

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-1-03] 慢性血栓塞栓性肺高血圧症患者における低侵襲・高精度な肺動脈平均圧の予測モデル作成

*吉井 隆浩¹、松尾 秀俊²、高橋 真依²、西尾 瑞穂²、河野 淳²、谷口 悠³、平田 健一³、村上 卓道²（1. 神戸大学医学部医学科, 2. 神戸大学大学院医学研究科 内科系講座放射線診断学分野 放射線診断学部門, 3. 神戸大学大学院医学研究科 内科学講座 循環器内科学分野）

*Takahiro Yoshii¹, Hidetoshi Matsuo², Mai Takahashi², Mizuho Nishio², Atsushi Kono², Yu Taniguchi³, Kenichi Hirata³, Takamichi Murakami²（1. 神戸大学医学部医学科, 2. 神戸大学大学院医学研究科 内科系講座放射線診断学分野 放射線診断学部門, 3. 神戸大学大学院医学研究科 内科学講座 循環器内科学分野）

キーワード：Machine Learning, Regression, Right Heart Catheterization, CTEPH, Mean Pulmonary Artery Pressure

【背景・目的】慢性血栓塞栓性肺高血圧症(CTEPH)の診断には右心カテーテル検査による肺動脈平均圧(m PAP)の測定が必要であり、予後予測にも用いられている。右心カテーテル検査は侵襲度が高いため、昨今は心エコー検査における三尖弁圧較差(TRPG)による推定値を用いて経過観察されている例もある。本研究では心エコー検査以外の低侵襲検査を加えることで、mPAPを低侵襲かつ高精度に予測できるモデルの作成を試みた。【方法】神戸大学医学部附属病院にて、2005年4月 から2020年1月 の間で、mPAPを測定した CTEPH患者のデータ(167症例)の mPAPについて後方視的に検討した。低侵襲に得られる対象患者の年齢・性別・血液検査所見(BNP値)・心エコー検査所見(TRPG)・胸部レントゲン検査所見(心胸郭比・右第2弓比・無血管野の有無)の計7項目を使用することとし、内部で勾配ブースティング決定木を使用する LightGBMを用いてモデルを作成した。対照群として、上記7項目を用いた重回帰モデル、TRPGのみを用いた単回帰モデルを作成し、予測精度を比較した。データは訓練用(80%)、評価用(20%)に分割し、mPAPの予測性能の評価には評価用データに対する決定係数を用いた。【結果】作成したモデルの mPAP予測性能(決定係数)は、TRPGのみを用いた単回帰モデルではそれぞれ0.288であった。TRPGを含む7項目を用いた重回帰モデルでは0.429であった。さらに、LightGBMモデルでは0.516であった。【結論】CTEPH患者の mPAPの予測モデルに関して、TRPG以外の低侵襲検査を加えることで、より高精度に予測可能なモデルの作成を行うことができた。さらに、症例データを元にした新たな説明変数の作成、また LightGBMを用いることで、さらなる精度の向上を実現することができた。

慢性血栓塞栓性肺高血圧症患者における 低侵襲・高精度な肺動脈平均圧の予測モデル作成

吉井 隆浩^{*1}、松尾 秀俊^{*2}、高橋 真依^{*2}、西尾 瑞穂^{*2}、河野 淳^{*2}、谷口 悠^{*3}、平田 健一^{*3}、村上 卓道^{*2}

^{*1} 神戸大学医学部医学科、

^{*2} 神戸大学大学院医学研究科 内科系講座放射線診断学分野 放射線診断学部門、

^{*3} 神戸大学大学院医学研究科 内科学講座 循環器内科学分野

Accurate prediction model of pulmonary artery mean pressure using minimally invasive examinations in chronic thromboembolic pulmonary hypertension patients

Takahiro Yoshii^{*1}, Hidetoshi Matsuo^{*2}, Mai Takahashi^{*2}, Mizuho Nishio^{*2},

Atsushi Kono^{*2}, Yu Taniguchi^{*3}, Kenichi Hirata^{*3}, Takamichi Murakami^{*2}

^{*1} Kobe University School of Medicine, Kobe, Japan,

^{*2} Department of Radiology, Kobe University Graduate School of Medicine, Kobe, Japan,

^{*3} Division of Cardiovascular Medicine, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe, Japan

[Background and purpose] Pulmonary artery mean pressure (mPAP) measurement by right heart catheterization is required for the diagnosis of chronic thromboembolic pulmonary hypertension (CTEPH) and is also used to predict the prognosis of CTEPH. However, due to the invasiveness of right heart catheterization, many patients are followed up using the tricuspid valve pressure gradient (TRPG) estimated by echocardiography. This study attempted to build a highly accurate model to predict mPAP by adding minimally invasive examinations other than echocardiography. [Methods] We retrospectively examined the mPAP of the data of CTEPH patients (167 cases) between April 2005 and January 2020 at Kobe University Hospital. A total of seven items were used in the study: age, gender, blood test finding (BNP), echocardiographic finding (TRPG), and chest radiographic findings (cardiothoracic ratio, right second arc ratio, and presence of avascular area) of the patients, and the model was built using multiple regression analysis. A single regression model using just TRPG was created as baseline, and the prediction accuracy was compared. The data were divided into training (80%) and test (20%) sets, and the coefficient of determination for the test set was used to evaluate the predictive performance of mPAP. [Results] The mPAP predictive performance (adjusted coefficient of determination) was 0.257 for the single regression model using only TRPG, and 0.304 for the multiple regression model using seven items, including TRPG. In order to further improve the results, a numerical transformation of explanatory variables and selection of explanatory variables were performed, and the result was 0.392. [Conclusion] By adding minimally invasive examinations other than TRPG to the prediction model of mPAP in CTEPH patients, we were able to create a model that can predict mPAP with higher accuracy. Further improvement in accuracy was achieved by numerical transformation and selection of explanatory variables.

Keywords: Machine Learning, Regression, Right Heart Catheterization, CTEPH, Mean Pulmonary Artery Pressure

1. 緒論

慢性血栓塞栓性肺高血圧症 (chronic thromboembolic pulmonary hypertension: CTEPH) は、肺動脈内の器質化血栓が原因となって肺動脈が狭窄・閉塞して肺高血圧症を慢性的に発症する疾患であり、予後不良の難病として知られている¹⁾。

肺高血圧症は右心カテーテル検査における肺動脈圧の平均値(mean PAP, 以後 mPAP)が 25mmHg を超える状態として定義されており²⁾、CTEPH はその原因の一つである。肺動脈平均圧(mPAP)は CTEPH の診断や重症度分類、予後の予測に用いられている他、手術適応基準としても知られており、本疾患において重要な指標の一つである³⁾。mPAP 測定には右心カテーテルが用いられるが、高コストで侵襲度が高く、感染症や穿刺部位の局所出血、その他死亡を含む重篤な有害事象が起こりうる高リスクな検査である³⁾。そこで、右心カテーテルを用いずに mPAP を非侵襲的に推定するために、心エコー検査や CT、MRI を用いた方法が研究されている⁴⁾。中でも心エコー検査は、右心カテーテルはもちろんのこと、CT、MRI と比較しても簡便で低コストである。

mPAP の推定には心エコー検査のカラードップラー法によ

る三尖弁圧較差(TRPG)を利用する方法が知られており、臨床現場でよく用いられている⁵⁾。しかし、心エコー検査単独での予測は不正確であり⁶⁾⁷⁾⁸⁾、低侵襲・簡便・低コストを維持しながら、高精度に mPAP を予測する方法が求められている。

そこで、本研究では、心エコー検査に加えて、患者の年齢・性別といった基本情報、胸部単純撮影検査や血液検査といった臨床現場において日常的に実施されている検査の所見を用いることにより低侵襲・簡便・低コストを維持した上で、高精度な予測モデルの作成を試みた。

2. 目的

本研究では、CTEPH 患者の mPAP 予測において、従来から用いられてきた心エコー検査に、他の低侵襲検査を加えることで、より高精度な予測モデルの作成を目的とする。

3. 方法

本研究は神戸大学医学部倫理委員会の承認(受付番号:B210112)の下で実施した。プログラム言語として Python 3.7.6 を使用し、Python の機械学習ライブラリとして scikit-learn 0.24.1 を用いた。

3.1 対象とデータ抽出

神戸大学医学部附属病院で CTEPH と診断された症例のうち、2005 年 4 月 28 日から 2020 年 1 月 27 日までの期間で、右心カテーテルで mPAP を測定した症例(167 症例)を対象とした。その内、欠損値を含む症例を除外し、最終的に 135 症例をモデルの作成に用いた。

3.2 説明変数と目的変数

説明変数として、患者基本情報(年齢、性別)、血液検査所見(脳性ナトリウム利尿ペプチド(BNP)値)、心エコー検査所見(TRPG)、胸部単純撮影検査所見(心胸郭比・右第 2 弓比・無血管野の有無)を用いた。胸部単純撮影検査の所見は放射線科医 1 名により測定された。

目的変数として、右心カテーテルで測定した mPAP を用いた。

3.3 データ分割

欠損値を含む症例を除外した合計 135 症例を、訓練用(80%)、評価用(20%)になるように分割した。試行毎に別の乱数シードを設定することで、訓練用と評価用の症例の組み合わせが毎回異なるようにした。

3.4 機械学習手法

TRPG のみを説明変数として作成したモデルでは単回帰分析、TRPG に他の低侵襲検査を加えて作成したモデルでは重回帰分析を用いた。

3.5 説明変数の数値変換と選択

説明変数の BNP 値と心胸郭比について、底をネイピア数 e とする対数に変換した。

説明変数の選択は、変数減少法により行なった。つまり、最初に全 7 変数(年齢、性別、BNP 値、TRPG、心胸郭比、右第 2 弓比、無血管野の有無)を用いて重回帰分析モデルを作成し、その後、説明変数を影響度が低い方から 1 変数ずつ除外してモデルを作成し、精度が向上するかを調べた。モデルにおける説明変数の影響度の指標として、標準偏回帰係数の絶対値を用いた。標準偏回帰係数は、説明変数と目的変数を標準化して作成した重回帰分析モデルにおける各説明変数の偏回帰係数とした。乱数シードを変化させながら 100 回試行を行い、それぞれにおける予測モデルの標準偏回帰係数を算出してその平均を求め、その絶対値を最終的な結果とした。

3.6 評価

自由度調整済み決定係数をモデルの予測精度の指標とした。標準偏回帰係数の場合と同様に、乱数シードを変化させながら 100 回試行を行い、それぞれにおいて予測モデルを作成した。各モデルについて、評価用データの自由度調整済み決定係数を算出し、全モデル 100 回分の平均を最終的な結果とした。

4. 結果

4.1 患者背景

モデルの作成に用いた 135 症例(欠損値を含む症例は除外)における説明変数と目的変数の統計値を求めた(表 1)。

表 1 症例の統計値

性別 (男性 %)	26.7
年齢 (year, mean(SD))	66.2 (13.0)
BNP (pg/mL, mean(SD))	265.5 (554.0)
TRPG (mmHg, mean(SD))	58.2 (18.2)
心胸郭比 (mean(SD))	0.54 (0.07)
右第 2 弓比(mean(SD))	0.16 (0.04)
無血管野の有無 (有 %)	1.5
mPAP (mmHg, mean(SD))	37.7 (10.2)

モデルの作成に用いた 135 症例の統計値を示す。量的変数は平均と標準偏差(SD)を、2 値の質的変数は一方の割合を記載した。

4.2 説明変数 7 変数を用いた予測

説明変数の全 7 変数(年齢、性別、BNP 値、TRPG、心胸郭比、右第 2 弓比、無血管野の有無)を用いて、mPAP の予測モデルを作成したところ、自由度調整済み決定係数は 0.304 となり、TRPG のみを用いて作成した予測モデルの自由度調整済み決定係数 0.257 に比べ、大きかった(表 2)。

4.3 説明変数の数値変換

さらに精度を向上させるために、BNP 値と心胸郭比を対数変換したところ、自由度調整済み決定係数は 0.357 となり、対数変換前のモデル(自由度調整済み決定係数 0.304)から上昇した(表 2)。

4.4 各説明変数の影響度

数値変換後の説明変数の全 7 変数を用いた予測モデルにおいて、各説明変数の影響度を比較するために、標準偏回帰係数の絶対値を求めた(図 1)。標準偏回帰係数の絶対値が大きい(目的変数の予測における影響度が高い)方から順に、TRPG、年齢、BNP 値、心胸郭比、右第 2 弓比、無血管野の有無、性別であった。

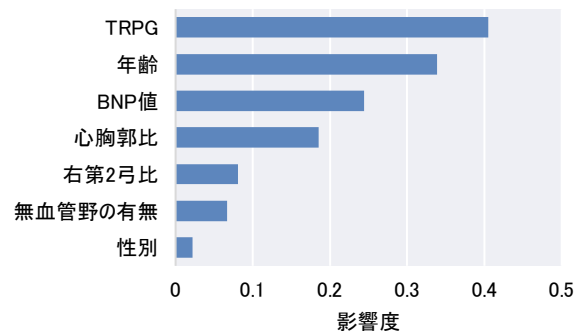


図 1 予測における説明変数の影響度

数値変換後の説明変数の影響度を示す。標準偏回帰係数の絶対値を説明変数の影響度とした。

4.5 説明変数の選択

説明変数の全 7 変数(数値変換後)を用いた予測モデルから、標準偏回帰係数の絶対値(目的変数の予測における影響度)が低い方から 1 変数ずつ除外してモデルを作成し、精度が向上するか調べた。その結果、性別、無血管野の有無、右第 2 弓比、心胸郭比までの除外では、自由度調整済み決定係数は上昇したが、さらに BNP 値を除外すると低下した。標準偏回帰係数の絶対値が BNP 値以上の 3 変数(TRPG、年齢、BNP 値)を用いて作成した重回帰分析モデルにおける

自由度調整済み決定係数は 0.392 となった(表 2)。

表 2 mPAP の予測精度

説明変数の種類	自由度調整済み決定係数 (標準偏差)
TRPG 単独	0.257 (0.145)
7 変数(対数変換前)	0.304 (0.189)
7 変数(対数変換後)	0.357 (0.164)
3 変数	0.392 (0.148)

自由度調整済み決定係数の平均と標準偏差を示す。7 変数は、年齢、性別、BNP 値、TRPG、心胸郭比、右第 2 弓比、無血管野の有無を指す。対数変換後は 7 変数のうち BNP 値、心胸郭比を対数変換したものである。3 変数は、TRPG、年齢、BNP 値を指す。

5. 考察

本研究では、神戸大学医学部附属病院で CTEPH と診断された症例におけるデータ 7 項目(年齢、性別、BNP 値、TRPG、心胸郭比、右第 2 弓比、無血管野の有無)を用いることで、TRPG 単独での予測よりも高精度な予測モデルの作成が可能となった。さらに、説明変数の数値変換と選択を行うことで、さらなる精度の向上を実現できた。これにより、低侵襲・簡便・低コストを維持した上で、精度向上を実現することができた。

精度向上の要因は、追加した変数が CTEPH の病態である肺動脈の狭窄・閉塞と関連が大きかったことが考えられる。特に、TRPG、年齢、BNP 値の 3 変数は mPAP の予測の精度向上に大きく寄与していた。TRPG はガイドラインにおいて重症度分類の基準の一つであり¹⁾、mPAP を推定する上で従来から用いられてきた変数である⁵⁾。BNP 値は、心室への負荷に応じて上昇し、心不全の病態を反映する指標であり、mPAP との相関が示唆されている⁹⁾。本研究に於いては年齢が大きな影響をもつ一方で、性別の影響は余り見られなかった。

今後のさらなる精度向上のためには、説明変数に新たに検査データを加える必要がある。低侵襲・簡便・低コストを維持した上での検査として、血液検査所見の D-ダイマーと CRP、またパルスオキシメーターを用いた経皮的動脈血酸素飽和度などが候補として考えられる。既に D-ダイマーと CRP は CTEPH 患者における有意な予後予測因子であるという研究があり¹⁰⁾、動脈血酸素飽和度は肺動脈性肺高血圧症発症の予測因子として有効であるという研究がある¹¹⁾。D-ダイマーと CRP は血液検査項目として臨床において日常的に計測されており、パルスオキシメーターによる経皮的動脈血酸素飽和度の測定も簡便に行うことが出来る。これらの検査データを説明変数に追加することで、さらなる予測精度が得られる可能性がある。

本研究の限界として、データが全て同一施設に由来するものである点が挙げられる。一方で、CTEPH は患者数が少ない疾患であり、平成 28 年度の指定難病患者数は 3200 人である¹²⁾。そのため、症例の収集は困難であるが、今後実用化を見据えた場合、他施設でのさらなる検証は必要と考えられる。

6. 結論

本研究により、CTEPH 患者の mPAP の予測モデルに関して、TRPG 以外の低侵襲検査(年齢、性別、BNP 値、TRPG、心胸郭比、右第 2 弓比、無血管野の有無)を加えることで、より高精度に予測可能なモデルの作成を行うことができた。さらに、説明変数の数値変換と選択を行うことで、さらなる精度の

向上を実現できた。

本機械学習モデルは CTEPH 患者の mPAP の予測に有用となる可能性が考えられた。

参考文献

- 1) 肺血栓塞栓症および深部静脈血栓症の診断、治療、予防に関するガイドライン(2017 年改訂版). 日本循環器学会, 2018. [https://j-circ.or.jp/old/guideline/pdf/JCS2017_ito_h.pdf (cited 2021-Aug-5)]
- 2) 肺高血圧症治療ガイドライン(2017 年改訂版). 日本循環器学会, 2018. [https://www.j-circ.or.jp/old/guideline/pdf/JCS2017_fukuda_h.pdf (cited 2021-Aug-24)]
- 3) Marius M Hoeper, Stephen H Lee, Robert Voswinckel, Massimiliano Palazzini, Xavier Jais, Alessandro Marinelli, et al. Complications of right heart catheterization procedures in patients with pulmonary hypertension in experienced centers. J Am Coll Cardiol. 2006; 48 (12):2546-52.
- 4) Waqas Ullah, Artem Minalyan, Sameer Saleem, Nayab Nadeem, Hafez M Abdullah, Abdelmohaymin Abdalla, et al. Comparative accuracy of non-invasive imaging versus right heart catheterization for the diagnosis of pulmonary hypertension: A systematic review and meta-analysis. Int J Cardiol Heart Vasc. 2020; 29:100568.
- 5) Naoko Sawada, Takayuki Kawata, Masao Daimon, Tomoko Nakao, Masaru Hatano, Hisataka Maki, et al. Detection of Pulmonary Hypertension with Systolic Pressure Estimated by Doppler Echocardiography. Int Heart J. 2019; 60 (4):836-844.
- 6) Nazzareno Galiè, Marius M Hoeper, Marc Humbert, Adam Torbicki, Jean-Luc Vachiery, Joan Albert Barbera, et al. Guidelines for the diagnosis and treatment of pulmonary hypertension: the Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS), endorsed by the International Society of Heart and Lung Transplantation (ISHLT). Eur Heart J. 2009; 30 (20):2493-537.
- 7) Micah R Fisher 1, Paul R Forfia, Elzbieta Chamara, Traci Houston-Harris, Hunter C Champion, Reda E Girgis, et al. Accuracy of Doppler echocardiography in the hemodynamic assessment of pulmonary hypertension. Am J Respir Crit Care Med. 2009; 179 (7):615-21.
- 8) Jonathan D Rich, Sanjiv J Shah, Rajiv S Swamy, Anna Kamp, Stuart Rich. Inaccuracy of Doppler echocardiographic estimates of pulmonary artery pressures in patients with pulmonary hypertension: implications for clinical practice. Chest. 2011; 139 (5):988-993.
- 9) Hidenobu Takagi, Hideki Ota, Koichiro Sugimura, Katharina Otani, Junya Tominaga, Tatsuo Aoki, et al. Dual-energy CT to estimate clinical severity of chronic thromboembolic pulmonary hypertension: Comparison with invasive right heart catheterization. Eur J Radiol. 2016; 85 (9):1574-80.
- 10) Nika Skoro-Sajer, Christian Gerges, Mario Gerges, Adelheid Panzenböck, Johannes Jakowitsch, Annabella Kurz, et al. Usefulness of thrombosis and inflammation biomarkers in chronic thromboembolic pulmonary hypertension-sampling plasma and surgical specimens. J Heart Lung Transplant. 2018; 37 (9):1067-1074.
- 11) Han Ki Park, Hong Ju Shin, Young Hwan Park, Bo Gyoung Ma. The importance of preoperative oxygen saturation as a predictor of pulmonary arterial hypertension after surgery of atrial septal defects. Interact Cardiovasc Thorac Surg. 2016; 23 (3):424-30.

- 12) 慢性血栓塞栓性肺高血圧症(CTEPH)診療ガイドライン. 日本肺高血圧・肺循環学会, 2018.
[http://jpcphs.org/pdf/guideline/cteph_guideline.pdf (cited 2021-Aug-19)]