの VS 情報を時系列の新しい順に(1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>)とすると、非重症化データの 1<sup>st</sup>と 2<sup>nd</sup>, 2<sup>nd</sup>と 3<sup>rd</sup>の平均時間差は、それぞれ 6.1 時間, 6.1 時間であった。重症化データ(12-36 時間以内モデル)の 1<sup>st</sup>と 2<sup>nd</sup>, 2<sup>nd</sup>と 3<sup>rd</sup>の平均時間差は、それぞれ 3.4 時間, 3.2 時間であった。重症化データ(24-48 時間以内モデル)の 1<sup>st</sup>と 2<sup>nd</sup>, 2<sup>nd</sup>と 3<sup>rd</sup>の平均時間差は、それぞれ 3.7 時間, 4.0 時間であった。

ランダムフォレスト法の構築には、Python のオープンソースの機械学習ライブラリの一つである sklearn の sklearn.ensemble.RandomForestClassifier を利用した。予備実験を踏まえて、ランダムフォレスト法のパラメータは、木の深さ 5, 木の数 300, さらに、非重症化データ数が重症化データ数より多いことを考慮し、class\_weight を“balanced”に設定した。その他は初期設定と同様に設定した。表 2 は、各手法が実験で利用した教師データ数とデータを作成した際に利用した患者数である。

##### 5.2 実験結果

NEWS と ML-EWS の予測精度を確認するため、Area Under the Receiving Operating Curve (AUROC)を用いた。学習した予測モデルがどの程度有用なのかを解析する統計学的方法で、曲線の下部分の面積を用いて、0 から 1 までの値で表現し、1 に近いほど判別性能が高いモデルである。k-分割交差検証(cross-validation)(k=3) <sup>7)</sup>を利用した平均パフォーマンス値の結果を示した。(結果の算出方法の詳細は、参考文献を参照。)実験では、重症化を陽性、非重症化を陰性とした。表 3 より 12-36 時間以内重症化予測での NEWS, ML-EWS の各 AUROC 値は、0.79, 0.88 であった。24-48 時間以内重症化予測では、NEWS, ML-EWS の各 AUROC 値は、0.77, 0.89 であった。AUROC の指標を用いた場合、ML-EWS は NEWS より良い結果を示したことを確認した。さらに、

Youden Index を利用した場合の各モデルの平均の感度と特異度の結果を表 4 で示した. ML-EWS が NEWS より感度・特異度の両方で良い結果を示したことから, 提案手法がより効率的な重症化予測が可能であることを確認した.

次に, ML-EWS による重要度分析を確認した. 図 1 は, 12-36 時間以内と 24-48 時間以内重症化予測での重要度分析の結果を示している. 12-36 時間以内重症化予測では, SpO<sub>2</sub> の重要度が高く, 24-48 時間以内重症化予測では, SpO<sub>2</sub>, 及び体温(BT)と呼吸数(RR)の重要度が比較的高い結果を示した.

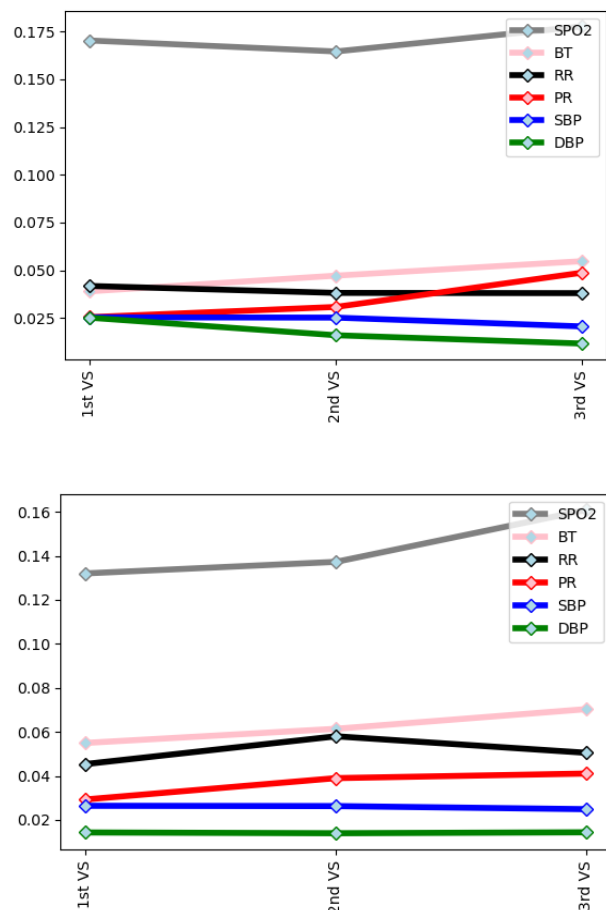


図 1: ML-EWS による重要度分析  
(上: 12-36 時間以内モデル, 下: 24-48 時間以内モデル)

## 6. 考察と今後の課題

AUROC 値及び感度・特異度の結果より, ML-EWS が NEWS より効率的に重症化予測可能であることを確認した. 重要度分析により, 特に 24-48 時間以内の予測では SpO<sub>2</sub>, 体温, 呼吸数が比較的重要であるという知見を得た. また, 変数重要度の低い項目を観察することも重要になってくる. 今回の結果では, 拡張期血圧(DBP)と収縮期血圧(SBP)の重要度が, 低い結果を得たことから, その 2 つの項目は利用しないような変数選択を行うことで, より汎用的で精度の良いモデル構築も検討可能である.

追加実験として, SpO<sub>2</sub> の重要度が非常に高い結果から, ML-EWS を少し変更した, SpO<sub>2</sub> のみを用いた予測モデル ML-EWS(SpO<sub>2</sub>)を構築した. AUROC 値では, ML-EWS は

ML-EWS(SpO<sub>2</sub>)より良い結果を示したことを確認した.

ML-EWS を病院内でのスクリーニング用途で利用するためには, より良い感度と特異度が必要になる. 今後の課題として, ML-EWS の感度と特異度を同時に改善する必要があり, 患者の年齢, 病名情報等も追加利用することで, 予測精度を改善する余地がある. より多くの情報(血液検査結果等)を利用する予測モデル<sup>8,9)</sup>は, 予測精度が改善されると報告されているが, 同時にモデルの汎用性を失うことにも繋がる. 汎用性を失わないモデルの設定を行う必要があると考える.

過去のデータを用いて作成した重症化予測モデルが, 新しく出現する新型コロナウイルスの変異株の陽性患者に対して有効か否かも検証する必要がある. 第 4 波(2021 年 3 月からとされる)では, 変異株が流行しており, その感染力は極めて高く, 重症化する可能性も高いという報告がされている. 本研究で作成した 12-36 時間以内・24-48 時間以内重症化予測モデルが, 第 4 波の新型コロナウイルス陽性者の重症化予測に適用可能か否かを調査する予定である.

## 7. 結論

NEWS の早期重症化予測精度を改善するために, VS 情報の時系列を考慮した機械学習早期重症化予測モデル ML-EWS を提案した. 本研究のデータセットを利用した数値実験の結果では, 感度・特異度と共に, ML-EWS が NEWS より良い予測結果を示した. ML-EWS の重要度分析によって, 予測を行う際に, SpO<sub>2</sub>, 及び体温と呼吸数が比較的重要であるという知見を得た. 今後の課題となる感度・特異度をより改善する方法については, 検討を続けていく予定である.

## 参考文献

- 1) Baker F. K, Hanrath T. A, Loeff S. I, Kay J. L, Back J, Duncan JA. C. National Early Warning Score 2 (NEWS2) to identify inpatient COVID-19 deterioration: a retrospective analysis. *Clinical medicine (London, England)* 2021; 21(2): 84-89.
- 2) Smith B. G, Prytherch R. D, Meredith P, Schmidt E. P, Featherstone I. P. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. *Resuscitation* 2013; 84(4): 465-470.
- 3) Gardner-Thorpe J, Love N, Wrightson J, Walsh S, Keeling N. The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. *Annals of the Royal College of Surgeons of England*. 2006; 88(6): 571-575.
- 4) Kwon M. J, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest. *Journal of the American Heart Association* 2018; 7(13).
- 5) Lee S. Y, Choi W. J, Park H. Y, et al. Evaluation of the efficacy of the National Early Warning Score in predicting in-hospital mortality via the risk stratification. *Journal of Critical Care* 2018; 47: 222-226.
- 6) Breiman L. Random Forests. *Machine Learning*, 2001; 45: 5-32.
- 7) Fedotenkova M. Extraction of multivariate components in brain signals obtained during general anesthesia [dissertation], Univ. of Lorraine 2016.
- 8) Arvind V, Kim S. J, Cho, H. B, Geng E, Cho K. S. Development of a machine learning algorithm to predict intubation among hospitalized patients with COVID-19. *Journal of Critical Care* 2021; 62 :25-30.
- 9) Youssef A, Kouchaki S, Shamout F, et al. Development and validation of early warning score systems for COVID-19 patients. *Healthcare Technology Letters* 2021, 1-12.