

一般口演 | 医療支援

一般口演2 医療支援

2021年11月19日(金) 09:10 ~ 11:10 F会場 (2号館2階224)

[2-F-1-06] 院内感染症診断支援ツール開発に向けた AIモデルの構築

*齋藤 翔¹、早川 佳代子^{1,2}、石井 雅通³、山元 佳¹、野本 英俊¹、田島 太一²、松永 展明²、相馬 健人⁴、小戸 司⁴、美代 賢吾³、大曲 貴夫^{1,2} (1. 国立国際医療研究センター 国際感染症センター, 2. 国立国際医療研究センター AMR臨床リファレンスセンター, 3. 国立国際医療研究センター 医療情報基盤センター, 4. NECソリューションイノベータ株式会社 医療ソリューション事業部)

*Sho Saito¹, Kayoko Hayakawa^{1,2}, Masamichi Ishii³, Kei Yamamoto¹, Hidetoshi Nomoto¹, Taichi Tajima², Nobuaki Matsunaga², Kento Soma⁴, Tsukasa Odo⁴, Kengo Miyo³, Norio Ohmagari^{1,2} (1. 国立国際医療研究センター 国際感染症センター, 2. 国立国際医療研究センター AMR臨床リファレンスセンター, 3. 国立国際医療研究センター 医療情報基盤センター, 4. NECソリューションイノベータ株式会社 医療ソリューション事業部)

キーワード : Nosocomial infection, Diagnosis, Artificial intelligence

【背景】 院内感染症対策は医療施設の重要なテーマであり、不適切な対応は予後悪化に繋がる。本研究は院内感染症の約半数を占めるカテーテル関連血流感染症（CRBSI）と尿路感染症（UTI）に対する診断支援ツール開発のためのAIモデル構築を目的とした。【方法】 2010年9月から2018年12月における当センターのData Warehouse、医事データ、感染症データを収集した。CRBSIの解析用コホートを「中心静脈カテーテル挿入後2日以上が経過している入院3日目以降」かつ「血液培養採取から5日以内に37.5℃以上の発熱がある」患者とし、患者群を「血液培養陽性かつ確定診断のついた患者」とした。またUTIの解析用コホートは「入院3日目以降」かつ「尿培養採取から1日以内に37.5℃以上の発熱がある」患者とし、患者群を設定した。電子カルテ上の患者基本情報、検査値、看護記録などを説明変数として用いた。学習用データと評価用データの割合は8:2とし、NECのAIエンジンである異種混合学習技術を用いて解析した。【結果】 CRBSIの解析には150例の患者群、185例の対照群から患者群と同件数となるように無作為に抽出した150例を用い、正解率80.0%、適合率76.4%、再現率86.6%、F値81.2であった。感染の特徴量として体温、消化器官用薬、入院期間、白血球数が挙げられた。またUTIは78例の患者群、397例の対照群から78例を解析に用い正解率81.2%、適合率85.7%、再現率75.0%、F値80.0であった。感染の特徴量として年代、SpO2値、呼吸数、心拍数、体温が挙げられた。【考察】 患者群と対照群が同数であり実際の症例頻度と異なる Limitationはあるが、本研究のモデルをもとに院内感染症の診断支援ツール開発を進める。【結語】 院内感染症を診断するためのAIモデルを構築した。

院内感染症診断支援ツール開発に向けた AI モデルの構築

齋藤翔^{*1}、早川佳代子^{*1,2}、石井雅通^{*3}、山元佳^{*1}、野本英俊^{*1}、田島太一^{*2}、松永展明^{*2}、相馬健人^{*4}、小戸司^{*4}、美代賢吾^{*3}、大曲貴夫^{*1,2}

*1 国立国際医療研究センター 国際感染症センター、*2 国立国際医療研究センター AMR 臨床リファレンスセンター、*3 国立国際医療研究センター 医療情報基盤センター、*4 NEC ソリューションイノベータ株式会社 医療ソリューション事業部

Construction of Artificial Intelligence model for development of nosocomial infection diagnosis support tool

Sho Saito^{*1}, Kayoko Hayakawa^{*1,2}, Masamichi Ishii^{*3}, Kei Yamamoto^{*1}, Hidetoshi Nomoto^{*1}, Taichi Tajima^{*2}, Nobuaki Matsunaga^{*2}, Kento Soma^{*4}, Tsukasa Odo^{*4}, Kengo Miyo^{*3}, Norio Ohmagari^{*1,2}

*1 Disease Control and Prevention Center, National Centre for Global Health and Medicine, *2 AMR Clinical Reference Center, National Centre for Global Health and Medicine, *3 Healthcare Intelligence Technology Center, National Centre for Global Health and Medicine, *4 Medical Solution Business Unit, NEC Solution Innovators

Catheter-related bloodstream infection (CRBSI) and urinary tract infection (UTI), the leading causes of nosocomial infections, may worsen patient outcomes if improperly treated. The single-center retrospective cohort study explored diagnostic models using heterogeneous mixture learning technologies with data from 150 patients with CRBSI and 150 controls, and 78 patients with UTI and 78 controls. A model of CRBSI achieved 80.0% accuracy, 76.4% precision, 86.6% recall, and the 81.2 f-measure; a model of UTI has 81.2%, 85.7%, 75.0%, and 80.0, respectively. The features of CRBSI were body temperature, number of hospital visits, neutrophils, length of hospital stay, gastrointestinal medications, cardiovascular medications, and vitamin preparations; The UTI were age, sex, degree of requiring care, and stool frequency, oxygen saturation, C-reactive protein, and respiration rate. The results suggest that the models built in the study can be utilized for developing diagnostic support tools for nosocomial infections.

Keywords: Nosocomial infection, Diagnosis, Artificial intelligence

1. 緒論

医療機関において入院から 48 時間以降に発生した感染症は院内感染症と定義され¹⁾、多くの病院にとって対策が必要な重要な課題である。院内感染症に対しては感染症専門スタッフが対応することが理想であるが、各施設の人員は限られており、すべての病院が感染症専門スタッフを有しているわけではないため、現実的には困難である。また感染症は症状の進行が速く、迅速な方針決定を必要とするため他の医療機関へのコンサルテーションが困難であることも多い。診断が適切に行われず、不適切な治療が行われた場合、患者予後が悪化する可能性があるため院内感染症の診断支援ツールは多くの医療従事者の不安と負担を軽減し、より適切な医療を患者に提供できる可能性を有している。

2. 目的

本研究は院内感染症の約半数を占めるカテーテル関連血流感染症 (Catheter related blood stream infection: CRBSI) と尿路感染症 (Urinary tract infection: UTI) に対する診断支援ツール開発のための AI モデル構築を目的とした²⁾

3. 方法

CRBSI、UTI それぞれの対象患者を定義し、対象患者のデータを収集し加工したデータを用いて機械学習で各感染症の有無を判定するモデルを構築した

3.1. 対象患者

本研究は 749 床を有する感染症指定医療機関である国立国際医療研究センター (National Center for Global Health and Medicine: NCGM) に 2010 年 9 月から 2018 年 12 月にかけて入院した患者を対象に単施設後ろ向きコホート研究とし

て実施した。すべての患者データは解析前に匿名化をおこなひ、倫理委員会の承認を得て実施している (NCGM-G-003261-01)。

3.2. データの抽出

本研究に用いるデータは NCGM における Data Warehouse、医事データ、感染症サーベイランスデータから抽出し収集した。CRBSI の解析用コホートを「中心静脈カテーテル挿入後 2 日以上が経過している入院 3 日目以降」かつ「血液培養採取から 5 日以内に 37.5°C 以上の発熱がある」患者とし、患者群を「血液培養陽性かつ感染症サーベイランスデータにおいて CRBSI の確定診断のついた」患者とし、対照群を「中心静脈カテーテルから血液培養採取がある、解析用コホートにおける患者群以外」の患者とした。また UTI の解析用コホートは「入院 3 日目以降」かつ「尿培養採取から 1 日以内に 37.5°C 以上の発熱がある」患者とし、患者群を「入院中に尿路感染症の病名 (ICD10 コード N390 に準じる) が登録され、肺炎 ((ICD10 コード J690, J18 に準じる) の病名が登録されていない) かつ「尿培養結果で尿路感染症の起原菌として頻度の高い菌が 105CFU/ml 以上の菌が検出された」患者、対照群を「入院中に尿路感染症の病名が登録されていない」かつ「尿培養結果が 103CFU/ml 未満または陰性」の患者とした。尿路感染症の起原菌の判定は感染症専門医が行った。

3.3. 活用した説明変数

CRBSI、UTI においてそれぞれに関与すると考えられる説明変数を収集し、説明変数としてデータを加工した。

3.3.1. CRBSI における活用データ

CRBSI 解析のための説明変数は、患者基本 (性別、年齢、生活習慣)、入院期間、薬剤 (各薬剤処方回数、注射回数)、

歯科処置、検体検査(血液、生化学)、看護実施情報(バイタルサイン)を収集した。本研究では、中心静脈カテーテル挿入前後における各特徴量と CRBSI との関係性を解析するために、下記の 4 期間に分けて各期間における統計量(平均値、最大値、最小値、分散値)を算出し説明変数として活用した(表1)。

表 1:CRBSI 解析に用いた項目

期間	詳細	項目
期間 1	入院時	患者基本情報
期間 2	入院から中心静脈カテーテル挿入まで	薬剤(各薬剤処方回数、注射回数)、 歯科処置、 検体検査値、 看護実施情報
期間 3	中心静脈カテーテル挿入から血液培養採取日(感染を疑った日)まで	検体検査値、 看護実施情報
期間 4	血液培養採取日(感染を疑った日)	看護実施情報

また、期間 4 における看護実施情報のうち体温、SPO2、呼吸数、収縮期血圧、心拍数を収集した後に、AI の解釈性を高めるために National early warning score(NEWS)³⁾を基にカテゴリ化して解析に活用した。NEWS をもとにしたカテゴリ化ルールを表 2 に示す。

3.3.2. UTI における活用データ

UTI 解析のための説明変数は、入院から尿培養採取日(感染を疑った日)における患者基本(性別、年齢、看護度)、入院期間、検体検査(血液、生化学)、看護実施(食事、便排出量)情報を収集した後に統計量(平均値、最大値、最小値、分散値)を算出し活用した。看護度に関しては、入院時に生活の自由度を基にカテゴリ化されたものを説明変数として解析に用いた。看護度の分類を表 3 に示す。

表 3:看護度の分類(生活の自由度)

カテゴリ	生活の自由度
I	常に寝たまま
II	ベッドで身体を起せる(自力で身体を起こせる)
III	病室内歩行ができる
IV	日常生活は、ほとんど不自由がない

UTI において、下記の 3 期間に分けて各期間における統計量(平均値、最大値、最小値、分散値)を算出し説明変数として活用した(表 4)。また、期間 3'の看護実施情報を、NEWS をもとにしたカテゴリ化ルール(表 2)を基にカテゴリ化して解析に用いた。

表 4:UTI 解析に用いた項目

期間	詳細	項目
期間 1'	入院時	患者基本情報
期間 2'	入院から尿培養採取日(感染を疑った日)	検体検査値、 看護実施情報

	まで	
期間 3'	尿培養採取日(感染を疑った日)	看護実施情報

3.4. データの前処理

データに欠損値が存在する場合には、中央値により補間処理を行った。また、欠損率が 40%を超える項目に関しては活用対象外とした。

検査結果値などの異常値に関しては、データの分布から外れ値に対して、臨床上存在し得ない(死に至っている等)のみ、誤入力データとして除外処理を行った。

3.5. 活用した機械学習ツール

本研究では、異種混合学習技術を用いて異種混合予測モデルを探索した⁴⁾。異種混合予測モデルとは、入力データを決定木式のルールによって場合分けし、各場合で異なる説明変数を組み合わせた線形モデルで予測するモデルである。異なる説明変数の組み合わせ(異種)による予測モデルを組み合わせて(混合)予測する。このモデルでは、予測式の組み合わせ、説明変数の組み合わせ、データの場合分けを行う。異種混合学習技術とは、因子化情報基準と因子化漸近ベイズ推論による異種混合モデルの学習技術の総称である。因子化情報基準は、予測モデルの良さを判定する。一般的には、赤池情報基準、ベイズ情報基準が用いられるが、これらは単一のモデルで予測する。異種混合モデルは、複数のモデルを切り替えながら予測することから、独自の情報基準を導出している。因子化情報基準は、予測精度が低い、あるいはデータの場合分けや説明変数の組み合わせが複雑なモデルは値が低くなり、シンプルで予測精度の高いモデルは値が高くなる特徴を持つ。また、因子化情報基準を最大化するモデルを探索するアルゴリズムは、因子化漸近アルゴリズムを使用する。これは、①データの場合分け条件の最適化、②説明変数の組み合わせの最適化、③不要な予測モデルの削除の3ステップを繰り返すことで因子化情報基準を最大にするモデルを探索するアルゴリズムである。このアルゴリズムを用いることで、各ステップで必ず因子化情報基準の値を改善し、高速に最適な異種混合予測モデルの発見が可能となる。

4. 結果

CRBSI の解析においてデータ件数は 150 例の感染症患者群と、185 例から無作為に抽出した対照群 150 例を用いて解析した。UTI の解析においてデータ件数は 78 例の患者群と、397 例から無作為に抽出した対照群 78 例を用いて解析した。

4.1. 基礎分析結果

本研究で活用したデータに偏りの有無を確認するために、年齢と性別に対し基礎分析を行った。

4.1.1. CRBSI における基礎分析結果

CRBSI において患者基本情報である年齢に関して、平均年齢は CRBSI 患者群が 68.9 歳であったのに対し、対照群は 63.2 歳であった。棄却域を 5%として Welch の t 検定を両側で行った結果、 $p>0.025$ となり、年齢に有意差がないことが証明された。

また、性別に関して CRBSI 患者群と対照群のそれぞれの件数を表 5 に示す。独立性の検定のためカイ二乗検定を行った結果、カイ二乗値は 0.4926 であった。自由度 1 のカイ二乗値の上側 5%点 は 3.84 であるため、独立性は棄却されず、有意差はみられなかった。

表 5 CRBSI 分析における性別の基礎分析結果

	男性	女性
患者群	90	60
対照群	84	66

4.1.2. UTI における基礎分析結果

次に、UTI において患者基本情報である年齢に関して、平均年齢は感染症患者群が 73.4 歳であったのに対し、非感染症群は 68.3 歳であった。棄却域を 5%として Welch の t 検定を両側で行った結果、 $p>0.025$ となり、年齢に有意差がないことが証明された。また、性別に関して UTI 患者群と対照群のそれぞれの件数を表 6 に示す。独立性の検定のためカイ二乗検定を行った結果、カイ二乗値は 17.38 であった。自由度 1 のカイ二乗値の上側 5% 点は 3.84 であるため、独立性は棄却され、UTI 患者群は有意に女性が多いことが証明された。

表 6 UTI 分析における性別の基礎分析結果

	男性	女性
患者群	28	50
対照群	54	24

4.2. 解析モデル結果

予測モデルの学習と評価には、クロスバリデーション(CV)を用いた。学習用データと評価用データの割合は 8:2 とし、組み合わせを変えてモデルを構築し学習と評価を行った。

4.2.1. CRBSI 解析モデル結果

150 例の CRBSI 患者群と、対照群 150 例からそれぞれ 120 例を学習データとしてモデルを作成し、残りの 30 例で評価を行った。学習データの CV 平均として正解率 83.0%、適合率 81.9%、再現率 84.6%、F 値 83.2、最良モデルは正解率 86.2%、適合率 81.7%、再現率 93.3%、F 値 87.1 であった。最良モデルにおいて、AI が判別した結果と実際の結果との混合行列を表 7 に示す。

表 7 学習における AI 判別分析結果の混合行列

		実際の結果	
		CR BSI	非感染症
AI による判別結果	CR BSI	112	25
	非感染症	8	95

また、評価データの CV 平均は正解率 77.0%、適合率 76.1%、再現率 78.6%、F 値 77.2、最良モデルは正解率 80.0%、適合率 76.4%、再現率 86.6%、F 値 81.2 であった。最良モデルにおいて、AI が判別した結果と実際の結果との混合行列を表 8 に示す。

表 8 評価における AI 判別分析結果の混合行列

		実際の結果	
		CR BSI	非感染症
AI による判別結果	CR BSI	26	8
	非感染症	4	22

4.2.2. UTI 解析モデル結果

UTI の解析には 78 例の患者群と、397 例から無作為に抽出した対照群 78 例を用いた。学習データの CV 平均として正

解率 84.2%、適合率 84.5%、再現率 83.9%、F 値 84.2、最良モデルは正解率 83.8%、適合率 87.5%、再現率 79.0%、F 値 83.0 であった。最良モデルにおいて、AI が判別した結果と実際の結果との混合行列を表 9 に示す。

表 9 学習における AI 判別分析結果の混合行列

		実際の結果	
		UTI	非感染症
AI による判別結果	UTI	49	7
	非感染症	13	55

評価データの CV 平均は正解率 78.8%、適合率 82.3%、再現率 74.2%、F 値 77.6、最良モデルは正解率 81.2%、適合率 85.7%、再現率 75.0%、F 値 80.0 であった。最良モデルにおいて、AI が判別した結果と実際の結果との混合行列を表 10 に示す。

表 10 評価における AI 判別分析結果の混合行列

		実際の結果	
		UTI	非感染症
AI による判別結果	UTI	12	2
	非感染症	4	14

4.3. 解析モデルの特徴量

4.3.1. CRBSI 解析モデルの特徴量

感染の特徴量として体温、過去の外來回数、好中球、循環器官用剤、消化器官用剤、入院期間が挙げられた。異種混合学習予測モデルにより作成した、決定木式のルール内容と、各ルールに該当する学習データと評価データの件数を図 1 に示す。また、各場合において作成された線形モデルの詳細を表 11 に示す。ここで、表中の特徴量の強さは、正の値であれば感染症に影響し、負の値であれば非感染症に影響していることを意味する。

表 11 CRBSI 解析において作成した異種混合予測モデルの詳細

線形モデル	特徴量	期間	統計量	特徴量の強さ	
①	体温	期間 4	最大値	1.09	
	②	体温	期間 4	最大値	1.23
		過去の外來回数	-	-	-0.77
		好中球	期間 3	最小値	-1.11
		循環器官用剤	期間 2	平均値	1.74
		消化器官用剤	期間 2	平均値	0.87
③	ビタミン剤	期間 2	平均値	-1.22	
	④	過去の外來回数	-	-7.38	

	入院期間	期間 1	-	-8.26
	好中球	期間 1	最小値	6.12
④	消化器官 用薬	期間 1	平均値	40.36

4.3.2. UTI の解析結果

感染の特徴量として年代、性別、看護度、便回数、SpO2 値、CRP、呼吸数が挙げられた。異種混合学習予測モデルにより作成した、決定木式のルール内容と、各ルールに該当する学習データと評価データの件数を図 2 に示す。また、各場合において作成された線形モデルの詳細を表 12 に示す。

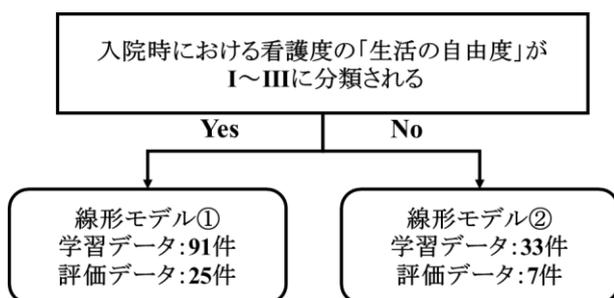


図 2: UTI 解析において作成した異種混合予測モデルの決定木式ルール

表 12 UTI 解析において作成した異種混合予測モデルの詳細

線形モデル	特徴量	期間	統計量	特徴量の強さ
①	年齢	-	-	1.02
	性別	-	-	2.47
	看護度	期間 1'	-	-1.97
	便回数	期間 2'	平均値	-1.41
	SPO2 値	期間 3'	最小値	1.69
	CRP	期間 2'	平均値	-0.86
②	呼吸数	期間 3'	最大値	0.81

5. 考察

本研究においては異種混合学習技術を用いて院内感染症の約半数を占める CRBSI と UTI の有無を判定するモデルを構築した。

CRBSI においては年齢、性別ともに患者群、対照群に差を認めず、UTI においては患者群、対照群に年齢差は認めず、女性の割合が有意に高かった。UTI は女性の発症率が高いため、患者コホートを反映していると考えられた⁵⁾。

体温、好中球数は CRBSI を発症した際には変動が大きい値であるため、特徴量として出現したと考えられる。またカルシウム値が決定木式ルールに組み込まれているが、カルシウム値の CRBSI への関与については今後詳細な検討が必要であると考えられる。UTI においては、患者の全身状態を表す看護度により決定木式ルールが構築され、UTI のリスクとして報告されている性別と年齢が特徴量に含まれている⁶⁾。ま

た腸管内の細菌が逆行性に膀胱や腎臓に到達し尿路感染症を引き起こすため、便回数はこれらの事象を反映していると考えられる⁷⁾。SpO2 値は尿路感染症と鑑別が必要となる肺炎において低下する値であるため、特徴量として出現した可能性がある。

Limitation として患者群と対照群が同数であり実際の症例頻度より高く設定されているため、正解率に影響を及ぼす可能性がある。しかし CRBSI, UTI ともに本研究における正解率は 80%を超えているため、本研究のモデルをもとに院内感染症の診断支援ツール開発をさらに進めることが可能であると考えられた。

6. 結論

院内感染症の診断支援ツールを開発するために有用であると考えられる AI モデルを構築した。今後、汎用性を高めることにより全国の医療従事者の負担を軽減し、より適切な医療を患者に提供するための院内感染症の診断支援ツールの開発が必要であると考えられる。

文献

- 1) Revelas A. Healthcare - associated infections: A public health problem. Niger Med J 2012; 53 (2) :59-64.
- 2) Weinstein RA. Nosocomial infection update. Emerging infectious diseases. 1998;4(3):416-20
- 3) Royal College of Physicians. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS. 2017 [https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2 (cited 2021-Aug-30)]
- 4) 藤巻 遼平, 森永 聡, 江藤 力, 本橋 洋介, 菅野 亨. 異種混合学習技術とビッグデータ分析ソリューションの研究開発. 2015 [http://www.fbi-award.jp/sentan/jusyou/2015/7.pdf (cited 2021 Aug 30)]
- 5) 山本 新吾, 石川 清仁, 速見 浩士ら. JAID/JSC 感染症治療ガイドライン2015 一尿路感染症・男性性器感染症一. 2016;90 (1): 1-30.
- 6) Centers for Disease Control and Prevention. Urinary Tract Infection. August 23, 2021 [https://www.cdc.gov/antibiotic-use/uti.html (cited 2021 Aug 30)]
- 7) Gupta K, Hooton TM, Naber KG, et al. International clinical practice guidelines for the treatment of acute uncomplicated cystitis and pyelonephritis in women: A 2010 update by the Infectious Diseases Society of America and the European Society for Microbiology and Infectious Diseases. Clin Infect Dis 2011;1;52(5):e103-20.

表 2 NEWS を活用した看護実施情報のカテゴリ化ルール

		カテゴリ						
		-3	-2	-1	0	1	2	3
看護 実施 情報	体温	35.0 以下		35.1~36.0	38.1~39.0	39.1 以上		
	SPO2	91 以下	92~93	94~95				
	呼吸数	8 以下		9~11		21~24	25 以上	
	収縮期血圧	90 以下	91~100	101~110				
	心拍数	40 以下		41~50	91~110	111~130	131 以上	

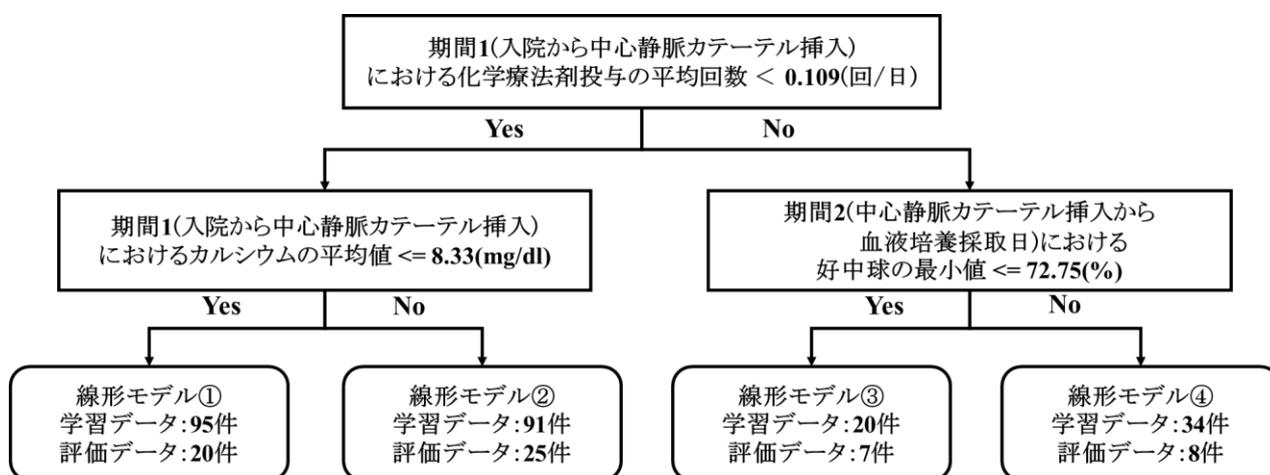


図 1: CRBSI解析において作成した異種混合予測モデルの決定木式ルール