

令和 5年 3月31日

各報道機関 御中

国立大学法人山梨大学
国立大学法人千葉大学

機械学習を用いた川崎病における
新たな免疫グロブリン静注 (IVIG) 療法不応予測

研究成果のポイント

1. 機械学習による IVIG 不応予測の中で最も精度が高かったのは多数の決定木に基づいたアンサンブル学習の一つである light gradient boosting machine (Light GBM) 法であり、Light GBM 法での AUC (area under the ROC curve)、感度、特異度はそれぞれ 0.78、0.50、0.88 であった。
2. Light GBM 法での IVIG 不応予測に影響した上位 20 項目を SHAP (SHapley Additive exPlanations) で評価した。IVIG 不応予測に重要な上位 3 項目は治療開始日、CRP 値、総コレステロール値であった。過去に報告されている IVIG 不応予測スコアの中で、総コレステロール値が含まれているスコアはないものの、近年の研究から総コレステロール値は川崎病の病態生理へ関与すると言われている。
3. SHAP の結果から構築した新たなスコア (山梨スコア) は、過去に報告されている IVIG 不応予測スコアと同等以上の予測精度であった。
4. 多数の決定木に基づいたアンサンブル学習は、多重共線性を考慮する必要がなく、予測変数を大幅に増加させても解析が成立する。今後、解析規模を拡張することで、従来の解析では予想しがたい川崎病の病態生理に関連する因子の発見、より有用な予測モデルの作成が期待される。

【概要】

山梨大学医学部小児科学講座の須長祐人臨床助教と犬飼岳史教授らと、千葉大学大学院医学研究院人工知能（AI）医学/理化学研究所の川上英良教授らの研究グループは、山梨県内の小児入院施設を中心に構成される山梨川崎病研究グループ（代表：長谷部洋平 山梨大学小児科学講座 特任助教）の協力により収集した2010年から2020年の川崎病臨床データから機械学習を用いてIVIG不応予測を行いました。またその機械学習の結果を踏まえて、臨床応用が可能で簡便な新たなスコアを作成いたしました。本研究の成果は、Springer Nature 出版の *Clinical Rheumatology* に、1月11日付けでオンライン掲載されました。

・ 掲載 URL: <https://doi.org/10.1007/s10067-023-06502-1>

・ 論文 掲載誌 : *Clinical Rheumatology*

・ 掲載日時 : 2023 年 1 月 11 日

・ 論文タイトル :

A simple scoring model based on machine learning predicts intravenous immunoglobulin resistance in Kawasaki disease

・ 著者 :

Yuto Sunaga¹, Atsushi Watanabe¹, Nobuyuki Katsumata^{1,2}, Takako Toda^{1,3}, Masashi Yoshizawa¹, Yosuke Kono¹, Yohei Hasebe¹, Keiichi Koizumi^{1,4}, Minako Hoshiai^{1,5}, Eiryō Kawakami^{6,7,8}, Takeshi Inukai¹,

1) 山梨大学 医学部 小児科学講座 2) 山梨県立中央病院 新生児科 3) 国立循環器病センター 4) 富士吉田市立病院 小児科 5) 山梨県立中央病院 小児科 6) 千葉大学大学院医学研究院人工知能（AI）医学 7) 治療学人工知能（AI）研究センター 8) 理化学研究所理化学研究所 先端データサイエンスプロジェクト

【背景】

川崎病（KD）は、乳幼児期に好発する全身の血管炎症候群であり、発熱の持続、眼球結膜の充血、口唇の紅潮、不定形発疹、四肢末端の変化や頸部リンパ節の腫脹などの症状を特徴とします。冠動脈をはじめとする中型動脈の炎症を中心とした動脈炎を生じて、冠動脈瘤を形成する可能性があります。一般的に中型血管の炎症から動脈炎、汎血管炎に至り発症10日以降で動脈の拡張を認めます。

川崎病の標準的な初期治療は、大量免疫グロブリン静注（IVIG）療法が確立されていますが、約20%の症例で発熱や炎症所見が改善せず、IVIG不応として追加治療を必要とします。IVIG不応症例は、冠動脈病変を合併するリスクが高いです。近年、治療前にIVIG不応リスク層別化を行い、高リスク群に対する初期治療をIVIG+他の抗炎症薬とすることで冠動脈病変を含めた川崎病の予後を改善する可能性があることが報告されています。冠動脈病変の発症を抑えるため、治療前のIVIG不応リスクをより正確に予測することは重要な課題です。

機械学習は、近年、転帰予測、診断、画像解釈など医学の分野で広く応用されつつあります。過去に川崎病症例を対象に機械学習を施行した報告はありますが、機械学習の結果を踏まえて簡便なスコアを作成した報告はありません。本研究では、機械学習を使用し診断時の臨床的特徴からIVIG不応を予測した上で、その結果を踏

まえ、IVIG不応に関連する最も重要な3つの項目から臨床応用が可能で簡便な新たなスコアを作成しました。

【研究の方法】

2010年6月から2020年12月の間に山梨大学およびその関連病院で診断・治療されたKD症例1,002例を対象とし、後方視的コホート研究を行いました。機械学習はpython、scikit learnで行いました。機械学習の説明変数は、診断時における臨床的特徴および検査データの計30項目とし、目的変数はIVIG反応性としました。KD症例1002例のうち、227例(22.7%)が初回IVIG不応でした。801例(約80%)をトレーニングデータ、201例(約20%)をテストデータとし、多数の決定木に基づいたアンサンブル学習アルゴリズムであるRandom Forest法、eXtreme Gradient Boosting (XGBoost)法とLight Gradient Boosting Machine (Light GBM)法を用いて行いました。また予測モデルに影響を与え、IVIG反応性の高い説明変数をSHAP (SHapley Additive exPlanation)法で特定しました。SHAP法は、協力ゲーム理論における報酬を貢献度に応じて分配する手法を機械学習に応用した手法であり、それぞれの説明変数(項目)および実際の数値が、どの程度予測に影響を与えたかを具体的に算出することができます。それらの結果から各項目のカットオフを決定し、機械学習に基づく簡便なスコア(山梨スコア)を構築しました。本研究の患児を対象に新たなスコア(山梨スコア)の曲線下面積(AUC)、感度および特異度を評価し、過去に報告されているロジスティック回帰分析によってつくられた代表的な3つのスコア[群馬スコア(小林ら、2006)、久留米スコア(江上ら、2006)、大阪スコア(佐野ら、2007)]と比較しました。

【主要な研究結果】

① 機械学習によるIVIG不応予測結果

各モデルにおけるIVIG不応予測の中で最も精度が高かったのは、Light GBM法でした。Light GBM法での精度、感度、特異度、陽性予測値、陰性予測値、陽性尤度比、陰性尤度比はそれぞれ0.78 (95%信頼区間: 0.72 - 0.84)、0.50 (0.36 - 0.64)、0.88 (0.82 - 0.93)、0.59 (0.43 - 0.74)、0.83 (0.77 - 0.89)、4.14 (2.48 - 6.90) および0.57 (0.43 - 0.75) でした。

② IVIG不応予測に影響した変数

Light GBM法での予測に影響した上位20項目をSHAPで評価しました(図2b)。SHAP値の絶対値が高いほど、IVIG不応予測をする上でその項目がより重要であることを示しています。IVIG不応予測に重要な上位3項目は、治療開始日、CRP値と総コレステロール値でした。過去に報告されているIVIG不応予測スコアの中で、総コレステロール値が含まれているスコアはないものの、近年の研究から、総コレステロール値は川崎病の病態生理へ関与すると言われています。

③ 新たなスコアの構築

SHAPの結果からIVIG不応予測モデルに影響した上位3項目をスコアの項目としました。各項目のSHAP dependence plotを解析し、その結果から開始日 ≤ 4 日、CRP ≥ 10 および7 mg/dL、総コレステロール ≤ 131 mg/dLという各項目のカットオフ値を設定しました(図3)。また開始日 ≤ 4 日目、CRP ≥ 10 mg/dLを2点、CRP ≥ 7

mg/dL(および<10 mg/dL)、総コレステロール \leq 131 mg/dL を 1 点として、この 3 項目からなる簡易スコアリングモデル(山梨スコア)を作成しました。当施設の患児を対象とし新たなスコア(山梨スコア)を、群馬スコア(小林ら、2006)、久留米スコア(江上ら、2006)、大阪スコア(佐野ら、2007)と比較したところ、3 項目から構成される新たなスコアは群馬スコアとほぼ同じ精度で IVIG 不応を予測し、久留米、大阪スコアよりさらに高い確率で予測できました(図 4)。

【今後の展望】

多数の決定木に基づいたアンサンブル学習アルゴリズムは、多重共線性を考慮する必要がないため、予測変数を大幅に増加させても解析が成立します。今後、解析規模を拡張することで、従来の解析からは予想しがたい川崎病の病態生理に関連する因子の発見、より有用な予測モデルの作成が期待されます。また本研究では IVIG 不応を予測しましたが、川崎病の予後を改善するためには、遠隔期の冠動脈病変の有無を予測できることが望まれます。

SHAP を適用することでランダムフォレスト法の説明変数の意義を明確化し、新規の簡便なスコアリングシステムを構築することができました。しかし、他の人種、集団での評価ができておらず、今後、他の人種や集団に適用するためにはさらなる検証が必要です。

【図の説明】

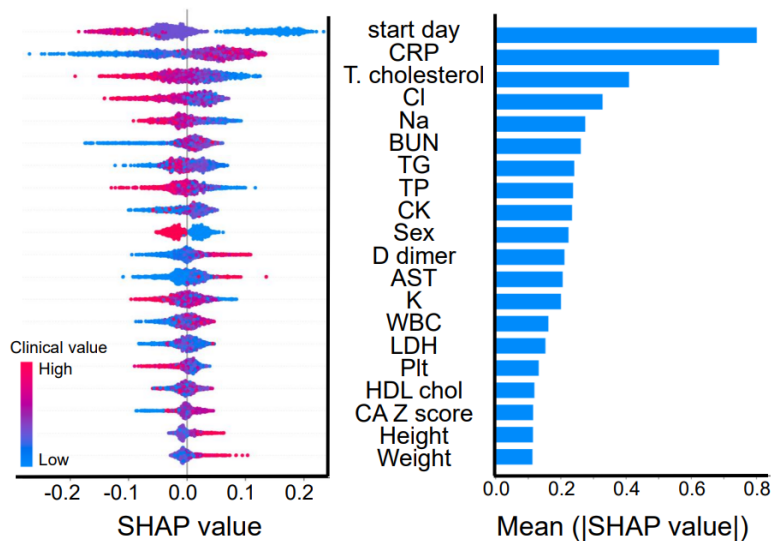


図 2b. SHAP の結果、Light Gradient Boosting Machine (Light GBM)法での IVIG 不応予測に重要な上位 20 項目を示しています。重要な項目であるほど上位となります。左の図の横軸は、SHAP 値を示しており、SHAP 値は、正の値つまり高値であるほど、IVIG 不応であることの予測に関連し、負の値つまり低値であるほど、IVIG 不応でなく有効であることの予測に関連します。ちなみにドットは各項目の実際の数値を示しており赤になるほど高値、青になるほど低値を示しています。右の図の横軸は、SHAP 値の絶対値の平均を示しており、SHAP 値の絶対値が高いほど、IVIG 不応予測をする上でその項目がより重要であることを示しています。IVIG 不応予測に重要な上位 3 項目は、治療開始日、CRP 値と総コレステロール値でした。

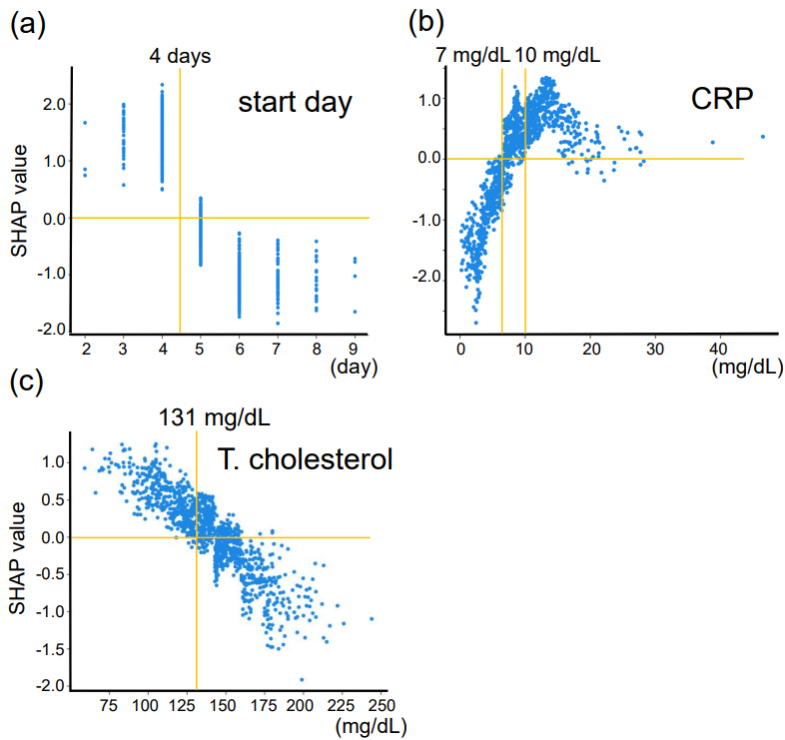


図3. 治療開始日、CRP値と総コレステロール値のSHAP dependence plotを示しています。SHAP dependence plotは、それぞれの項目がランダムフォレストモデルの予測にどのように影響するか、また、プロット内でその値が変化すると、SHAP値がどのように変化するかを示しています。図のY軸の値はSHAP値を示し、X軸の値は項目の実際の数値を示しており、SHAP値が0を超えると、IVIG不応のリスクが高まることを示しています。

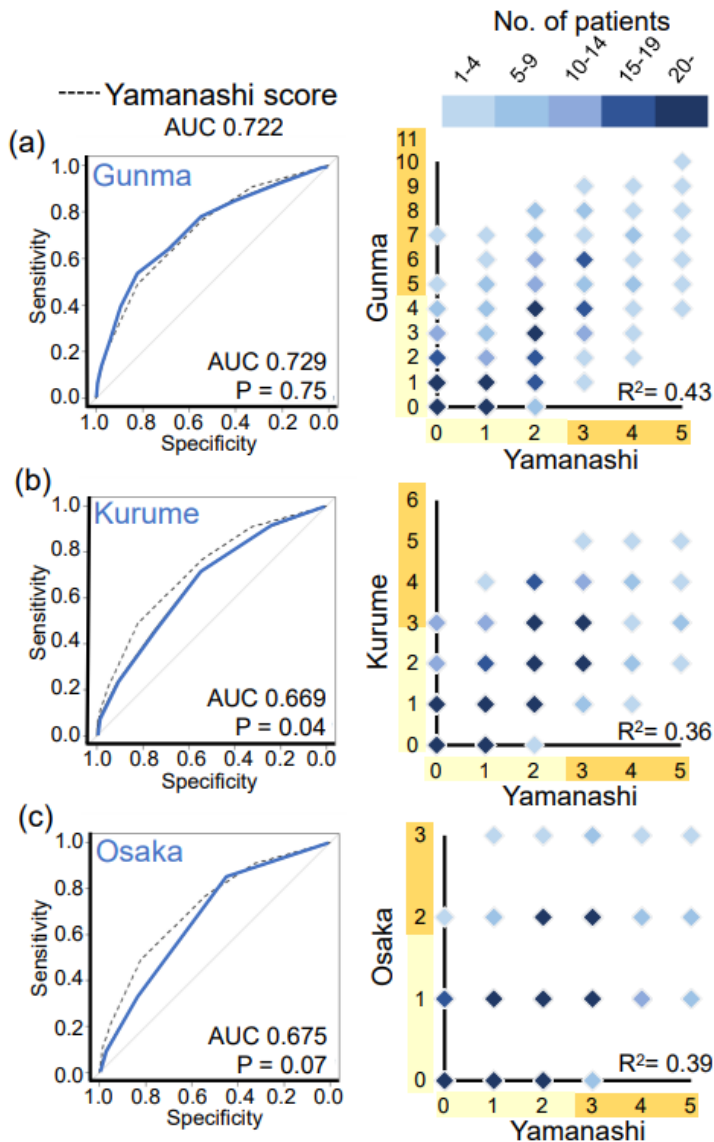


図 4. 山梨スコアと過去に報告されている 3 つの IVIG 不応予測スコア（群馬、久留米、大阪スコア）との比較を示しています。当施設の患児を対象とし、左の図は ROC 曲線の比較を右の図は各スコアの相関関係を示しています。3 項目から構成される山梨スコア [AUC: 0.72 (95%CI: 0.67 - 0.77)] は、7 項目で構成される (a) 群馬スコア [AUC: 0.73 (95%CI: 0.67 - 0.79)] とほぼ同等の予測精度であり、(b) 久留米スコア [0.67 (95%CI: 0.61 - 0.73)]、(c) 大阪スコア [0.68 (95%CI: 0.62 - 0.73)] よりさらに高い確率で予測できました。また山梨スコアと各スコアはいずれも相関していました。

【謝辞】

山梨川崎病研究グループのデータ収集にご協力いただきましたすべての先生方に厚く御礼申し上げ、感謝の意を表します。

<研究についての問い合わせ先>

山梨大学 医学部 小児科学講座 犬飼岳史、須長祐人

TEL : 055-273-9606

E-mail : yamanashi.shounika-tr@yamanashi.ac.jp

千葉大学亥鼻地区事務部総務課企画係

TEL : 043-226-2946

E-mail : inohana-plan@chiba-u.jp

<広報についての問い合わせ先>

山梨大学企画部広報企画課

TEL : 055-220-8005, 8006 FAX : 055-220-8799

E-mail : koho@yamanashi.ac.jp

千葉大学広報室

TEL : 043-290-2018 FAX : 043-284-2550

E-mail : koho-press@chiba-u.jp

川崎病における機械学習を用いた 新たな免疫グロブリン静注(IVIG)療法不応予測



A Simple Scoring Model Based on Machine Learning Predicts
Intravenous Immunoglobulin Resistance in Kawasaki Disease

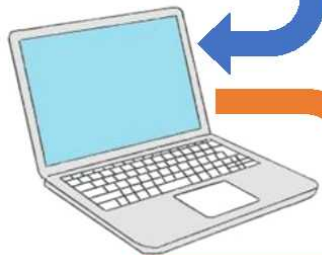
Y Sunaga, A Watanabe, N Katsumata, Toda T, et al. Clin Rheumatol. 2023 Jan 11;1-11. DOI: 10.1007/s10067-023-06502-1

- 川崎病の標準的な初期治療は、大量免疫グロブリン静注(IVIG)療法が確立されている。
- 約20%の症例で発熱や症状が改善せず、IVIG不応として追加治療を必要とする。そのような症例では冠動脈病変を合併するリスクが高いと言われている。
- 近年、治療前にIVIG不応リスク層別化を行い、高リスク群に対する初期治療を強化することで冠動脈病変を含めた川崎病の予後を改善する可能性があることが報告されており、治療前のIVIG不応リスクをより正確に予測することがより重要となっている。

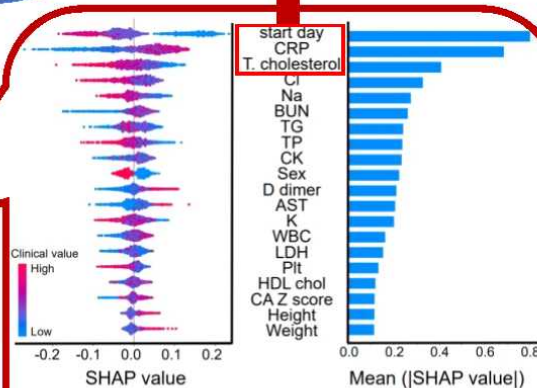


山梨県内の小児入院施設を中心に構成される山梨川崎病研究グループの協力により収集した2010年から2020年の川崎病臨床データから機械学習を使用し、IVIG不応を予測した。その結果を踏まえてIVIG不応に関連する最も重要な3つの項目から臨床応用が可能で簡便な新たなスコアを作成した。

2010-2020年の川崎病
臨床データ (1002例)



機械学習を用いて
IVIG不応を予測

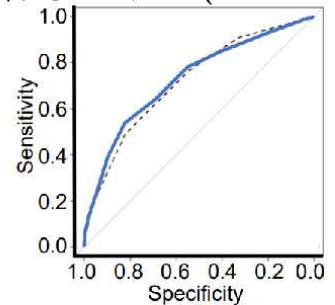


SHAPを用いて明らかとなった
IVIG不応予測に影響があった項目

山梨スコア Cutoff score ; 3 points

- 2 point : 治療開始日 ≤ 4 days
- 2 point : CRP ≥ 10 mg/dL
- 1 point : 10 > CRP ≥ 7 mg/dL
- 1 point : T. cholesterol ≤ 131 mg/dL

山梨スコア --- (AUC=0.72)
群馬スコア — (AUC=0.73)



AUC(area under the ROC curve)

研究成果のポイント

- 機械学習によるIVIG不応予測の中で最も精度が高かったのは多数の決定木に基づいたアンサンブル学習の一つであるlight gradient boosting machine (Light GBM)法であり、Light GBM法でのAUC (area under the ROC curve)、感度、特異度はそれぞれ0.78、0.50、0.88であった。
- Light GBM法でのIVIG不応予測に影響した上位20項目をSHAP(Shapley additive explanation)で評価した。IVIG不応予測に重要な上位3項目は治療開始日、CRP値、総コレステロール値であった。過去に報告されているIVIG不応予測スコアの中で、総コレステロール値が含まれているスコアはないものの、近年の研究から総コレステロール値は川崎病の病態生理へ関与すると言われている。
- SHAPの結果から構築した新たなスコア(山梨スコア)は、過去に報告されているIVIG不応予測スコアと同等以上の予測精度であった。
- 多数の決定木に基づいたアンサンブル学習は、多重共線性を考慮する必要がなく、予測変数を大幅に増加させても解析が成立する。今後、解析規模を拡張することで、従来解析では予想しがたい川崎病の病態生理に関連する因子の発見、より有用な予測モデルの作成が期待される。