

電子カルテデータ解析における欠測値補完法の一検討

多田 亜彩美^{†1,a)} Le Hieu Hanh^{†1,b)}

概要：電子カルテデータの二次利用への期待が高まる一方で、それらのデータには欠測が含まれている場合がある。そのようなデータを用いて信頼性の高い解析を行うためには、欠測に対する適切な処理が必要である。欠測を含むデータへのアプローチの一つに、欠測値の補完を行うという手法がある。よく知られた補完手法として LOCF (Last Observation Carried Forward) 法や平均値代入法などの単変量代入法があり、適用が簡単であるため、多くの研究で用いられてきた。本研究では、単変量代入法について、補完値から判断される異常性に着目し、その適用可能割合も含めて評価した。結果は、欠測箇所より前・後いずれの値も考慮した平均値代入法と線形補間法の精度が高く、適用可能割合は平均値代入法が最も高かった。

An Evaluation of Methods for Missing Data Imputation in Electronic Medical Records

1. はじめに

1.1 研究の背景

様々な分野でデジタル・トランスフォーメーション (DX) の必要性が訴えられている。医療分野では厚生労働省が「医療 DX」を掲げ、電子カルテデータの共有や標準化、医療等情報の二次利用に向けた検討などに取り組んでいる [1]。

電子カルテデータの二次利用への期待が高まる一方で、それらのデータには欠測が含まれている場合がある。一例として、米国で収集された入院患者のデータベースである MIMIC-IV[2], [3], [4] において検体検査結果のデータを調べると、全体では1割程度の欠測が発生している。そのようなデータを用いて信頼性の高い解析を行うためには、欠測に対する適切な処理が必要である。

欠測を含むデータへのアプローチの一つに、「欠測値に何らかの値を代入することで『あたかも欠測値がなかったかのような』疑似完全データを作成する“代入法”」があげられる [5]。代入法はさらに、Jazayeri ら [6] が述べている通り、単変量の場合と多変量の場合の2つに分けられる。

単変量代入法は、欠測箇所を持つ変量と同じ変量のみを考慮して補完する方法であり、LOCF (Last Observation Carried Forward) 法や平均値代入法などがある。適用が

簡単であり、多くの研究で用いられてきた。

また多変量代入法は、欠測箇所を持つ変量に加え他の変量も考慮して補完する方法であり、様々な手法が提案されている。電子カルテデータにおける補完手法としては、兵頭ら [7] による、蛋白分画検査と患者情報を用いた機械学習的な手法による血液検査の補完や、Xu ら [8] による、時間情報に加え複数の検査間の情報を特徴量に組み入れた機械学習による血液検査の補完、Jazayeri ら [6] による、複数の検査結果をもとにした患者の類似性を考慮した補完などが提案されている。

1.2 本研究の目的

単変量代入法は、多変量代入法に比べ、適用が簡単であるという特徴がある。先に述べた電子カルテデータにおける各種の多変量代入法では、複数の検査結果を変量として用いることに対し、単変量代入法では、欠測箇所のある1つの変量のみを考慮するだけでよい。このような適用の容易さから、単変量代入法は多くの研究で用いられてきたと考えられる。例えば Tanaka ら [9] が2013年に行った調査によると、日本の製薬メーカーでの臨床試験において、欠測処理のために最もよく用いられた統計的手法は LOCF 法であった。

また電子カルテデータにおいて、先に例として挙げた検体検査結果は、「病名の診断、次に行う医療行為の決定、必要に応じて行う次の検査の決定時において、患者の状態を

^{†1} 現在、お茶の水女子大学
Presently with Ochanomizu University
^{a)} k2500004@edu.cc.ocha.ac.jp
^{b)} le@is.ocha.ac.jp

知るために極めて重要な情報」となる [10]。検査結果が異常を示せば、その後の医療行為等に影響を及ぼすと考えられる。したがって、電子カルテデータ解析において欠測箇所を補完する場合には、その補完値から判断される異常性が重要となる。

以上を踏まえて本研究では、従来から用いられている適用が簡単な単変量代入法について、補完値から判断される異常性に着目し、その適用可能割合も含めて評価することとした。具体的には、従来の手法をもとに6つの手法を検討し、実際の電子カルテデータである MIMIC-IV へ各手法を適用して、手法の適用可能割合と、再現率・適合率といった精度を評価した。

1.3 本稿の構成

本稿では、第2節にて関連研究について述べる。次に第3節にて本研究で検討した補完手法について説明する。第4節にて評価実験に用いたデータセット、評価対象として用いた手法、評価方法とその結果について述べる。最後に第5節にてまとめを行う。

2. 関連研究

2.1 単変量代入法

単変量代入法には次のようなものがある。

LOCF 法は、ある変数において、観測された最後の値を欠測箇所に代入する手法である。

平均値代入法は、ある変数において、他の観測値から計算された平均値を欠測箇所に代入する手法である。

線形補間法は、ある変数において、欠測箇所の前後の観測値2点を通る直線を求め、欠測箇所の値は、その前後の観測値2点で作る線分上にあるものとして、直線の式から求めた値を欠測箇所に代入する手法である。

また、一般的には多変量代入法に分類されるが、回帰代入法という手法がある。関連する何らかの観測値を説明変数として回帰モデルを構築し、モデルから予測された値を欠測箇所に代入する手法である。一方、補完手法に限らず、時系列データに対しては、線形回帰モデルを用いて、過去の値を説明変数として目的変数である未来の値を予測する分析手法が用いられることがある。この手法で補完を行う場合には、欠測箇所を持つ1つの変数のみを考慮するため、単変量代入法と言える。

2.2 電子カルテデータにおける多変量代入法

多変量代入法について、電子カルテデータに関する補完手法としては、次のようなものが提案されている。

兵頭ら [7] は、血液検査結果を補完するため、蛋白分画検査という異なる検査結果と患者情報を用いた補完手法を提案した。

Luo[11] は、自身がまとめ役となった、時系列を持つ臨床

データの欠測補完の試みに関する共有タスク [12] での成果を報告している。人工的な欠測を含む13種の血液検査結果に対し、12のチームが様々な手法で欠測補完を試みた。

その中で総合して最も高い精度を達成したのが、Xuら [8] が提案した手法であった。時間情報に加え、13種の血液検査間の情報から構築される特徴量を用いた機械学習による補完手法を提案した。

また Jazayeri ら [6] は、13種の血液検査結果をもとに類似性を考慮した他の患者の重み付けを行い、他の患者の検査結果値から加重平均により補完する手法を提案した。

このように電子カルテデータのための様々な新しい補完手法が提案されている。一方、1.2節にて述べた通り、単変量代入法は適用が簡単で多くの研究で用いられてきた。そのような背景を踏まえ、本研究では、単変量代入法に着目しその精度等を評価することとした。

3. 検討手法

本研究では、次の6つの補完手法の検討を行った。2.1節にて紹介した4つの手法と、電子カルテデータの特徴として時系列を持つデータであることがあげられるが、平均値代入法については時系列を考慮した2つの手法を加え、合計6つの手法を検討した。各手法の詳細を以下に述べる。

3.1 LOCF 法

患者ごとにデータをグループ化し、欠測箇所を持つグループに対して、欠測箇所より前の値のうち最も新しい検査結果値を欠測箇所に代入した。欠測箇所が2つ以上連続する場合は、連続して同一の値を代入した。

3.2 欠測箇所より前の値のみを用いた平均値代入法

患者ごとにデータをグループ化し、欠測箇所を持つグループに対して、欠測箇所より前の値をすべて用いて平均値を求め、欠測箇所に代入した。

3.3 欠測箇所より前の値のみを用いた重み付き平均値代入法

患者ごとにデータをグループ化し、欠測箇所を持つグループに対して、値の時間的な重みを算出し、重み付き平均値を求め、欠測箇所に代入した。

重みの算出方法は次の通りである。

r 行目の欠測箇所に対し、当該行の時間情報 $time_r$ と、欠測箇所より i 個前の $r-i$ 行目の時間情報 $time_{r-i}$ との差を秒単位で求めた。これをパラメータ t で割り小数点以下を切り上げ、逆数を取ることで、 r 行目の欠測箇所に対して i 行前の値の時間的な重み w_i を求めた。導出式は以下の通りである。

$$w_i = \frac{1}{\lceil \frac{time_r - time_{r-i}}{t} \rceil}$$

表 1 パラメータ t

Table 1 Parameter t

hour	$t = 60 \times 60$
day	$t = 60 \times 60 \times 24$

表 2 パラメータ d

Table 2 Parameter d

回数 1	$d = 1$
回数 3	$d = 3$
回数 5	$d = 5$
回数 10	$d = 10$

パラメータ t は、表 1 の通り、重みを時間単位で求める場合 *hour* と、日単位で求める場合 *day* に設定した。

次に、欠測箇所より前の各値と、上記の通り求めた各値の重みを用いて、重み付き平均値を求め、欠測箇所に代入した。

3.4 線形回帰代入法

ある時点での検査結果値を目的変数 y とし、その検査結果値より d 行前の値を説明変数 x_d とし、以下のような線形回帰モデルを構築した。

$$y = a_0 + a_1x_1 + \dots + a_dx_d$$

使用する行数をパラメータ回数 d として、表 2 の通り、1 行前の値を用いた回数 1 の場合、1~3 行前の値を用いた回数 3 の場合、1~5 行前の値を用いた回数 5 の場合、1~10 行前の値を用いた回数 10 の場合を検討した。

3.5 当該患者が持つ値をすべて用いた平均値代入法

患者ごとにデータをグループ化し、欠測箇所を持つグループに対して、グループ内のすべての値を用いて平均値を求め、欠測箇所に代入した。

3.6 線形補間法

患者ごとにデータをグループ化し、欠測箇所を持つグループに対して、グループごとに手法を適用した。欠測箇所の前後の値 2 点を通る直線を求め、欠測箇所の値は 2 点で作る線分上にあるものとして、直線の式から求めた値を欠測箇所に代入した。

欠測箇所の前後の値 2 点および欠測箇所の値の点は、線上に等間隔に並んでいるものとした。欠測箇所が 2 つ以上連続する場合は、一連の欠測箇所の前後の値 2 点を用いて、その 2 点で作る線分上に欠測箇所が等間隔に並んでいるものとして補完値を求めた。欠測箇所がグループの最終行にあるなど、欠測箇所の後の値が存在しない場合は、前の値をそのまま代入することとした。

4. 評価実験

本節では、評価実験に用いたデータセットについて説明し、評価対象として用いた手法、評価方法とその結果について述べる。

表 3 実験に用いたデータセットの各列の概要

Table 3 Columns Used in Dataset

列名	説明
labevent_id	テーブル内の各行にユニークな id
subject_id	患者ごとにユニークな id
itemid	検査項目ごとにユニークな id
storetime	検査結果が利用可能になった時間
valuenum	検査結果を示す数値データ
valueuom	検査結果値に対する単位
ref_range_lower	検査結果の基準値における下限
ref_range_upper	検査結果の基準値における上限
flag	検査結果の異常を示す文字列
comments	匿名化された自由記述のコメント。 結果やその解釈に関する注意や、 場合によっては検査結果そのものが コメントに含まれる場合がある。

4.1 データセット

4.1.1 MIMIC-IV の概要

本研究では、米国で収集された実際の入院患者のデータベースである MIMIC-IV より、検体検査結果を記録した labevents テーブルを用いた。labevents テーブルが持つ列のうち、実験に用いたデータセットの各列の概要は表 3 の通りである。表 3 は、MIMIC documentation[13] をもとに筆者が作成した。

検査項目ごとにユニークな itemid を用いて、目的の検査項目に関する結果を絞り込むことができる。また患者ごとにユニークな subject_id を保持しており、検査が行われた患者が識別できる。時間情報として、検査結果が利用可能になった時間 storetime を保持している。検査結果を示す数値データは valuenum 列に、その結果値に対する単位は valueuom 列に格納されている。valuenum が欠測している場合があり、それらを本実験での補完対象とした。

ref_range_lower 列、ref_range_upper 列には、当該の検査項目の基準範囲の下限値と上限値が格納されている。検査結果が異常を示す場合は、flag 列に abnormal の文字列が記録されている。異常がない場合は NULL である。また自由記述のコメントを持つ comments 列が存在する。

4.1.2 使用検査項目

labevents テーブルには様々な種類の検体検査の結果が格納されているが、本実験では、C 反応性蛋白とグルコースの検査結果を用いた。C 反応性蛋白は、感染症などの炎症性疾患で増加する物質で、炎症の活動性の指標となる検査項目である [14]。グルコースは、糖尿病や、その他高血糖・低血糖を示す疾患の診断等に用いられる検査項目である [14]。

4.1.3 正解セット

flag 列に記録される異常について、検査結果値を格納する valuenum 列が NULL である場合にも flag 列に abnormal

表 4 C 反応性蛋白 comments 列の文字列
Table 4 Comments of CRP

内容	個数	割合
>300*.	1,558	41.33%
GREATER THAN 300.	740	19.63%
GREATER THAN 300. LOW RISK <1.0, AVERAGE RISK 1.0-3.0, HIGH RISK >3.0 (BUT <10).	725	19.23%
---	196	5.20%
>300.	125	3.32%
<0.1.	116	3.08%
>300*. Verified By Replicate Analysis.	61	1.62%
>300. LOW RISK <1.0, AVERAGE RISK 1.0-3.0, HIGH RISK >3.0 (BUT <10).	61	1.62%
<0.1. LOW RISK <1.0, AVERAGE RISK 1.0-3.0, HIGH RISK >3.0 (BUT <10).	48	1.27%
<0.1. Low risk <1.0, Average risk 1.0-3.0, High risk >3.0 (but <10).	41	1.09%
GREATER THAN 300. VERIFIED BY REPLICATE ANALYSIS. LOW RISK <1.0, AVERAGE RISK 1.0-3.0, HIGH RISK >3.0 (BUT <10).	33	0.88%
GREATER THAN 300*. LOW RISK <1.0, AVERAGE RISK 1.0-3.0, HIGH RISK >3.0 (BUT <10).	26	0.69%
GREATER THAN 300. VERIFIED BY REPLICATE ANALYSIS.	16	0.42%
GREATER THAN 300.0. LOW RISK <1.0, AVERAGE RISK 1.0-3.0, HIGH RISK >3.0 (BUT <10).	12	0.32%
GREATER THAN 300. VERIFIED BY DILUTION. LOW RISK <1.0, AVERAGE RISK 1.0-3.0, HIGH RISK >3.0 (BUT <10).	11	0.29%
NULL	1	0.03%
合計	3,770	100.00%

と記録されている場合があった。詳細を確認するため、valuenum 列が NULL である場合の comments を抽出すると、C 反応性蛋白およびグルコースそれぞれについて、表 4, 表 5 のような文字列が comments 列に記録されていることが分かった。

表 3 に示した通り、comments 列には、結果やその解釈に関する注意や、場合によっては検査結果そのものがコメントに含まれる場合がある。C 反応性蛋白の基準値は 0～5mg/L であるが、表 4 に示す comments の内容からは、その基準値を大きく超える結果が推測される場合がある。同様に、グルコースの基準値は、患者により異なる場合があるが、概ね 70～100mg/dL であり、表 5 に示す comments の内容からは、その基準値を大きく超える結果が推測さ

表 5 グルコース comments 列の文字列
Table 5 Comments of Glucose

内容	個数	割合
---	965	52.70%
GREATER THAN 500. VERIFIED.	533	29.11%
GREATER THAN 500.	182	9.94%
>500. VERIFIED.	81	4.42%
>500.	25	1.37%
GREATER THAN 500.		
QNS TO VERIFY.	15	0.82%
GREATER THAN 800. VERIFIED.	13	0.71%
GREATER THAN 500*.	11	0.60%
NULL	6	0.33%
合計	1,831	100.00%

表 6 データセットの概要
Table 6 Outline of Dataset

検査項目名	全データ数	欠測数	欠測率
C 反応性蛋白	177,842	3,573	2.009%
グルコース	3,903,505	860	0.022%

れる。

これらの状況から、何らかの理由により valuenum 列に記録されるべき数値としての結果は得られず NULL となっているが、異常を示唆する結果から、flag 列に abnormal が記録される場合があることが推測された。そこで本実験では、この flag 列を正解セットとして用いることとした。評価方法の詳細については、4.3 節で述べる。

なお、comments 列の記述より結果が基準値範囲外であることが読み取れるが、flag 列には何も記入されていない場合もあった。そのような場合は abnormal として扱うこととした。

また、完全に匿名化されたコメントは 3 つのアンダースコア「___」で表され、コメントが記入されなかった場合は NULL となる [13]。そのような場合は comments 列から異常の有無を読み取ることができないため、valuenum が NULL である行のうち、comments が「___」または NULL の場合は、データセットから除外した。

4.1.4 データセットの抽出と前処理

以上の検討を踏まえ、次のようにデータセットの抽出と前処理を行った。

まず labevents テーブルに対し、C 反応性蛋白またはグルコースに該当する itemid で行を絞り込んだデータをそれぞれ抽出した。多くの場合、一人の患者に対し複数の検査結果が記録されており、検査結果が利用可能になった時間を示す storetime 列から時間の情報を得ることができる。患者ごとに検査結果を時系列順に並べるため、subject_id と storetime によりデータを昇順に並べた。

次に、正解セットを適切に構築するため、valuenum が NULL である行のうち、comments が「___」または NULL

の場合は、データセットから除外した。

以上の通り用意した、C 反応性蛋白およびグルコースのデータセットの概要は表 6 の通りである。

4.2 検討手法の一覧

本研究で検討した補完手法は以下の通りとなる。

- (1) LOCF 法
- (2) 欠測箇所より前の値のみを用いた平均値代入法 (AVG 法)
- (3) 欠測箇所より前の値のみを用いた重み付き平均値代入法 (wAVG 法) *hour, day*
- (4) 線形回帰代入法 (LR 法) 次数 1, 次数 3, 次数 5, 次数 10
- (5) 当該患者が持つ値をすべて用いた平均値代入法 (allAVG 法)
- (6) 線形補間法 (LI 法)

(2)～(6) の手法については、これ以降、括弧内に記載した略称を用いる。また時系列の観点では、(1)～(4) は欠測箇所より前の値を考慮した補完手法であり、(5)、(6) は欠測箇所の前・後いずれの値も考慮した補完手法である。

4.3 評価方法

各手法の適用可能割合を評価値として求めた。例えば (1)LOCF 法であれば、欠測箇所が先頭行にある場合などは、手法を適用することができない。表 6 で示した欠測数に対し、各手法が適用できた割合を求めた。

また、実際の治療においては、検査結果の異常の有無により次の処置等が変わる。そこで本実験では、検査結果の異常性に着目し、補完値に基づき判定した異常の再現率・適合率・F 値を求めた。

4.1.3 節に示した通り、本実験では flag 列を正解セットとして用いることとした。各手法適用後、補完値と、*ref_range_lower* 列、*ref_range_upper* 列に格納された基準値を比較し、基準値範囲外であれば、新たに作成した列に異常を示す *abnormal* を記録した。このように判定した補完値に基づく flag と、オリジナルの flag とを比較し、再現率・適合率・F 値を求めた。

また、(4)LR 法については、患者ごとに、欠測箇所より前の値を *d* 個以上連続して持つものからデータセットを作成した。そのうち 90 % を学習用データセット、10 % をテスト用データセットとしてランダムに分割し、学習用データセットを用いて線形回帰モデルを構築した。評価値として、テスト用データセットを用いて、モデルから得られる予測値との平均二乗誤差を求めた。

4.4 実験環境

本実験は、Python (バージョン 3.12.3) を用いて行った。またデータ処理のためのライブラリとして、pandas (バー

表 7 C 反応性蛋白 適用結果

Table 7 Results of CRP

手法	パラメータ	適用割合	再現率	適合率	F 値
(1)LOCF	-	52.8%	0.889	0.990	0.937
(2)AVG	-	52.8%	0.937	0.973	0.955
(3)wAVG	<i>hour</i>	52.8%	0.939	0.978	0.958
	<i>day</i>	52.8%	0.939	0.977	0.957
(4)LR	次数 1	37.0%	1.000	0.913	0.955
	次数 3	18.3%	1.000	0.902	0.949
	次数 5	11.0%	1.000	0.881	0.937
	次数 10	4.9%	1.000	0.834	0.910
(5)allAVG	-	82.5%	0.963	0.981	0.972
(6)LI	-	52.8%	0.960	0.989	0.974

表 8 グルコース 適用結果

Table 8 Results of Glucose

手法	パラメータ	適用割合	再現率	適合率	F 値
(1)LOCF	-	85.8%	0.921	1.000	0.959
(2)AVG	-	85.8%	0.976	1.000	0.988
(3)wAVG	<i>hour</i>	85.8%	0.977	1.000	0.988
	<i>day</i>	85.8%	0.978	1.000	0.989
(4)LR	次数 1	79.0%	0.957	1.000	0.978
	次数 3	58.1%	1.000	1.000	1.000
	次数 5	53.5%	1.000	1.000	1.000
	次数 10	44.8%	1.000	1.000	1.000
(5)allAVG	-	99.7%	0.994	1.000	0.997
(6)LI	-	85.8%	0.991	1.000	0.995

表 9 C 反応性蛋白
線形回帰モデルの評価

Table 9 Evaluation of
Linear Regression
Model of CRP

パラメータ	平均二乗誤差
次数 1	1,859.2
次数 3	1,484.8
次数 5	1,309.7
次数 10	1,109.6

表 10 グルコース
線形回帰モデルの評価

Table 10 Evaluation of
Linear Regression
Model of Glucose

パラメータ	平均二乗誤差
次数 1	5,442,404.9
次数 3	4,556.8
次数 5	4,447.6
次数 10	444,621.4

ジョン 2.2.3), numpy (バージョン 2.2.6) を、線形回帰モデル構築のためのライブラリとして scikit-learn (バージョン 1.7.0) を用いた。

4.5 実験結果

C 反応性蛋白への各手法の適用結果は表 7 に、グルコースへの各手法の適用結果は表 8 に示すとおりである。各評価項目で最も高い値を示したものは太字にし、2 番目に高い値を示したものには下線を引いた。

なお、グルコースの適用結果において適合率がすべて 1.000 となっている。これは各手法が適用できた欠測箇所のオリジナルの flag 列の値がすべて *abnormal* であったため、補完値から *abnormal* と予測したものはすべて正解と

なったからである。

(4)LR 法における平均二乗誤差による線形回帰モデルの評価結果は、表 9、表 10 に示すとおりである。

表 7、表 8 より、各手法の適用可能割合をみると、(5)allAVG 法が最も高い割合で適用することができた。欠測箇所より前に値が存在している必要がある他の手法と異なり、(5)allAVG 法は、当該患者の持つ値がすべて NULL でない限り適用可能である。このような手法の特徴から、(5)allAVG 法が最も適用可能割合が高くなった。

次に F 値をみると、いずれの手法でも 1.000 に近い結果が出た。(4)LR 法以外で比較すると、C 反応性蛋白では (5)allAVG 法と (6)LI 法が 0.97 を超えた。同様にグルコースでも (5)allAVG 法と (6)LI 法では 0.99 を超え、中でも比較的よい結果となった。

実際のデータを確認すると、異常が続いていたが一転して正常となる、あるいは正常が続いていたが一転して異常となる等、欠測箇所を境に正常・異常の傾向が変化している患者が一定数存在することが分かった。(5)allAVG 法、(6)LI 法では、欠測箇所より前・後いずれの値も用いて補完値を計算しているため、このような場合にも比較的正確に異常を判定することができたと考えられる。

表 9、表 10 より、(4)LR 法のモデルの評価をみると、いずれも予測値の誤差が非常に大きい結果となった。またグルコースにおける (4)LR 法の次数 1 の場合を除き、再現率が 1.000 となった。これはモデルによる予測値がほぼすべて abnormal と判定されたためである。適合率を表 7 の C 反応性蛋白の結果から確認すると、(4)LR 法は他に比べて低い値となっており、実際には異常と記録されなかったものについても abnormal と判定している割合が高いと分かる。(4)LR 法については、モデルの選択や、学習に用いることができたデータセットに偏りがいないか等、今後の課題として検討を行う。

5. おわりに

5.1 まとめ

本研究では、単変量代入による補完手法について、補完値から判断される異常性に着目し、各手法の適用可能割合、再現率・適合率・F 値に関する精度を評価した。結果は、欠測箇所より前・後いずれの値も考慮する (5)allAVG 法、(6)LI 法の精度が高く、適用可能割合は (5)allAVG 法が最も高いことが分かった、

5.2 今後の課題

今後の課題としては、以下のようなものがある。

(4)LR 法についてはモデル選択が適切であったか等検討を行う。また (3)wAVG 法以外の手法では、時間間隔まで考慮しなかったため、今後の課題とする。さらに、今回は C 反応性蛋白とグルコースの 2 つへ手法を適用したが、検

査項目により時系列的な傾向が異なる可能性もある。他の検査項目にも手法を適用し、評価を行う。また、今回の評価では検査結果の異常性に着目したが、今後、補完値の絶対値に対する評価も行う。

参考文献

- [1] 厚生労働省：医療 DX について，<https://www.mhlw.go.jp/stf/iryoudx.html>. 2025.7.29 アクセス.
- [2] Johnson, A., Bulgarelli, L., Pollard, T., Gow, B., Moody, B., Horng, S., Celi, L. A. and Mark, R.: MIMIC-IV (version 3.1), PhysioNet (2024).
- [3] Johnson, A. E. W., Bulgarelli, L., Shen, L., Gayles, A., Shammout, A., Horng, S., Pollard, T. J., Hao, S., Moody, B., Gow, B., Lehman, L. H., Celi, L. A. and Mark, R. G.: MIMIC-IV, A Freely Accessible Electronic Health Record Dataset, *Scientific Data*, Vol. 10, No. 1 (2023).
- [4] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., Mietus, J. E., Moody, G. B., Peng, C. K. and Stanley, H. E.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals., *Circulation [Online]*, Vol. 101, No. 23, pp. e215–e220 (2000).
- [5] 高井啓二，星野崇宏，野間久史：欠測データの統計科学—医学と社会科学への応用，岩波書店 (2016).
- [6] Jazayeri, A., Liang, O. S. and Yang, C. C.: Imputation of Missing Data in Electronic Health Records Based on Patients' Similarities, *Journal of Healthcare Informatics Research*, Vol. 4, pp. 295–307 (2020).
- [7] 兵頭勇己，久原太助，檜山麻里子，安井繁宏，畠山 豊，奥原義保：蛋白分画検査の波形情報を使用した血液検査の欠測値補間の試み，日本医療情報学会 第 41 回医療情報学連合大会（第 22 回日本医療情報学会学術大会），No. 3-G-1-01 (2021).
- [8] Xu, X., Liu, X., Kang, Y., Xu, X., Wang, J., Sun, Y., Chen, Q., Jia, X., Ma, X., Meng, X., Li, X. and Xie, G.: A Multi-directional Approach for Missing Value Estimation in Multivariate Time Series Clinical Data, *Journal of Healthcare Informatics Research*, Vol. 4, pp. 365–382 (2020).
- [9] Tanaka, S., Fukinbara, S., Tsuchiya, S., Suganami, H. and Ito, Y. M.: Current Practice in Japan for the Prevention and Treatment of Missing Data in Confirmatory Clinical Trials: A Survey of Japanese and Foreign Pharmaceutical Manufacturers, *Therapeutic Innovation & Regulatory Science*, Vol. 48, pp. 717–723 (2014).
- [10] 横田治夫：電子カルテデータ解析 医療支援のためのエビデンス・ベースド・アプローチ，共立出版 (2022).
- [11] Luo, Y.: Evaluating the state of the art in missing data imputation for clinical data, *Briefings in Bioinformatics*, Vol. 23, No. 1, pp. 1–9 (2022).
- [12] ICHI2019: IEEE ICHI Data Analytics Challenge on Missing data Imputation (DACMI), <https://ewh.ieee.org/conf/ichi/2019/challenge.html>. 2025.7.31 アクセス.
- [13] MIMICdocumentation: labevents, <https://mimic.mit.edu/docs/iv/modules/hosp/labevents/>. 2025.7.29 アクセス.
- [14] 櫻林郁之介，熊坂一成（監修）：最新 臨床検査項目辞典，医歯薬出版 (2008).