

一般口演 | 医療データ解析

一般口演7

医療データ解析

2019年11月22日(金) 14:50 ～ 16:20 D会場 (国際会議場 2階中会議室201)

[2-D-2-01] ランダムフォレストを用いた潜在的臨床判断値の探索

○松尾 亮輔¹、小川 泰右¹、山崎 友義¹、荒木 賢二¹、外山 比南子²（1. 宮崎大学医学部附属病院IR部, 2. 医療データサイエンス研究所）

キーワード：Latent Clinical Decision Limits, Random Forests, Knowledge Discovery

【背景】臨床検査値には、特定の病態を考慮しない健常者における基準範囲と、特定の病態に関して、その診断・予防や治療・予後の判定基準のための臨床判断値が定められている。説明変数の目的変数に対する重要度が得られるランダムフォレストが特徴選択に広く用いられているが、説明変数の潜在的な判断値を探索する手法があれば、膨大な臨床検査データを用いた検査値に関する新たな臨床知識の発見が期待される。

【目的】ランダムフォレストの分岐情報を活用して、患者の病態に基づいた潜在的な臨床検査の判断値を探索するための方法を提案する。

【方法】始めに、ある病態の患者群内で検査項目の基準範囲を推定し、検査項目ごとに推定した基準範囲内の症例を抽出する。次に、抽出した症例を用いて、複数の検査項目を説明変数として予測モデルをランダムフォレストにより構築する。最後に、症例抽出に用いた特定の検査項目が重要度ランキングで上位の場合に、ランダムフォレストの分岐情報を集約して該当検査値と目的変数の関係性を可視化する。

【結果】宮崎大学医学部附属病院の2014年4月から2018年3月の検査データを用いた。DPCコードの6桁と手術あり、化学療法の有無に基づくDPCクラスタごとで、特定の検査項目の推定基準範囲内の症例に対して、入院時の複数の検査項目の値を用いて、在院日数の予測モデルをランダムフォレストにより構築した。可視化結果から、症例抽出に用いた血小板数、平均赤血球容積、平均血小板容積の3項目において、定義上の基準範囲内で、ある閾値を超えた場合または以下の場合に総じて在院日数が長くなるDPCクラスタが3ケースみられた。

【結論】本研究は潜在的に臨床上有用な検査の判断値を探索するための方法を提案した。本実験により、在院日数に影響を及ぼす潜在的な判断値の候補が得られたことから、提案手法が探索的に潜在的臨床判断値を発見できることが示唆された。

ランダムフォレストを用いた潜在的臨床判断値の探索

松尾亮輔^{*1}、小川泰右^{*1}、山崎友義^{*1}、荒木賢二^{*1}、外山比南子^{*2}

*1 宮崎大学医学部附属病院IR部、*2 医療データサイエンス研究所

The exploration of latent clinical decision limits by using random forests

Ryosuke Matsuo^{*1}, Taisuke Ogawa^{*1}, Tomoyoshi Yamasaki^{*1}, Kenji Araki^{*1}, Hinako Toyama^{*2}

*1 Department of Hospital Institutional Research, Faculty of Medicine, University of Miyazaki Hospital, *2 Institute of Medical Data Sciences

We propose a method for exploring latent clinical decision limits (LCDLs) by using random forests. The proposed method consists of two phases: case finding and data analysis. In the first phase, the reference ranges for laboratory tests in data of specific patients are estimated by minimizing the standard deviation of laboratory tests. For each laboratory test, we extract the patients whose laboratory test results are within the estimated reference ranges. In the second phase, the relationship between the laboratory test and a dependent variable is visualized by aggregating the selected branch information (the thresholds, conditions, and the delta values of the dependent variable) in the learned model of random forests using the extracted patients. We exploratory found the candidates of LCDLs for AMY, ALP, TC, MCV, and MPV, which affect long hospital length of stay (HLOS) from the visualization results. There were no characteristics between HLOS and the ages and laboratory test results of patients who correspond to the LCDLs. However, they tend to take particular medications such as psychotropic and hypolipidemic depending on the patient's condition. This suggests that the proposed method could discover valuable LCDLs by combining systematic causal analyses between LCDLs and dependent variables.

Keywords: Latent Clinical Decision Limits, Random Forests, Knowledge Discovery.

1. 緒論

臨床検査値には、特定の病態を考慮しない健常者における基準範囲 (RIs) と、特定の病態に関して、その診断・予防や治療・予後の判定基準のための臨床判断値 (CDLs) が定められている¹⁾。RIs と CDLs は、臨床病理結果の解釈をサポートするために不可欠な臨床検査情報とされている²⁾。

RIs を定めるには、事前に参照集団からサンプルを集める従来の direct approach と、スクリーニングや診断、モニタリングを目的に集められたデータを用いる indirect approach がある³⁾。RIs のデータ駆動型研究では、The Stanford Translational Research Integrated Database Environment (STRIDE) のデータを用いて、RIs を抽出するための教師なし学習の手法が提案されている⁴⁾。この手法は、Hampel identifier により臨床検査の極端値を検出し、極端値と関連の高い ICD9 コードを Fisher's exact test および Bonferroni's correction により特定する。特定した ICD9 コードに該当する患者を除外する処理を繰り返し行い、極端値と関連の高い ICD9 コードがみられなくなった際の患者集団から RIs を求める。RIs は検査結果の中央 95% に含まれ、下限と上限のカットオフはそれぞれ 2.5, 97.5 パーセンタイルである。ICD9 コードのリストアップにより、RIs の定義としての集団特徴を明確化することは、International Federation for Clinical Chemistry and Laboratory Medicine に準拠しているとされる。

CDLs を定めるには、ベイジアンアプローチ、疫学的アプローチ、生理学的アプローチの3つがある⁵⁾。CDLs のデータ駆動型研究では、血漿ミエロペルオキシダーゼ (MPO) の最適なカットオフ値を決定するために、ロジスティック回帰モデルを活用する研究がある⁶⁾。MPO のカットオフ値を 250 から 650 の間で 1 ずつ値を増加させていくグリッドサーチを用いて、最大の AUC を導く値を最適なカットオフとしている。ロジスティック回帰モデルは、選択的冠動脈造影を受けている患者 1895 名の 3 年間の prospective follow-up study のデータを用いた

主要心血管イベントのリスク予測に適用している。最適なカットオフ値は統計分析のための群分け基準に活用されている。

目的変数と説明変数の 2 つの要素を用いる機械学習の教師あり学習は、目的変数に対する説明変数の重要度を捉えられるが、CDLs においても、臨床検査値を説明変数、特定の病態に関する診断・予防や治療・予後をも目的変数として教師あり学習の問題に取り組むことができると考える。

目的変数に対する重要な説明変数を選択する方法は特徴選択と呼ばれ、ある説明変数の集合から、目的変数と関係性の高い説明変数で構成されるサブセットを選択する。特徴選択は、オーバーフィッティングの回避やモデルパフォーマンスの改善、高速かつ費用対効果の高いモデルの実現、データの深い洞察を可能にする⁷⁾。特徴選択の 3 つの分類⁷⁾: フィルター法、ラッパー法、組み込み法のうち、組み込み法は、予測や分類といった機械学習の過程で特徴選択が行われ、目的変数に対する説明変数の重要度が得られる。組み込み法は、元々の説明変数の表現形と目的変数の関係で重要な変数を探し出すことが容易である。この説明変数の表現を分析目的に応じて意味的に変換することで、目的変数との関係の中での知識発見につながると考えられる。本研究では、連続値の説明変数における潜在的な判断値を捉えられる表現形への変換を、ランダムフォレストにより構築される予測モデル内の説明変数の情報を活用したデータマイニングにより行う。

ランダムフォレスト⁸⁾は、組み込み法に分類される機械学習手法の1つで、複数の決定木を作り、その多数決で予測のためのモデルを構築する。モデルの構築過程で、説明変数ごとで目的変数への重要度が得られる。複数の決定木で構成されるランダムフォレストのモデルは、潜在的な判断値を探索するために必要な要素を含んでいると考える。具体的には、ランダムフォレストのモデルの分岐情報における、分岐の閾値、条件 (以下 $\leq$, より大きい $>$)、分岐の閾値と条件に一致した場合の目的変数の変化量の 3 つを選択する。ランダムフ

オレストは複数の決定木に基づくモデルであることから、上述の3項目の情報を収集することが可能である。収集した情報を活用して、潜在的な判断値を捉えるための説明変数の表現変換として、説明変数と目的変数の関係性を可視化する。これらの選択・収集・可視化の3つのプロセスにより、潜在的な判断値を探索する。

ランダムフォレストを用いて構築された予測モデルにより得られる結果の解釈性を高める研究として、ボクセル単位の神経画像の解釈性改善のため、個別のボクセル単位の特徴重要度からグループ重要度を特徴重要度の合計・平均・最大の3つの統計量から求め、グループ重要度を統計的に解釈可能な指標に変換する手法が提案されている⁹⁾。この手法は、FDG-PET スキャンのデータセットを用いたアルツハイマー病の予後に関わる脳領域の識別に適用されている。また、結果の解釈に特徴の向き(プラス・マイナス)を考慮する研究がある¹⁰⁾。この研究では、二値のクラスラベルごとで、全テストサンプルに対して、正しく予測をした全ての決定木内において、ある特徴がある閾値により二値のクラスラベルのいずれかに予測される時に、閾値より以上(有)または以下(無)の割合を求めている。この方法は、Stanford FEATURE¹¹⁾システムの分子生物学データに適用されている。しかしながら、ランダムフォレストによるモデル内の分岐における閾値、条件および目的変数の変化量を選択、収集し、収集した情報を活用して可視化することで潜在的判断値を探索する方法はみられない。

2. 目的

本研究の目的は、ランダムフォレストによる予測モデル内の分岐情報を活用して、患者の病態に基づいた潜在的な臨床検査の判断値を探索する方法を提案する。説明変数の潜在的な判断値を探索する方法があれば、膨大な臨床検査データによる検査値に関する新たな臨床知識の候補獲得が期待される。

3. 方法

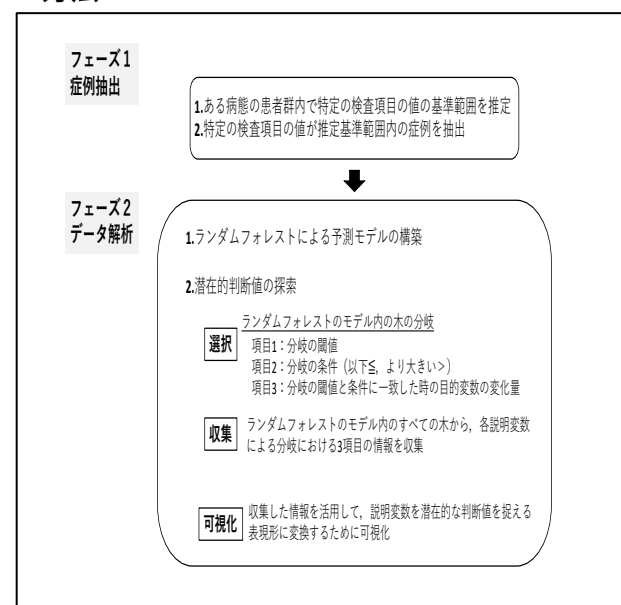


図1 提案手法の概要図

提案手法は、症例抽出とデータ解析の2つのフェーズで構成される。症例抽出のフェーズでは、ある病態の患者群内で各検査項目の基準範囲を推定し、検査項目ごとに推定した基準範囲内の症例を抽出する。解析フェーズでは、抽出した

症例を用いて、複数の検査項目を説明変数として予測モデルをランダムフォレストにより構築する。そして、症例抽出に用いた特定の検査項目が特徴重要度のランキングで上位に位置する場合に、選択および収集された分岐の情報を活用して、該当検査値と目的変数の関係性を可視化する。図1は提案手法の概要図である。以降でそれぞれのフェーズの詳細を述べる。

3.1 フェーズ1: 症例抽出

患者の病態に基づいて潜在的な判断値を探索するにあたり、ある病態の患者群内で特定の検査項目におけるアウトライヤー症例を除外する。そのため、特定の検査項目ごとに、設定する病態の患者群内での検査値の基準範囲を推定し、推定した基準範囲内の症例を抽出する。推定基準範囲は、検査値が平均値±2SD 外の患者数を段階的に除外していき、標準偏差が変化しない最小の標準偏差により求める。最小化された標準偏差を用いた推定基準範囲は以下の通りである。

$$\begin{aligned} \text{推定基準範囲} &= \bar{x}_{\min} - 2 \cdot SD_{\min} \leq \text{検査値} \\ &\leq \bar{x}_{\min} + 2 \cdot SD_{\min} \end{aligned}$$

$\bar{x}_{\min}$ は標準偏差が最小の時の該当患者群の平均検査値で、 $SD_{\min}$ は該当患者群の検査値の最小標準偏差値である。

3.2 フェーズ2: データ解析

3.2.1 ランダムフォレストによる予測モデルの構築

フェーズ1により抽出した症例を用いて、ランダムフォレストにより予測モデルを構築する。訓練データから、ブートストラップ法により訓練データと同じ数のサンプルを復元抽出する。抽出したサンプルを用いて、ルートノードを起点として、すべての説明変数を対象に、最も良く分割ができる説明変数とその閾値を選択していくことで木を構成する。この選択を設定する木の深さまで、そして設定する木の数おこなう。最も良く分割できる説明変数とその閾値は、分割された2つの子ノードの不純度をそれぞれのサンプル数の割合で加重平均した値を最小化する組である。

3.2.2 潜在的判断値の探索

ランダムフォレストにより構築した予測モデルを活用して、選択・収集・可視化の3つのプロセスにより、潜在的な臨床判断値を探索する。

始めに、潜在的な判断値を捉えるために必要な項目として、ランダムフォレストにより構築したモデル内の木の分岐の閾値、分岐の条件(以下≤, より大きい>), 分岐の閾値と条件に一致した時の目的変数の変化量の3項目を選択する。目的変数の変化量は、あるノード k のデータ N_k における目的変数の平均値 $\bar{y}_k$ から、分割後のノード k_{left} と k_{right} のデータにおける目的変数の平均値 $\bar{y}_{k_{\text{left}}}$ と $\bar{y}_{k_{\text{right}}}$ をそれぞれ引いた2通りの値とする。

次に、判断値としての傾向を捉えるために、各説明変数による分岐における3項目の情報をランダムフォレストにより構築したモデル内のすべての木から収集する。

最後に、収集した情報を活用して、潜在的な判断値を捉える表現形に変換するための可視化を行う。説明変数ごとで可視化するグラフの横軸は、ランダムフォレストにより構築したモデル内の分割されるノードが該当の説明変数である場合のすべての分岐の閾値である。縦軸は、構築されたモデル内で、該当説明変数の値が分岐の閾値と条件に一致した時の目的変数の変化量であり、グラフの値はその変化量で分岐の条件

(以下 $\leq$, より大きい $>$)ごとに色分けして表示される. 分岐の条件が以下 $\leq$ の場合は黒, より大きい $>$ 場合は灰色とする.

本研究では上述の潜在的判断値探索法を臨床検査データに適用し, 潜在的臨床判断値(LCDLs)の探索を試みる.

3.2 実験設定

本研究では, 宮崎大学医学部附属病院のデータウェアハウスに格納されている2014年4月から2018年3月までの臨床検査データを用いる. 特に, 患者ごとの入院時(入院前後3日)の検査項目の値を予測モデルの説明変数とする. 患者の病態は, DPCコードの6桁と, 手術あり, 化学療法の有無に基づくDPCクラスタにより捉える. 本研究では, 年間の入院件数が50件以上のクラスタに限定する. 症例抽出の際は, DPCクラスタごとで, 入院前後3日で検査値の件数が該当DPCクラスタ内の患者数の30%以上の検査項目を対象に基準範囲を推定する. 推定基準範囲内の抽出した患者数が該当DPCクラスタ患者数の30%以上の検査項目において, 抽出症例の入院時の複数の検査項目の値を用いて, 在院日数の予測モデルをランダムフォレストにより構築する. 入院時の複数の検査項目の値は, フェーズ1の症例抽出で用いた標準偏差の最小化により, 推定された基準範囲内の場合のみを用いる. 入院前後3日に複数の推定基準範囲内の検査値が記録されている場合は最大値を用いる. 本研究では, 予防のための知見を得られるようにするため, 検査値を術前の検査値に限定する. 入院前後3日の間に手術の記録がない場合, 入院前後3日の検査値を術前検査値とし, 手術の記録がある際は, 入院前後3日の検査値の日付が手術日より前(同日含まない)の場合を術前検査値とする. 入院前後3日の間で複数の手術が記録されている場合, 最も点数が高い手術を選択する. 同じ点数で異なる手術が記録されている場合は, 最も入院日に近い手術を選択する.

本実験では, 訓練データを7割, テストデータを3割とする. ランダムフォレストの木の数をも100個, 木の深さの最大をも3とする. 本実験は, PythonのScikit-learn¹²⁾を用いる. 検査項目が特徴重要度のランキングで3位以内の場合で, かつ患者数が100名以上の場合を潜在的判断値探索の対象とする.

4. 結果

LCDLsの探索による可視化結果から, 症例抽出に用いたAMY, ALP, TC, MCV, MPVの5つの検査項目において, 定義上の基準範囲内で, ある閾値以下の場合またはより大きい場合に総じて在院日数が長くなるDPCクラスタが5ケースみられた. 該当する5つのDPCクラスタは化学療法なしであった. 表1は, 在院日数が長くなることに影響するLCDLsをまとめたもので, 検査項目名とDPCクラスタ, カットオフ値, カットオフの条件(閾値以下 $\leq$ または閾値より大きい $>$), 該当検査項目の基準範囲, カットオフ値とその条件に一致する訓練データ内のLCDLs該当患者数で構成される. 図2-6は5ケースの可視化結果である.

検査項目がAMYで斜視(外傷性・癒着性を除く.)のクラスタでは, 患者数256名, 説明変数の検査項目数は43個, 特徴重要度のランクは3位であった. 図2から, AMYが97.0 U/Lより大きいと在院日数が長くなる傾向にあり, LCDLs 該当患者数は19名であった.

検査項目がALPで肘, 膝の外傷(スポーツ障害等を含む.)のクラスタでは, 患者数128名, 説明変数の検査項目数は42個, 特徴重要度のランクは2位であった. 図3から, ALPが242.0 U/Lより大きいと在院日数が長くなる傾向にあり, LCDLs 該当患者数は12名であった.

検査項目がTCで膵臓, 脾臓の腫瘍のクラスタでは, 患者数154名, 説明変数の検査項目数は65個, 特徴重要度のランクは2位であった. 図4から, TCが202.0 mg/dLより大きいと在院日数が長くなる傾向にあり, LCDLs 該当患者数は12名であった.

表 1 在院日数が長くなることに影響する LCDLs

検査項目名	DPC クラスタ	カットオフ値	カットオフ条件	基準範囲	患者数
AMY	斜視（外傷性・癒着性を除く.）	97.0 U/L	$>$	44 - 132 U/L ¹³⁾	19名
ALP	肘, 膝の外傷（スポーツ障害等を含む.）	242.0 U/L	$>$	106 - 322 U/L ¹³⁾	12名
TC	膵臓, 脾臓の腫瘍	202.0 mg/dL	$>$	142 - 248 mg/dL ¹³⁾	12名
MCV	緑内障	86.1 fL	$\leq$	83.6 - 98.2 fL ¹³⁾	8名
MPV	黄斑, 後極変性	8.75 fL	$\leq$	7.0 - 11.0 fL ¹⁴⁾	7名

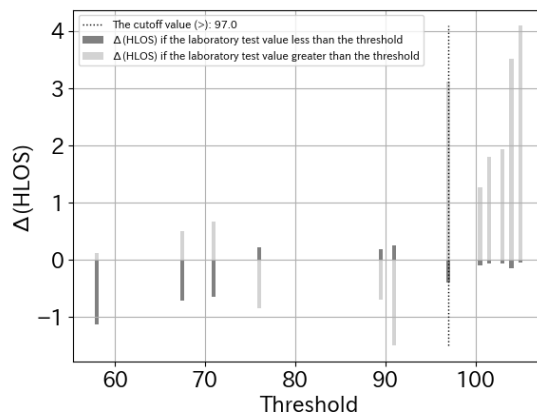


図 2 AMY:斜視(外傷性・癒着性を除く.)の可視化結果

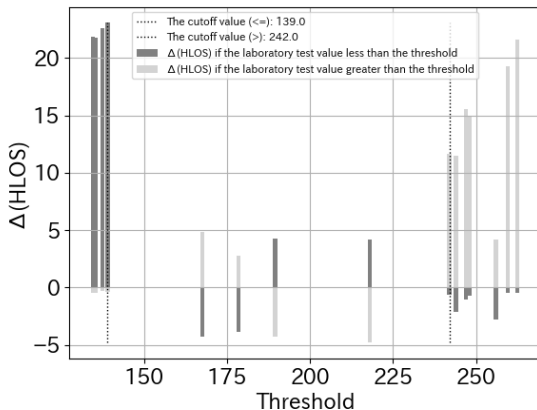


図 3 ALP:肘, 膝の外傷(スポーツ障害等を含む.)の可視化結果

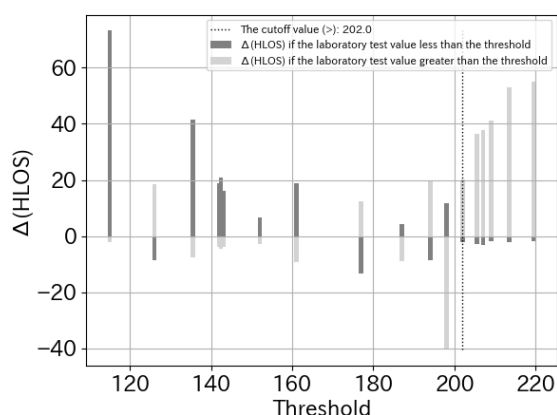


図4 TC:脾臓, 脾臓の腫瘍の可視化結果

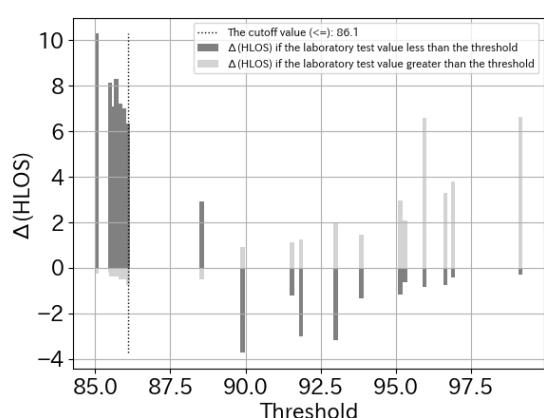


図5 MCV:緑内障の可視化結果

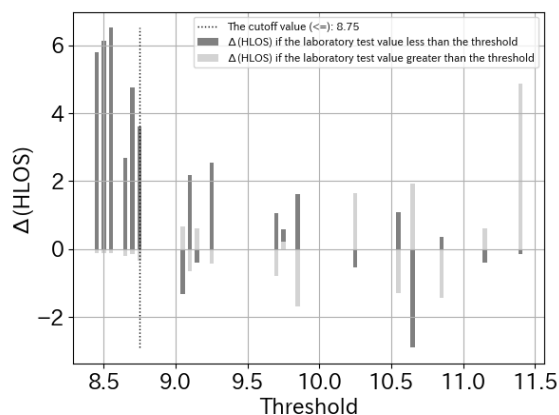


図6 MPV:黄斑, 後極変性の可視化結果

検査項目が MCV で緑内障のクラスタでは、患者数 155 名、説明変数の検査項目数は 39 個、特徴重要度のランクは 2 位であった。図 5 から、MCV が 86.1 fL 以下であると在院日数が長くなる傾向にあり、LCDLs 該当患者数は 8 名であった。

検査項目が MPV で黄斑、後極変性のクラスタでは、患者数 159 名、説明変数の検査項目数は 37 個、特徴重要度のランクは 3 位であった。図 6 から、MPV が 8.75 fL 以下であると在院日数が長くなる傾向にあり、LCDLs 該当患者数は 7 名であった。

5. 考察

探索的に得られた LCDLs の在院日数への影響の原因を究明するため、2通りの分析を行う。まず、目的変数の在院日数と、LCDLs にあたる検査項目の検査値および患者基本情報の年齢から原因を究明する。そのため、LCDLs 該当患者において、平均在院日数より長い群と短い群で年齢と検査値の差を分析する。次に、診療行為から原因を究明するため、LCDLs 該当患者と非該当患者の 2 群で、入院前後 3 日の手術と薬剤投与の差を分析する。

5.1 在院日数と年齢、検査値による原因究明

図 7-9 は、5 ケースの中で LCDLs 該当患者が 10 名以上の 3 ケースを対象とした分布図で、縦軸が在院日数、横軸が年齢または検査値である。平均在院日数は、該当する DPC クラスタ内の患者の在院日数の平均としている。赤点が該当 DPC クラスタの訓練データ内の LCDLs 該当患者、青点がそれ以外の該当 DPC クラスタの全患者のデータである。

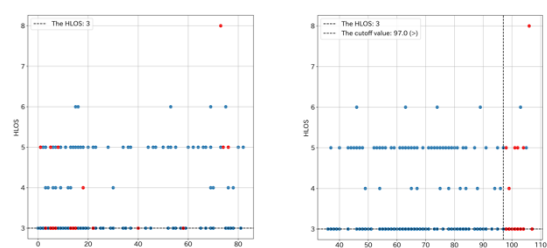


図7 AMY:斜視(外傷性・癒着性を除く.)の在院日数と年齢・検査値分布

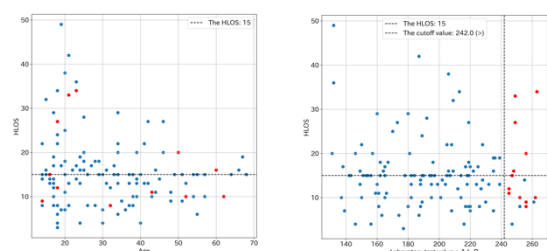


図8 ALP:肘, 膝の外傷(スポーツ障害等を含む.)の在院日数と年齢・検査値分布

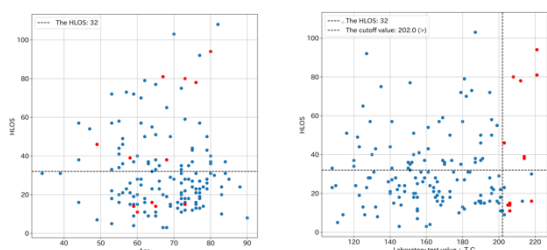


図9 TC:脾臓, 脾臓の腫瘍の在院日数と年齢・検査値分布

検査項目が AMY で斜視(外傷性・癒着性を除く.)のクラスタの LCDLs 該当患者では、平均在院日数より短い患者 12 名の平均年齢は 16.1 歳、平均検査値は 100.92 U/L であった。平均在院日数より長い患者 7 名の平均年齢は 36.4 歳、平均検査値は 101.71 U/L であった。マン-ホイットニーの U 検定の結果から、平均年齢と平均検査値の p 値はそれぞれ 0.35, 0.549 であったことから、年齢と検査値に有意な差はみられなかった。

検査項目がALPで肘、膝の外傷(スポーツ障害等を含む。)のクラスタのLCDLs 該当患者では、平均在院日数より短い患者7名の平均年齢は33.9歳、平均検査値は251.86 U/Lであった。平均在院日数より長い患者5名の平均年齢は34.4歳、平均検査値は253 U/Lであった。マン-ホイットニーのU検定の結果から、平均年齢と平均検査値のp値はそれぞれ0.807, 0.622であったことから、年齢と検査値に有意な差はみられなかった。

検査項目がTCで膵臓、脾臓の腫瘍のクラスタのLCDLs 該当患者では、平均在院日数より短い患者5名の平均年齢は64.2歳、平均検査値は208.2 mg/dLであった。平均在院日数より長い患者7名の平均年齢は67.3歳、平均検査値は213.29 mg/dLであった。マン-ホイットニーのU検定の結果から、平均年齢と平均検査値のp値はそれぞれ0.464, 0.189であったことから、年齢と検査値に有意な差はみられなかった。

検査項目がMCVで緑内障のクラスタのLCDLs 該当患者では、平均在院日数より短い患者2名の平均年齢は62歳、平均検査値は85.55 fLであった。平均在院日数より長い患者6名の平均年齢は59.3歳、平均検査値は85.22 fLであった。マン-ホイットニーのU検定の結果から、平均年齢と平均検査値のp値はそれぞれ0.868, 0.396であったことから、年齢と検査値に有意な差はみられなかった。

検査項目がMPVで黄斑、後極変性のクラスタのLCDLs 該当患者では、平均在院日数より短い患者4名の平均年齢は73.5歳、平均検査値は8.55 fLであった。平均在院日数より長い患者3名の平均年齢は79.67歳、平均検査値は8.57 fLであった。マン-ホイットニーのU検定の結果から、平均年齢と平均検査値のp値はそれぞれ0.372, 1.000であったことから、年齢と検査値に有意な差はみられなかった。

在院日数と年齢・検査値の関係分析から、LCDLs 該当患者において、在院日数が長くなることと年齢が強く関連していないことがわかった。検査値においても、在院日数より短い群と長い群の間で平均検査値に有意な差がみられなかったことから、LCDLsは検査値が低いまたは高く、在院日数が長い患者の影響を特に受けていることが示唆された。

5.2 診療行為の情報をういた原因究明

表2-10は、LCDLs 該当患者の特徴的な診療行為と統計的に有意な診療行為を示す有意差検定結果である。特徴的な診療行為の表は、診療行為名、診療行為の出現頻度、診療行為がされた該当患者数、全患者数、該当患者の割合で構成される。この表では、LCDLs 該当患者8割以上に行われた特徴的な診療行為が示されている。有意差検定結果の表は、診療行為名、p値、LCDLs 該当患者で診療行為のある人数、非該当患者で診療行為のある人数、該当患者で診療行為無の人数、非該当患者で診療行為無の人数、オッズ比で構成される。この表では、p値が0.05未満の統計的に有意な診療行為のみ示されている。有意差検定は、カイ二乗検定、症例数が少ない場合はフィッシャーの正確確率検定を実施した。検査項目がALPで肘、膝の外傷(スポーツ障害等を含む。)のクラスタでは統計的に有意な差はみられなかった。

表2-10の結果から、DPCクラスタによらず、概して解熱鎮痛消炎剤や抗生剤を投与される患者の割合が多かったことから、感染症や炎症を起こしていたことで、入院期間が長くなってしまっていたことが示唆される。しかしながら、LCDLs 該当患者と非該当患者で有意差を検定すると、感染症や炎症に関して統計的に有意な差はみられなかった。一方、検査項目がAMYで斜視(外傷性・癒着性を除く。)のクラスタでは、

精神神経用剤の投与に有意差があり、オッズ比が3.538以上であったことから、LCDLs 該当患者に精神神経用剤が投与されやすい傾向にあることがわかった。また、検査項目がMCVで緑内障のクラスタでは、高脂血症用剤の投与に有意差があり、オッズ比が12.222であったことから、LCDLs 該当患者に高脂血症用剤が投与されやすい傾向にあることがわかった。このことから、DPCクラスタで識別される患者の病態によって、在院日数が長くなることに影響する因子が異なることが示唆された。

表2 AMY:斜視(外傷性・癒着性を除く。), 特徴的な診療行為

診療行為名	出現頻度	該当患者数	全患者数	該当患者割合
薬剤_局所麻酔剤	25	19	19	1.00
薬剤_眼科用剤	94	19	19	1.00
薬剤_血液代用剤	58	19	19	1.00
手術_斜視手術	19	18	19	0.95
薬剤_解熱鎮痛消炎剤	41	18	19	0.95
薬剤_主としてグラム陽性・陰性菌に作用するもの	39	18	19	0.95

表3 AMY:斜視(外傷性・癒着性を除く。), 有意差検定結果

診療行為名	p値	該当患者_診療行為有	非該当患者_診療行為有	該当患者_診療行為無	非該当患者_診療行為無	オッズ比
薬剤_全身麻酔剤	0.006	14	151	5	9	0.167
薬剤_下剤, 浣腸剤	0.009	3	2	16	158	14.813
薬剤_精神神経用剤	0.018	7	21	12	139	3.861
薬剤_ビタミンA及びD剤	0.030	2	1	17	159	18.706

表4 ALP:肘、膝の外傷(スポーツ障害等を含む。), 特徴的な診療行為

診療行為名	出現頻度	該当患者数	全患者数	該当患者割合
薬剤_全身麻酔剤	15	12	12	1.00
薬剤_局所麻酔剤	35	12	12	1.00
薬剤_血液代用剤	71	12	12	1.00
薬剤_その他の血液・体液用薬	12	12	12	1.00
薬剤_主としてグラム陽性・陰性菌に作用するもの	25	12	12	1.00
薬剤_解熱鎮痛消炎剤	44	11	12	0.92
薬剤_精神神経用剤	12	10	12	0.83
薬剤_下剤, 浣腸剤	20	10	12	0.83

表5 TC:膵臓、脾臓の腫瘍、特徴的な診療行為

診療行為名	出現頻度	該当患者数	全患者数	該当患者割合
薬剤_消化性潰瘍用剤	34	11	12	0.92
薬剤_催眠鎮静剤, 抗不安剤	16	10	12	0.83

表6 TC:膵臓、脾臓の腫瘍、有意差検定結果

診療行為名	p値	該当患者_診療行為有	非該当患者_診療行為有	該当患者_診療行為無	非該当患者_診療行為無	オッズ比

薬剤_その他の 滋養強壮薬	0.033	2	1	10	94	18.800
------------------	-------	---	---	----	----	--------

表 7 MCV:緑内障, 特徴的な診療行為

診療行為名	出現 頻度	該当患 者数	全患 者数	該当患者 割合
薬剤_眼科用剤	108	8	8	1.00
薬剤_主としてグラム陽性・陰性菌 に作用するもの	18	7	8	0.88

表 8 MCV:緑内障, 有意差検定結果

診療行為名	p 値	該当 患者_ 診療 行為 有	非該 当患 者_診 療行 為有	該当 患者_ 診療 行為 無	非該 当患 者_診 療行 為無	オッズ 比
薬剤_高脂血症用剤	0.002	5	12	3	88	12.222
薬剤_外皮用殺菌消 毒剤	0.043	2	3	6	97	10.778

表 9 MPV:黄斑, 後極変性, 特徴的な診療行為

診療行為名	出現 頻度	該当患 者数	全患 者数	該当患者 割合
薬剤_眼科用剤	75	7	7	1.00
薬剤_その他の診断用薬(体外診断 用医薬品を除く。)	23	7	7	1.00
薬剤_解熱鎮痛消炎剤	8	6	7	0.86
薬剤_主としてグラム陽性・陰性菌に 作用するもの	14	6	7	0.86

表 10 MPV:黄斑, 後極変性, 有意差検定結果

診療行為名	p 値	該当 患者_ 診療 行為 有	非該 当患 者_診 療行 為有	該当 患者_ 診療 行為 無	非該 当患 者_診 療行 為無	オッズ 比
薬剤_その他の血 液・体液用薬	0.031	2	3	5	101	13.467

6. 結論

本研究は目的変数に対する潜在的な判断値を探索するための手法を提案し、提案手法を臨床検査データに適用し、潜在的臨床判断値(LCDLs)の探索を試みた。提案手法は、ランダムフォレストにより構築した予測モデル内の分岐情報を活用して、潜在的な判断値を捉えられる表現形に説明変数を変換した。LCDLsの探索による可視化結果から、目的変数として設定した在院日数に影響を及ぼす潜在的な判断値の候補がAMY、ALP、TC、MCV、MPVの5つの検査項目で得られた。これらの検査項目では、定義上の基準範囲内であるが、閾値以下・より大きい場合に、在院日数が長くなるDPCクラスターが5ケースみられた。

LCDLs 該当患者がなぜ在院日数が長くなってしまったかの原因究明を試みると、在院日数と患者基本情報の年齢、在院日数とLCDLsにあたる検査項目の検査値では特徴があらわれなかった。手術や薬剤投与の診療行為を深掘りすると、DPC クラスターに依存せず、LCDLs 該当患者は解熱鎮痛消炎剤や抗生剤を投与されている傾向にあった。そのため、感染症や炎症を起こしていたことで、在院日数が延びてしまっていたことが示唆された。しかしながら、LCDLs 該当患者と非該当患者で感染症や炎症に統計的有意差はみられなかった。一方で、精神神経用剤や高脂血症用剤といったある特定の薬剤が、特定の LCDLs の該当患者に投与されやすい傾向にあったことから、DPC クラスターによって、在院日数が長くなることに影響する因子が異なることが示唆された。

このようなことから、LCDLs の探索による可視化から目的変数に対する説明変数の重要度の解釈性を高め、探索的に

LCDLs の候補を獲得できることが示唆されたが、LCDLs と目的変数との間の因果関係を医学的に評価するには、1施設だけでなく他施設のデータを用いた上で、方法論化された原因究明を行うことが必要であると考ええる。また、目的変数として在院日数だけでなく、合併症や死亡といった患者のリスクの観点から LCDLs の探索に取り組む。

謝辞

本研究は JSPS 科研費 JP18K09948 の助成を受けたものです。

参考文献

1) 一般社団法人日本臨床検査医学会. 臨床検査のガイドライン JSLM2018, 基準範囲・臨床判断値, 2019. [https://www.jslm.org/books/guideline/2018/03.pdf (cited 2019-Aug-29)].

2) Ozarda Y, Sikaris K, Streichert T, Macri J, IFCC Committee on Reference intervals and Decision Limits (C-RIDL). Distinguishing reference intervals and clinical decision limits-A review by the IFCC Committee on Reference Intervals and Decision Limits. Critical Reviews in Clinical Laboratory Sciences 2018 ; 55(6) : 420-431.

3) Jones GRD, Haeckel R, Loh TP et al. Indirect methods for reference interval determination-review and recommendations. Clinical Chemistry and Laboratory Medicine 2018 ; 57(1) : 20-29.

4) Poole S, Schroeder LF, Shah N. An unsupervised learning method to identify reference intervals from a clinical database. Journal of Biomedical Informatics 2016 ; 59 : 276-284.

5) Ceriotti F, Henny J. “Are my Laboratory Results Normal?” Considerations to be Made Concerning Reference Intervals and Decision Limits. EJIFCC 2008 ; 19(2) : 106-114.

6) Tang WHW, Wu Y, Nicholls SJ, Hazen SL. Plasma myeloperoxidase predicts incident cardiovascular risks in stable patients undergoing medical management for coronary artery disease. Clinical Chemistry 2011 ; 57(1) : 33-39.

7) Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. Bioinformatics 2007 ; 23(19) : 2507-2517.

8) Breiman L. Random Forests. Machine Learning 2001 ; 45(1) : 5-32.

9) Wehenkel M, Sutera A, Bastin C, Geurts P, Phillips C. Random Forests Based Group Importance Scores and Their Statistical Interpretation: Application for Alzheimer’s Disease. Frontiers in Neuroscience 2018 ; 12 : 411.

10) Petkovic D, Altman RB, Wong M, Vigil A. Improving the explainability of Random Forest classifier-user centered approach. Pacific Symposium on Biocomputing 2018 ; 23 : 204-215.

11) Wei L, Altman RB. Recognizing complex, asymmetric functional sites in protein structures using a Bayesian scoring function. Journal of Bioinformatics and Computational Biology 2003 ; 1(1) : 119-138.

12) Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 2011 ; 12(Oct) : 2825-2830.

13) 国立がん研究センター. 臨床検査基準値一覧, 2016. [https://www.ncc.go.jp/jp/ncch/division/clinical_laboratory/ken sa.pdf (cited 2019-Aug-29)].

14) 一般社団法人 日本血栓止血学会. 平均血小板容積(MPV)・血小板分布幅(PDW). [http://www.jsth.org/glossary_detail/?id=254 (cited 2019-Aug-29)].