



Title	構造が複雑なデータに対する統計モデルの性能比較：検査の診断性能とリスク比推定の評価への応用
Author(s)	稲尾, 翼
Description	配架番号：2974
Degree Grantor	北海道大学
Degree Name	博士(医学)
Dissertation Number	甲第16881号
Issue Date	2026-03-25
DOI	https://doi.org/10.14943/doctoral.k16881
Doc URL	https://hdl.handle.net/2115/99726
Type	doctoral thesis
File Information	INA0_Tasuku.pdf, 全文



学 位 論 文

構造が複雑なデータに対する統計モデルの性能比較:
検査の診断性能とリスク比推定の評価への応用

(Comparison of statistical model performance with
complex data structure: Evaluating diagnostic accuracy
and risk ratio estimation)

2026年3月

北 海 道 大 学

稲 尾 翼

Tasuku Inao

学 位 論 文

構造が複雑なデータに対する統計モデルの性能比較:
検査の診断性能とリスク比推定の評価への応用

(Comparison of statistical model performance with
complex data structure: Evaluating diagnostic accuracy
and risk ratio estimation)

2026年3月

北 海 道 大 学

稲 尾 翼

Tasuku Inao

目次

1	発表論文目録および学会発表目録	1
2	要旨	2
3	略語表	5
4	序文	6
5	陽性の頻度が稀な診断結果データに対するゴールドスタンダードな検査法がない場合の感度・特異度推定	8
5.1	序文	8
5.2	ベイズ流潜在クラスモデリング	10
5.3	シミュレーション実験 1: 事前情報の強さに対する感度・特異度の推定	12
5.3.1	実験 1 の設定	12
5.3.2	実験 1 の結果	13
5.4	シミュレーション実験 2: 感度の真値を与えた下での感度・特異度の推定	16
5.4.1	実験 2 の設定	16
5.4.2	実験 2 の結果	16
5.5	シミュレーション実験 3: 既存の頻度流手法との比較	19
5.5.1	実験 3 の設定	19
5.5.2	実験 3 の結果	19
5.6	考察	20
6	慢性心不全患者における死亡または入院イベントと動物性・植物性タンパク質摂取量の関連	25
6.1	序文	25
6.2	対象コホートの説明	26
6.3	稀でないイベント発生に対する回帰モデル	27
6.3.1	対数二項回帰と修正ポアソン回帰	27
6.3.2	回帰モデル間の性能比較 - シミュレーション実験の設定	29
6.3.3	回帰モデル間の性能比較 - シミュレーション実験の結果	30
6.4	解析の手法	31
6.5	結果	33
6.5.1	各タンパク質摂取量群に対するデータの特性	33

6.5.2	解析結果 1: 総エネルギー摂取量に占める特定のタンパク質の割合とイベントの関連	38
6.5.3	解析結果 2: タンパク質の摂取量とイベントの関連	46
6.6	考察	50
7	全体の結論	56
8	謝辞	58
9	利益相反	59
10	参考文献	60
11	付録	68
11.1	パラメーター空間を制限しない場合の MCMC の結果	68
11.2	実験 1 における他の有病率の設定での結果	68
11.3	実験 2 における他の有病率の設定での結果	68
11.4	比較手法の数理的背景	78
11.5	対数二項回帰と修正ポアソン回帰の性能比較の結果.	80
11.6	動物性・植物性タンパク質の摂取量の組み合わせごとの結果	80

1 発表論文目録および学会発表目録

List of publication

1. Inao, T., Okada, K., Yang, Y., & Yokota, I. Simulation studies on the estimation of sensitivity and specificity in the absence of a gold standard and with rare positive frequencies. *Communications in Statistics-Simulation and Computation*, 2024; 53(12): 5742-51.

List of presentation

1. Inao, T., Okada, K., Yang, Y., & Yokota, I. Simulation studies on the estimation of sensitivity and specificity in the absence of a gold standard and with rare positive frequencies. 39th Annual Research Meeting on the Japanese Classification Society, November 07-08, 2020, Kyoto, Japan.

2 要旨

[背景と目的]

多種多様なデータが取得される現代において、データの構造自体が複雑化しており、これに対処した統計的手法の性能評価が重要な課題となっている。特に臨床医学や公衆衛生の分野では、疾患の有無を正確に判別する「検査の診断性能評価」と、疾患の発症や予後を予測する「リスク推定値の評価」という二つの主要な応用領域において、どの統計モデルが最も妥当かつ正確な結論を導き出すかを厳密に検証することが必要不可欠である。本論文はこれら二つの応用領域に焦点を当て、構造が複雑なデータ環境下での統計モデルの性能を系統的に比較し、その知見を応用することを目的とする。

疾患の有無を判定する検査の診断性能評価において、完全に正確なゴールドスタンダードが存在しないか、倫理的・費用的理由から利用が困難な場合(有病率が低い集団スクリーニング等)が少なくない。このような状況下では、従来の検査法では感度・特異度の客観的かつ安定的な推定が困難であり、統計的妥当性に問題が残ることが指摘されている。第5章の研究(研究1)では、潜在変数を設定した確率モデルであるベイズ流潜在クラスモデルに着目し、特に陽性の頻度が稀な診断結果データに対する感度・特異度の推定性能について、事前分布の強さや有効サンプルサイズといった要因の影響を系統的に調査し、推定の安定性を評価することを行なった。

公衆衛生政策の取り組みとして重要視される客観的根拠に基づく意思決定は栄養疫学の分野でも進められているが、タンパク質摂取量と心不全イベントの関連に関する研究は総タンパク質に着目したものが主であり、組成が異なる動物性・植物性タンパク質に分けての関連性の詳細な検討は不十分である。また、心不全患者における死亡または入院イベントは比較的稀ではない(本データでは10%を超える)ため、単純なオッズ比ではなくリスク比を正確に推定するための回帰モデル(対数二項回帰、修正ポアソン回帰)を選定することが好ましく、そのモデル間の性能を比較することも必要である。第6章の研究(研究2)では、慢性心不全患者における動物性・植物性タンパク質の摂取量および総エネルギーに対する割合と死亡または入院イベントとの関連性について、リスク比を用いて正確に検討した。

[対象と方法]

研究1: 2種類の検査結果を対象とし、モデルの識別不可能性の課題に対しBeta分布を事前分布に設定したベイズ流潜在クラスモデルの推定方法を採用した。事前情報の強さ(有効サンプルサイズ N_{ESS}), 事前情報を与えるパラメータ, 有病率の真値(0.5%, 10%など), 感度の真値を変えた3種類のシミュレーション実験を設定した。推定には、パラメータ空間の制限した下でのマルコフ連鎖モンテカルロ法(MCMC)のブロック化メト

ロポリス・ヘイスティング (M-H) アルゴリズムを適用した。評価指標は、MCMC の収束割合、絶対バイアス、平均二乗誤差の平方根 (RMSE) とし、既存の頻度流手法 (Hui-Walter モデル, Lu らのモデル) とも比較した。

研究 2: 国内 10 施設に登録された慢性心不全患者 146 名の栄養疫学コホートデータを用いた。アウトカムは「追跡開始から 1 年以内の全死亡または全入院イベント」(イベント All) および「心血管死亡または心不全増悪による入院イベント」(イベント CHF) の 2 種類とした。解析に先立ち、イベント発生割合が稀ではないことから、リスク比を推定する対数二項回帰と修正ポアソン回帰の性能を比較するシミュレーションを行った。実データ解析では、欠測値を多重代入法で補完し、交差検証法による AIC, AUC に基づくモデル選択結果から調整変数を決定した。各種タンパク質 (総タンパク質, 動物性, 植物性) の摂取量および総エネルギーに対する割合とイベントとの関連性について、制限つき三次スプライン曲線で描画し、その傾きから高摂取群と低摂取群の閾値を設定した。最終モデルに修正ポアソン回帰を適用し、リスク比と 95% 信頼区間を推定した。

[結果と考察]

研究 1: ベイズ流潜在クラスモデルは、従来の頻度流手法と比較して、陽性の頻度が稀な状況 (有病率 0.5%) でも推定値が計算できた割合が高く、RMSE も小さく、推定の安定性で優れていた。MCMC の収束には N_{ESS} が 100 以上であることが望ましいことが示唆された。感度 Se2 の推定において、もう一方の検査の感度 Se1 のみに情報を入れる戦略 (シナリオ 1-2) では N_{ESS} をどれだけ大きくしても絶対バイアスが改善されないという特徴的な振る舞いが確認された。一方、Se1 と特異度 Sp1 の両方に事前情報を設定する戦略 (シナリオ 3) では、 N_{ESS} の上昇に伴い絶対バイアスがゼロに近づいた。有病率が低い状況では感度が過小評価される傾向があり、バイアスをゼロに近づけるためには適切な事前情報の設定が極めて重要であることが示された。

研究 2: 回帰モデルの性能比較シミュレーションの結果、修正ポアソン回帰は、イベント発生割合、共変量間の相関、リスク比の真値に関わらず、対数二項回帰よりも収束割合や推定値のバイアス・RMSE の点で優れており、小標本かつ共変量間に相関がある状況でも頑健な性質を持つことが示された。栄養疫学解析の結果、総エネルギーに占めるタンパク質の割合の観点では、動物性タンパク質の割合が増加するとイベントリスクが上昇傾向にあり、特に高摂取割合群 (10.3% 以上) ではイベント CHF のリスク比が約 1.81~1.93 倍と高い傾向が示された。対照的に、植物性タンパク質の割合が増加するとリスクが下降傾向にあり、低摂取割合群 (6.0% 未満) ではイベント CHF のリスク比が約 1.35~1.41 倍と高くなる傾向が示された。また、動物性タンパク質の摂取割合と炭水化物摂取割合には負の相関が認められ、タンパク質の種類の選択が他の 3 大栄養素の摂取バランスと関連していることが示唆された。

[結論]

研究 1: ゴールドスタンダードがない状況での 2 種類の検査の診断精度推定において、ベイズ流潜在クラスモデルは従来の頻度論的手法と比較して推定値の収束割合が高く、推定結果が安定しているという利点を示した。特に、陽性の頻度が稀な疾患を想定した場合、ベイズ流潜在クラスモデルによる推定を推奨する根拠を提供する。ただし、もう一方の検査の感度のみで事前情報を入れる戦略はバイアスが改善されないため避けるべきである。この知見は、今後、別の分野においてゴールドスタンダードがない状況での検査結果しか計測できない場合においてもベイズ流潜在クラスモデルの使用を推奨する一助となる。

研究 2: 広範なシミュレーションを通じて、修正ポアソン回帰がリスク比推定において対数二項回帰よりも優れていることを実証した。慢性心不全患者コホートの栄養疫学データ解析では、総エネルギーにおける動物性タンパク質の割合が多いと心血管死・心不全入院イベントのリスクが高くなり、植物性タンパク質の割合が大きいとリスクが低くなる可能性が示された。この結果は、心不全患者に対する栄養指導で、タンパク質の中でも特に植物性タンパク質の割合を増やすといった明瞭な指導を提供する根拠となり得る。

本研究で得られた知見は、複雑なデータからより信頼性の高い科学的結論を導くための具体的な指針を提供し、医学研究の発展に寄与すると考えられる。

3 略語表

本文中および図中で使用した略語は以下のとおりである.

略語	正式名称
ACE	Angiotensin-converting enzyme
AIC	Akaike information criterion
ARB	Angiotensin II receptor blocker
AUC	Area under the curve
BDHQ	Brief-type self-administered diet history questionnaire
BMI	Body mass index
BNP	B-type natriuretic peptide
COVID-19	Coronavirus disease 2019
CP	Convergence of probability
eGFR	estimated Glomerular filtration rate
ESS	Effective sample size
Hb	Hemoglobin
HbA1c	Hemoglobin A1c
HFpEF	Heart Failure with preserved Ejection Fraction
HFrfEF	Heart Failure with reduced Ejection Fraction
hs-CRP	high-sensitivity C-reactive protein
IQR	Interquartile Range
LOOCV	Leave-One-Out Cross-Validation
LVEF	Left ventricular ejection fraction
MCMC	Markov Chain Monte Carlo
MLHFQ	Minnesota Living with Heart Failure Questionnaire
MRA	Mineralocorticoid receptor antagonist
MSE	Mean squared error
NYHA	New York Heart Association
PHQ-9	Patient Health Questionnaire-9
QOL	Quality of life
RMSE	Root mean squared error
ROC	Receiver operating characteristic

4 序文

多種多様なデータが取得されるようになり、データの構造自体が複雑になっているケースがある。これに対処した統計的手法が開発される中で、特に臨床医学や公衆衛生の分野では、二つの応用領域において、モデルの性能評価が不可欠となっている。一つは、疾患の有無を正確に判別するための「検査の診断性能評価」であり、もう一つは疾患の発症や予後を予測するための「リスク推定値の評価」である。これらの領域では、わずかな推定誤差やモデルの選択ミスが、患者の診断や治療方針に影響を及ぼす。そのため、複雑な構造を持つデータに対して、どの統計モデルが最も妥当かつ正確な結論を導き出すかを厳密に比較・検証することが、統計学的方法論における重要な課題である。高度な解析手法の適用が必ずしも最善の結果をもたらすわけではない。統計モデルを適用するに先立ち、データにおける変数間の関係性、欠測値の有無、データのスケールなどデータの構造をあらゆる角度から吟味する姿勢が、妥当な統計的結論を導くための前提となる。本論文はこれら二つの主要な応用領域に焦点を当て、構造が複雑なデータに対して統計モデルの性能を系統的に比較し、その知見を応用することを目的とするものである。

本稿では、「構造が複雑なデータに対する統計モデルの性能比較」というメインテーマに基づき、二つのテーマに取り組んだ成果を報告する。一つ目のテーマは、ゴールドスタンダードがない状況における診断精度の推定のための統計モデルの性能評価である。ゴールドスタンダードとは、「最も正確で信頼性が高い」と臨床医学や研究で広く認められている基準のことであり、疾患や状態が実際に存在するかどうかを判断するための絶対的な指標である。疾患有無のような二値の結果においては、ゴールドスタンダードがあるような状況では診断精度の評価として AUC を使うことで評価が可能である。しかし、十分な診断精度を持つ診断方法がない状況や実用性の観点から限定的な状況でしか使用できない場合では、ゴールドスタンダードを確立することは困難である。こうした状況は、患者からの自己申告が診断の要になっているような分野ではしばしば見られる問題である (Nelson et al., 2010; DeTure and Dickson, 2019)。ゴールドスタンダードな診断がないような状況でも決定論的に診断を行うために、複数の検査間で診断結果が異なった時にのみ追加の検査を行う「不一致分析 (discrepant analysis)」や複数の検査結果を組み合わせて一つの診断結果に統合する「複合参照基準 (composite reference standards)」といった方法がある (Rutjes et al., 2007; Walsh, 2018)。しかし、これらの方法は誤分類の問題があることに加え、統計的な妥当性については問題が残ることが指摘されている (Sun and Zhou, 2025)。そのため、本研究では潜在変数を設定した確率モデルである潜在クラスモデルを利用し、感度・特異度の推定を行う方針とした。

二つ目のテーマは、慢性心不全患者の死亡または入院イベントの発生と動物性・植物

性タンパク質の摂取量の関連性の検討である。公衆衛生政策の取り組みとして重要視されている考え方の一つに「Evidence-based decision making」という概念がある。これは客観的根拠に基づく意思決定のことであり、ここで提示された根拠をもとに公衆衛生政策の決定や新しい疫学研究の誕生を促す (Jenicek, 1997; Sackett, 1997)。栄養疫学の分野においてもこの概念に基づいて進められる研究があり、炭水化物や脂質といった栄養素の摂取量と特定の疾患に関するイベントとの関連性が調査されている (Franz et al., 2002; Willett, 2012; Apekey et al., 2022)。炭水化物や脂質と並ぶ三大栄養素であるタンパク質の摂取量と心不全イベントに関する研究もいくつかあるが、多くは総タンパク質摂取量に着目して研究されている (Streng et al., 2022; Saijo et al., 2023)。そのため、本研究では単に総タンパク質だけでなく、組成が異なる動物性タンパク質と植物性タンパク質に分けて、死亡または入院のイベント発生との関連性を詳細に検討した。栄養データは、多くの栄養素が互いに複雑に関連しあっており、統計解析においては全エネルギーに対する各種タンパク質の占める割合を回帰モデルに含めることで、栄養バランスの観点からイベント発生について考察する。さらにこの研究では、イベントの発生件数が比較的稀ではなく、単純なオッズ比ではなくリスク比を正確に推定するため、データ解析前の検討として候補となる回帰モデルの性能を推定値のバイアスなどから比較によって選定を行うことも検討する。

以上で紹介した二つの研究テーマは、根本において「構造が複雑なデータ環境下でどの統計モデルが高い推定性能を持つか」という共通の問いに基づいている。このとき、単にモデルの当てはまりの良さが高いだけでなく、「推定のバイアスが小さいこと」や「結果の解釈性」といった様々な観点から、最適なモデルを選択するための指針を確立することが極めて重要となる。一つ目のテーマでは、ゴールドスタンダードの不確実性という「観測誤差による複雑な構造」を持つ診断評価データに対し、異なる流派の潜在クラスモデルの推定性能を比較した。一方、二つ目のテーマでは、栄養データの「組成的な複雑性」とリスク比推定における「回帰モデルの選択の妥当性」に焦点を当て、推定値のバイアスを最小限に抑え、より正確な推定性能を持つモデルの選択規準を探索した。このように、本論文は二つの異なる臨床的課題を事例としつつも、その根幹において「統計モデルの性能比較」という統一的なテーマを追求し、データ構造の複雑さが統計的結論に与える影響を多角的に検証することを試みている。本論文で得られた知見は、臨床研究や疫学研究において、複雑なデータからより信頼性の高い科学的結論を導くための具体的な指針を提供するものである。研究者や臨床家が、直面する複雑なデータ構造に対して、どのような統計モデルを適用し、その結果をどのように解釈すべきかという実践的な問いに対して一つの指針を与える。

5 陽性の頻度が稀な診断結果データに対するゴールドスタンダードな検査法がない場合の感度・特異度推定

5.1 序文

医療現場における正確な診断は、適切な治療方針の決定や公衆衛生対策の立案に不可欠である。興味のある検査の診断評価は、真の疾患有無の状態あるいはそれらに限りなく近い「ゴールドスタンダード」と呼ばれる基準に則った結果と興味のある検査の診断結果の比較によって伝統的に行われてきた。例えば、疾患の有無を判定する二値検査では、感度と特異度を算出することで性能が評価される。これはゴールドスタンダードが存在することを前提としている。同様に、血糖値や腫瘍マーカーのような連続値によって判定する検査および疾患重症度カテゴリのような結果に順序性のあるカテゴリ値によって判定する検査の性能も ROC 曲線の曲線下面積である AUC を用いて評価されるが、その際もゴールドスタンダードとの比較が行われる。

一部の疾患や病態において、完全に正確なゴールドスタンダードは存在しないか、あるいは倫理的・費用的・技術的な理由から利用が困難な場合が少なくない。例として、新型コロナウイルス感染症 (COVID-19) における検査を取り上げる。2020 年 6 月から 7 月にかけて、新型コロナウイルス感染者の濃厚接触者と入国者を対象に唾液を用いた診断法の有用性を確かめる調査が全国の保健所および空港で実施された (Yokota et al., 2021a,b,c)。唾液は自己採取であるために、従来の鼻咽頭液診断法と比較して医療従事者の二次感染リスクが低くなる上に、検査対象者にとっても採取の負担が軽減されることが期待されていたことが調査の動機である。もし、鼻咽頭液を用いた診断法をゴールドスタンダードと見なすことができれば、唾液診断法の感度や特異度が計算可能なので、診断性能を測ることができる。しかし、曝露から発症までの時間差があるため各検査の時点で患者が感染していた確証はなく、COVID-19 を発症した患者の鼻咽頭液から新型コロナウイルスが検出されない場合がある。そのため、鼻咽頭液診断法の感度は 100% と断言できない。このような真の状態を示すゴールドスタンダードが存在しない環境下では、診断検査の感度および特異度を客観的かつ安定的に推定するための統計学的アプローチの確立が喫緊の課題である。

ゴールドスタンダードがない状況における診断性能の評価について、統計学的アプローチが複数提案されている (Sun and Zhou, 2025)。連続値を使った診断では ROC 曲線の曲線下面積 (AUC) によって性能を測られる場面が多く、その AUC を推定するための統計学的手法などが提案されている (Wang et al., 2006; Branscum et al., 2008; Ladouceur et al., 2011)。また、順序性のあるカテゴリによって診断される検査の場合にも変量効果モデルを利用した手法 (Albert, 2007; Kim et al., 2021) やノンパラメトリック解析によ

る手法などが存在する (Zhou et al., 2005; Fosgate et al., 2007). 当然ではあるが, 単純に疾患有無を判定するという目的で, 二値の検査に対する統計学的手法も多く研究されている. 先述の COVID-19 の検査のように, 複数の二値検査結果が得られた場合, 感度と特異度を潜在クラスモデルに基づいて推定することができる. 潜在クラスモデルを利用した先行研究 (Hui and Walter, 1980; Joseph et al., 1995) では, 潜在変数を条件として診断が検査間で独立であるという「条件付き独立性の仮定」を置いている. 図 1 は, COVID-19 検査における潜在クラスモデルを表している. この例では, 潜在変数 (感染有無) で分けたときに各状態に当てはまる確率が感度と特異度の単純な積で表現されており, この積で記述する表現が条件付き独立性である. 条件付き独立性は単純な表現で記述できる一方で, モデルの誤特定には頑健でないという脆弱性もある (Schofield et al., 2021). 条件付き独立性の仮定を置かない方法として, 二つ以上の潜在クラスを用いる方法 (Rindskopf and Rindskopf, 1986), 変量効果を用いる方法 (Qu et al., 1996), 階層モデルを用いる方法 (Dendukuri and Joseph, 2001; Dendukuri et al., 2009; Wang et al., 2020) なども提唱されている.

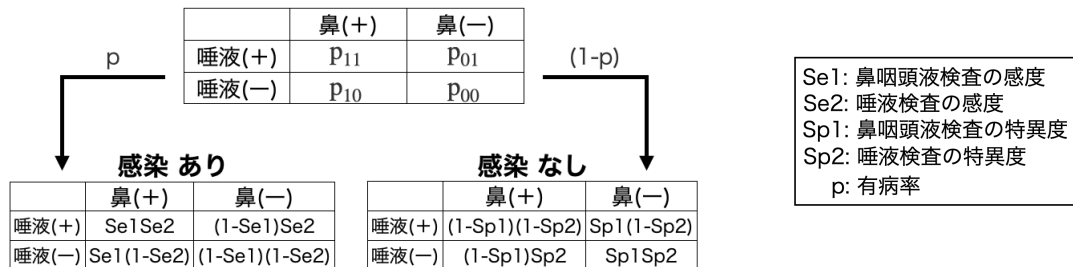


図 1: 条件付き独立性の仮定を置いた場合の鼻咽頭液検査と唾液の検査における潜在クラスモデル. (+), (-) はそれぞれ陽性と陰性を表す. $p_{11}, p_{01}, p_{10}, p_{00}$ はそれぞれのセルに存在する確率を表す.

潜在クラスモデルを利用する上で, パラメータの識別可能性が問題となる (Goodman, 1974; Jones et al., 2010). つまり, モデルの自由度とパラメータの数が一致していない場合は, パラメータの推定値は一意に決まらない. 先述の COVID-19 の検査の例ではパラメータは鼻咽頭液・唾液それぞれの感度・特異度・有病率の 5 つであるが, モデルの自由度は 3 であり識別不可能な状態にある. 識別不可能であるため, モデルの尤度関数から一意な最尤推定値を計算することはできない. この状況を打破する方法の一つに, パラメータに情報のある事前分布を与えたベイズ流の推定方法 (以下, ベイズ流潜在クラスモデル) が挙げられる (Collins and Huynh, 2014; Kostoulas et al., 2017).

先述の唾液を用いた診断法の有用性を確かめる調査では, 鼻咽頭液と唾液による同時集団スクリーニングのデータを条件付き独立性の下で潜在クラスモデルが適用された

表 1: 観察人数 (a) と条件付き独立性の下での確率モデル (b).

(a) 観察人数		
検査 1	検査 2	
	陽性	陰性
陽性	y_{11}	y_{10}
陰性	y_{01}	y_{00}

(b) 条件付き独立性の下での確率モデル.

感染			非感染	
検査 1	検査 2		検査 2	
	陽性	陰性	陽性	陰性
陽性	$Se1Se2$	$Se1(1-Se2)$	$(1-Sp1)(1-Sp2)$	$(1-Sp1)Sp2$
陰性	$(1-Se1)Se2$	$(1-Se1)(1-Se2)$	$Sp1(1-Sp2)$	$Sp1Sp2$

(Yokota et al., 2020). このスクリーニングでは, 鼻咽頭液および唾液の検査の両方で, 優れた感度とほぼ完璧な特異度が示された. 新型コロナウイルス感染症スクリーニングのように有病率が低い場合, 陽性頻度は稀である. データがスパースであるため, 事前情報の量がパラメーター推定に比較的大きな影響を与える. 本研究では, 一つの集団に対して二つの独立した検査を適用し, 陽性結果の頻度が稀である状況において, 潜在クラスモデルにおける感度と特異度の推定に対する事前分布の強さの影響を調査した.

5.2 ベイズ流潜在クラスモデリング

本モデルの適用は, 検査が 2 種類の場合に限定する. 検査 1 の感度と特異度をそれぞれ $Se1 \cdot Sp1$, 検査 2 の感度と特異度をそれぞれ $Se2 \cdot Sp2$ とする. また有病率は p とおく. ゴールドスタンダードが観察できない状況において, 表 1 のように感染者と非感染者に分割することができる. 条件付き独立の仮定が成立する下で, 陽性・陰性の組み合わせはそれぞれ単純な積の形で書き表すことができる. 二つの検査において観察される事象が観察される確率をそれぞれ $\pi_{11}, \pi_{10}, \pi_{01}, \pi_{00}$ とおくと, 式 (1) のように感染者・

非感染者の確率の和で書き表すことができる.

$$\begin{aligned}
\pi_{11} &= pSe1Se2 + (1-p)(1-Sp1)(1-Sp2) \\
\pi_{10} &= pSe1(1-Se2) + (1-p)(1-Sp1)Sp2 \\
\pi_{01} &= p(1-Se1)Se2 + (1-p)Sp1(1-Sp2) \\
\pi_{00} &= p(1-Se1)(1-Se2) + (1-p)Sp1Sp2
\end{aligned} \tag{1}$$

ここで, 実際の観察人数を $y_{11}, y_{10}, y_{01}, y_{00}$ とおくと, 対数尤度関数を式 (2) のように構成することができる.

$$\log L = y_{11}\log(\pi_{11}) + y_{10}\log(\pi_{10}) + y_{01}\log(\pi_{01}) + y_{00}\log(\pi_{00}). \tag{2}$$

上記の確率モデルにおけるパラメータは $Se1, Sp1, Se2, Sp2, p$ の 5 つであるため, 解を求めるためには少なくとも 5 つの独立な方程式が求められる. しかし, $\sum_{i,j \in (1,0)} \pi_{ij} = 1$ であることを考慮すると, 本モデルの自由度は 3 である. つまり, パラメータの数が自由度よりも多いため, パラメータの値を推定するために追加の仮定が必要となる. この課題はベイズ流推測によって, 解決することができる.

本研究で取り上げたベイズ流潜在クラスモデルでは, 5 つのパラメータの事前分布として Beta 分布をそれぞれ設定している. Beta 分布は形状パラメータの値で確率の値をコントロールすることができることに加え, 尤度関数が二項分布である場合の共役事前分布であるため事後分布は Beta 分布と決まり, 形状パラメータを定めると事後確率が一意に定まるという特徴を有する (Joseph et al., 1995). ベイズ流潜在クラスモデルの性能を評価するための三種類のシミュレーションを 5.3 節から 5.6 節にかけて説明する. これらのシミュレーションでは事後分布を構成するために, マルコフ連鎖モンテカルロ法 (MCMC) のメトロポリス・ヘイスティングス (M-H) アルゴリズムを採用した (Metropolis et al., 1953; Hastings, 1970). M-H アルゴリズムはいくつかある MCMC のアルゴリズムの中でも, 特に事後分布の形状が解析的に定められない状況において事後平均などの統計量が求められるという利点がある (Gilks et al., 1995). 本手法では Beta 分布と二項分布が共役の関係にあることから事後分布は Beta 分布と決まっているため, M-H アルゴリズムである必要はないが, M-H アルゴリズムの適用範囲をカバーしているため, この手法を採用した. MCMC を管理するための設定として burn-in 期間, thinning の二つがあり, 5.3 節以降で説明する実験でもこれらの値を固定値で設定している (Gelman and Rubin, 1992). Burn-in 期間とは, 初期のサンプルを破棄するための期間である. マルコフ連鎖の初期段階では, 定常分布 (目的とする事後分布) に収束せず初期値の影響を強く受けている可能性があり, 初期の不安定な期間のサンプルを取り除くことで推定の精度を高めることが期待できる. 一方, thinning とはマルコフ連鎖で生成された連続す

るサンプルから一定間隔でサンプルを抽出する処理である。MCMCによって得られるサンプルは、自己相関が高いことが多いため、この相関が高い全サンプルを用いた場合、実質的なサンプルサイズ(有効サンプルサイズ)が見かけよりも小さくなり、推定の効率が低下する可能性がある。そのため、thinningによってサンプル間の自己相関を低減し、独立なサンプルを得ることで事後分布の推定の効率と精度の向上が期待される。

5.3 シミュレーション実験1: 事前情報の強さに対する感度・特異度の推定

5.3.1 実験1の設定

シミュレーション実験1では、事前情報の強度およびその与え方が感度および特異度の推定に及ぼす影響を定量的に調査した。集団スクリーニングなどの適用場面において、疾患の有病率が低い場合、観測される陽性結果の頻度が極めて稀になる。このようなスパースなデータの状況下では、推定量の小標本バイアスが増大する可能性がある。さらにベイズ推定の性質上、事後分布が事前情報によって相対的に強く影響を受けることが予想される。

本実験では大きく分けて5つのシナリオを設定した。それぞれのシナリオは表2の通りである。説明のために、有効サンプルサイズ N_{ESS} 、感度の真値 η 、特異度の真値 ζ 、有病率の真値 θ を導入する。それぞれのシナリオが表す状況について説明する。シナリオ1は感度に事前情報を含む状況を想定したシナリオであり、シナリオ1-1とシナリオ1-2の二通りに分けられる。シナリオ1-1では $Se1, Se2$ が形状パラメータ $(\alpha, \beta) = (\eta N_{ESS}, (1-\eta)N_{ESS})$ であるBeta分布に従うと設定し、 $Sp1, Sp2, p$ は無情報事前分布に従うものとした。シナリオ1-2では $Se1$ にのみ同様の事前情報を仮定し、 $Se2$ は他の三つのパラメータと同様に無情報事前分布を仮定している。シナリオ1の設定と対照的に、シナリオ2では特異度のみ事前情報を含む状況を想定しており、シナリオ1と同様にシナリオ2-1とシナリオ2-2に分かれる。シナリオ2-1では $Sp1, Sp2$ が形状パラメータ $(\alpha, \beta) = (\zeta N_{ESS}, (1-\zeta)N_{ESS})$ であるBeta分布に従い、 $Se1, Se2, p$ は無情報事前分布に従うと設定した。シナリオ2-2では $Sp1$ にのみ同様の事前情報を仮定し、 $Sp2$ は無情報事前分布を仮定している。シナリオ1とシナリオ2ではそれぞれ感度のみまたは特異度のみに事前情報を入れていたが、シナリオ3では $Se1, Sp1$ に事前情報を入れ、検査1の感度・特異度の情報をパラメータ推定に全て用いた状況を想定した。また、シナリオ4では感度・特異度のパラメータには事前情報を入れず、有病率にのみ事前情報を入れた。さらにコントロールとして、5つのパラメータのどれにも事前情報を入れないシナリオ0を用意した。

シナリオ全体を通して、感度の真値 η および特異度の真値 ζ を $(\eta, \zeta) = (0.7, 0.99)$ と固定した。有病率の真値 θ については、0.1%, 0.5%, 1.0%, 10% の四つのパターンを設定した。また、有効サンプルサイズ N_{ESS} の値は 0, 5, 10, 20, 30, 50, 100, 200, 300, 500 の十通りのほか、参照値として N_{ESS} を無限大に飛ばした時の極限における状況（事前情報に真値を与えるという状況と同等）も設定した。MCMC の設定として、サンプリング数を 50000, burn-in 期間を 5000, thinning を 1 と設定した。さらに、シミュレーション回数は 10000, サンプルサイズは 2000 としている。実験 1 における評価は MCMC の収束割合の評価と 5 つのパラメータの点推定値の評価の 2 つに分けられる。MCMC の収束割合は、10000 回のシミュレーションの中でラグを 500 とした自己相関係数の値が 0.1 未満となった割合と定義した。また、点推定値の精度は、絶対バイアスの値および平均二乗誤差の平方根 (RMSE) で評価した。シミュレーションは SAS®9.4 (SAS Institute, Cary, NC) で実行し、結果の可視化は R (version 4.2.3; R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria) で行った。

本実験において特に注意を払っている点として、パラメータ空間の制限 (Jones et al., 2010) と M-H アルゴリズムのブロック化 (Roberts and Sahu, 1997) がある。前者については、パラメータに対する追加条件として同じ検査の感度と特異度の和は 1 以上という仮定をすべてのシナリオの事前分布に反映させた。後者のブロック化は MCMC 実行時にパラメータの更新に関わる問題であり、本実験では 5 つのパラメータを感度・特異度・有病率の三つのブロックに区切り、このブロックごとにパラメータ更新を考えるとという方針をとった。この工夫については、5.6 節で再度詳しく触れる。

5.3.2 実験 1 の結果

シナリオ 1-1 とシナリオ 1-2 の相違点が $Se2$ への事前情報の有無であることおよびシナリオ 2-1 と 2-2 の相違点が $Sp2$ への事前情報の有無であることから、 $Se2$ と $Sp2$ における結果を中心にベイズ流潜在クラスモデルを評価する。有病率が 0.01%, 0.1%, 1.0% における結果については、付録 (図 19 から図 21) に掲載した。

有病率が 0.5% および 10% における $Se2$ の絶対バイアス、平均二乗誤差、MCMC の収束割合を図 2 に示す。まずは収束割合について結果を比較する。感度のみ事前情報を入れるシナリオ 1 と有病率にのみ事前情報を入れるシナリオ 4 は、有病率の値・有効サンプルサイズの大きさに関わらず全ての MCMC サンプルが収束した。これに対して、シナリオ 2 とシナリオ 3 では有効サンプルサイズが 100 未満の状況では収束割合が 1.0 を満たさなかった。特にシナリオ 2-1 に関しては、有効サンプルサイズが 5 であるときと比べて、有効サンプルサイズが 10 と 20 の状況の方が収束した MCMC サンプルが少なかった。続いて、絶対バイアスについて評価する。シナリオ 1-2 における値は他のシ

表 2: 各シナリオ別の事前情報の与え方.

シナリオ	事前情報により変化	変化しない
0		$Se1, Se2, Sp1, Sp2, p \sim \text{Uniform}(0, 1)$
1-1	$Se1 \sim \text{Beta}(\eta N_{\text{ESS}}, (1 - \eta) N_{\text{ESS}})$ $Se2 \sim \text{Beta}(\eta N_{\text{ESS}}, (1 - \eta) N_{\text{ESS}})$	$Sp1, Sp2, p \sim \text{Uniform}(0, 1)$
1-2	$Se1 \sim \text{Beta}(\eta N_{\text{ESS}}, (1 - \eta) N_{\text{ESS}})$	$Se2 \sim \text{Uniform}(0, 1)$ $Sp1, Sp2, p \sim \text{Uniform}(0, 1)$
2-1	$Sp1 \sim \text{Beta}(\zeta N_{\text{ESS}}, (1 - \zeta) N_{\text{ESS}})$ $Sp2 \sim \text{Beta}(\zeta N_{\text{ESS}}, (1 - \zeta) N_{\text{ESS}})$	$Se1, Se2, p \sim \text{Uniform}(0, 1)$
2-2	$Sp1 \sim \text{Beta}(\zeta N_{\text{ESS}}, (1 - \zeta) N_{\text{ESS}})$	$Sp2 \sim \text{Uniform}(0, 1)$ $Se1, Se2, p \sim \text{Uniform}(0, 1)$
3	$Se1 \sim \text{Beta}(\eta N_{\text{ESS}}, (1 - \eta) N_{\text{ESS}})$ $Sp1 \sim \text{Beta}(\zeta N_{\text{ESS}}, (1 - \zeta) N_{\text{ESS}})$	$Se2 \sim \text{Uniform}(0, 1)$ $Sp2 \sim \text{Uniform}(0, 1)$ $p \sim \text{Uniform}(0, 1)$
4	$p \sim \text{Beta}(\theta N_{\text{ESS}}, (1 - \theta) N_{\text{ESS}})$	$Se1, Se2, Sp1, Sp2 \sim \text{Uniform}(0, 1)$

ナリオと比較して、有効サンプルサイズを大きくしても絶対バイアスの値が変わらないという特徴が確認された。つまり、もう一方の検査である検査1にどれだけ情報を与えたとしても（真値を入れたとしても）、バイアスが改善されなかったということである。シナリオ2-1, 2-2における絶対バイアスは、有効サンプルサイズが同じ下で同程度の値になった。シナリオ3もシナリオ2と同様の絶対バイアスの値を持つが、真値を与えたシナリオについてはシナリオ2の場合と比較して僅かに大きい値となった。また、 $Se2$ に事前情報を与えていないシナリオ（シナリオ1-1を除いた状況）では、有病率が低い状況（陽性の頻度が低い状況）において、10%を超える絶対バイアスが観測された。一方、シナリオ1-1では、有病率が低いにもかかわらず、事前情報が強くなるにつれてバイアスはゼロに近づいた。

次に、有病率が0.5%および10%における $Sp2$ の絶対バイアス、平均二乗誤差、MCMCの収束割合を図3に示す。収束割合について、 $Se2$ における結果と同様に、感度のみ事前情報を入れるシナリオ1と有病率にのみ事前情報を入れるシナリオ4は、有病率の値・有効サンプルサイズの大きさに関わらず全てのMCMCサンプルが収束した。また、シナリオ2とシナリオ3において有効サンプルサイズが100未満の状況では収束割合が1.0を満たさなかった点も $Se2$ と同様の結果である。続いて、絶対バイアスについて評価する。検査1にのみ事前情報を入れているシナリオ2-2では、どれだけ $Sp1$ に事前

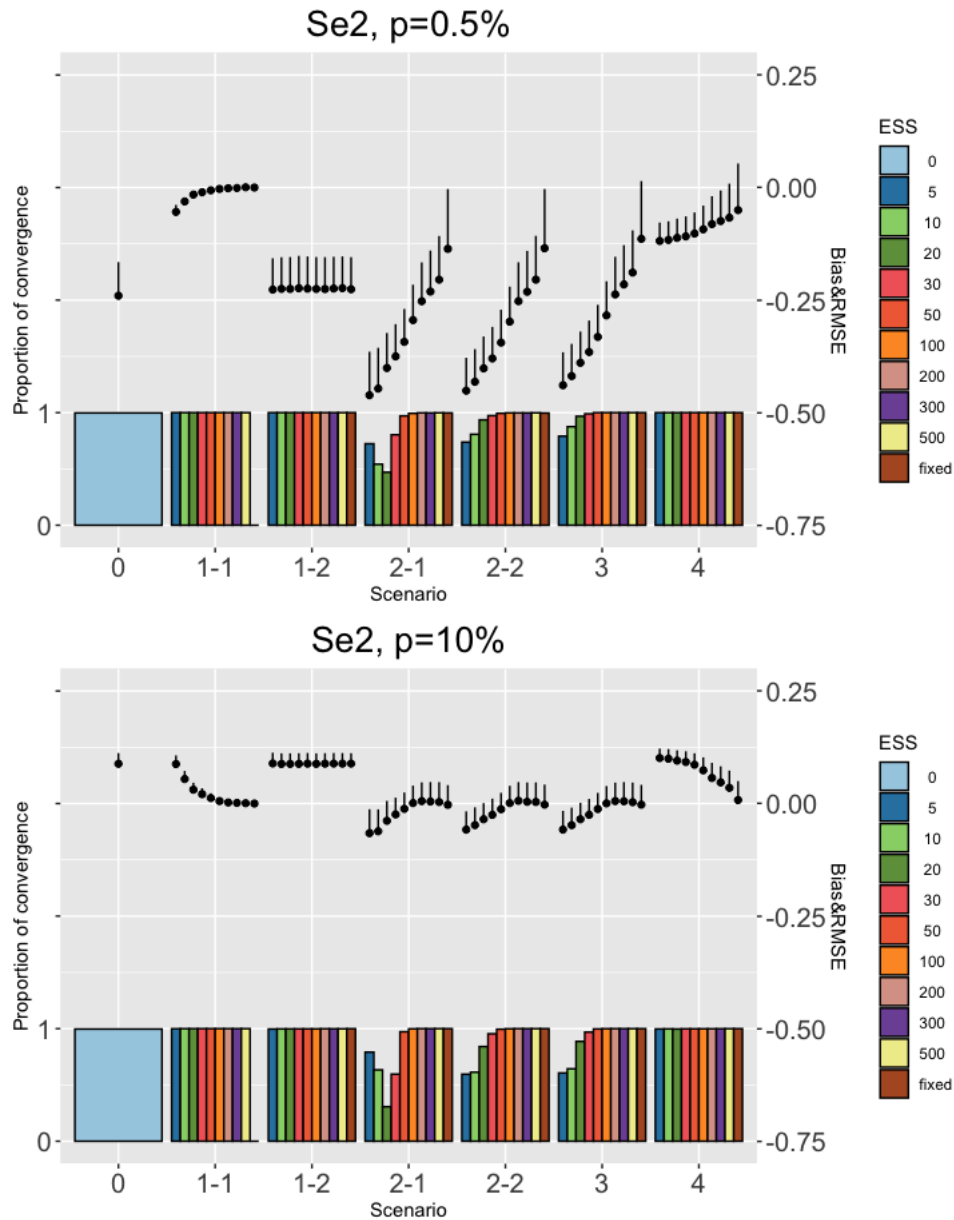


図 2: シミュレーション実験 1 における Se_2 の結果. 有病率 p が 0.5% の場合 (上パネル) と 10% の場合 (下パネル) の結果を比較している. 黒色の点は Se_2 のバイアスを表し, 実線の長さは Se_2 の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ (ESS) に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している. ESS の “Fixed” は, 事前情報に真値を与えた状況を示している.

情報を入れたとしても $Sp2$ の絶対バイアスは改善されなかった。しかし、それ以外のシナリオでは有効サンプルサイズが増えるにつれ、絶対バイアスはゼロに近づいた。

有病率の大きさの観点から $Se2, Sp2$ の結果を俯瞰する。有病率が 0.5% の状況は有病率が比較的低く、非有病者が多くを占める状況であるため、有病率が 10% の場合と比較すると $Sp2$ の絶対バイアスは全体的に小さい。対して、 $Se2$ に関しては有病率が 0.5% の時の方が 10% の場合と比較して、絶対バイアスが大きくなる傾向にある。どちらの場合においても事前情報を入れることで有効サンプルサイズの上昇とともに絶対バイアスをゼロに近づけることができることが確認できるが、 $Se2$ におけるシナリオ 1-2 のように一方の検査の同じパラメータにしか情報を入れていなかった場合にはバイアスが改善されないという特徴的な振る舞いが確認された。

5.4 シミュレーション実験 2: 感度の真値を与えた下での感度・特異度の推定

5.4.1 実験 2 の設定

5.3.2 節で確認したように、シナリオによっては感度の推定精度が N_{ESS} を増やすことで改善される場合とそうでない場合がある。実験 1 では 2 種類の検査の感度・特異度の真値は等しいと仮定していたが、実験 2 では感度の真値を変更することで感度の推定精度がどの程度変わるか確認する。実験 2 では実験 1 における“Fixed”のシナリオ(事前情報に真値を与えるシナリオ)のみを対象とし、感度の真値 η について、0.5, 0.6, 0.7, 0.8, 0.9, 1.0 の 6 パターンでの設定を行い、シナリオ 1-2, 2-1, 2-2, 3, 4 における $Se2$ の精度を絶対バイアスと RMSE で評価した。MCMC の設定およびサンプルサイズやシミュレーション回数は、実験 1 と同様の値を設定した。

5.4.2 実験 2 の結果

感度の真値を変化させた場合の $Se2$ の絶対バイアスと RMSE を図 4 に示す。有病率が 0.5% の場合、シナリオ 1-2 における絶対バイアスは他のシナリオと比較して大きかった。つまり、陽性頻度が稀な状況ではすべての設定で感度が過小評価されている。特に、 $Se1$ のみが真の値に固定されているシナリオ 1-2 は、従来の方法で観察されたものと比較して深刻なバイアスを呈した。有病率が 10% の状況では、特異度(シナリオ 2) または有病率(シナリオ 4) に事前情報が与えられている場合、バイアスは概ね小さかったものの、シナリオ 1-2 は真の感度に応じて正の方向と負の方向の両方にバイアスがあった。有病率が 0.1%, 1.0% における結果は、付録(図 22)に掲載している。

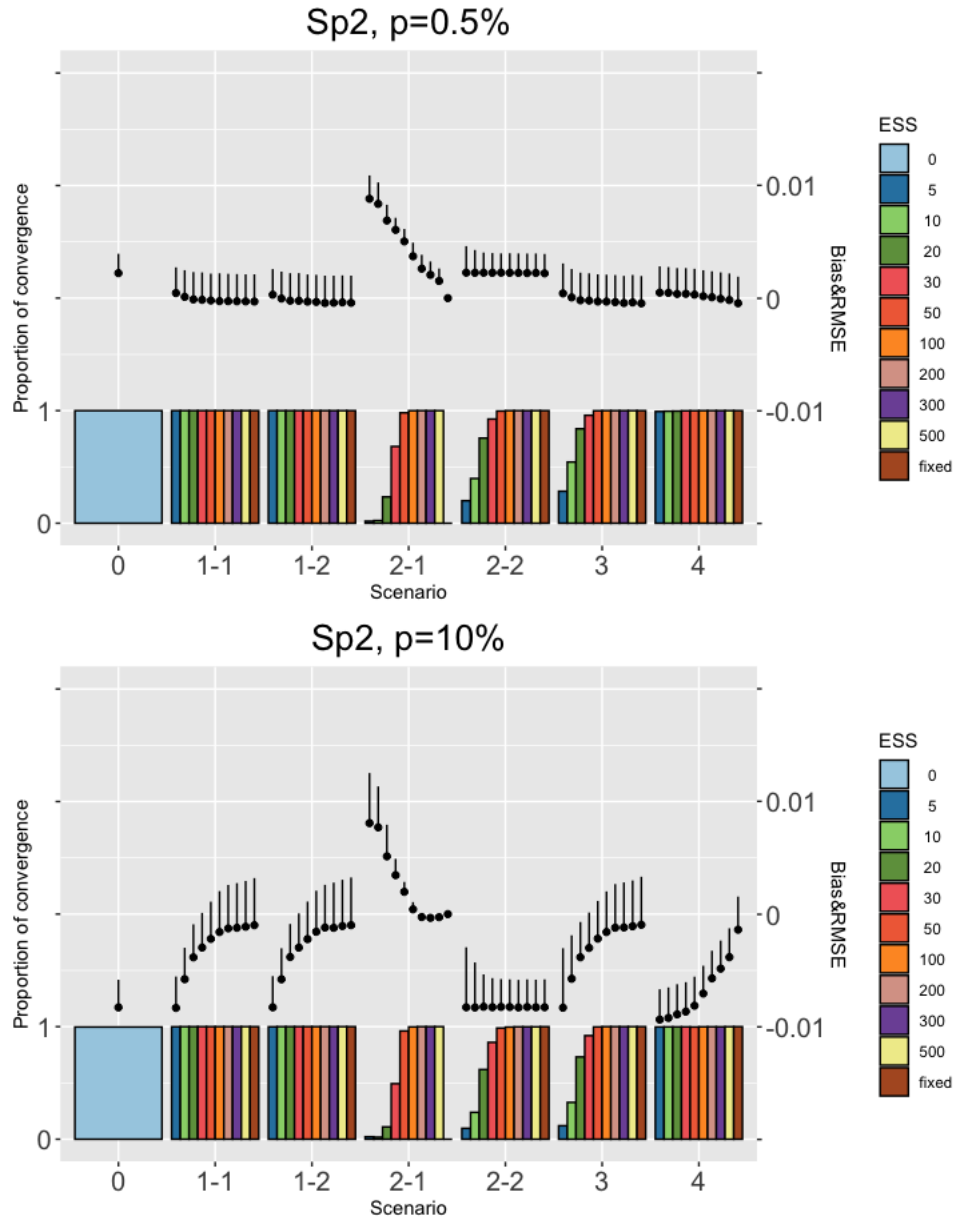


図 3: シミュレーション実験 1 における $Sp2$ の結果. 有病率 p が 0.5% の場合 (上パネル) と 10% の場合 (下パネル) の結果を比較している. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している. ESS の “Fixed” は, 事前情報に真値を与えた状況を示している.

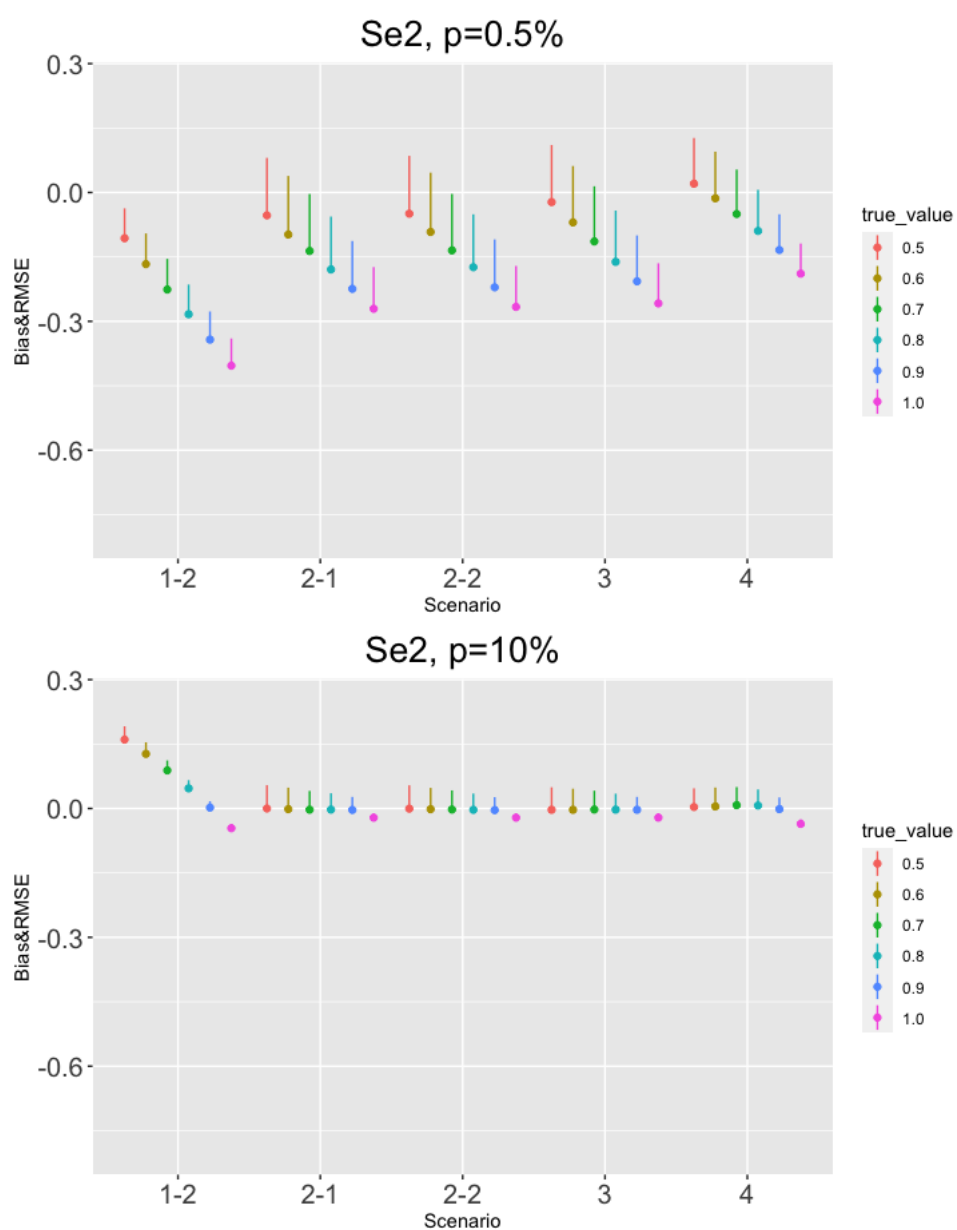


図 4: シミュレーション実験2における Se_2 の結果. 有病率 p が 0.5% の場合 (上パネル) と 10% の場合 (下パネル) の結果を比較している. 点は Se_2 のバイアスを表し, 実線の長さは Se_2 の平均二乗誤差の平方根 (RMSE) を表す. 真値に応じて色付けしている.

5.5 シミュレーション実験3: 既存の頻度流手法との比較

5.5.1 実験3の設定

シミュレーション実験1および実験2では、ベイズ流潜在クラスモデルにおける実験の設定がパラメータの精度に与える影響に焦点を当てていた。追加の検討事項として、他の推定方法論との性能比較を行った。シミュレーション実験3では Hui and Walter (1980) で提案された推定方法および Lu et al. (2018) で提案された推定方法とベイズ流潜在クラスモデルを比較した。1800 人と 200 人の 2 集団に対して、2 つの検査の感度と特異度のバイアスと RMSE を計算した。感度の真値は 0.5 から 0.9 まで設定し、特異度の真値は 0.99 で固定した。ベイズ流潜在クラスモデルにおける推定手法では、 $N_{ESS} = 200$ の事前情報を両方の特異度に設定した。この比較は一度目の投稿時に、査読者から頂いたコメントに沿って追加したシミュレーションである。これらの手法とベイズ流潜在クラスモデルの比較は、条件付き独立性の仮定が推定結果に与える影響を検証するために行われた。

Hui and Walter (1980) では、異なる有病率をもつ二集団に対してゴールドスタンダードがない状況における二つの検査の感度と特異度を推定する問題を取り上げている。必要な仮定として条件付き独立性の他に、二つの検査の感度と特異度は同じであることが求められる。Lu et al. (2018) の方法は、条件付き独立性の仮定ではなく経験的条件づけアプローチによって推定を行なっている。これらの方法の数理的背景は 11.4 節に記した。

5.5.2 実験3の結果

Se_2 における絶対バイアス、平均二乗誤差の平方根 (RMSE)、および推定値が計算できた割合を図 5 に示す。Hui-Walter モデルでは Se_2 の推定値が $[0,1]$ の区間外に現れる状況が 90% 近くあり、これらは推定不能として扱った。有病率 0.5% の場合、真値が大きくなるとベイズ流潜在クラスモデルと Lu らのモデルではバイアスの大きさが負の方向に大きくなるのに対し、Hui-Walter モデルでは 0 に近づくという結果が示された。このとき、Hui-Walter モデルで推定した Se_2 の絶対バイアスは他の手法よりも小さく見えるが、棒グラフの結果から推定値が計算できた回数自体は少なく、推定値が求められた数少ない結果をまとめているだけで正確な性能を測っているとは断言できない。有病率 0.5% における結果と有病率 10% における結果を比較すると、陽性例の頻度が高い 10% の有病率の場合においてベイズ流潜在クラスモデルによる推定値の絶対バイアスの方が小さく抑えられた。Lu らのモデルによる推定値も同様のバイアスの傾向を示した。しかし、このモデルによる推定ではいずれかのパラメータの推定値が 0 または 1 の

時に尤度が最大化されるという状況があり、極端な値をとる推定値が計算される状況が20%近く確認された。全体を通して、ベイズ流潜在クラスモデルによる推定値のRMSEは他の手法のRMSEよりも小さく、バイアスの値だけでなく分散も加味して考えると、ベイズ流潜在クラスモデルは他の二つのモデルと比較して推定がより安定していることを示唆している。

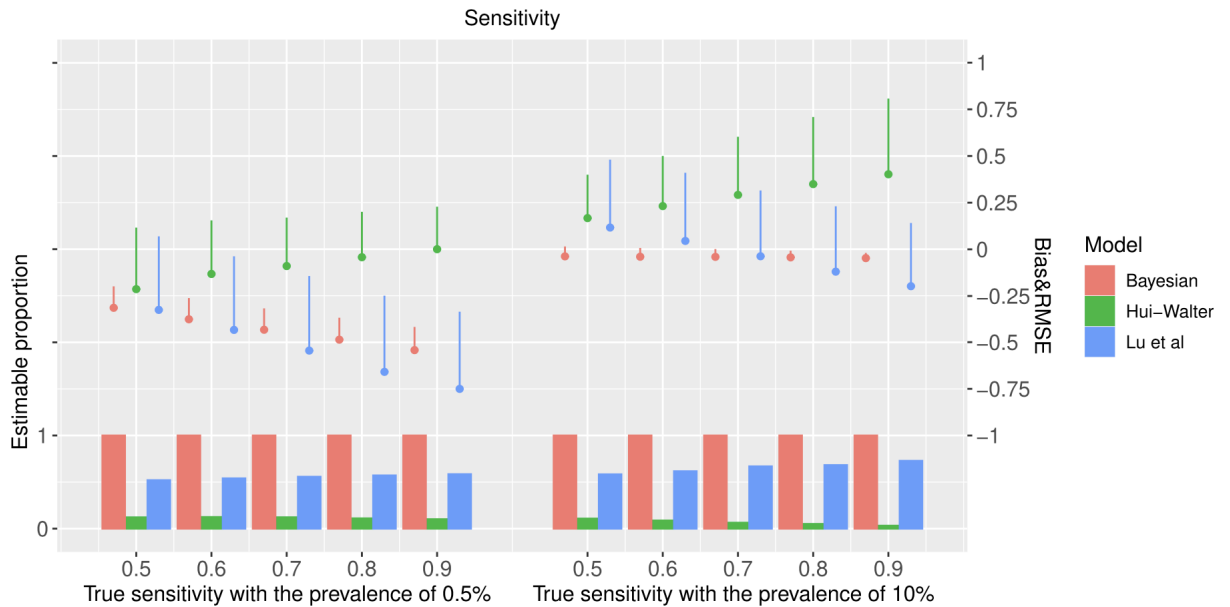


図 5: シミュレーション実験3における Se_2 の結果. 有病率 p が 0.5% の場合 (左パネル) と 10% の場合 (右パネル) の結果を比較している. 点は Se_2 のバイアスを表し、実線の長さは Se_2 の平均二乗誤差の平方根 (RMSE) を表す. 推定手法に応じて色付けした棒グラフは、 Se_2 の推定値が計算できた割合を表している.

5.6 考察

条件付き独立性を仮定した下でゴールドスタンダードがない状況において感度・特異度・有病率を推定する問題において、ベイズ流潜在クラスモデルはパラメータの数が自由度よりも多くなる状況においても事前情報を入れることで推定可能であるという点が利点である. 本研究では事前情報の強さや有効サンプルサイズ、パラメータの真値などあらゆる観点からベイズ流潜在クラスモデルの性能を確かめた. また、既存の頻度

流手法とも比較することで、特に推定可能な場面が多いことを示した。一方で、条件付き独立性の仮定を入れたとしても、有限標本では絶対バイアスがある程度存在することは注目すべき点である。

シミュレーション実験1の感度 $Se2$ の推定結果について、注目すべき点をまとめる。一点目に、MCMCの収束割合の結果から事前情報として入れる有効サンプルサイズは100以上が好ましい点である。シナリオ2, 3では有効サンプルサイズが50以下ではMCMCの収束割合が1に満たない状況があり、限られた推定値で評価することになる。実際、そのような状況では有病率が0.5%の場合でも10%の場合でも比較的バイアスが大きいの。そのため、100程度の有効サンプルサイズを確保することが感度の推定に重要であることが示唆される。二点目に、 $Se1$ のみに情報を入れた時(シナリオ1-2)には有効サンプルサイズをどれほど大きくしても絶対バイアスは小さくならないが、 $Sp1$ と $Se1$ に情報を入れる(シナリオ3)と有効サンプルサイズの上昇とともに絶対バイアスの値はゼロに近づく。対となる感度 $Se1$ だけでは $Se2$ の推定性能の改善にはつながらず、むしろ特異度 $Sp1$ が $Se2$ の推定性能を改善するということである。この様相は有病率が0.5%の場合と10%の場合でもともに見られるが、有病率が10%の状況では有効サンプルサイズの値に関係なく $Se1, Sp1$ ともに情報を入れた方が $Se1$ のみに情報を入れるよりも絶対バイアスが小さくなるのに対して、有病率が0.5%の状況では有効サンプルサイズが300以上の状況でバイアスがより小さくなった。これは先述のMCMCの収束割合の低さとは別に何らかの理由があると考えられるが、原因は未だ不明である。有病率がある程度高い状況においては $Se1$ だけでなく $Sp1$ にも事前情報を入れることを推奨するが、有病率が低い場合においては特に有効サンプルサイズが300程度確保できない状況で $Se1$ と $Sp1$ の両方に事前情報を入れて $Se2$ を推定することは、絶対バイアスの観点からはかえって推定精度が悪化している。三点目に、シナリオ2-1, 2-2, 3の比較によって絶対バイアスの値および変化は類似している点である。これは $Se2$ の推定精度に対して、 $Se1$ や $Sp2$ よりも $Sp1$ が比較的大きな影響を及ぼしていることを示している。四点目に、有病率が10%の状況と比較して0.5%における絶対バイアスが大きく、また負の方向に存在している。ここから、陽性の頻度が稀な状況では $Se2$ の値は過小評価されていることがわかる。

続いて、シミュレーション実験1の特異度 $Sp2$ の推定結果について、注目すべき点をまとめる。一点目に、MCMCの収束割合の結果から事前情報として入れる有効サンプルサイズは100以上が好ましい点である。この点は、 $Se2$ において指摘した点と同様である。二点目に、有病率0.5%におけるシナリオ2-1とシナリオ2-2を比較すると、有効サンプルサイズが200未満の状況では $Sp2$ にのみ情報を入れる場合よりも $Sp1, Sp2$ の両方に情報を入れた方が、バイアスは大きくなるという点である。一方、有病率10%のときは $Sp1, Sp2$ の両方に情報を入れた方が $Sp1$ のみに入れた時よりもバイアスの大

きさは小さくなっている。また、有効サンプルサイズが比較的小さい状況に絞ると、シナリオ 2-1 では正のバイアスからゼロに向かっているが、シナリオ 2-2 では一律に負のバイアスとなっている。つまり、 $Sp1, Sp2$ ともに事前情報を入れた場合では $Sp2$ は過大評価となっており、 $Sp1$ のみに情報を入れる場合は過小評価となっている。三点目に、シナリオ 1-1, 1-2, 3 のバイアスの軌跡および値がほぼ同等であるという点である。これは $Se2$ に対する三点目の考察と対照になっており、 $Sp2$ の推定精度は $Se2$ と $Sp1$ と比較して $Se1$ が与える影響が大きいということがわかる。四点目に、有病率が 10% の状況と比較して 0.5% における絶対バイアスが小さくなっている。ここから、陽性の頻度が稀な状況では $Sp2$ の値は精度良く推定されており、シナリオ 2 を除いてそれほど有効サンプルサイズの大きさも気にしなくていいことが示唆される。

シミュレーション実験 2 では感度の真値を変更したシミュレーションを実施した。もう一方の検査の感度に情報を入れるシナリオ 1-2 では、有病率が 10% の状況では真値を変化させるとバイアスは正の値から負の値をとった。極端な状況 (真値が 1.0) を除いて、推定値は平均値で考えると過大評価になっている可能性が高い。また、特異度や有病率のパラメータに情報を入れた方がバイアスが小さいという点から、有効サンプルサイズが十分に用意できる状況であるならば、もう一方の検査の感度だけに情報を入れるよりも特異度や有病率に入れた方が推定精度は良いといえることができる。一方で、有病率 0.5% におけるシナリオ 1-2 では真値が大きくなるとよりバイアスが大きくなる。真値の値に限らず負のバイアスになっていったため、推定値は平均値で考えると過小評価になっている。また、その他のシナリオも多くの場合には負のバイアスで、真値を大きくするとバイアスは大きくなっているが、シナリオ 1-2 の場合と比べると比較的小さいという点是有病率 10% の場合と同様の傾向であった。

既存の頻度流手法との比較を行ったシミュレーション実験 3 では、特に推定値が値として計算できるかという点で違いがあった。Hui-Walter モデルを利用した解析で推定値が計算できなかったのは、二つの検査と二つの集団という状況でパラメータの数がモデルの自由度を上回った影響で Jacobian が計算できなかったことが原因である。ベイズ流潜在クラスモデルは Hui-Walter モデルや Lu らのモデルとは異なり、有病率が高い状況だけでなく低い状況においても全てのシミュレーションで推定値が計算された。有病率 10% の状況ではバイアスの値や RMSE の値が有病率の真値の値に関わらずほとんど 0 であるという点から、陽性の頻度が稀でない疾患に対する感度・特異度推定では、状況を限定せずベイズ流潜在クラスモデルによる推定を推奨する。一方、有病率 0.5% の状況におけるベイズ流潜在クラスモデルのバイアスの大きさは、Lu らのモデルと比較すると真値に関わらず小さいが Hui-Walter モデルの値と比べると大きい。しかし、5.5.2 節でも指摘したように、Hui-Walter モデルから導かれる推定値は 0 から 1 の範囲に収まらないという致命的な欠点があり、統計モデルとしての妥当性が疑われる。頻度流手法

では尤度の最大化が推定値を求める方法となっており,このような不安定な推定結果が導出される懸念がある.一方,ベイズ流潜在クラスモデルでは尤度に事前情報が掛け合わされており,事前分布の平均値に縮小されるため安定な推定に繋がっている.計算可能な点に加えて, RMSE は他 2 つのモデルよりも小さく,推定値の分散という点で考えたとしてもベイズ流潜在クラスモデルによる推定を強く推奨したい.

5.3.1 節で触れたブロック化 M-H 法とパラメータ領域の制限について取り上げる.通常の MCMC ではすべてのパラメータを 1 つのベクトルとして,採択確率について更新するか否かを判断する.このときパラメータ同士の相関が高いような組み合わせがある状況では,自己相関も高くなるために収束したと判定される MCMC サンプルが少なくなることが問題点である (Roberts and Sahu, 1997). 実際,研究初期に行ったシミュレーション実験 1 では,パラメータに対するブロック化は設定せず, Se_2 と Sp_2 の収束割合を計上していた. その結果については, 図 14 から図 18 に掲載している. Se_2 の有病率 0.5% の結果などをみると,収束割合が大きくないということがわかる. この結果を受けて,感度・特異度・有病率でそれぞれブロック化を行い,感度の更新と特異度の更新と有病率の更新を分けて行なった. そのようにして得られた 2 では,自己相関の値が改善したことにより収束割合も上昇した. 次に,感度または特異度の事前分布に対して取りうる値を制限したことについてであるが,これは推定値の値に関係する. こちらも研究初期の時点では,事前情報を与えないパラメータに関してはすべて一律に無情報事前分布を与えたシミュレーションをしていた. その時の各シミュレーションの点推定値の分布をみると,感度と特異度の和が 1 未満の領域に落ちる推定値がいくつか存在した. そのような領域は,診断能としてはあまり高い診断精度をもつ検査とは言えないということになる. このようなことが起きた原因は,「label-switching」という現象によるものである (Jones et al., 2010). これは条件つき独立の仮定が成り立つもとでの尤度 (式 1 と式 2) を考えたとき, p と $1-p$, Se_1 と $1-Sp_1$, Se_2 と $1-Sp_2$ を入れ替えても尤度の式自体は変わっていないため,推定したいパラメータとは別のパラメータが推定されてしまう問題である. そのため,事前情報を変化させない Se あるいは Sp に対して制限を課している.

最後に,本研究で新たに得られた 3 つの知見と課題についてまとめる. 一つ目の知見は,感度・特異度の推定においてそのパラメータ自身に事前情報を入れなければ正負に限らず両方向にバイアスが現れる点である. 二つ目の知見は,もう一方の検査の同じ種類のパラメータに情報を入れたとき,その情報をどれだけ多く入れたとしてもバイアスは改善されない点である. 三つ目の知見は,尤度最大化を目的とした頻度流推定では推定値が計算できない状況があるが,ベイズ流潜在クラスモデルでは推定値が計算可能であるという長所を持ち合わせている点である. 本研究では 5 つの推定パラメータのうち一つまたは二つにのみ情報的な事前分布を置く限定的な状況しか評価していない.

また, 本研究は二つの検査に対する結果であるため, 三つ以上の検査において条件付き独立性の仮定の有無で, ベイズ流潜在クラスモデルによる推定値のバイアスがどの程度あるのかという調査もモデルの性能を測る上で重要であろう.

6 慢性心不全患者における死亡または入院イベントと動物性・植物性タンパク質摂取量の関連

6.1 序文

心不全はあらゆる心疾患の末期像であり、増悪を繰り返すことにより QOL の低下を招く。心不全を抱える患者集団においては、栄養不良の状態が一般的に観察されることが指摘されている (Lv and Ru, 2021)。例えば、低体重である状態が心不全の進行や生命予後に影響を及ぼすことが明らかにされており、BMI が高値である集団と比べて低値である集団の予後が悪いという“肥満パラドックス”と呼ばれる特徴がある (Amundson et al., 2010; Hamaguchi et al., 2010)。心不全における急激な体重減少は心臓悪液質として知られる低栄養状態であり、基礎代謝量の上昇と摂取熱量の減少が原因とされている (Akner and Cederholm, 2001)。一方、どのようなものを摂取するかについても関心が向けられている。主要な栄養素の 1 つである脂肪酸の摂取量と心不全の発症が関連しているという報告もあり、心不全の罹患や心機能低下と関連することが知られている (Soukoulis et al., 2009)。脂肪酸の興味深い点は、n-3 系不飽和脂肪酸の摂取は心不全の発症抑制効果につながる一方で、n-6 系不飽和脂肪酸の摂取が炎症反応の増強を引き起こしたり、飽和脂肪酸が活性酸素を介してアポトーシスを誘導するなど、脂肪酸の中でもその種類によって病態への影響が大きく異なっている点である (Mozaffarian and Wu, 2011; Matsui et al., 2012; Wu et al., 2014)。栄養の摂取は主に食事であるが、食事は複数の栄養素が含まれている複数の食品を組み合わせるため、栄養状態が臨床状態とどのように関連しているのかを解明することは容易ではない。さらに厄介な点として、心不全の合併症である抑うつなどの精神疾患は食事の摂取を妨げることがあり、精神状態が交絡因子として働いている可能性が指摘されている (Tsuchihashi-Makaya et al., 2009; Bauer et al., 2012)。栄養状態と疾患との関係性を評価するためには、毎日の食事パターン (何をどれくらい摂取しているか) と心理状態が重要であり、複雑に絡み合った栄養疫学データを紐解くことで心不全の予防にもつながることが期待される。

ヒトの生命活動の維持、適切な成長、発達、生殖機能の確保、そして生存そのものにとって、食事由来のタンパク質を十分量摂取することは、本質的に極めて重要である (Wu, 2016; Budhathoki et al., 2019)。これまでの栄養疫学的な研究報告では、慢性心不全の患者においてタンパク質摂取量が低いことは臨床的な経過や予後予測において好ましくない結果をもたらすことが報告されている (Streng et al., 2022; Saijo et al., 2023)。心不全患者の病態とタンパク質摂取に関しての研究はいくつかされているが、特定のタンパク質が心不全患者の転帰にどのような影響を及ぼすかという点については、未だ十分な知見が得られていない。ヒトが日常的に摂取する動物性タンパク質と植物性タン

パク質は、そのアミノ酸組成が根本的に異なっており (Wu, 2016), 食事におけるタンパク質源の選択は同時に摂取される他の主要栄養素, 微量栄養素および多様な生理活性物質の構成にも影響を与える。この複合的な要因が、結果として心臓代謝系の健康に対して異なる影響を及ぼし、様々な転帰をもたらすと考えられている (Mariotti, 2019)。本研究では、慢性心不全を有する日本人患者における動物性タンパク質および植物性タンパク質の摂取と死亡および入院との関連性を明らかにすることを目的として開始した。

特定の要素における数値の大小と興味のあるイベント発生有無の関連性を評価する際、最初に検討する統計量の一つにロジスティック回帰モデルから導かれるオッズ比がある (Cornfield, 1951; Hosmer Jr et al., 2013)。オッズ比の特徴は、疾患発生イベントが稀であるときにリスク比の近似値となることである。イベント発生のような二値アウトカムに対して、適用する場面が数多く存在する。しかし、Greenland (1987) でも指摘されているように、オッズ比の値自体が解釈可能な範囲は限定的であり、積極的に使用を推奨されるものではない。近年では CONSORT のガイドラインにおいてもリスク比・リスク差などの直接解釈をすることができる指標を用いることを推奨している (Schulz et al., 2010)。ロジスティック回帰に代わる二値アウトカムに対する回帰モデルとして、対数二項回帰 (Agresti, 2013) と修正ポアソン回帰 (Zou, 2004) がある。本研究では、実際に解析に使用する慢性心不全栄養疫学データと類似した状況で2つのモデル間の性能に差がどの程度あるのか。また、どちらの解析手法を選択すべきかということテーマに、データ解析の前に数値シミュレーションを行う。

本研究で取り上げる栄養疫学データを解析する上で、特に慎重に吟味した点として「タンパク質の摂取量 (g) の値を用いてイベントに対するリスク比を推定した解析」と「総エネルギー摂取量に対するタンパク質の摂取割合 (%) の値を用いた場合のイベントに対するリスク比を推定した解析」の2つの側面で考察したという点が挙げられる。これには二つの理由がある。一点目に、各タンパク質の高摂取群と低摂取群を分割する閾値を絶対量と割合の観点から与えるためである。栄養データが複雑であることを考慮して、単純に中央値などの分位点で設定することはせず、絶対量や割合の値の取りうる範囲におけるリスク比の変化を見て閾値を設定することを実行した。二点目に、絶対量・割合がイベントに与える影響をリスク比の値から考察するためである。実際、イベントを予防するという観点からタンパク質の絶対量と全エネルギーに対するタンパク質摂取割合のどちらに注意を向ければ良いかということに対する一つの指針になるだろう。

6.2 対象コホートの説明

国内十施設 (北見赤十字病院・市立釧路総合病院・帯広協会病院・溪和会江別病院・北海道大学病院・小樽協会病院・国立病院機構函館病院・彦根市立病院・鳥取大学・

済生会福岡総合病院) の多施設共同栄養疫学研究に登録された慢性心不全患者を対象としたコホートである。2012 年 12 月から 2014 年 9 月の間に施設にて心不全の治療を行っており, 前回の心不全による退院から一ヶ月以上経過した 20 歳以上の患者 156 名が組み入れられている。ネフローゼ症候群, 肝硬変, 進行性の悪性腫瘍, コントロール不良の糖尿病 (HbA1c の値が 7% を超える症例) と登録時に診断され, 治療中であった患者は除外している。生体試料 (早朝尿の採取と空腹時採血), 質問票を用いた情報収集 (BDHQ (Kobayashi et al., 2011), PHQ-9, MLHFQ) がこの研究のために収集された。また, 診療録から疾患情報や服用する薬なども登録されている。患者は最大 3 年間追跡され, 一年追跡するごとに質問票による再調査と予後調査を行った。予後調査は, 全死亡・心血管死亡・全ての入院・心不全増悪による入院・心血管死疾患による入院が記録された。また本コホートを用いた既報の研究には, 柑橘類の果物を摂取することによる患者の精神状態への影響を調べた研究 (Kakutani et al., 2022), 複数の栄養スコア指標と実際の臨床症状からスコアの診断性能を調査した研究 (Kinugasa et al., 2022), 総摂取カロリー量と全死亡または入院イベントとの関連性を調べた研究 (Obata et al., 2021) がある。

6.3 稀でないイベント発生に対する回帰モデル

6.3.1 対数二項回帰と修正ポアソン回帰

ロジスティック回帰はリンク関数がロジットの二項回帰モデルであるが, 対数二項回帰・修正ポアソン回帰はリンク関数が対数の回帰モデルである。2 値アウトカムを $Y_1, \dots, Y_n$ とし, n 人の対象者それぞれの p 次元共変量ベクトルを $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$, $i = 1, 2, \dots, n$ とおく。このとき, アウトカムのベクトル $\mathbf{Y}$ とデザイン行列 $\mathbf{X}$ を定義する。

$$\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T, \quad \mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \cdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

ロジスティック回帰は式 (3) で与えられる (対象者 i に対する式か全体に対する式かの違いで実質同じ)。 $\mathbb{E}[Y_i|\mathbf{x}_i]$ は $\mathbf{x}_i$ を与えた下での Y_i の期待値であるが, Y_i は 2 値しか取らないので, 実質的には $Y_i = 1$ となる確率 π と同値である。

$$\begin{aligned} \text{logit}(\mathbb{E}[\mathbf{Y}|\mathbf{x}_i]) &= \log \left(\frac{\mathbb{E}[\mathbf{Y}|\mathbf{x}_i]}{1 - \mathbb{E}[\mathbf{Y}|\mathbf{x}_i]} \right) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \\ \text{logit}(\mathbb{E}[\mathbf{Y}|\mathbf{X}]) &= \log \left(\frac{\mathbb{E}[\mathbf{Y}|\mathbf{X}]}{1 - \mathbb{E}[\mathbf{Y}|\mathbf{X}]} \right) = \beta_0 + \beta_1 \mathbf{x}_1 + \dots + \beta_p \mathbf{x}_p \end{aligned} \tag{3}$$

オッズの定義は両辺の指数を取ることで達成される以下の式 (4) である.

$$\frac{\pi}{1-\pi} = \frac{\mathbb{E}[\mathbf{Y}|\mathbf{X}]}{1-\mathbb{E}[\mathbf{Y}|\mathbf{X}]} = \exp(\beta_0 + \beta_1 \mathbf{x}_1 + \dots + \beta_p \mathbf{x}_p) \quad (4)$$

ここで, ある定数 a を用いて $\mathbf{x}_1 = a + 1$ と代入した式を $\mathbf{x}_1 = a$ を入れた式で割ると,

$$\frac{\exp(\beta_0 + \beta_1(a+1) + \dots + \beta_p \mathbf{x}_p)}{\exp(\beta_0 + \beta_1 a + \dots + \beta_p \mathbf{x}_p)} = \exp(\beta_1)$$

となり, これがオッズ比である. 一方, 対数二項回帰と修正ポアソン回帰は, 式 (5) で表現され, こちらではリスク比が計算される.

$$\begin{aligned} \log(\mathbb{E}[\mathbf{Y}|\mathbf{x}_i]) &= \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \\ \log(\mathbb{E}[\mathbf{Y}|\mathbf{X}]) &= \beta_0 + \beta_1 \mathbf{x}_1 + \dots + \beta_p \mathbf{x}_p \end{aligned} \quad (5)$$

取り上げた三つの回帰モデルは一般化線形モデルと呼ばれる系の一部である. 一般化線形モデルの枠組みで回帰係数ベクトル $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ の最尤推定値を求めるには, 推定方程式の解を求めればよい (Liang and Zeger, 1986; Agresti, 2013). 推定方程式は

$$\begin{aligned} U(\boldsymbol{\beta}) &= \sum_{i=1}^n \mathbf{D}_i^T \mathbf{V}_i^{-1} (Y_i - \mu_i) \\ &= \sum_{i=1}^n \frac{\partial \mu_i}{\partial \boldsymbol{\beta}} v(\mu_i) (Y_i - \mu_i) = 0 \end{aligned} \quad (6)$$

で表される方程式である. μ_i, V_i はそれぞれ平均関数, 分散関数と呼ばれる. 上記 3 モデルの平均関数と分散関数は以下の表 3 の通りである. 推定方程式の特徴として, 平均関数と分散関数の二つさえ指定すれば, 確率分布に置かれる仮定 (誤差の正規性など) は必要なく推定値を求められるという点がある. 推定方程式から計算される推測手法は, 擬似尤度法と呼ばれる (Wedderburn, 1974; Tsiatis, 2006).

表 3: 二値アウトカムに対する回帰モデルごとの平均関数と分散関数.
ロジスティック回帰 対数二項回帰 修正ポアソン回帰

	$\exp(\boldsymbol{\beta} \mathbf{x}_i)$	$\exp(\boldsymbol{\beta} \mathbf{x}_i)$	$\exp(\boldsymbol{\beta} \mathbf{x}_i)$
平均関数	$\frac{\exp(\boldsymbol{\beta} \mathbf{x}_i)}{1 + \exp(\boldsymbol{\beta} \mathbf{x}_i)}$		
分散関数	$\mu_i(1 - \mu_i)$	$\mu_i(1 - \mu_i)$	μ_i^1

¹ 作業モデルとして仮定される分散関数である. あくまでも計算安定のために使用.

$\boldsymbol{\beta}$ の推定値は, Newton-Raphson アルゴリズムによって計算される. その推定手続きは

1. β の初期値 β_0 を設定する.

2. Hessian 行列 $H(\beta^{(k)}) = -\frac{\partial^2 U(\beta^{(k)})}{\partial \beta^T}$ を計算する

3. 更新式による値の更新を行う: $\beta^{(k+1)} = \beta^{(k)} + H(\beta^{(k)})^{-1} U(\beta^{(k)})$

4. 収束判定を行う

$|\beta^{(k+1)} - \beta^{(k)}| < \epsilon$ なら $\beta^{(k+1)}$ が推定値. この不等式を満たさなければ, ステップ 2 に戻って再度更新する. ϵ は解析者が設定する十分小さな数 (10^{-6} など).

というように繰り返し計算によって実行される. 対数二項回帰と修正ポアソン回帰では分散関数の設定が異なるが, これは Newton-Raphson アルゴリズムの収束に影響を与えている. 対数二項回帰の分散関数の設定では, μ_i が $[0, 1]$ の範囲に収まらず β の推定値が得られない状況がある, という指摘がある (Wallenstein and Bodian, 1987; McNutt et al., 2003).

β の点推定値は上記のアルゴリズムで得られるが, 95%信頼区間をどう構成するか. 分散関数が正しく設定されている場合は, Fisher 情報行列の逆行列から推定量の共分散行列を推定することができる. しかし, 修正ポアソン回帰の場合は分散関数が作業モデルとして仮定される分散関数であるため, この指定が誤っていれば Fisher 情報から求められる分散にはバイアスが入る (野間久史, 宇野慧, 石井亮太, 2025). そこで, モデルが誤特定されている場合でも, 真の漸近分散に一致するように計算されるサンドイッチ分散 $\hat{V}$ を使った Wald 信頼区間を構成することが標準である.

$$\hat{V}(\hat{\beta}) = I(\hat{\beta})^{-1} J(\hat{\beta}) I(\hat{\beta})^{-1}$$

$$I(\beta) = -\mathbb{E} \left[\frac{\partial^2 U(\beta)}{\partial \beta^2} \right], \quad J(\beta) = \mathbb{E} [U(\beta) U(\beta)^T]$$

6.3.2 回帰モデル間の性能比較 - シミュレーション実験の設定

慢性心不全患者コホートの多施設共同栄養疫学研究データを解析する上で, 対数二項回帰モデルと修正ポアソン回帰モデルのどちらを採用すべきか. 両者の性能を比較するため, シミュレーション実験を行った. イベント発生率を p , 共変量の数を C とする. この C は興味のある変数 X_1 と $C-1$ 個の調整変数 $\{X_2, \dots, X_C\}$ から成ると想定している. 共変量のシミュレーションデータは平均 0, 分散 Σ の多変量正規分布から発生させた. 分散・共分散行列 Σ は対角成分を 1 とし, 興味のある変数 1 つと調整変数 1 つ

の間に相関 r を導入した. また C 個の共変量のうちランダムに 30% 分の変数を選択し, 0 以上の値であれば 1, 0 未満の値なら 0 を返すように設定し, 二値共変量として変換した. 共変量の集合 $\{1, X_1, X_2, \dots, X_C\}$ に対して, ベクトル $\{\beta_0, \beta_1, \dots, \beta_C\}$ との内積をとり, 線形予測子を構成した. β_0 には $\log(p)$ を設定し, β_1 にはリスク比の真値 RR の対数を入れた. $\{\beta_2, \beta_3, \dots, \beta_C\}$ については, 一様分布 $U(-0.5, 0.5)$ から発生させた. 線形予測子を指数で変換した値を二項確率としたベルヌーイ分布から各個人の観測値を生成した.

生成したデータに対して対数二項回帰モデルと修正ポアソン回帰モデルを適用し, X_1 におけるリスク比の点推定値のバイアスと RMSE, そしてモデルの収束割合を評価した. シミュレーション回数は 1000 回と設定した. p は 10%, 30%, 50% の三種類, r は 0, 0.1, 0.4, 0.8 の四種類, C は 5, 10, 15 の三種類, RR は 0.7, 1.0, 2.0 の三種類を用意し, それぞれの組み合わせでシミュレーションを実行した. またサンプルサイズは多施設共同栄養疫学研究データとほぼ同等の 150 人とした.

6.3.3 回帰モデル間の性能比較 - シミュレーション実験の結果

共変量を 5 個とした場合のシミュレーション実験の結果を表 4 に示す. 紙面の都合上, バイアスについては 100 倍した値を掲載している. 注目すべき点はリスク比の推定値の収束割合 (CP) であり, リスク比の真値, イベント発生割合, 二変数間の相関に関わらず, 対数二項回帰に当てはめるよりも修正ポアソン回帰に当てはめた場合の方が優れていた. さらに, 修正ポアソン回帰ではどのような状況においても推定値が収束した割合が 1.0 であり, 収束性の面では安定していた. また, バイアスの絶対値と RMSE についても一部の状況 ($RR=1.0$ かつ $r=0$) を除いて, 修正ポアソン回帰の方が小さい値となっており, 推定性能の面でも対数二項回帰と比べて優れていた. リスク比の真値 (RR) と二変数間の相関の値 r を固定したとき, $RR=1.0$ における修正ポアソン回帰でのバイアスの絶対値はイベント発生率 p の上昇とともに小さくなったが, 他二つの状況では大きくなっている. 一方, 対数二項回帰においては p の上昇とともにバイアスの絶対値が概ね上昇していた. RMSE に関しては一定方向の上昇・下降といった特徴的な傾向は見られなかった. また, p と r を固定して RR 間で比較すると, 修正ポアソン回帰については $p=0.1$ の状況で RR の上昇とともにバイアスが増大した. ところが, その他の状況ではバイアスは正の値から負の値に転じており, 点推定値は RR の値によって過大評価にも過小評価にもなり得ることが確認された. 次に RR と p の値を固定したもとの r の値についてバイアスを見ると, バイアスの値が正であるときは r の値が大きくなるほどバイアスの値は概ね上昇していた. 一方, $RR=2.0$ の時のようなバイアスの値が負である場合は, r の上昇とともに正の方向に受かっていることがわかる. 共変量が 10 個の場合と 15 個の場合については, 11.5 節 (表 10 と表 11) に掲載した. 共変量の個数が増えて

も上記と同様の傾向が確認された。

6.4 解析の手法

6.4 節では実際にタンパク質の摂取量と主要イベントの関連性を調べるために行った解析手法を順に紹介する。タンパク質は動物性タンパク質・植物性タンパク質・総タンパク質(動物性タンパク質と植物性タンパク質の和)の三種類を用意し、主要イベントは「追跡開始から一年以内の全死亡または全入院の有無」・「追跡開始から一年以内の心血管死または心不全増悪による入院の有無」の2種類を設定した。まずは、修正ポアソン回帰モデルに投入する調整変数を決めるため、モデル選択を行った。候補変数は「年齢・性別・BMI・NYHA 心機能分類・Hb・HbA1c・eGFR・LVEF・空腹時アルブミン・hs-CRP・血漿BNPの対数値・6分間歩行距離・MLHFQスコア」である。これらの変数の選定は、「日本人集団において動物性・植物性タンパク質と全死亡の関連性を調査した過去の報告」や「エネルギー摂取量が心不全による死亡または入院に及ぼす影響を調査した先行研究」をもとに行った(Budhathoki et al., 2019; Obata et al., 2021; Kinugasa et al., 2022)。候補となる変数群の中には欠測値が含まれているため、モデル選択の前に多重代入法による数値代入を行った(Rubin, 1987; Little and Rubin, 2019)。この多重代入法では、1 データセットあたり 20 回の代入を行い、合計 200 データセットを作成した。このとき、欠測仮定として Missing at random を仮定している。200 個ある多重代入データセットの各データセットに対して、フォールド数 5 の交差検証法(クロスバリデーション)を行った(Stone, 1974)。各 fold において訓練データの AIC(Akaike, 2003) とテストデータに対する AUC を計算し、全てのフォールド、そして全ての多重代入データセットの AIC の平均値と AUC の平均値を算出した。つまり 1 モデルあたり、1000 個分の AIC と AUC が算出され、それら各々の平均値を取ったということである。候補モデルについては Best subset selection 法の適用、つまり考えられる全てのモデルに対してこの計算を行った。最終的なモデルの決定は、まず AIC が最小となるモデルの AIC の値+2 までの範囲のモデルを抽出し、その中で AUC が高いモデルを選んだ。次に、各種タンパク質摂取量と主要イベントの関係を捉えるため、摂取量と制限付き三次スプライン曲線を描画した(Gauthier et al., 2020)。制限付き三次スプライン曲線は最終モデルで選ばれた修正ポアソン回帰モデルを当てはめた。曲線の節点の数は四つで、各種タンパク質の 5%点, 35%点, 65%点, 95%点に配置した。このスプライン曲線の傾きが大きくなるときのタンパク質摂取量の値を高摂取群と低摂取群を分ける閾値として設定した。最後に、多重代入データセットに対して最終的に選んだモデルを適用して、リスク比の点推定値と 95%信頼区間を導出した。なお、イベントが欠測していた症例について、多重代入法による欠測値埋めのステップでは組み入れていたが、スプライン曲線の描画

表 4: 対数二項回帰と修正ポアソン回帰の性能を比較した結果 (共変量が 5 個の場合).

			$r = 0$		$r = 0.1$		$r = 0.4$		$r = 0.8$	
RR	指標	p	mp	lb	mp	lb	mp	lb	mp	lb
0.7	Bias	0.1	0.90	5.29	1.14	20.42	1.60	23.30	3.79	26.18
		0.3	2.30	49.15	1.80	48.10	1.92	46.63	2.94	45.62
		0.5	5.96	45.67	5.89	45.79	5.74	45.88	6.07	45.91
	RMSE	0.1	0.17	0.18	0.17	0.41	0.19	0.46	0.29	0.56
		0.3	0.08	0.77	0.08	0.75	0.08	0.73	0.13	0.72
		0.5	0.08	0.71	0.08	0.71	0.09	0.71	0.11	0.71
	CP	0.1	1.00	0.38	1.00	0.66	1.00	0.74	1.00	0.75
		0.3	1.00	0.59	1.00	0.62	1.00	0.63	1.00	0.64
		0.5	1.00	0.64	1.00	0.64	1.00	0.64	1.00	0.65
1.0	Bias	0.1	3.08	1.32	3.47	4.37	4.02	4.17	7.16	6.97
		0.3	1.38	18.93	1.48	17.29	1.96	17.04	3.11	15.64
		0.5	0.09	15.79	0.14	15.88	0.06	15.81	1.00	15.87
	RMSE	0.1	0.26	0.26	0.26	0.31	0.28	0.33	0.42	0.44
		0.3	0.12	0.60	0.12	0.59	0.13	0.59	0.19	0.57
		0.5	0.07	0.57	0.07	0.57	0.08	0.57	0.12	0.57
	CP	0.1	1.00	0.58	1.00	0.73	1.00	0.75	1.00	0.77
		0.3	1.00	0.62	1.00	0.63	1.00	0.63	1.00	0.65
		0.5	1.00	0.64	1.00	0.64	1.00	0.65	1.00	0.65
2.0	Bias	0.1	7.73	-80.22	7.85	-82.46	7.97	-78.74	15.72	-70.14
		0.3	-22.82	-81.07	-21.99	-82.71	-21.21	-82.96	-20.00	-84.29
		0.5	-43.59	-84.21	-43.37	-84.12	-42.91	-84.19	-41.00	-84.13
	RMSE	0.1	0.47	0.93	0.50	0.94	0.53	0.97	0.84	1.03
		0.3	0.31	0.99	0.30	1.00	0.32	1.00	0.40	1.01
		0.5	0.46	1.00	0.46	1.00	0.46	1.00	0.47	1.00
	CP	0.1	1.00	0.58	1.00	0.73	1.00	0.74	1.00	0.78
		0.3	1.00	0.62	1.00	0.63	1.00	0.63	1.00	0.65
		0.5	1.00	0.64	1.00	0.64	1.00	0.65	1.00	0.65

¹ バイアスについては 100 倍した値を記載

² RR: リスク比の真値. CP: リスク比の推定値の収束割合. p : イベント発生率. r : 二変数間の相関の値.
mp: 修正ポアソン回帰, lb: 対数二項回帰.

および修正ポアソン回帰によるリスク比の推定時点では集団から除いた (Von Hippel, 2007; Austin et al., 2021). また, 各種タンパク質の総エネルギーに占める割合についても上記とほとんど同様の解析を行った. さらに, 対数二項回帰モデルを利用して求めた結果についても検討した. 解析は R(ver.4.4.1; R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria) で行った.

6.5 結果

6.5.1 各タンパク質摂取量群に対するデータの特性

登録された慢性心不全患者 156 名のうち, 追跡不能症例 7 例と BDHQ の欠測があった 3 例の合計 10 例を除外して, 146 例を解析対象者とした. 表 5 は, 対象集団のベースライン特性である. 総エネルギーにおける各タンパク質の割合について, 閾値で高摂取群と低摂取群に分けて集計した. 総タンパク質における閾値は 16.0%, 動物性タンパク質における閾値は 10.3%, 植物性タンパク質における閾値は 6.0%と設定した. この閾値の値は, 後述する制限つき三次スプライン曲線の傾きから決定した. 総タンパク質における低摂取群は 76 人, 高摂取群は 70 人であった. 動物性タンパク質における低摂取群は 89 人, 高摂取群は 57 人であった. 植物性タンパク質における低摂取群は 61 人, 高摂取群は 85 人であった. また, LVEF, eGFR, hs-CRP, 血漿 BNP, 6 分間歩行距離については欠測値が見られたが, 欠測値は除いて集計した. 集団全体の年齢の中央値 [IQR] は 68 歳 [60 歳-77 歳] であり, どの各タンパク質群でも高齢者が多くを占めている. 性別を見ると, 女性の総タンパク質高摂取群, 動物性タンパク質高摂取群の占める割合は低摂取群と比較して多く, 15%から 20%近く差があった. LVEF は 144 人について集計しており, どの群も中央値は 45%であった. NYHA 心機能分類では最重度 class IV の患者は 0 人で, class III が全体の 10%程度を占めた. また, 心不全の主要な原因となる疾患の既往歴・ベースライン時点の検査値・日常的に服用する薬についても集計した. 心疾患の中では虚血性心疾患と拡張型心筋症による心不全の症例が多く, どちらも全体で 46 症例観察された. 血漿 BNP の中央値 [IQR] は, 156pg/mL[76pg/mL-374pg/mL] であった. 6 分間歩行距離の中央値 [IQR] は 484m[346m-490m], MLHFQ スコアの中央値 [IQR] は 15[5-27] であった. 動物性タンパク質の摂取群と植物性タンパク質の摂取群を組み合わせた 4 群におけるそれぞれのベースライン特性については, 表 12 に示している.

表 6 は, 1 日あたりの食事摂取量を BDHQ の項目に沿って集計した結果である. タンパク質・脂質・炭水化物においては, 総エネルギー摂取量に対する割合で集計した. 総エネルギー摂取量の中央値 [IQR] は 1624[1273-1989] であり, 動物性タンパク質の高摂取割合群では 1699[1294-1991], 低摂取割合群では 1580[1272-1991] であった. 植物性タンパ

表 5: 対象集団のベースライン特性

特性	全体 (N=146)	総タンパク質		動物性タンパク質		植物性タンパク質	
		低摂取群 (N=76)	高摂取群 (N=70)	低摂取群 (N=89)	高摂取群 (N=57)	低摂取群 (N=61)	高摂取群 (N=85)
年齢 (歳)	68 (60-77)	64 (56-73)	71 (63-79)	66 (58-74)	70 (62-80)	66 (56-74)	69 (63-77)
女性	46 (32)	17 (22)	29 (41)	22 (25)	24 (42)	17 (28)	29 (34)
BMI (kg/m ²)	22.9 (20.6-25.6)	23.2 (20.9-25.4)	22.5 (20.2-25.7)	23.3 (20.9-25.6)	22.5 (19.8-25.5)	23.1 (20.1-25.4)	22.7 (20.8-25.7)
LVEF (%)	45 (30-56)	45 (31-56)	45 (29-61)	47 (33-57)	43 (27-53)	48 (30-55)	43 (32-59)
NYHA class III	15 (10)	7 (9)	8 (11)	6 (7)	9 (16)	8 (13)	7 (8)
心不全の主な原因:							
虚血性心疾患	46 (32)	22 (29)	24 (34)	27 (30)	19 (33)	14 (23)	32 (38)
拡張型心筋症	46 (32)	26 (34)	20 (29)	29 (33)	17 (30)	23 (38)	23 (27)
その他の心疾患	54 (37)	28 (37)	26 (37)	33 (37)	21 (37)	24 (39)	30 (35)
高血圧	79 (54)	35 (46)	44 (63)	46 (52)	33 (58)	26 (43)	53 (62)
糖尿病	38 (26)	14 (18)	24 (34)	22 (25)	16 (28)	10 (16)	28 (33)
脂質異常症	103 (71)	55 (72)	48 (69)	67 (75)	36 (63)	40 (66)	63 (74)
検査値:							
ヘモグロビン (g/dl)	13.2 (11.9-14.3)	13.7 (12.4-14.6)	12.2 (11.5-13.7)	13.6 (12.2-14.5)	12.2 (11.6-13.7)	13.4 (12.2-14.5)	12.9 (11.7-14.2)
血清アルブミン (g/dl)	4.2 (3.9-4.4)	4.2 (4.0-4.4)	4.1 (3.9-4.3)	4.2 (3.9-4.4)	4.1 (4.0-4.4)	4.2 (3.9-4.4)	4.2 (4.0-4.3)
eGFR (ml/min/1.73m ²)	54.3 (38.4-67.6)	54.9 (39.1-71.5)	54.3 (37.9-63.8)	55.8 (40.6-71.4)	49.0 (35.7-63.0)	58.9 (42.5-71.0)	51.3 (37.1-64.6)
HbA1c (%)	5.8 (5.6-6.2)	5.8 (5.6-6.2)	5.9 (5.7-6.3)	5.8 (5.6-6.3)	5.8 (5.6-6.2)	5.7 (5.5-6.1)	5.8 (5.7-6.3)
hs-CRP (mg/dl)	484 (250-1465)	471 (214-1290)	590 (279-1640)	477 (223-1298)	550 (263-1630)	545 (296-1293)	456 (242-1710)
血漿 BNP (pg/ml)	156 (76-374)	134 (53-271)	200 (91-426)	133 (51-358)	203 (106-498)	132 (74-239)	190 (77-410)
6 分間歩行試験 (m)	432 (346-490)	442 (356-490)	414 (347-499)	434 (361-490)	417 (305-498)	431 (335-497)	432 (350-485)
MLHFQ	15 (5-27)	15 (5-23)	15 (4-29)	15 (4-24)	13 (5-28)	15 (6-25)	16 (4-28)
服用薬:							
ACE 阻害薬または ARB	111 (76)	57 (75)	54 (77)	64 (72)	47 (82)	44 (72)	67 (79)
β 遮断薬	129 (88)	71 (93)	58 (83)	80 (90)	49 (86)	53 (87)	76 (89)
MRA	85 (58)	42 (55)	43 (61)	48 (54)	37 (65)	37 (61)	48 (56)

中央値 (第 1 四分位数-第 3 四分位数) または n(%) で示す。コホート全体を総タンパク質 (閾値:16.0%), 動物性タンパク質 (閾値:10.3%), 植物性タンパク質 (閾値:6.0%) の高摂取群と低摂取群に分けた。すべての閾値は制限付き三次スプライン曲線の傾きの値に基づいて決定した。

ク質においては高摂取割合群では1619[1276-1924], 低摂取割合群では1645[1246-2114]であった。脂質摂取量の総エネルギー摂取量に対する割合について、集団全体の中央値[IQR]は25.2[21.9 - 29.8]であった。また、動物性タンパク質の高摂取割合群と低摂取割合群ではそれぞれ29.2[24.9-32.7], 23.3[18.4-26.1]であり、植物性タンパク質の高摂取割合群と低摂取割合群ではそれぞれ25.2[22.0-29.2], 25.2[21.8-29.9]であった。炭水化物摂取量の総エネルギー摂取量に対する割合について、集団全体の中央値[IQR]は54.6[49.4 - 59.6]であった。また、動物性タンパク質の高摂取割合群と低摂取割合群ではそれぞれ49.6[44.2-54.5], 57.8[53.1-61.7]であり、植物性タンパク質の高摂取割合群と低摂取割合群ではそれぞれ55.8[51.2-60.2], 52.2[46.5-58.4]であった。さらに、食品群として肉類、魚介類、果物類などについても集計を行った。中央値で比較すると動物性タンパク質においても植物性タンパク質においても高摂取群の方が低摂取群よりも多く摂取されており、動物性タンパク質では魚介類や牛・乳製品、植物性タンパク質では大豆製品や野菜の摂取量が群間で差が大きかった。

追跡開始から12ヶ月間の有害事象を集計した結果が、表7である。集団全体の4%にあたる6例の全死亡が発生した。6例のうち、心不全の悪化または不整脈による心血管死が4例(全体の3%), がんまたは脳卒中による非心血管死が2例(集団全体の1%)であった。集団全体の19%にあたる28例の患者が何らかの原因で入院した。心不全による入院が15例(集団全体の10%), 不整脈・感染症・がん・脳卒中・腎不全などの非心不全による入院が13例(集団全体の9%)であった。本研究の主要評価項目について、31例の全死亡および全入院(集団全体の21%)と18例の心血管死および心不全による入院(集団全体の12%)が記録された。タンパク質摂取量に基づく結果の比較、全死亡および全入院の割合は、動物性タンパク質の高摂取群が低摂取群と比較して(それぞれ32%と15%), また植物性タンパク質の低摂取群が高摂取群と比較して(それぞれ26%と18%), より高かった。同様に、心血管死および心不全による入院の割合は、動物性タンパク質の高摂取群が低摂取群と比較して(それぞれ19%と8%), また植物性タンパク質の低摂取群が高摂取群と比較して(それぞれ15%と11%), より高かった。動物性タンパク質と植物性タンパク質の摂取群を組み合わせた4群における集計結果については、表13に示した。

ここで、各種タンパク質の値と総エネルギー摂取量、脂質摂取量、炭水化物摂取量について描いた散布図を図6から図9に示す。図6と図7では、タンパク質(4kcal/g)・脂質(9kcal/g)・炭水化物(4kcal/g)のカロリー消費量を値として使用して主要評価項目別に描画しており、図8と図9では、タンパク質・脂質・炭水化物については総エネルギー摂取量との比を値として使用して主要評価項目別に描画している。省略のため、以降では文中における「全死亡または再入院イベント」をイベント All, 「心血管死死亡または心不全増悪による入院イベント」をイベント CHF と呼ぶこととする。図8と図9の

表 6: BDHQ によって算出された 1 日あたりの食事摂取量

栄養素または食品項目	全体 (N=146)		総タンパク質		動物性タンパク質		植物性タンパク質	
	(N=146)		低摂取群 (N=76)	高摂取群 (N=70)	低摂取群 (N=89)	高摂取群 (N=57)	低摂取群 (N=61)	高摂取群 (N=85)
総エネルギー (kcal/日)	1624 (1273-1989)		1558 (1266-2001)	1663 (1303-1968)	1580 (1272-1985)	1699 (1294-1991)	1645 (1246-2114)	1619 (1276-1924)
総タンパク質 (%)	15.9 (13.7-18.4)		13.8 (12.5-14.6)	18.4 (17.2-20.7)	14.2 (12.7-15.5)	18.7 (17.5-21.0)	14.6 (12.8-16.9)	16.6 (14.5-18.5)
動物性タンパク質 (%)	9.3 (7.5-11.7)		7.6 (6.3- 8.7)	11.9 (10.3-14.4)	7.9 (6.5- 9.1)	12.4 (11.3-14.9)	9.1 (7.6-11.5)	9.4 (7.5-11.9)
植物性タンパク質 (%)	6.2 (5.6-6.8)		6.0 (5.5-6.5)	6.4 (5.9-7.2)	6.2 (5.6-7.0)	6.2 (5.5-6.5)	5.5 (5.0-5.7)	6.6 (6.4-7.3)
総脂質 (%)	25.2 (21.9-29.8)		22.9 (17.9-26.1)	28.4 (24.9-31.9)	23.3 (18.4-26.1)	29.2 (24.9-32.7)	25.2 (21.8-29.9)	25.2 (22.0-29.2)
動物性脂質 (%)	12.0 (9.9-15.4)		10.2 (7.6-11.9)	14.8 (12.3-18.4)	10.4 (8.0-11.9)	15.4 (13.4-18.9)	12.9 (10.2-16.3)	11.5 (9.8-14.7)
植物性脂質 (%)	12.2 (9.9-15.3)		11.1 (9.0-14.9)	12.9 (10.9-15.5)	12.2 (9.4-15.3)	12.2 (10.5-15.3)	11.1 (8.8-14.3)	12.9 (10.8-15.6)
飽和脂肪酸 (%)	6.6 (5.5-8.2)		5.8 (4.7-7.3)	7.3 (6.2-8.7)	5.9 (4.9-7.2)	8.0 (6.5-9.1)	7.2 (5.8-8.6)	6.3 (5.4-7.7)
一価不飽和脂肪酸 (%)	8.8 (7.4-10.6)		7.8 (6.3-9.3)	9.9 (8.3-11.4)	8.2 (6.4-9.6)	10.0 (8.4-11.5)	8.6 (7.3-10.7)	8.9 (7.4-10.4)
多価不飽和脂肪酸 (%)	6.1 (4.9-7.1)		5.0 (4.2-6.5)	6.8 (5.9-7.9)	5.4 (4.4-6.9)	6.6 (5.8-7.9)	5.6 (4.7-6.9)	6.4 (5.3-7.1)
炭水化物 (%)	54.6 (49.4-59.6)		58.5 (53.0-63.0)	51.3 (46.2-55.5)	57.8 (53.1-61.7)	49.6 (44.2-54.5)	52.2 (46.5-58.4)	55.8 (51.2-60.2)
総食物繊維 (g/1000 kcal)	6.6 (5.1-8.3)		5.8 (4.5-6.8)	7.8 (6.4-8.9)	6.0 (4.9-8.0)	7.5 (5.5-8.6)	5.2 (4.2-6.6)	7.6 (6.2-9.0)
水溶性食物繊維 (g/1000 kcal)	1.6 (1.2-2.1)		1.3 (1.0-1.7)	1.9 (1.5-2.3)	1.5 (1.1-2.0)	1.9 (1.4-2.2)	1.3 (0.9-1.7)	2.0 (1.5-2.3)
不溶性食物繊維 (g/1000 kcal)	4.7 (3.7-6.0)		4.2 (3.4-5.0)	5.5 (4.6-6.4)	4.4 (3.5-5.8)	5.2 (4.1-6.1)	3.8 (3.1-4.6)	5.5 (4.5-6.4)
食塩 (g/1000 kcal)	5.9 (5.0-6.8)		5.3 (4.6-5.9)	6.7 (6.1-7.9)	5.5 (4.7-5.9)	6.8 (6.1-8.3)	5.7 (4.7-6.6)	6.0 (5.3-6.8)
豚肉・牛肉 (g/1000 kcal)	13.3 (7.2-21.8)		11.5 (7.2-18.9)	16.8 (7.6-24.6)	11.8 (7.1-19.8)	16.5 (7.6-27.5)	13.4 (7.6-21.6)	13.1 (6.6-21.9)
鶏肉 (g/1000 kcal)	9.1 (4.7-18.6)		8.0 (4.4-14.0)	14.0 (6.2-22.6)	8.3 (4.4-14.5)	15.1 (6.8-23.3)	7.6 (4.4-16.3)	11.9 (6.0-20.6)
加工肉 (g/1000 kcal)	2.6 (1.4-5.6)		2.7 (1.5-5.1)	2.6 (1.1-5.8)	2.9 (1.5-5.8)	2.5 (1.1-5.0)	2.6 (1.4-5.1)	2.7 (1.4-5.8)
魚介類 (g/1000 kcal)	52.3 (35.3-75.9)		37.6 (23.8-50.1)	75.7 (54.9-94.9)	42.3 (24.7-52.9)	78.8 (66.3-100.8)	52.2 (38.2-74.7)	52.3 (32.5-76.1)
卵 (g/1000 kcal)	16.7 (9.3-27.7)		14.6 (6.9-21.0)	23.8 (13.7-32.1)	15.3 (7.0-22.1)	25.2 (13.7-31.6)	16.9 (7.4-28.7)	16.5 (10.5-27.2)
牛乳・乳製品 (g/1000 kcal)	62.6 (13.9-102.3)		46.4 (5.2-92.3)	76.2 (18.2-121.3)	40.8 (0.0-92.2)	83.0 (40.5-132.2)	57.8 (14.3-107.5)	66.5 (10.4-102.0)
大豆製品 (g/1000 kcal)	91.4 (62.8-139.4)		79.2 (44.7-115.4)	99.9 (80.7-150.2)	87.3 (52.6-128.5)	97.6 (71.5-150.3)	76.3 (42.0-105.4)	100.0 (76.0-150.2)
野菜 (g/1000 kcal)	139.0 (87.6-189.0)		119.0 (61.9-161.5)	168.9 (118.8-213.5)	127.8 (73.6-183.1)	154.2 (111.7-202.3)	118.3 (61.9-150.6)	161.2 (110.6-211.9)
いも類 (g/1000 kcal)	19.9 (11.6-35.0)		15.8 (8.5-30.8)	25.6 (12.9-42.4)	19.4 (8.7-35.1)	22.0 (12.0-33.5)	14.8 (8.6-26.2)	27.7 (12.2-44.0)
果物 (g/1000 kcal)	47.4 (21.8-82.9)		32.9 (13.8-72.8)	63.3 (32.3-104.1)	39.5 (15.7-80.5)	56.9 (29.6-97.9)	34.0 (15.5-74.6)	56.9 (25.2-103.2)

中央値 (第 1 四分位数-第 3 四分位数) または n(%) で示す。1 日あたりの栄養素または食品摂取量の単位 (総エネルギー量を除く) は、総エネルギー量に対する百分率または総エネルギー量 1000kcal あたりの g 数で示されている。コホート全体を総タンパク質 (閾値:16.0%), 動物性タンパク質 (閾値:10.3%), 植物性タンパク質 (閾値:6.0%) の高摂取群と低摂取群に分けた。すべての閾値は制限付き三次スプライン曲線の傾きの値に基づいて決定した。単位が%の項目については、総エネルギー摂取量に対する割合で集計している。

表 7: 12 ヶ月の追跡期間中の有害事象

有害事象	全体 (N=146)	総タンパク質			動物性タンパク質			植物性タンパク質		
		低摂取群 (N=76)	高摂取群 (N=70)	低摂取群 (N=89)	高摂取群 (N=57)	低摂取群 (N=61)	高摂取群 (N=85)			
主要評価項目:										
全死亡または全入院	31 (21)	15 (20)	16 (23)	13 (15)	18 (32)	16 (26)	15 (18)			
心血管死または心不全入院	18 (12)	9 (12)	9 (13)	7 (8)	11 (19)	9 (15)	9 (11)			
副次評価項目:										
全死亡	6 (4)	3 (4)	3 (4)	1 (1)	5 (9)	4 (7)	2 (2)			
心血管死	4 (3)	3 (4)	1 (1)	1 (1)	3 (5)	3 (5)	1 (1)			
非心血管死	2 (1)	0 (0)	2 (3)	0 (0)	2 (4)	1 (2)	1 (1)			
全入院	28 (19)	13 (17)	15 (21)	12 (13)	16 (28)	15 (25)	13 (15)			
心不全入院	15 (10)	7 (9)	8 (11)	6 (7)	9 (16)	7 (11)	8 (9)			
非心不全入院	13 (9)	6 (8)	7 (10)	6 (7)	7 (12)	8 (13)	5 (6)			

中央値 (第 1 四分位数-第 3 四分位数) または n(%) で示す. コホート全体を総タンパク質 (閾値:16.0%), 動物性タンパク質 (閾値:10.3%), 植物性タンパク質 (閾値:6.0%) の高摂取群と低摂取群に分けた. すべての閾値は制限付き三次スプライン曲線の傾きの値に基づいて決定した.

結果から、タンパク質摂取量が増えると、総エネルギー摂取量、炭水化物摂取量、脂質摂取量も増加傾向にある。特に、総エネルギー摂取量と炭水化物に関しては、動物性タンパク質における結果よりも植物性タンパク質における結果でこの傾向が強く見られる。イベント発生が起きた対象者の分布を確認すると、イベント AII に関しては点在しており、特定のタンパク質摂取量で多く発生している様子は見られなかった。一方、イベント CHF に関しては植物性タンパク質摂取量が比較的少ないところで大部分が確認された。

次に、総エネルギー摂取量におけるタンパク質摂取量の割合の値を使って描いた散布図(図6と図7)を評価する。タンパク質摂取割合の増加と総エネルギー摂取量には特徴的な上昇・下降傾向は見られない。総タンパク質摂取割合と動物性タンパク質摂取割合が増加すると、炭水化物摂取割合は減少する傾向にあった。一方、植物性タンパク質の摂取割合の上昇とともに炭水化物摂取割合は増加の傾向が見られる。また、総タンパク質摂取割合と動物性タンパク質摂取割合が増加すると脂質の摂取割合に増加傾向が見られた。イベント AII について、植物性タンパク質摂取割合が低く、総エネルギー摂取量が低い領域に多く集まっていた。この特徴は総エネルギー摂取量だけでなく炭水化物摂取割合、脂質摂取割合にも見られた。イベント CHF に関しては点在しており、特定のタンパク質摂取量で多く発生している様子は見られなかった。

6.5.2 解析結果 1: 総エネルギー摂取量に占める特定のタンパク質の割合とイベントの関連

変数選択によって選ばれた修正ポアソン回帰モデルを使って、各タンパク質の高摂取群と低摂取群を分ける閾値を決めた結果を報告する。総エネルギー量に対するタンパク質の摂取割合と修正ポアソン回帰に当てはめ推定されたリスク比との関係性を制限つき三次スプライン曲線で描画した結果を図10に示す。総タンパク質の割合は両イベントともにリスク比が1.0周りでほぼ横ばいになっていた。動物性タンパク質の割合が約10%を超えると、イベント AII に対するリスク比が1.0を上回り、約5-10%程度では1.0を下回るが、約5%未満になるとまた上昇した。また、イベント CHF についても動物性タンパク質の割合が約10%を超えるとリスク比が1.0を超えた。さらに、植物性タンパク質の割合が約6%を下回るとリスク比が1.0を超えており、約6%以上摂取するとリスク比は1.0を下回った。同じイベント内でタンパク質摂取割合間の結果を比較すると、動物性タンパク質の割合が増加するとリスク比は上昇傾向にあり、植物性タンパク質の割合が増加するとリスク比は下降傾向にある。総タンパク質の割合は増減してもリスク比の値を大きく変化させていなかった。制限つき三次スプライン曲線の傾きが水平になる点あるいは傾きが最大となる点を参考に、総タンパク質低摂取割合群は16.0%未満、高摂取割合群は16.0%以上と定義した。また、動物性タンパク質低摂取割合群は

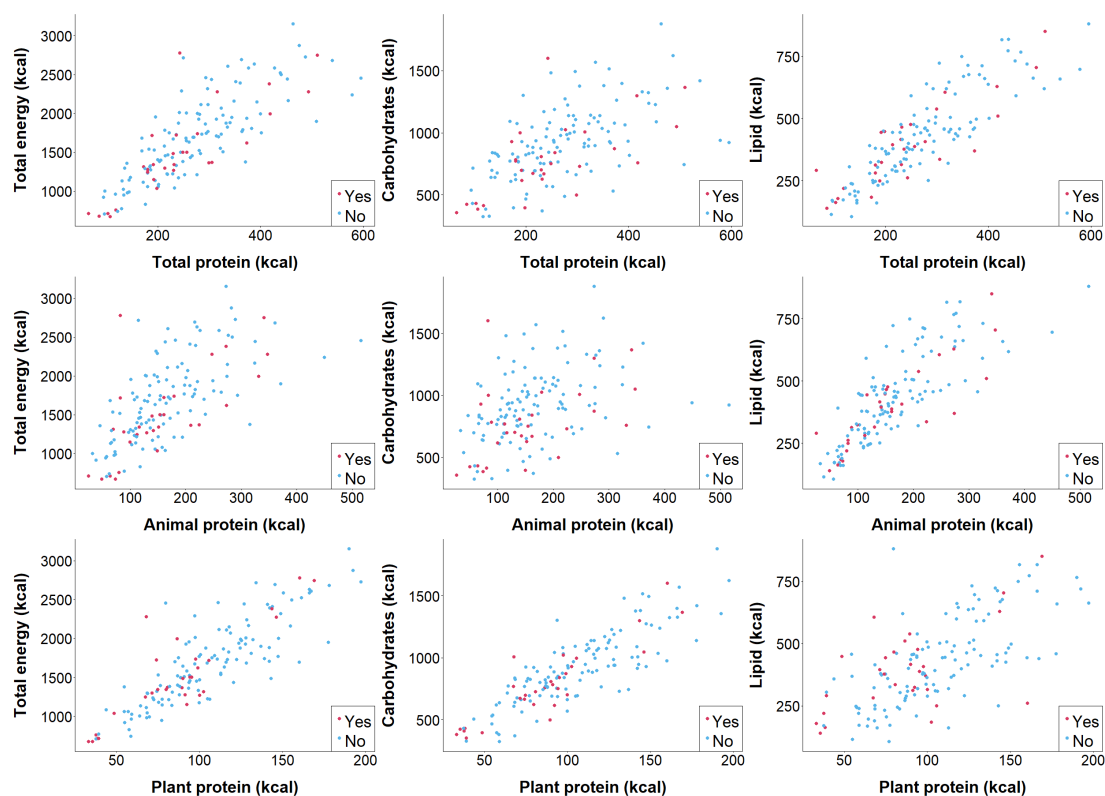


図 6: 各種タンパク質の値と総エネルギー摂取量 (Total Energy), 炭水化物摂取量 (Carbohydrates), 脂質摂取量 (Lipid) の値との散布図. 凡例は全死亡または全入院イベントの有無を指しており, ある場合は Yes, ない場合は No と色分けしている.

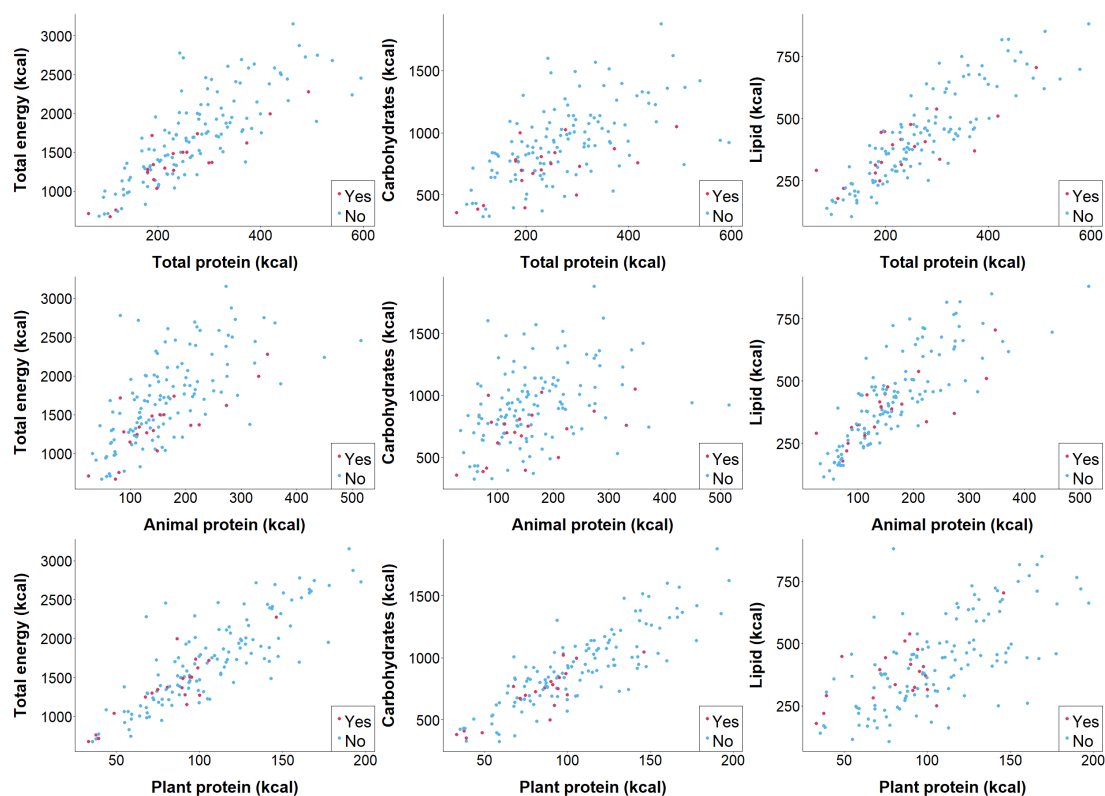


図 7: 各種タンパク質の値と総エネルギー摂取量 (Total Energy), 炭水化物摂取量 (Carbohydrates), 脂質摂取量 (Lipid) の値との散布図. 凡例は心血管死または心不全による入院イベントの有無を指しており, ある場合は Yes, ない場合は No と色分けしている.

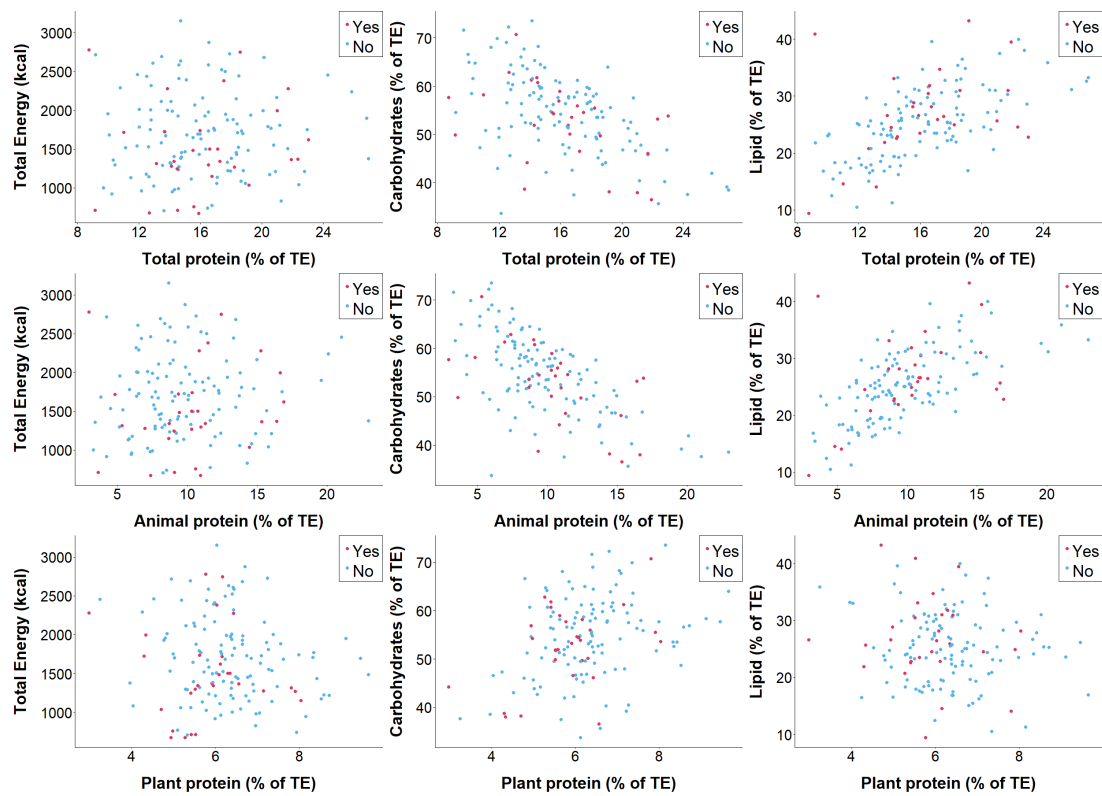


図 8: 各種タンパク質の総エネルギー摂取量 (TE) に対する割合と総エネルギー摂取量 (Total Energy), 炭水化物摂取量 (Carbohydrates), 脂質摂取量 (Lipid) の値との散布図. 凡例は全死亡または全入院イベントの有無を指しており, ある場合は Yes, ない場合は No と色分けしている.

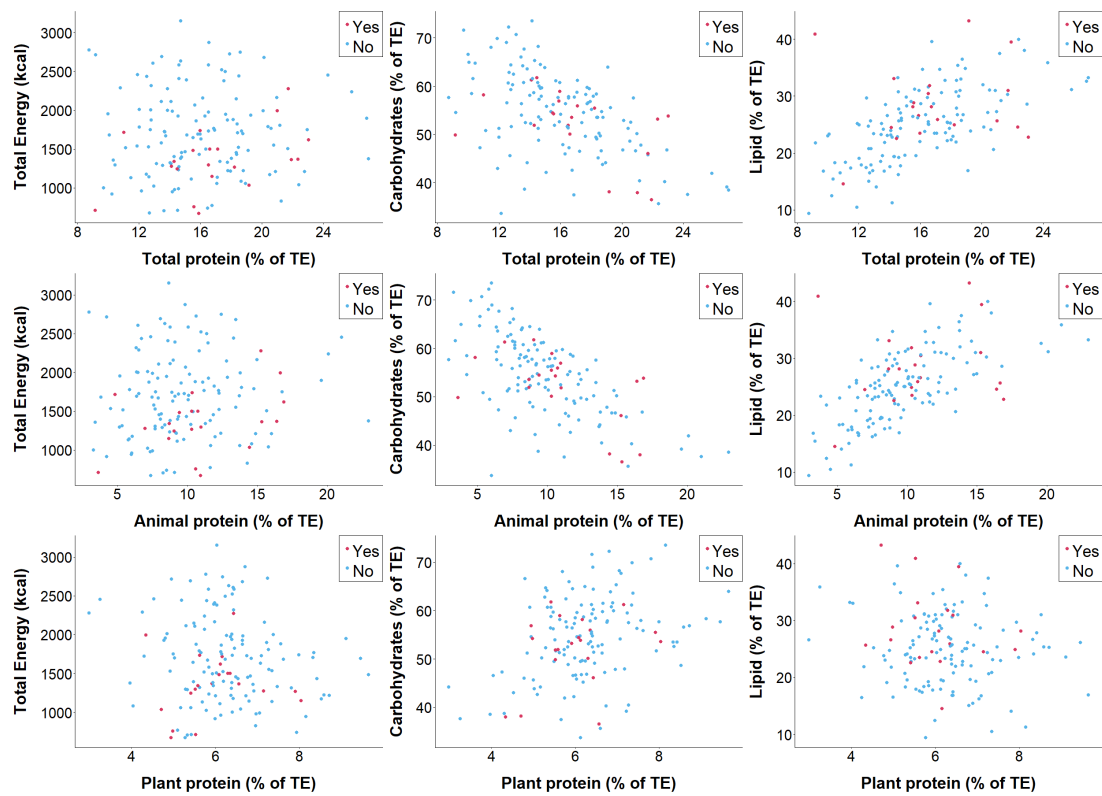


図 9: 各種タンパク質の総エネルギー摂取量 (TE) に対する割合と総エネルギー摂取量 (Total Energy), 炭水化物摂取量 (Carbohydrates), 脂質摂取量 (Lipid) の値との散布図. 凡例は心血管死または心不全によるイベントの有無を指しており, ある場合は Yes, ない場合は No と色分けしている.

10.3%未満, 高摂取割合群は 10.3%以上と定義した. さらに, 植物性タンパク質低摂取割合群は 6.0%未満, 高摂取割合群は 6.0%以上と定義した. これらの閾値は, 両イベントともに同じ値で設定した.

対数二項回帰に当てはめて推定されたリスク比との関係性を制限つき三次スプライン曲線で描画した結果を図 11 に示す. 動物性タンパク質は, 修正ポアソン回帰における結果と同様に, 割合が約 10%を超えると, イベント AII に対するリスク比が 1.0 を上回り, 約 5-10%程度では 1.0 を下回るが, 約 5%未満になるとまた上昇した. また, イベント CHF についても動物性タンパク質の割合が約 10%を超えるとリスク比が 1.0 を超えた. さらに, 植物性タンパク質の割合が約 6%を下回るとリスク比が 1.0 を超えており, 約 6%以上摂取するとリスク比は 1.0 を下回った. この植物性タンパク質に対する結果も修正ポアソン回帰における結果と同様である. 総タンパク質の割合は両イベントともにリスク比が 1.0 周りでほぼ横ばいになっていたが, 修正ポアソン回帰における結果と比べると取りうるリスク比の値の範囲が広いという結果が得られた.

表 8 は, タンパク質摂取割合の値によって 2 群に分けた場合の修正ポアソン回帰分析および対数二項回帰分析によって推定されたリスク比と 95%信頼区間の結果である. タンパク質ごとにモデルを 2 つ採用しており, 各イベントごとに結果をまとめた. まずはイベント AII の結果を述べる. 動物性タンパク質の低摂取割合群に対する高摂取割合群のリスク比は, 1.78[95%CI: 0.93–3.40](モデル 1), 1.71[95%CI: 0.90–3.24](モデル 2) であった. 一方, 植物性タンパク質の高摂取割合群に対する低摂取割合群のリスク比は, 1.56[95%CI: 0.79–3.09](モデル 1), 1.82[95%CI: 0.96–3.45](モデル 2) であった. 総タンパク質の低摂取割合群に対する高摂取割合群のリスク比は, 1.15[95%CI: 0.60–2.20](モデル 1), 1.18[95%CI: 0.60–2.31](モデル 2) であった. 続いて, イベント CHF の結果を述べる. 動物性タンパク質の低摂取割合群に対する高摂取割合群のリスク比は, 1.93[95%CI: 0.81–4.56](モデル 1), 1.81[95%CI: 0.81–4.07](モデル 2) であった. 一方, 植物性タンパク質の高摂取割合群に対する低摂取割合群のリスク比は, 1.41[95%CI: 0.62–3.18](モデル 1), 1.35[95%CI: 0.58–3.14](モデル 2) であった. 総タンパク質の低摂取割合群に対する高摂取割合群のリスク比は, 1.13[95%CI: 0.47–2.72](モデル 1), 1.19[95%CI: 0.49–2.92](モデル 2) であった.

表 8 では, 対数二項回帰に当てはめたリスク比とその 95%信頼区間を推定した結果も表示した. 多重代入データのうち 1 つ以上のデータセットでリスク比の推定ができた場合に, “*” のマークをつけた. また, どの多重代入データセットに対してもリスク比の推定値が収束しなかった場合は, “NA” とした. イベント AII の総タンパク質に対する結果は全ての多重代入データセットでリスク比の推定値が収束しており, リスク比は 1.00[95%CI: 0.52–1.91](モデル 2) であった. またモデル 1 は一部の多重代入データセットだけで計算した結果であるが, リスク比は 0.97[95%CI: 0.52–1.83] であった. イベント

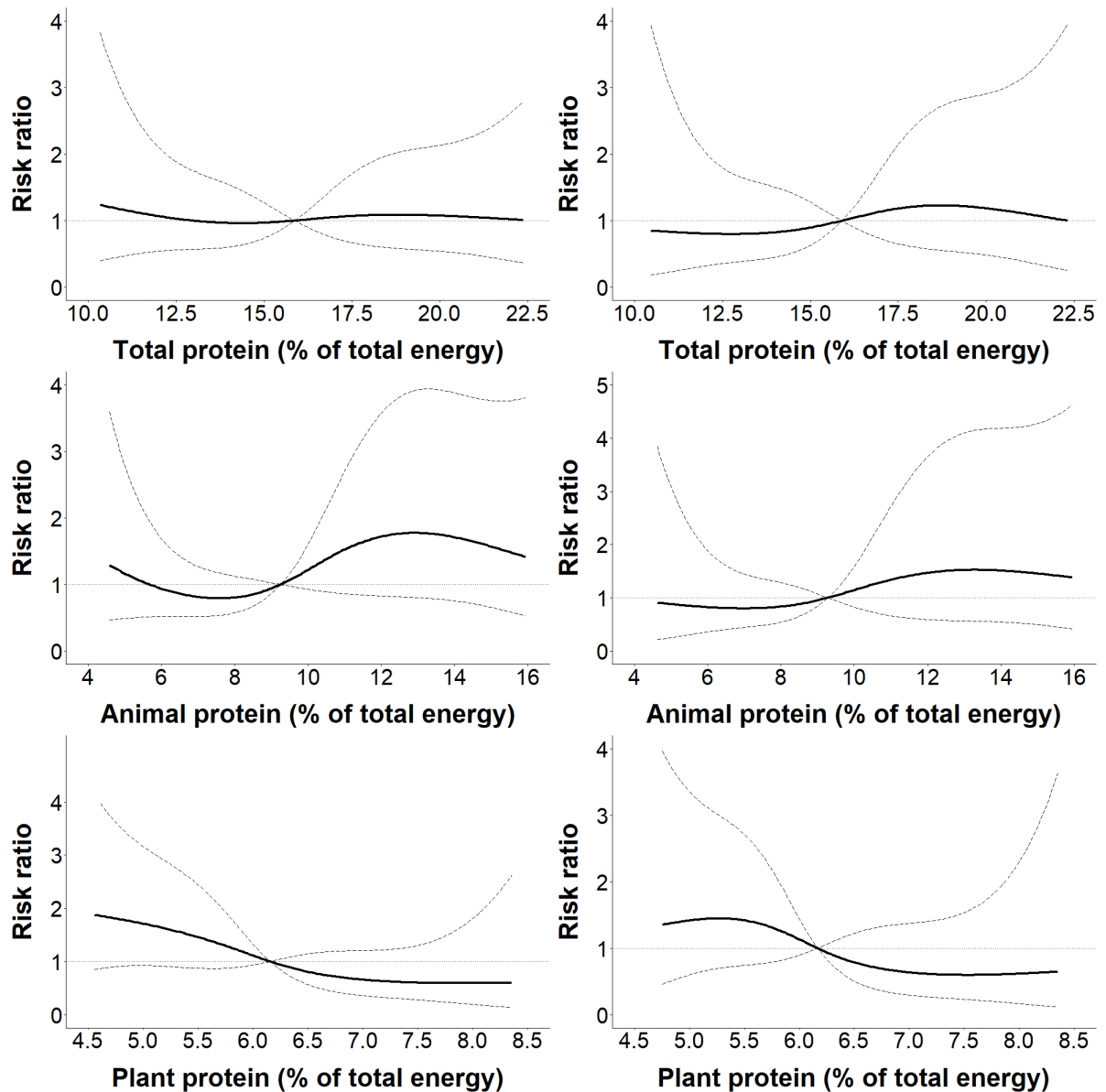


図 10: 総エネルギー量に占める各種タンパク質の割合に対するリスク比の制限つき三次スプライン曲線. 修正ポアソン回帰モデルに当てはめている. 点線は 95%信頼区間の上限値と下限値を表している. 左列は全死亡または全入院イベントに対する結果であり, 右列は心血管死死亡または心不全増悪による入院イベントに対する結果. モデル選択の結果を受けて, BNP の対数値で調整している.

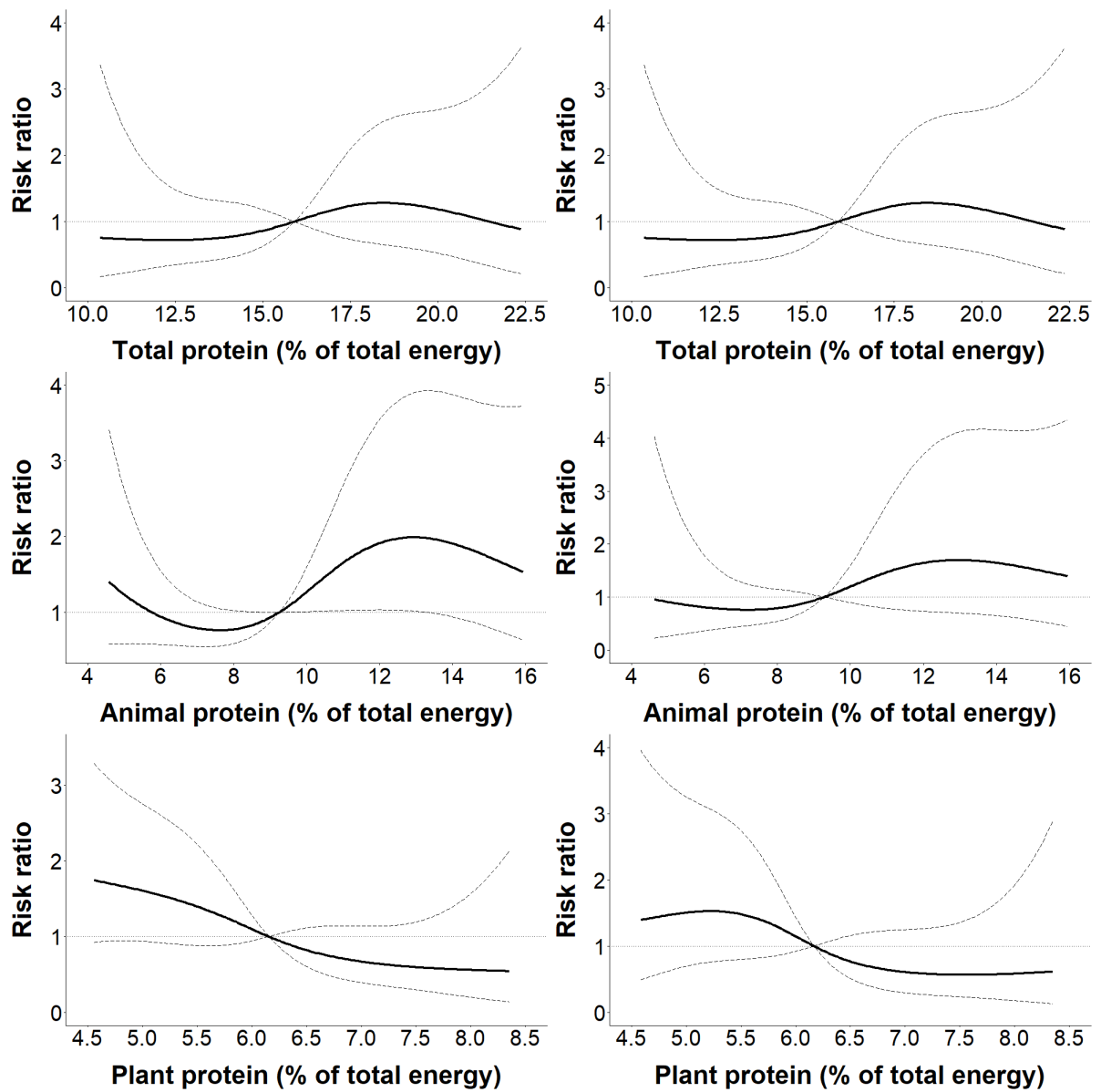


図 11: 総エネルギー量に占める各種タンパク質の割合に対するリスク比の制限つき三次スプライン曲線. 対数二項回帰モデルに当てはめている. 点線は95%信頼区間の上限値と下限値を表している. 左列は全死亡または全入院イベントに対する結果であり, 右列は心血管死死亡または心不全増悪による入院イベントに対する結果.

ト CHF についてはモデル2だけ推定結果が算出され、リスク比 1.12[95%CI: 0.48–2.60] となった。動物性タンパク質・植物性タンパク質に対する結果では全ての多重代入データセットで推定不能であった。

6.5.3 解析結果 2: タンパク質の摂取量とイベントの関連

図 12 は、タンパク質と修正ポアソン回帰に当てはめ推定されたリスク比との関係性を制限つき三次スプライン曲線で描画した結果である。総タンパク質は両イベントともにリスク比が 1.0 周りでほぼ横ばいになっていた。イベント AII に対する結果は、60g 未満の範囲においてリスク比はわずかに 1.0 を超えていた。イベント CHF では 1.0 周りでほぼ横ばいであった。動物性タンパク質における結果では、約 40g を超えるとイベント AII に対するリスク比がわずかに上昇したが、イベント CHF に対するリスク比は約 70g を境にリスク比が 1.0 を上回る結果になった。植物性タンパク質における結果では、両イベントともに約 25g を超えるとリスク比が 1.0 を下回り、タンパク質量の増加とともにリスク比は単調に減少した。制限つき三次スプライン曲線の傾きが最大となる点あるいは傾きの符号が変わる点を参考に閾値を設定した。イベント AII については、総タンパク質低摂取群は 85.5g 未満 (85.5g 以上を高摂取群)、動物性タンパク質低摂取群は 36.0g 未満 (36.0g 以上を高摂取群)、植物性タンパク質低摂取群は 22.6g 未満 (22.6g 以上を高摂取群) と定義した。また、イベント HF については、総タンパク質低摂取群は 81.5g 未満 (81.5g 以上を高摂取群)、動物性タンパク質低摂取群は 52.6g 未満 (52.6g 以上を高摂取群)、植物性タンパク質低摂取群は 29.6g 未満 (29.6g 以上を高摂取群) と定義した。

対数二項回帰に当てはめ推定されたリスク比との関係性を制限つき三次スプライン曲線で描画した結果を図 13 に示す。動物性タンパク質における結果では、約 40g を超えるとイベント AII に対するリスク比がわずかに上昇した。一方で、イベント CHF についてはリスク比の値が収束せずリスク比の推移を確認することができなかった。植物性タンパク質における結果では、両イベントともに約 25g を超えるとリスク比が 1.0 を下回り、タンパク質量の増加とともにリスク比は単調に減少した。この植物性タンパク質における結果は修正ポアソン回帰に当てはめた結果と大きな違いはなかった。総タンパク質については、修正ポアソン回帰に当てはめた時と比較してリスク比の取りうる範囲が広く、60g よりも摂取量が少ないとリスク比は 1 を上回っているが 60g より多く摂取しているとリスク比は 1.0 を下回った。

表 9 は、タンパク質摂取量によって 2 群に分けた場合の修正ポアソン回帰分析および対数二項回帰によって推定されたリスク比と 95%信頼区間の結果である。タンパク質ごとにモデルを 2 つ採用しており、各イベントごとに結果をまとめた。まずはイベント AII の結果を述べる。動物性タンパク質の低摂取群に対する高摂取群のリスク比は、

表 8: タンパク質摂取割合を使用した回帰分析で推定されたリスク比とその 95%信頼区間.

回帰モデル ⁰	リスク比 (95%信頼区間)			
	修正ポアソン回帰モデル		対数二項回帰モデル	
	Event All	Event CHF	Event All	Event CHF
総タンパク質: 高 (vs. 低)				
モデル 1 ^{1,2}	1.15 (0.60–2.20)	1.13 (0.47–2.72)	0.97 (0.52–1.83)*	NA
モデル 2 ^{3,4}	1.18 (0.60–2.31)	1.19 (0.49–2.92)	1.00 (0.52–1.91)	1.12 (0.48–2.60)*
動物性タンパク質: 高 (vs. 低)				
モデル 1 ^{5,6}	1.78 (0.93–3.40)	1.93 (0.81–4.56)	NA	NA
モデル 2 ^{7,8}	1.71 (0.90–3.24)	1.81 (0.81–4.07)	NA	NA
植物性タンパク質: 低 (vs. 高)				
モデル 1 ^{9,10}	1.56 (0.79–3.09)	1.41 (0.62–3.18)	NA	NA
モデル 2 ^{11,12}	1.82 (0.96–3.45)	1.35 (0.58–3.14)	NA	NA

⁰ 全ての解析において、総エネルギー摂取量による調整を行っている。* が付いた結果は多重代入データセットのうち、リスク比が推定できたデータセットのみで統合した結果を表す。また、NA は全てのデータセットで推定できなかった場合を表す。

¹ Event All に対する調整変数: 年齢, HbA1c, BNP の対数値 (修正ポアソン); eGFR, LVEF, アルブミン (対数二項)

² Event CHF に対する調整変数: BMI, アルブミン, 6MWT, BNP の対数値 (修正ポアソン); HbA1c, アルブミン, BNP の対数値 (対数二項)

³ Event All に対する調整変数: 年齢, eGFR, BNP の対数値 (修正ポアソン); eGFR, LVEF, hs-CRP (対数二項)

⁴ Event CHF に対する調整変数: 性別, BMI, 6MWT, BNP の対数値 (修正ポアソン); eGFR, LVEF (対数二項)

⁵ Event All に対する調整変数: 年齢, BNP の対数値 (修正ポアソン); Hb, eGFR, LVEF (対数二項)

⁶ Event CHF に対する調整変数: BMI, アルブミン, 6MWT, BNP の対数値 (修正ポアソン); eGFR, LVEF (対数二項)

⁷ Event All に対する調整変数: LVEF, BNP の対数値 (修正ポアソン); BMI, eGFR, LVEF (対数二項)

⁸ Event CHF に対する調整変数: BMI, Hb, LVEF, 6MWT, BNP の対数値 (修正ポアソン); eGFR, BNP の対数値 (対数二項)

⁹ Event All に対する調整変数: 年齢, BNP の対数値 (修正ポアソン); 年齢, BMI, HbA1c, アルブミン (対数二項)

¹⁰ Event CHF に対する調整変数: BMI, 6MWT, BNP の対数値 (修正ポアソン); 年齢, HbA1c, LVEF (対数二項)

¹¹ Event All に対する調整変数: BMI, BNP の対数値 (修正ポアソン); 年齢, HbA1c, アルブミン (対数二項)

¹² Event CHF に対する調整変数: BMI, LVEF, BNP の対数値 (修正ポアソン); HbA1c, eGFR, LVEF (対数二項)

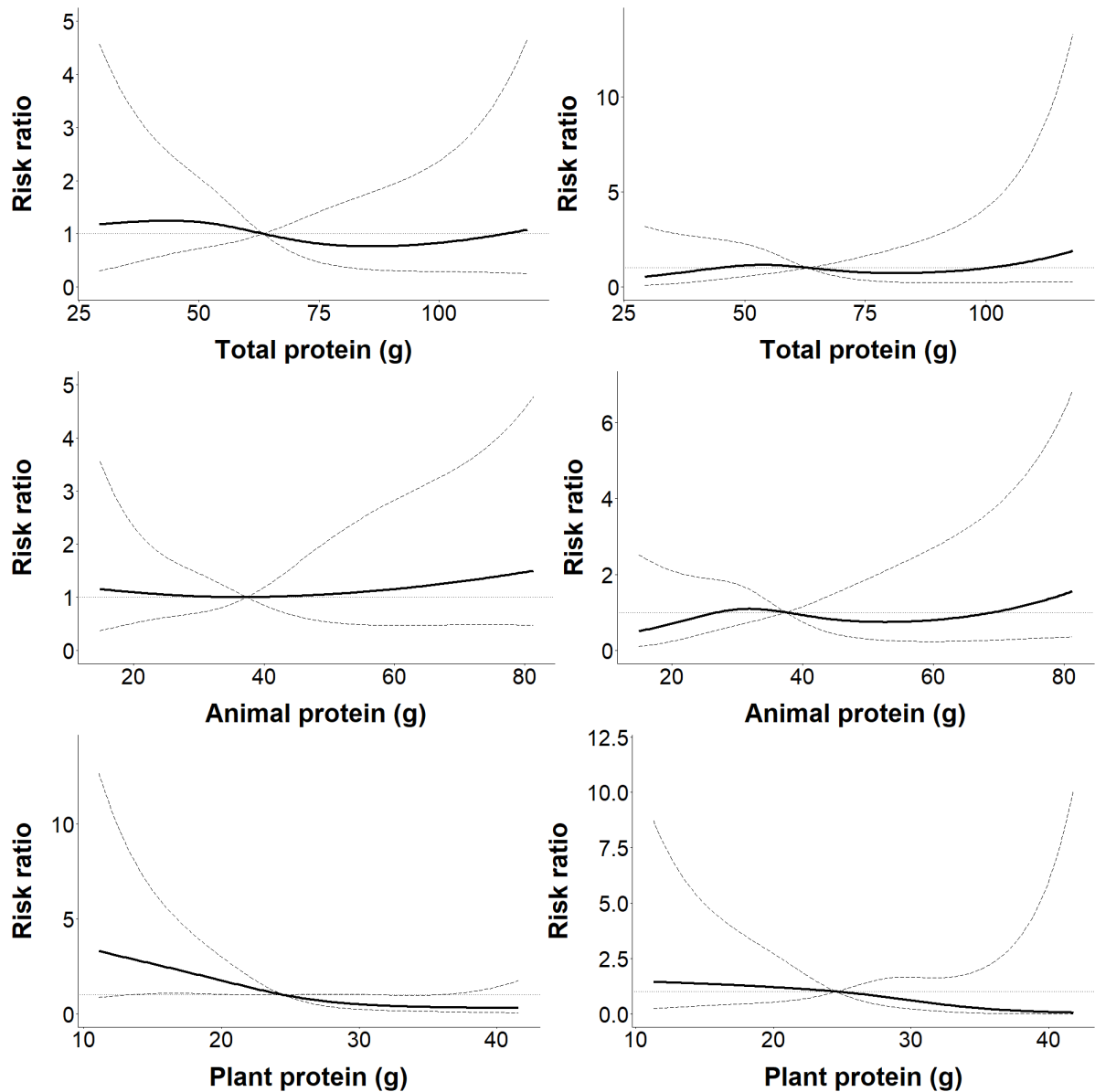


図 12: 各種タンパク質の値に対するリスク比の制限つき三次スプライン曲線. 修正ポアソン回帰モデルに当てはめている. 点線は 95%信頼区間の上限値と下限値を表している. 左列は全死亡または全入院イベントに対する結果であり, 右列は心血管死死亡または心不全増悪による入院イベントに対する結果. モデル選択の結果を受けて, BNP の対数値で調整している.

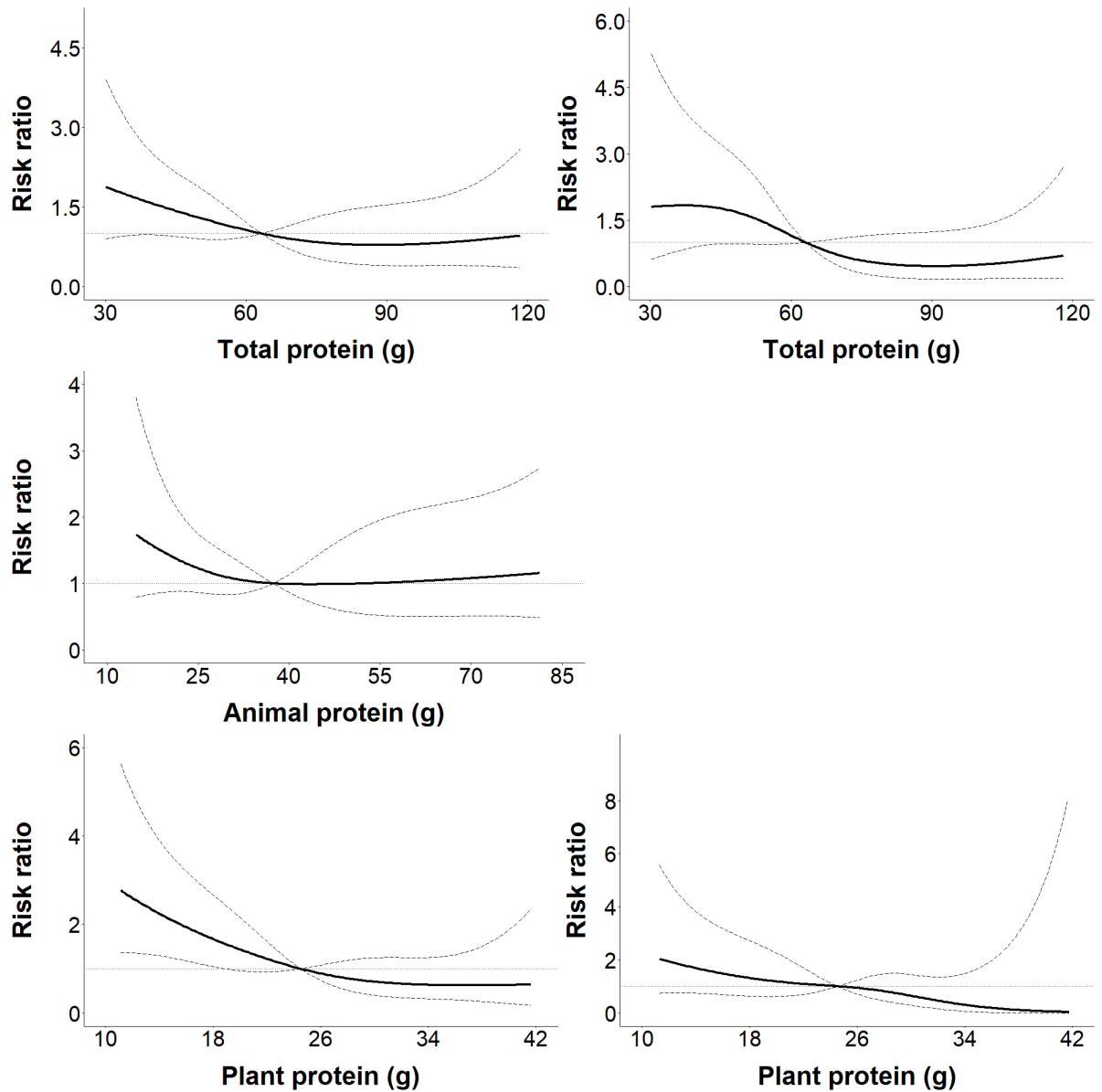


図 13: 各種タンパク質の値に対するリスク比の制限つき三次スプライン曲線. 対数二項回帰モデルに当てはめている. 点線は 95%信頼区間の上限値と下限値を表している. 左列は全死亡または全入院イベントに対する結果であり, 右列は心血管死死亡または心不全増悪による入院イベントに対する結果. イベント CHF における動物性タンパク質の結果は, リスク比の値が収束しなかったため描画できなかった.

1.20[95%CI: 0.51–2.80](モデル 1), 1.05[95%CI: 0.44–2.50](モデル 2) であった。一方、植物性タンパク質の高摂取群に対する低摂取群のリスク比は、1.49[95%CI: 0.51–4.40](モデル 1), 1.82[95%CI: 0.61–5.43](モデル 2) であった。総タンパク質の低摂取群に対する高摂取群のリスク比は、0.98[95%CI: 0.32–2.99](モデル 1), 0.99[95%CI: 0.32–3.05](モデル 2) であった。続いて、イベント CHF の結果を述べる。動物性タンパク質の低摂取群に対する高摂取群のリスク比は、1.56[95%CI: 0.43–5.67](モデル 1), 1.32[95%CI: 0.42–4.18](モデル 2) であった。一方、植物性タンパク質の高摂取割合群に対する低摂取割合群のリスク比は、1.56[95%CI: 0.36–6.76](モデル 1), 2.17[95%CI: 0.47–9.96](モデル 2) であった。総タンパク質の低摂取割合群に対する高摂取割合群のリスク比は、1.34[95%CI: 0.35–5.12](モデル 1), 1.24[95%CI: 0.31–4.91](モデル 2) であった。

表 9 では、対数二項回帰に当てはめリスク比とその 95%信頼区間を推定した結果も表示した。まずはイベント All の結果を述べる。動物性タンパク質の低摂取群に対する高摂取群のリスク比は、0.55[95%CI: 0.21–1.44](モデル 1), 0.51[95%CI: 0.20–1.31](モデル 2) であった。一方、植物性タンパク質の高摂取群に対する低摂取群のリスク比は、1.88[95%CI: 0.75–4.68](モデル 1), 1.70[95%CI: 0.69–4.15](モデル 2) であった。総タンパク質の低摂取群に対する高摂取群のリスク比は、1.51[95%CI: 0.58–3.91](モデル 1), 1.59[95%CI: 0.61–4.13](モデル 2) であった。続いて、イベント CHF の結果を述べる。動物性タンパク質の低摂取群に対する高摂取群のリスク比はモデル 2 では推定不能であり、モデル 1 では 1.21[95%CI: 0.43–3.42](モデル 2) であった。一方、植物性タンパク質の高摂取群に対する低摂取群のリスク比は、4.05[95%CI: 0.44–37.14](モデル 1), 3.92[95%CI: 0.43–35.97](モデル 2) であった。総タンパク質の低摂取群に対する高摂取群のリスク比はモデル 1 では推定不能であり、モデル 2 では 1.84[95%CI: 0.47–7.18](モデル 2) であった。対数二項回帰モデルで推定されたリスク比の統合推定値は、全て推定可能だった多重代入データセットのみで統合した結果である。

6.6 考察

慢性心不全患者が食生活で摂取する栄養と死亡または入院イベントとの関連について、タンパク質に着目して解析を実施した。これまで総タンパク質の摂取量と死亡・入院イベントとの関連性については議論されていたが、タンパク質の組成に注目して動物性タンパク質と植物性タンパク質それぞれについてイベントとの関連性を調査したという点が本研究の重要な貢献である。また、タンパク質の摂取量は単純なタンパク質摂取量と総エネルギーに対するタンパク質の割合を考慮し、それぞれの観点から高摂取群と低摂取群を分割する閾値を探索した。さらに、イベント有無という二値アウトカムだからと性急にロジスティック回帰によるオッズ比での要約を解析方法として選択する

表 9: タンパク質摂取量を使用した回帰分析で推定されたリスク比とその 95%信頼区間.

回帰モデル ⁰	リスク比 (95%信頼区間)			
	修正ポアソン回帰モデル		対数二項回帰モデル	
	Event All	Event CHF	Event All	Event CHF
総タンパク質: 高 (vs. 低)				
モデル 1 ^{1,2}	0.98 (0.32–2.99)	1.34 (0.35–5.12)	1.51 (0.58–3.91)*	NA
モデル 2 ^{3,4}	0.99 (0.32–3.05)	1.24 (0.31–4.91)	1.59 (0.61–4.13)*	1.84 (0.47–7.18)*
動物性タンパク質: 高 (vs. 低)				
モデル 1 ^{5,6}	1.20 (0.51–2.80)	1.56 (0.43–5.67)	0.55 (0.21–1.44)*	1.21 (0.43–3.42)*
モデル 2 ^{7,8}	1.05 (0.44–2.50)	1.32 (0.42–4.18)	0.51 (0.20–1.31)*	NA
植物性タンパク質: 低 (vs. 高)				
モデル 1 ^{9,10}	1.49 (0.51–4.40)	1.56 (0.36–6.76)	1.88 (0.75–4.68)*	4.05 (0.44–37.14)*
モデル 2 ^{11,12}	1.82 (0.61–5.43)	2.17 (0.47–9.96)	1.70 (0.69–4.15)*	3.92 (0.43–35.97)*

⁰ 全ての解析において、総エネルギー摂取量による調整を行っている。* が付いた結果は多重代入データセットのうち、リスク比が推定できたデータセットのみで統合した結果を表す。また、NA は全てのデータセットで推定できなかった場合を表す。

- ¹ Event All に対する調整変数: 年齢, HbA1c, BNP の対数値 (修正ポアソン); eGFR, LVEF, アルブミン (対数二項)
- ² Event CHF に対する調整変数: BMI, アルブミン, 6MWT, BNP の対数値 (修正ポアソン); Age, eGFR, LVEF(対数二項)
- ³ Event All に対する調整変数: BMI, BNP の対数値 (修正ポアソン); eGFR, LVEF, hs-CRP(対数二項)
- ⁴ Event CHF に対する調整変数: 性別, BMI, 6MWT, BNP の対数値 (修正ポアソン); 性別, eGFR, LVEF(対数二項)
- ⁵ Event All に対する調整変数: 年齢, BNP の対数値 (修正ポアソン); eGFR, LVEF, アルブミン (対数二項)
- ⁶ Event CHF に対する調整変数: BMI, アルブミン, 6MWT, BNP の対数値 (修正ポアソン); Hb, eGFR, LVEF, 6MWT, BNP の対数値 (対数二項)
- ⁷ Event All に対する調整変数: hs-CRP, BNP の対数値 (修正ポアソン); eGFR, LVEF, hs-CRP(対数二項)
- ⁸ Event CHF に対する調整変数: BMI, Hb, LVEF, 6MWT, BNP の対数値 (修正ポアソン); eGFR, 6MWT, hs-CRP, BNP の対数値 (対数二項)
- ⁹ Event All に対する調整変数: 年齢, BNP の対数値 (修正ポアソン), LVEF(対数二項)
- ¹⁰ Event CHF に対する調整変数: BMI, 6MWT, BNP の対数値 (修正ポアソン); eGFR, アルブミン (対数二項)
- ¹¹ Event All に対する調整変数: BMI, BNP の対数値 (修正ポアソン); hs-CRP(対数二項)
- ¹² Event CHF に対する調整変数: BMI, LVEF, BNP の対数値 (修正ポアソン); eGFR, 6MWT(対数二項)

のではなく、イベントの発生割合と統計モデルの性能比較の結果から修正ポアソン回帰によるリスク比での要約を選んだことで、オッズ比よりもさらに直感的に解釈ができるようにした。動物性タンパク質と植物性タンパク質で摂取量に対するリスク比の軌跡が、異なったことも特筆すべき点である。

回帰モデルの選択のために、対数二項回帰と修正ポアソン回帰の性能比較を行い、収束割合・バイアス・RMSE の観点から修正ポアソン回帰での解析を選択した。対数二項回帰と修正ポアソン回帰の性能を比較した既報はあるが、サンプルサイズが 500 から 1000 の範囲での検証に留まっており、慢性心不全コホートのように 150 程度のサンプルサイズでの検証は実施されていない (Chen et al., 2018)。加えて、本実験では栄養データの複雑性を反映させるため、変数間の相関を変数として加えたが、このような検討もされていなかった。本実験では小標本という状況の中で、相関の大きさにも影響を受けず頑健な性質を修正ポアソン回帰が有することを示すことができた。二値の結果変数に対する回帰モデルにおいて、ロジスティック回帰を用いてオッズ比で要約している研究が未だに多く見られるが、修正ポアソン回帰は推定値の収束の安定性・推定精度の良さ・リスク比という解釈の容易性のどの観点から見ても今後主流な回帰モデルの 1 つとして採用されるべきである。

各種タンパク質の摂取量と総エネルギー摂取量、炭水化物摂取量、脂質摂取量との比較では、タンパク質の摂取絶対量が多いほどどの摂取量も増加傾向にあった。特に炭水化物摂取量と植物性タンパク質、脂質摂取量と動物性タンパク質については正の強い相関がある。タンパク質を比較的多く摂取している患者さんは、他の三大栄養素摂取量も多く摂取しており、総エネルギー摂取量の増大にもつながっていると考えられる。一方、タンパク質の摂取割合での観点では、総タンパク質摂取割合、動物性タンパク質摂取割合が多い患者さんは、炭水化物摂取量が比較的少ない傾向にあった。また、総タンパク質摂取割合、動物性タンパク質摂取割合の増大が脂質摂取割合にもつながっている傾向にあった。一方、植物性タンパク質摂取割合の増加は炭水化物摂取割合を増加と弱い相関があり、脂質とはほとんど相関がない。このことから、タンパク質摂取割合の増減によって影響を受けるのは炭水化物であり、その影響の方向は動物性タンパク質と植物性タンパク質で反対になることが言える。

候補変数が 13 種類あるため、全通りのモデルを考える場合は、 $2^{13} = 8192$ 通りのモデルについて AIC と AUC を導出することになる。しかし、計算量が莫大である上に、最終的に選択するモデルは 2 つ程度に決めていたため、ほとんどのモデルの AIC と AUC の結果が破棄されるという懸念があった。そこで本実験では当初、試験的に最大 10 個まで変数を入れないモデルを考え、7 個以上の変数を投入したモデルの AIC は 6 個の変数を投入したモデルの AIC の最小値よりも 2 以上大きくなり、結果的に選ばれなくなるという経験的な性質を発見した。そこで、計算負荷の観点から本解析では最終的には最大

でも 6 個の変数を投入するモデルしか対象にしていない。また、6 個までの変数に絞ってモデル ($\sum_{x=1}^6 13C_x = 4095$ 通り) それぞれについて AIC と AUC を計算したとしても、R での実行時間は約 36 時間かかった。本解析では、計算時間の短縮のために並列処理を行えるよう解析プログラムを構成し、実行時間を 4 時間まで短縮させるということも成功させた。また、本研究におけるモデルの最終決定方法として、二段階の措置をとった。これは AIC の値で最小値からプラス 2 までのモデルを第一段階として抽出し、二段階目としてそのモデル群の中で AUC が高いモデルを採用している。AIC の値が小さいモデルではデータへの当てはまりが良いのは確かであるが、AIC 最小のモデルが必ずしもリスク比の推定値の標準誤差が小さいとは限らず、推定結果とは切り離して考えるべきである。AIC の式がパラメータ数と変数の差の 2 倍であるという定義から、本解析では差が 2 までは許容することとした (Anderson, 2008)。

各イベントに対するリスク比の 95%信頼区間にはどれも 1 が含まれ、統計的有意性という観点からは帰無仮説 (リスク比=1) を棄却できない。しかし、Amrhein et al. (2019) でも指摘されているように、統計的に有意な差が無かったことを関連性がないと結論づけるのは誤りで、今回の解析でのリスク比と関連性がない (リスク比=1) という状況が両立することを認めた上で、点推定値が今回のデータでは最も当てはまりが良いリスク比であると解釈すべきである。そのため、点推定値を主軸に考察を述べることにする。総タンパク質摂取割合と動物性タンパク質摂取割合は高摂取群の方がリスクが大きく、特に動物性タンパク質では 1.7 倍から 2.0 倍になった。一方、植物性タンパク質摂取割合は低摂取群の方が 1.3 倍から 1.9 倍ほど大きい。総タンパク質摂取割合はイベント All とイベント CHF であまり大きさに違いがないが、動物性タンパク質ではイベント CHF のリスク比の値がイベント All の値より大きく、動物性タンパク質摂取割合が 10.3%を超えるとイベント発生リスクが約 2.0 倍に上がることがわかる。反対に植物性タンパク質摂取割合が 6.0%を切ると、イベント CHF の発生率が約 1.5 倍になるため、この閾値を目安に食事バランスを保つことが望ましい。また、摂取量で見ると、動物性タンパク質摂取量の増加よりも植物性タンパク質摂取量の減少がイベント発生リスクを引き上げている。総タンパク質摂取量に関しては、イベント All のリスクは高摂取群と低摂取群の間であまり変わらず、イベント CHF のリスクは高摂取群の方が上昇している。総合すると、タンパク質摂取割合の観点から総エネルギー量に対して動物性タンパク質の摂取割合が多くならないよう注意し、摂取量の観点から植物性タンパク質の摂取量を多くすることを指導すると、イベントの発生予防になると期待できる。

シミュレーション実験だけでなく、対象コホートを使った実データに対しても修正ポアソン回帰による解析結果と対数二項回帰による解析結果を比較した。制限つき 3 次スプラインの結果を見ると、リスク比の軌跡の変化について両モデルに大きな違いは見られないが、イベント CHF に対する動物性タンパク質摂取量の結果のように、リスク比の

値が収束しないことによって描画できないような例が存在した。そのため、制限つき3次スプラインによる対数二項回帰を用いた閾値探索では、必ずしも閾値が決定できるものではない。また統合リスク比を求めた結果では、修正ポアソン回帰の場合ではどの多重代入データセットでもリスク比の値が収束しており安定して統合推定値を算出できたが、対数二項回帰の場合では一部の多重代入データセット(全体の約5%から約40%)でのみリスク比が計算できた事例が多く、そのためリスク比の統合推定値も修正ポアソン回帰における結果から得た考察とは真逆の考察になりうる危険性がある。また、制限つき3次スプライン描画前のモデル選択の段階において、AICとAUCが算出できなかったモデルの候補もあった。対数二項回帰に基づくモデル選択では、AICやAUCによる単純なモデル選択規準だけでなく、多重代入データセットのうち何%のデータセットにおいてリスク比の点推定値が求められたかといった別の規準も加えた方が推定安定性の観点から好ましいと考えられる。この規準をいくらに設定すれば良いかと言う議論は未だなされていない為に今後の新規研究として検討しても良いが、対数二項回帰での検討に拘らず修正ポアソン回帰を使えばこのような懸念は払拭されるであろう。

本研究で採用した統計解析手法でも十分新しい知見を得ることができたが、解析手法について2点議論する。まず一点目に、モデル選択時の交差検証法をフォールド数5でやった点を取り上げる。フォールド数の設定にエラー率と計算負荷の観点から論じた既報は存在するが、統一的なコンセンサスは得られておらず、フォールド数は10、5、あるいはサンプルサイズと同等(LOOCV)が経験則として用いられている(Marcot and Hanea, 2021; Oyedele, 2023)。本研究の初期段階では、データ全体に対する訓練データの割合を増やして回帰モデルの精度の向上を図る目的で、フォールド数を10と設定していた。しかし、その場合にはテストデータに全例イベント発生なしとなるセットができ、AUCを計算することができなかったため、フォールド数を5に変更した。フォールド数5で行った変数選択結果とフォールド数10で行った結果が変わるか検討することも重要である。このような状況では、データ全体のイベント発生割合と各フォールドでのテストデータ内のイベント発生割合を一致させる層化(10フォールド)交差検証法による解析が好ましい可能性が高い。各フォールドのテストデータにイベント有無が混在していないとAUCは計算できないので、層化交差検証法によるモデル選択も検討すべきであろう。二点目に、AICとAUCの要約の仕方である。本解析では、各フォールド・多重代入データを区別せず、計算されたAICとAUCの全平均をモデル選択のスコアとした。現状の方法では1つのデータセット内で計算されている各フォールドでのAIC(AUC)とデータセット間のAIC(AUC)を等価に扱っているため、この部分に重み等の考慮が必要ではないかと指摘される可能性がある。しかし、本解析では各データセットで各フォールドに入る症例の集合をランダムにしているため、系統的なバイアスは生じないと考えている。感度解析として、1つのデータセット内での平均AIC(AUC)の分

布 (データセット数だけ値がある) を描いて, データセット間で平均 AIC(AUC) の分散がどの程度あるのかを可視化することも重要である.

本研究の限界点として, 追跡期間の長さ, 対象集団の人種, 心不全の種類の3点を挙げる. 本研究では心不全による退院後1年以内のイベントを対象としているが, NYHA心不全分類におけるI度とII度の症例が多くを占めていたため, さらに長く追跡を行うことでイベントを観察できた可能性がある. サンプルサイズの観点からも146と大規模とは言えない大きさのため, 本研究はより大きなサンプルサイズで行うコホート研究のための足がかりとなるパイロット研究として位置づけられる. また, 日本人のみを対象としていたが, より一般化可能性を求めるのであれば, 食習慣が異なると予想される様々な人種の組み入れを行った研究が望ましい. さらに, 本研究ではHFrEFとHFpEFの両方の患者が含まれていなかったが, これを区別せずに集計・解析を行った. 心臓の収縮機能の低下(HFrEF)と心臓の拡張機能の異常(HFpEF)という意味で病態も治療方法も異なるため, より大規模な研究であれば区別して議論をすることが好ましい.

最後に, 本研究で新たに得られた2つの知見と将来の展望についてまとめる. 一点目に動物性タンパク質の総エネルギー量に占める割合に基づいて, 高摂取割合群と低摂取割合群に分割したとき, 高摂取割合群の方がイベント発生リスクが大きいこと. そして, 植物性タンパク質の場合は低摂取割合群の方がイベント発生リスクが高いことである. 二点目に, 動物性タンパク質摂取量と炭水化物摂取量は正の関係があるが, 摂取割合で見ると負の関係があったことである. 現行の解析では動物性タンパク質と植物性タンパク質は別々にモデルの構成をしていたが, 今後の展望としてタンパク質の組み合わせ(例えば, 植物性タンパク質高摂取群かつ動物性タンパク質高摂取群に入る患者を参照集団とした群分け)でイベント発生リスクを議論することで, 両タンパク質の複合的な影響がイベント発生に与えるメカニズムの解明に繋がると期待される.

7 全体の結論

本研究では、「ベイズ流潜在クラスモデルを利用したゴールドスタンダードがない状況に対する2つの検査精度の推定」という研究と「修正ポアソン回帰モデルを利用したタンパク質の摂取量と摂取割合という2つの側面から考察する心不全に関するイベントのリスクの推定」という研究を扱った。本研究から得られた新知見は以下の通りである。

第5章

- ゴールドスタンダードな検査方法がない状況で2種類の検査の感度と特異度を推定するにあたり、陽性の頻度が稀な疾患を想定してモデルの性能を検証した。
- ベイズ流潜在クラスモデルは、従来の頻度論による手法と比較して推定値の収束割合が高く、推定結果が安定していた。
- ベイズ流潜在クラスモデルでの感度推定で、もう一方の検査の感度にのみ情報を入れるという戦略は、有効サンプルサイズをどれだけ大きくしてもバイアスが改善されないため避けるべきである。

第6章

- 広範なシミュレーションにおいて、バイアス・MSE・収束割合の点で修正ポアソン回帰は対数二項回帰よりも優れていた。
- 慢性心不全多施設共同栄養疫学コホートにおいて、動物性タンパク質の摂取量と炭水化物摂取量は正の相関があったが、総エネルギー摂取量における動物性タンパク質が占める割合と炭水化物が占める割合には負の相関があった。
- 心血管死と心不全による入院イベントのリスクは、総エネルギーにおける動物性タンパク質の割合が多いと高くなり、植物性タンパク質の割合が大きいと低くなる可能性がある。

本研究では複雑なデータ構造を持つ医学データの中で、感染症のデータから2種類の検査の診断精度の推定と栄養疫学のデータから心不全イベント発生のリスクとタンパク質摂取量の関連性の検討を行なった。前者の研究はベイズ流潜在クラスモデルによる感度推定の収束安定性と有効サンプルサイズ増加に伴うバイアスの推移について知見

を提供した. 今後, 別の分野においてゴールドスタンダードがないような状況での検査結果しか計測できない状況であってもベイズ流潜在クラスモデルの使用を推奨する根拠として, 本研究の結果は一助となる. 後者の研究ではこれまで総タンパク質と心不全イベントとの関連性しかほとんど検討されていなかった中で, 動物性タンパク質と植物性タンパク質それぞれにおいての量と割合の観点からリスクに与える影響を考察した. 将来のより大規模な研究において各タンパク質が心不全イベントに与える影響を考察するための論拠となること, そして心不全の患者さんに対しての栄養指導で, タンパク質の中でも特に植物性タンパク質の割合を増やすといった明瞭な指導を提供する根拠として, 本研究の結果は医学研究の発展に寄与すると考えられる.

8 謝辞

本研究の指導ならびに議論をして下さった指導教員の横田勲准教授には、修士課程から6年間を通して、この分野について全く知識のなかった私に生物統計学の研究の在り方をご指導いただきました。心より感謝を申し上げます。また、研究室の先輩である岡田和史先生、楊一馳先生には、主に技術面で支えていただきました。御二方の研究に対する真摯な姿勢を間近で学ばせていただいたことは、大変貴重な経験でした。

本研究に対して、データ提供においてご尽力くださるだけでなく臨床の観点から助言・指導いただきました北海道大学病院医療・ヘルスサイエンス研究開発機構の横田卓特任講師、そして統計解析の観点から指導ならびに議論をいただきました同機構の伊藤陽一教授には、リアルワールドデータを利用した研究に対する姿勢についてご指南いただきました。深く御礼申し上げます。

博士課程に在籍中の身でありながら、博士課程1年次から3年次までは北海道大学病院、4年次には北海道大学医学研究院と遺伝子病制御研究所に勤務しておりました。研究と並行して業務が行えるよう調整して下さった、北海道大学病院データサイエンスセンター・医療情報企画部の皆様、遺伝子病制御研究所・大学院医学院分子神経免疫学分野の村上正晃教授およびメンバーの皆様には、最大限のご配慮をいただきました。

日頃のゼミで熱く議論を交わして下さり、時には私の息抜きにも付き合ってくださいました北海道大学大学院医学院社会医学分野医学統計学教室員のみなさまに深く感謝申し上げます。最後に、遠方から支えてくださった家族に感謝を表します。

なお、本学位論文の執筆過程において、Google Gemini(モデル名: Gemini Pro, Google)を研究のアウトライン構成案作成および論理展開の検証に関する執筆補助ツールとして利用したことをここに開示する。しかし、本論文の内容、学術的な妥当性及びすべての記述に対する最終的な責任は、著者自身が負うものとする。

9 利益相反

開示すべき利益相反状態はない。

10 参考文献

- Agresti, A (2013). *Categorical Data Analysis*. Wiley, Hoboken, 3rd edition.
- Akaike, H (2003). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.
- Akner, G and Cederholm, T (2001). Treatment of protein-energy malnutrition in chronic nonmalignant disorders. *The American Journal of Clinical Nutrition*, 74(1):6–24.
- Albert, P. S (2007). Random effects modeling approaches for estimating roc curves from repeated ordinal tests without a gold standard. *Biometrics*, 63(2):593–602.
- Amrhein, V, Greenland, S, and McShane, B (2019). Scientists rise up against statistical significance. *Nature*, 567(7748):305–307.
- Amundson, D. E, Djurkovic, S, and Matwiyoff, G. N (2010). The obesity paradox. *Critical Care Clinics*, 26(4):583–596.
- Anderson, D. R (2008). *Model based inference in the life sciences: a primer on evidence*. Springer.
- Apekey, T. A, Maynard, M. J, Kittana, M, and Kunutsor, S. K (2022). Comparison of the effectiveness of low carbohydrate versus low fat diets, in type 2 diabetes: systematic review and meta-analysis of randomized controlled trials. *Nutrients*, 14(20):4391.
- Austin, P. C, White, I. R, Lee, D. S, and van Buuren, S (2021). Missing data in clinical research: a tutorial on multiple imputation. *Canadian Journal of Cardiology*, 37(9):1322–1331.
- Bauer, L. K, Caro, M. A, Beach, S. R, Mastromauro, C. A, Lenihan, E, Januzzi, J. L, and Huffman, J. C (2012). Effects of depression and anxiety improvement on adherence to medication and health behaviors in recently hospitalized cardiac patients. *American Journal of Cardiology*, 109(9):1266–1271.
- Branscum, A. J, Johnson, W. O, Hanson, T. E, and Gardner, I. A (2008). Bayesian semiparametric roc curve estimation and disease diagnosis. *Statistics in Medicine*, 27(13):2474–2496.

- Budhathoki, S, Sawada, N, Iwasaki, M, Yamaji, T, Goto, A, Kotemori, A, Ishihara, J, Takachi, R, Charvat, H, Mizoue, T, et al. (2019). Association of animal and plant protein intake with all-cause and cause-specific mortality in a japanese cohort. *JAMA Internal Medicine*, 179(11):1509–1518.
- Chen, W, Qian, L, Shi, J, and Franklin, M (2018). Comparing performance between log-binomial and robust poisson regression models for estimating risk ratios under model misspecification. *BMC Medical Research Methodology*, 18(1):63.
- Collins, J and Huynh, M (2014). Estimation of diagnostic test accuracy without full verification: a review of latent class methods. *Statistics in Medicine*, 33(24):4141–4169.
- Cornfield, J (1951). A method of estimating comparative rates from clinical data. applications to cancer of the lung, breast, and cervix. *Journal of the National Cancer Institute*, 11(6):1269–1275.
- Dendukuri, N, Hadgu, A, and Wang, L (2009). Modeling conditional dependence between diagnostic tests: A multiple latent variable model. *Statistics in Medicine*, 28(3):441–461.
- Dendukuri, N and Joseph, L (2001). Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*, 57(1):158–167.
- DeTure, M. A and Dickson, D. W (2019). The neuropathological diagnosis of alzheimer’ s disease. *Molecular Neurodegener*, 14(1):32.
- Fosgate, G. T, Scott, H. M, and Jordan, E. R (2007). Development of a method for bayesian nonparametric roc analysis with application to an elisa for johne’s disease in dairy cattle. *Preventive Veterinary Medicine*, 81(1-3):178–193.
- Franz, M. J, Bantle, J. P, Beebe, C. A, Brunzell, J. D, Chiasson, J.-L, Garg, A, Holzmeister, L. A, Hoogwerf, B, Mayer-Davis, E, Mooradian, A. D, et al. (2002). Evidence-based nutrition principles and recommendations for the treatment and prevention of diabetes and related complications. *Diabetes Care*, 25(1):148–198.
- Gauthier, J, Wu, Q, and Gooley, T (2020). Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians. *Bone Marrow Transplantation*, 55(4):675–680.
- Gelman, A and Rubin, D. B (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

- Goodman, L. A (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61(2):215–231.
- Greenland, S (1987). Interpretation and choice of effect measures in epidemiologic analyses. *American Journal of Epidemiology*, 125(5):761–768.
- Hamaguchi, S, Tsuchihashi-Makaya, M, Kinugawa, S, Goto, D, Yokota, T, Goto, K, Yamada, S, Yokoshiki, H, Takeshita, A, Tsutsui, H, et al. (2010). Body mass index is an independent predictor of long-term outcomes in patients hospitalized with heart failure in japan—a report from the japanese cardiac registry of heart failure in cardiology (jcare-card)–. *Circulation Journal*, 74(12):2605–2611.
- Hastings, W. K (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109.
- Hosmer Jr, D. W, Lemeshow, S, and Sturdivant, R. X (2013). *Applied logistic regression*. John Wiley & Sons.
- Hui, S. L and Walter, S. D (1980). Estimating the error rates of diagnostic tests. *Biometrics*, pages 167–171.
- Jenicek, M (1997). Epidemiology, evidenced-based medicine, and evidence-based public health. *J Epidemiol*, 7(4):187–197.
- Jones, G, Johnson, W. O, Hanson, T. E, and Christensen, R (2010). Identifiability of models for multiple diagnostic testing in the absence of a gold standard. *Biometrics*, 66(3):855–863.
- Joseph, L, Gyorkos, T. W, and Coupal, L (1995). Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *American Journal of Epidemiology*, 141(3):263–272.
- Kakutani, N, Yokota, T, Fukushima, A, Obata, Y, Ono, T, Sota, T, Kinugasa, Y, Takahashi, M, Matsuo, H, Matsukawa, R, et al. (2022). Impact of citrus fruit intake on the mental health of patients with chronic heart failure. *Journal of Cardiology*, 79(6):719–726.
- Kim, C, Lin, X, and Nelson, K. P (2021). Measuring rater bias in diagnostic tests with ordinal ratings. *Statistics in Medicine*, 40(17):4014–4033.

- Kinugasa, Y, Sota, T, Kamitani, H, Nakayama, N, Nakamura, K, Hirai, M, Yanagihara, K, Kato, M, Ono, T, Takahashi, M, et al. (2022). Diagnostic performance of nutritional indicators in patients with heart failure. *ESC Heart Failure*, 9(4):2096–2106.
- Kobayashi, S, Murakami, K, Sasaki, S, Okubo, H, Hirota, N, Notsu, A, Fukui, M, and Date, C (2011). Comparison of relative validity of food group intakes estimated by comprehensive and brief-type self-administered diet history questionnaires against 16 d dietary records in japanese adults. *Public Health Nutrition*, 14(7):1200–1211.
- Kostoulas, P, Nielsen, S. S, Branscum, A. J, Johnson, W. O, Dendukuri, N, Dhand, N. K, Toft, N, and Gardner, I. A (2017). Stard-blem: standards for the reporting of diagnostic accuracy studies that use bayesian latent class models. *Preventive Veterinary Medicine*, 138:37–47.
- Ladouceur, M, Rahme, E, Bélisle, P, Scott, A. N, Schwartzman, K, and Joseph, L (2011). Modeling continuous diagnostic test data using approximate dirichlet process distributions. *Statistics in Medicine*, 30(21):2648–2662.
- Liang, K.-Y and Zeger, S. L (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22.
- Little, R. J and Rubin, D. B (2019). Statistical analysis with missing data. John Wiley & Sons.
- Lu, D, Zhou, C, Tang, L, Tan, M, Yuan, A, and Chan, L (2018). Evaluating accuracy of diagnostic tests without conditional independence assumption. *Statistics in Medicine*, 37(19):2809–2821.
- Lv, S and Ru, S (2021). The prevalence of malnutrition and its effects on the all-cause mortality among patients with heart failure: A systematic review and meta-analysis. *PLoS One*, 16(10):e0259300.
- Marcot, B. G and Hanea, A. M (2021). What is an optimal value of k in k-fold cross-validation in discrete bayesian network analysis? *Computational Statistics*, 36(3):2009–2031.
- Mariotti, F (2019). Animal and plant protein sources and cardiometabolic health. *Advances in Nutrition*, 10:S351–S366.

- Matsui, H, Yokoyama, T, Sekiguchi, K, Iijima, D, Sunaga, H, Maniwa, M, Ueno, M, Iso, T, Arai, M, and Kurabayashi, M (2012). Stearoyl-coa desaturase-1 (scd1) augments saturated fatty acid-induced lipid accumulation and inhibits apoptosis in cardiac myocytes. *PloS one*, 7(3):e33283.
- McNutt, L.-A, Wu, C, Xue, X, and Hafner, J. P (2003). Estimating the relative risk in cohort studies and clinical trials of common outcomes. *American Journal of Epidemiology*, 157(10):940–943.
- Metropolis, N, Rosenbluth, A. W, Rosenbluth, M. N, Teller, A. H, and Teller, E (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.
- Mozaffarian, D and Wu, J. H (2011). Omega-3 fatty acids and cardiovascular disease: effects on risk factors, molecular pathways, and clinical events. *Journal of the American College of Cardiology*, 58(20):2047–2067.
- Nelson, C. J, Cho, C, Berk, A. R, Holland, J, and Roth, A. J (2010). Are gold standard depression measures appropriate for use in geriatric cancer patients? a systematic evaluation of self-report depression instruments used with geriatric, cancer, and geriatric cancer samples. *Journal of Clinical Oncology*, 28(2):348–356.
- Obata, Y, Kakutani, N, Kinugawa, S, Fukushima, A, Yokota, T, Takada, S, Ono, T, Sota, T, Kinugasa, Y, Takahashi, M, et al. (2021). Impact of inadequate calorie intake on mortality and hospitalization in stable patients with chronic heart failure. *Nutrients*, 13(3):874.
- Oyedele, O (2023). Determining the optimal number of folds to use in a k-fold cross-validation: A neural network classification experiment. *Research in Mathematics*, 10(1):2201015.
- Qu, Y, Tan, M, and Kutner, M. H (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, 52(3):797–810.
- Rindskopf, D and Rindskopf, W (1986). The value of latent class analysis in medical diagnosis. *Statistics in Medicine*, 5(1):21–27.
- Gilks, W. R, Richardson, S, and Spiegelhalter, D (1995). *Markov chain Monte Carlo in practice*. CRC press.

- Roberts, G. O and Sahu, S. K (1997). Updating schemes, correlation structure, blocking and parameterization for the gibbs sampler. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 59(2):291–317.
- Rubin, D. B (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, New York.
- Rutjes, A, Reitsma, J, Coomarasamy, A, Khan, K, and Bossuyt, P (2007). Evaluation of diagnostic tests when there is no gold standard. a review of methods. *Health Technology Assessment*, 11(50):ix–51.
- Sackett, D. L (1997). Evidence-based medicine. In *Seminars in Perinatology*, volume 21, pages 3–5. Elsevier.
- Saijo, T, Yasumoto, K, Ryomoto, K, Momoki, C, and Habu, D (2023). Association of protein intake during hospitalization with readmission in older adult patients with heart failure at risk of malnutrition. *Nutrition in Clinical Practice*, 38(3):686–697.
- Schofield, M. R, Maze, M. J, Crump, J. A, Rubach, M. P, Galloway, R, and Sharples, K. J (2021). On the robustness of latent class models for diagnostic testing with no gold standard. *Statistics in Medicine*, 40(22):4751–4763.
- Schulz, K. F, Altman, D. G, Moher, D, Group, C, et al. (2010). Consort 2010 statement: updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, 63(8):834–840.
- Soukoulis, V, Dihu, J. B, Sole, M, Anker, S. D, Cleland, J, Fonarow, G. C, Metra, M, Pasini, E, Strzelczyk, T, Taegtmeyer, H, et al. (2009). Micronutrient deficiencies: an unmet need in heart failure. *Journal of the American College of Cardiology*, 54(18):1660–1673.
- Stone, M (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 36(2):111–147.
- Streng, K. W, Hillege, H. L, Ter Maaten, J. M, van Veldhuisen, D. J, Dickstein, K, Ng, L. L, Samani, N. J, Metra, M, Ponikowski, P, Cleland, J. G, et al. (2022). Clinical implications of low estimated protein intake in patients with heart failure. *Journal of Cachexia, Sarcopenia and Muscle*, 13(3):1762–1770.
- Sun, A and Zhou, X.-H (2025). Estimation of diagnostic test accuracy without gold standards. *Statistics in Medicine*, 44(3-4):e10315.

- Tsiatis, A. A (2006). *Semiparametric theory and missing data*. Springer.
- Tsuchihashi-Makaya, M, Kato, N, Chishaki, A, Takeshita, A, and Tsutsui, H (2009). Anxiety and poor social support are independently associated with adverse outcomes in patients with mild heart failure. *Circulation Journal*, 73(2):280–287.
- Von Hippel, P. T (2007). Regression with missing ys: an improved strategy for analyzing multiply imputed data. *Sociological Methodology*, 37(1):83–117.
- Wallenstein, S and Bodian, C (1987). Epidemiologic programs for computers and calculators. inferences on odds ratios, relative risks, and risk differences based on standard regression programs. *American Journal of Epidemiology*, 126(2):346–355.
- Walsh, T (2018). Fuzzy gold standards: approaches to handling an imperfect reference standard. *Journal of Dentistry*, 74:S47–S49.
- Wang, C, Lin, X, and Nelson, K. P (2020). Bayesian hierarchical latent class models for estimating diagnostic accuracy. *Statistical Methods in Medical Research*, 29(4):1112–1128.
- Wang, C, Turnbull, B, Gröhn, Y, and Nielsen, S (2006). Estimating receiver operating characteristic curves with covariates when there is no perfect reference test for diagnosis of johne’s disease. *Journal of Dairy Science*, 89(8):3038–3046.
- Wedderburn, R. W (1974). Quasi-likelihood functions, generalized linear models, and the gauss—newton method. *Biometrika*, 61(3):439–447.
- Willett, W. C (2012). Dietary fats and coronary heart disease. *Journal of internal medicine*, 272(1):13–24.
- Wu, G (2016). Dietary protein intake and human health. *Food and Function*, 7(3):1251–1265.
- Wu, J. H, Lemaitre, R. N, King, I. B, Song, X, Psaty, B. M, Siscovick, D. S, and Mozaffarian, D (2014). Circulating omega-6 polyunsaturated fatty acids and total and cause-specific mortality: the cardiovascular health study. *Circulation*, 130(15):1245–1253.
- Yokota, I, Sakurazawa, T, Sugita, J, Iwasaki, S, Yasuda, K, Yamashita, N, Fujisawa, S, Nishida, M, Konno, S, and Teshima, T (2021a). Performance of qualitative and quantitative antigen tests for sars-cov-2 using saliva. *Infectious Disease Reports*, 13(3):742–747.

- Yokota, I, Shane, P. Y, Okada, K, Unoki, Y, Yang, Y, Inao, T, Sakamaki, K, Iwasaki, S, Hayasaka, K, Sugita, J, Nishida, M, Fujisawa, S, and Teshima, T (2020). Mass screening of asymptomatic persons for sars-cov-2 using saliva. *Clinical Infectious Diseases*, 73(3):e559–e565.
- Yokota, I, Shane, P. Y, Okada, K, Unoki, Y, Yang, Y, Iwasaki, S, Fujisawa, S, Nishida, M, and Teshima, T (2021b). A novel strategy for sars-cov-2 mass screening with quantitative antigen testing of saliva: a diagnostic accuracy study. *The Lancet Microbe*, 2(8):e397–e404.
- Yokota, I, Shane, P. Y, and Teshima, T (2021c). Logistic advantage of two-step screening strategy for sars-cov-2 at airport quarantine. *Travel Medicine and Infectious Disease*, 43:102127.
- Zhou, X.-H, Castelluccio, P, and Zhou, C (2005). Nonparametric estimation of roc curves in the absence of a gold standard. *Biometrics*, 61(2):600–609.
- Zou, G (2004). A modified poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology*, 159(7):702–706.
- 野間久史, 宇野慧, 石井亮太 (2025). 修正ポアソン回帰・修正最小二乗回帰による 2 値アウトカムデータの多変量解析. *日本統計学会誌*, 55(1):137–157.

11 付録

11.1 パラメーター空間を制限しない場合の MCMC の結果

研究初期の実験で, label-switching の影響を加味せずに行ったシミュレーションの $Se2, Sp2$ の結果を図 14 から図 18 に示す.

11.2 実験 1 における他の有病率の設定での結果

有病率が 0.01%, 0.1%, 1.0% における $Se2, Sp2$ の結果を図 19 から図 21 に示す.

11.3 実験 2 における他の有病率の設定での結果

有病率が 0.1%, 1.0% における $Se2$ の結果を図 22 に掲載する.

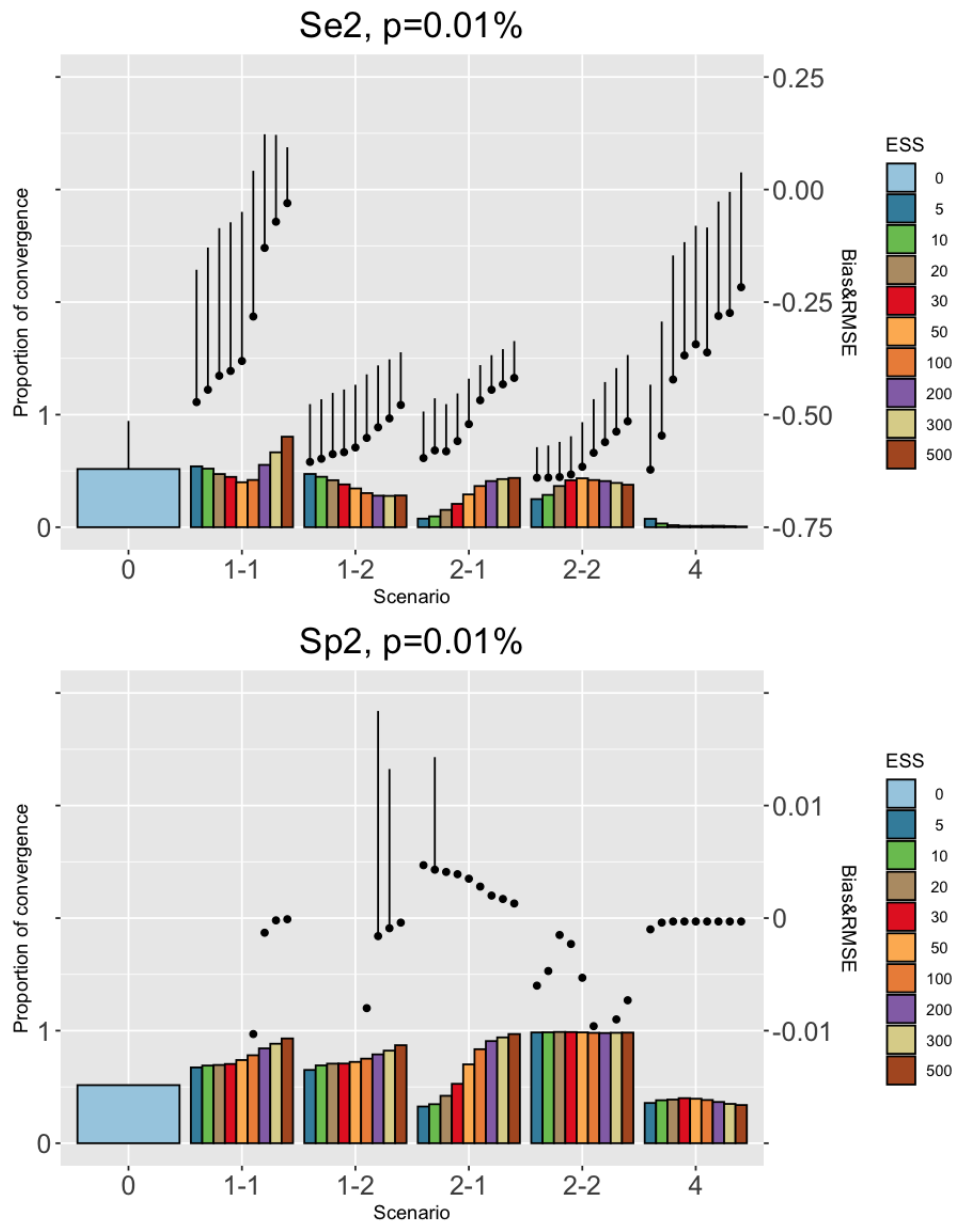


図 14: 有病率 p が 0.01% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している.

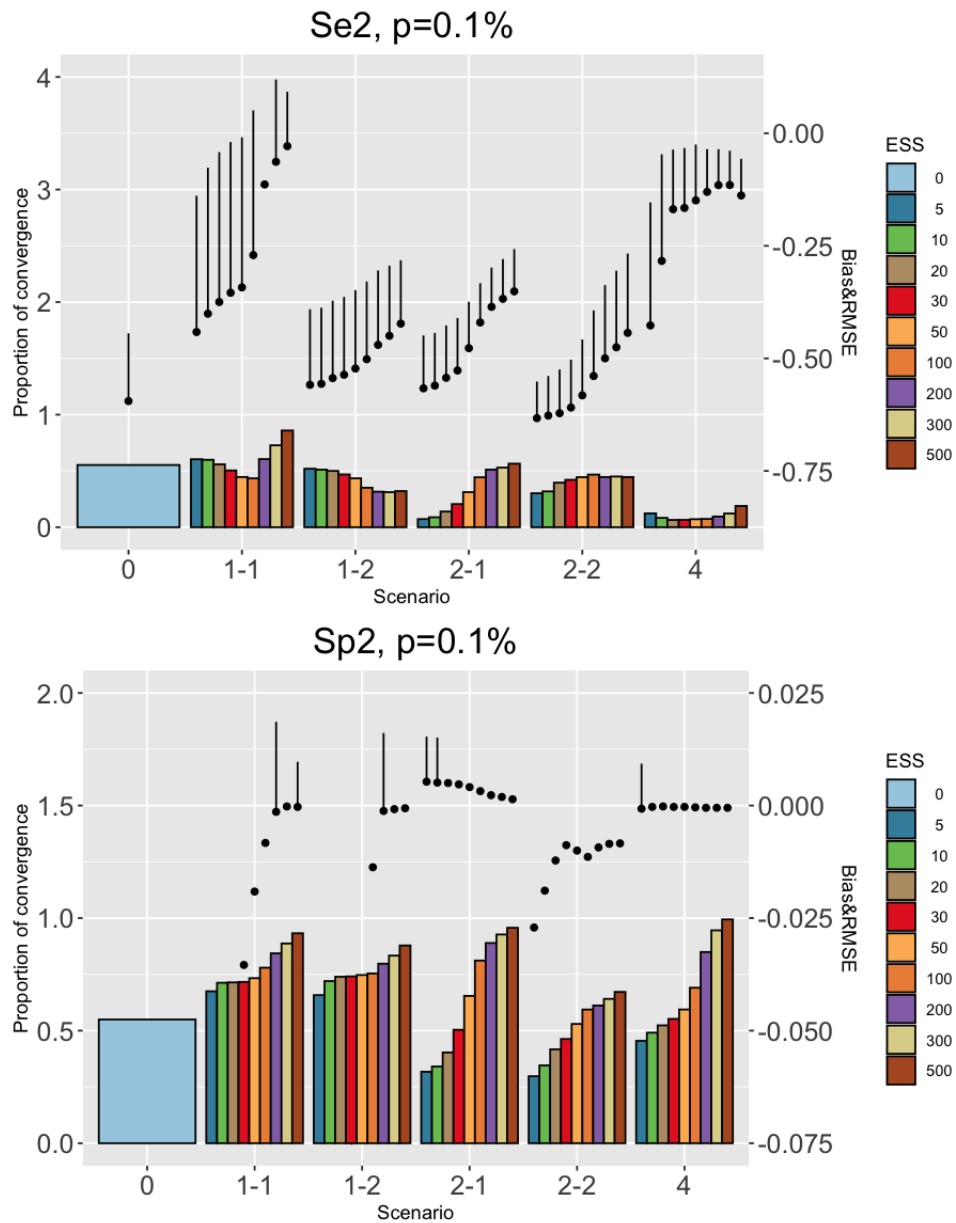


図 15: 有病率 p が 0.1% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している.

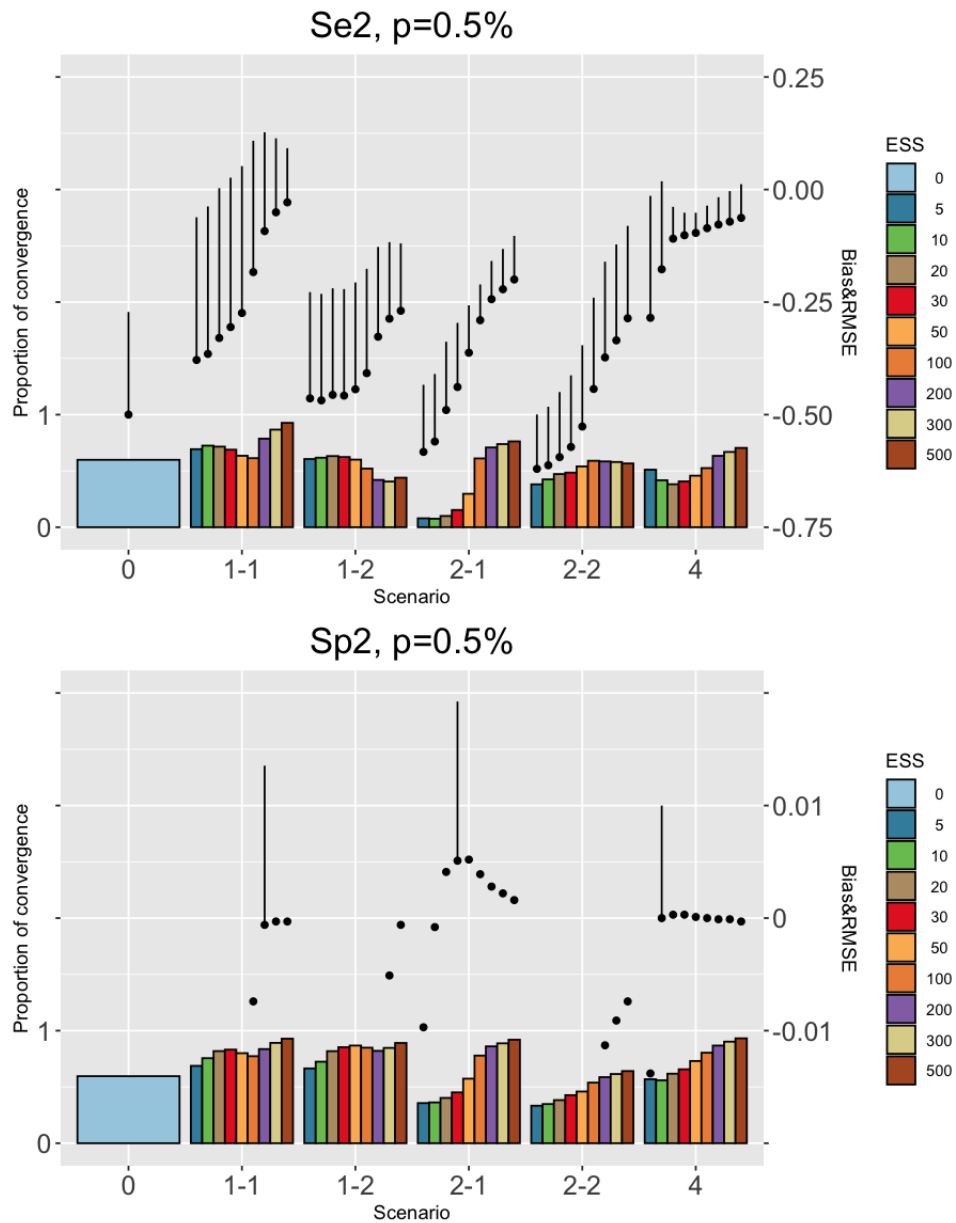


図 16: 有病率 p が 0.5% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している.

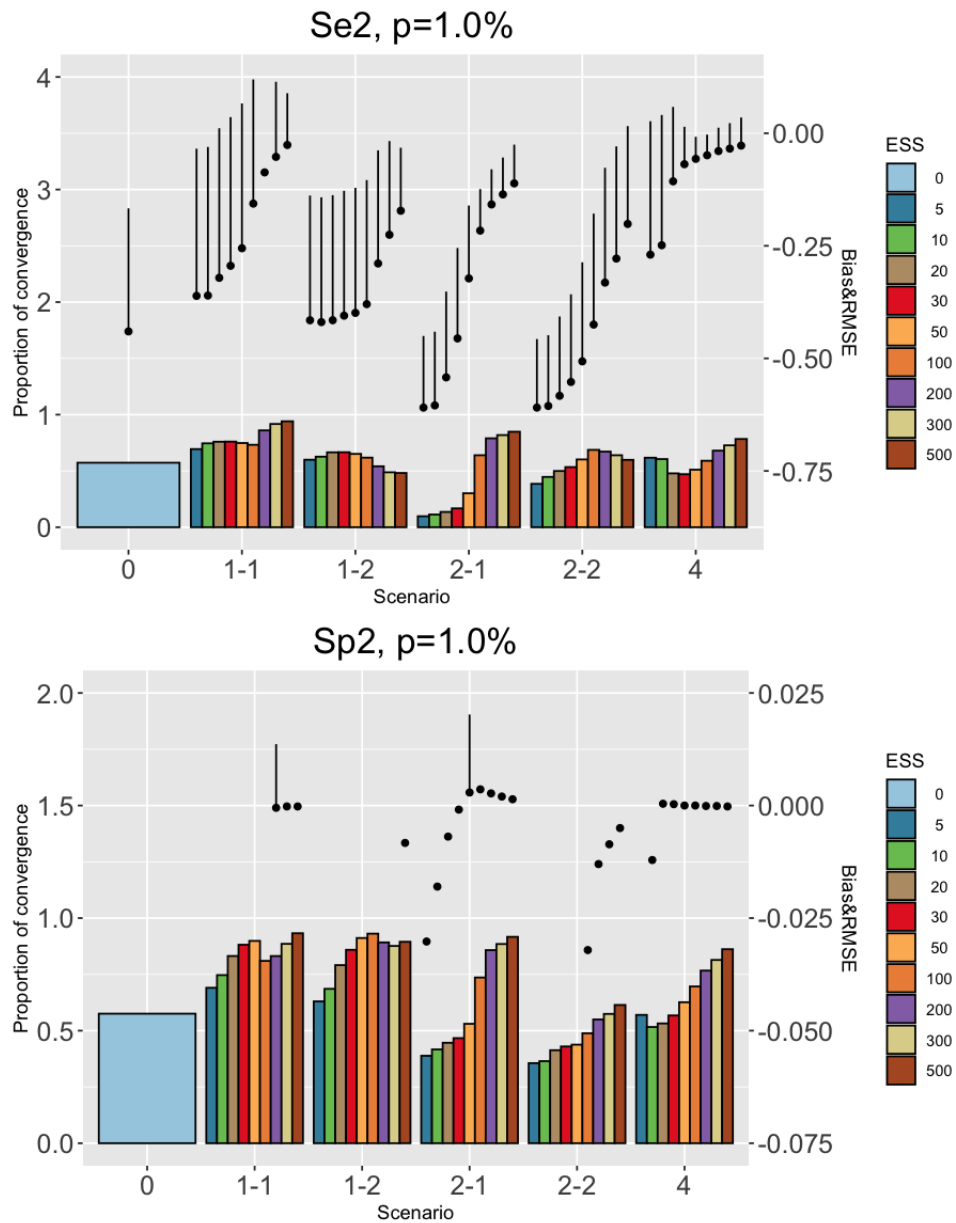


図 17: 有病率 p が 1.0% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している.

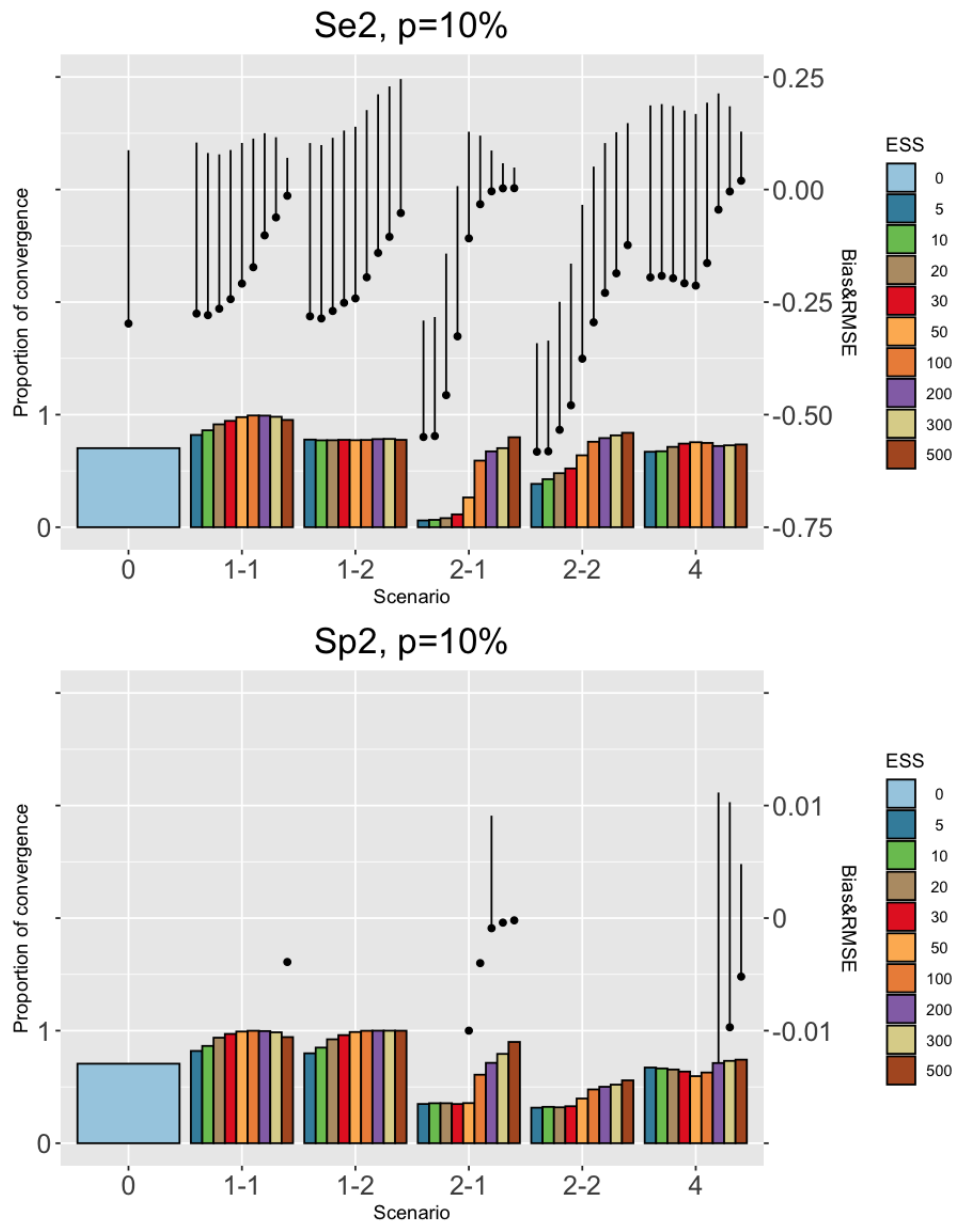


図 18: 有病率 p が 10% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している.

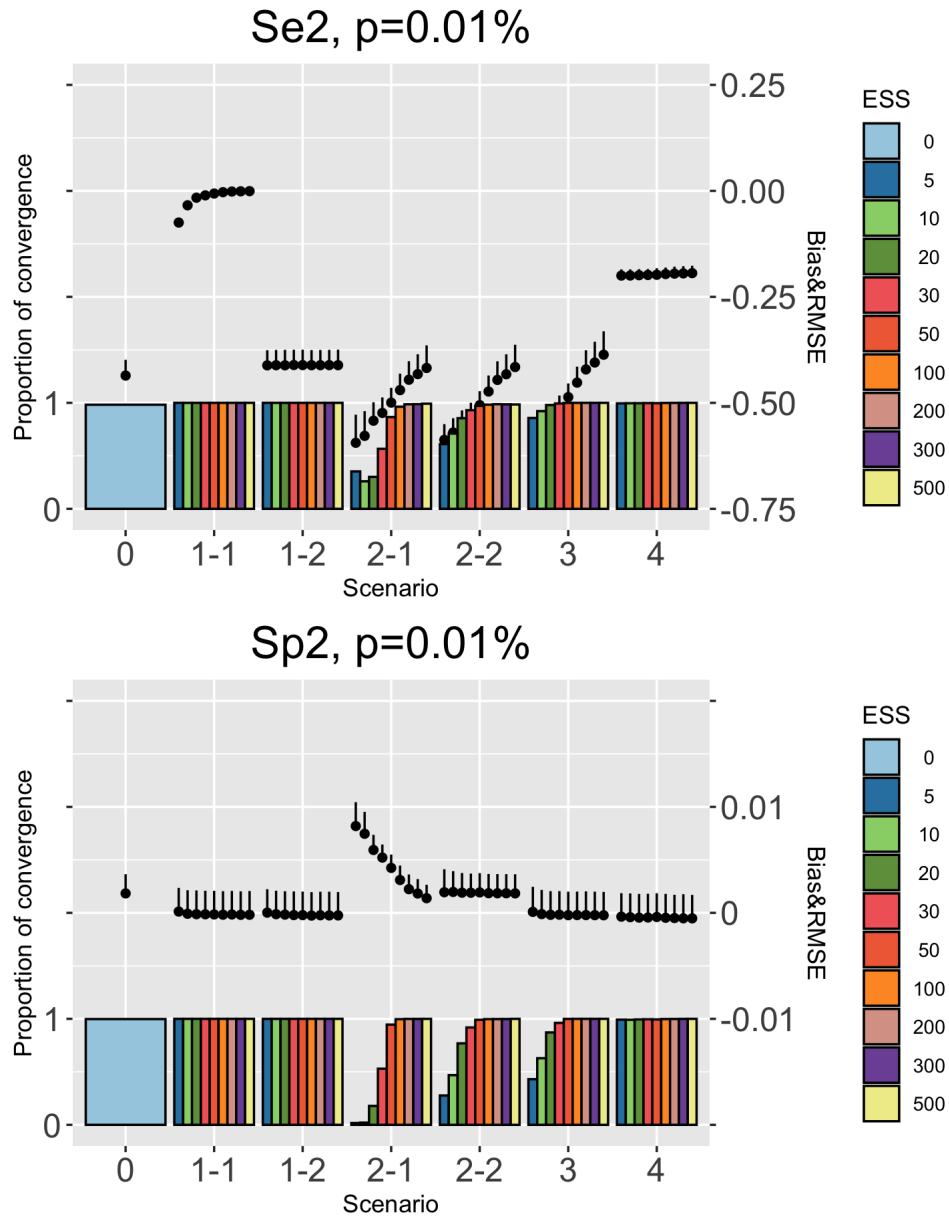


図 19: 有病率 p が 0.01% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している. ESS の “Fixed” は, 事前情報に真値を与えた状況を示している.

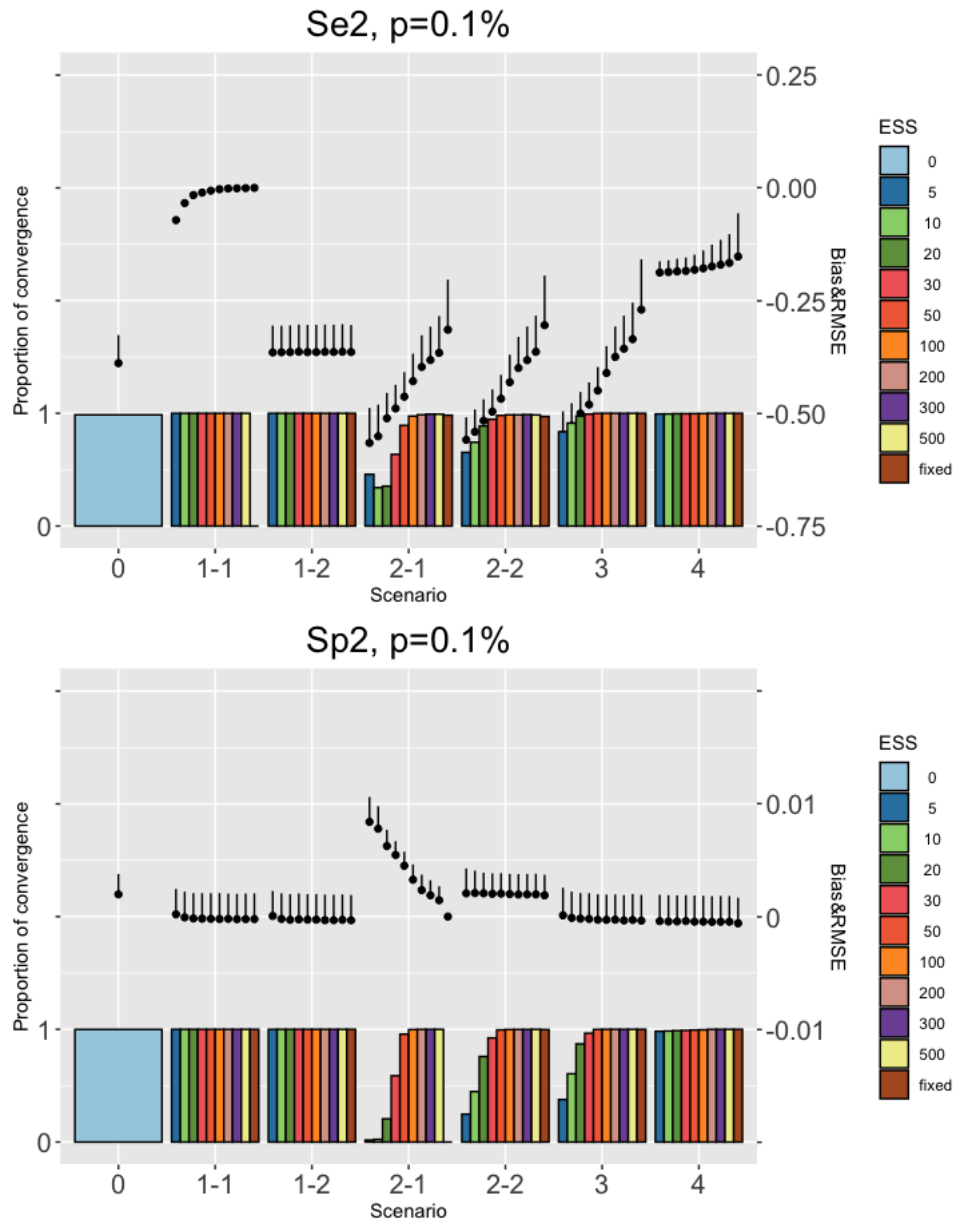


図 20: 有病率 p が 0.1% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している. ESS の “Fixed” は, 事前情報に真値を与えた状況を示している.

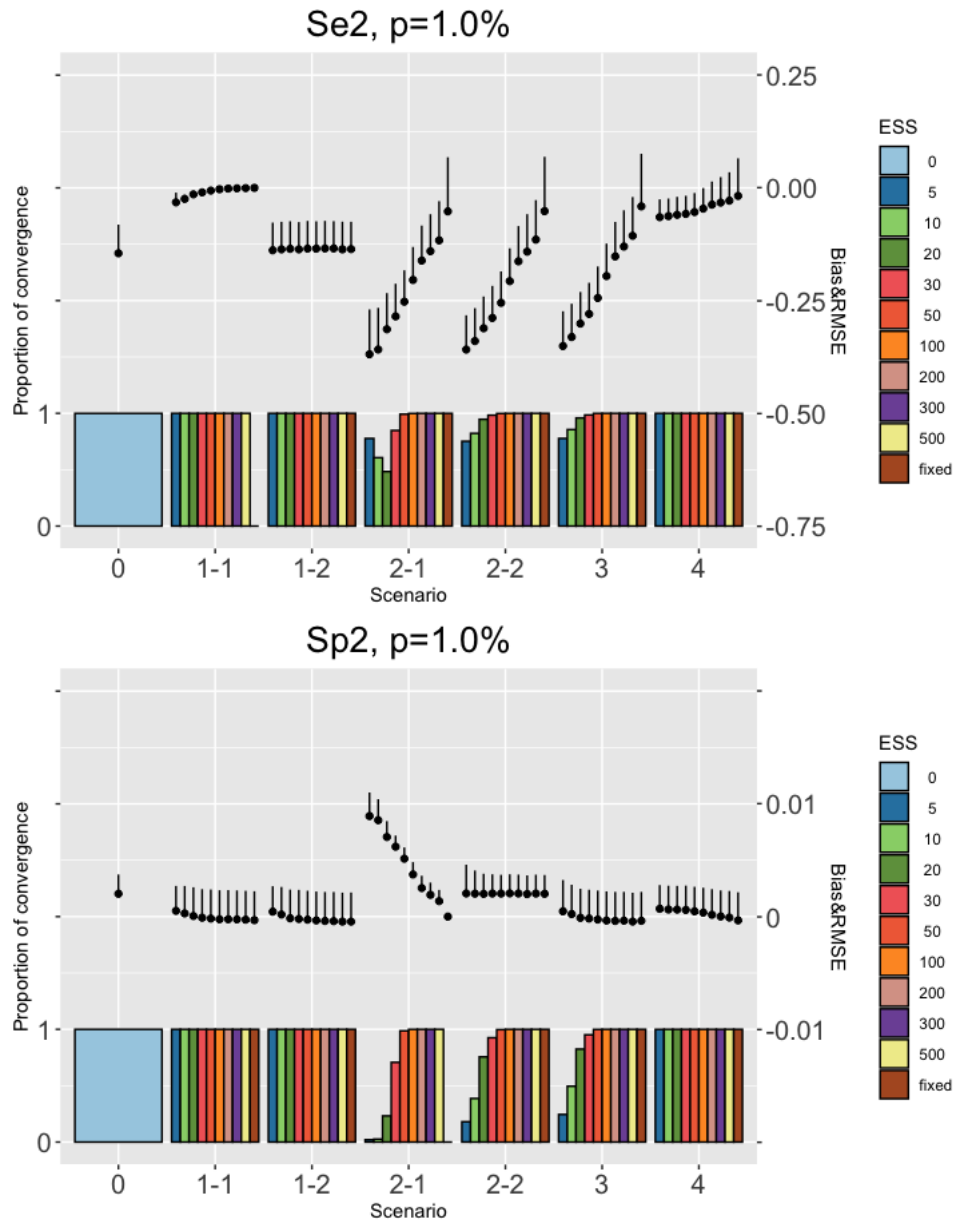


図 21: 有病率 p が 1.0% の場合のシミュレーション実験 1 における $Se2$ (上パネル) と $Sp2$ (下パネル) の結果. 黒色の点は $Sp2$ のバイアスを表し, 実線の長さは $Sp2$ の平均二乗誤差の平方根 (RMSE) を表す. 有効サンプルサイズ ESS に応じて色付けされた棒グラフは, マルコフ連鎖モンテカルロ法の収束割合を表している. ESS の “Fixed” は, 事前情報に真値を与えた状況を示している.

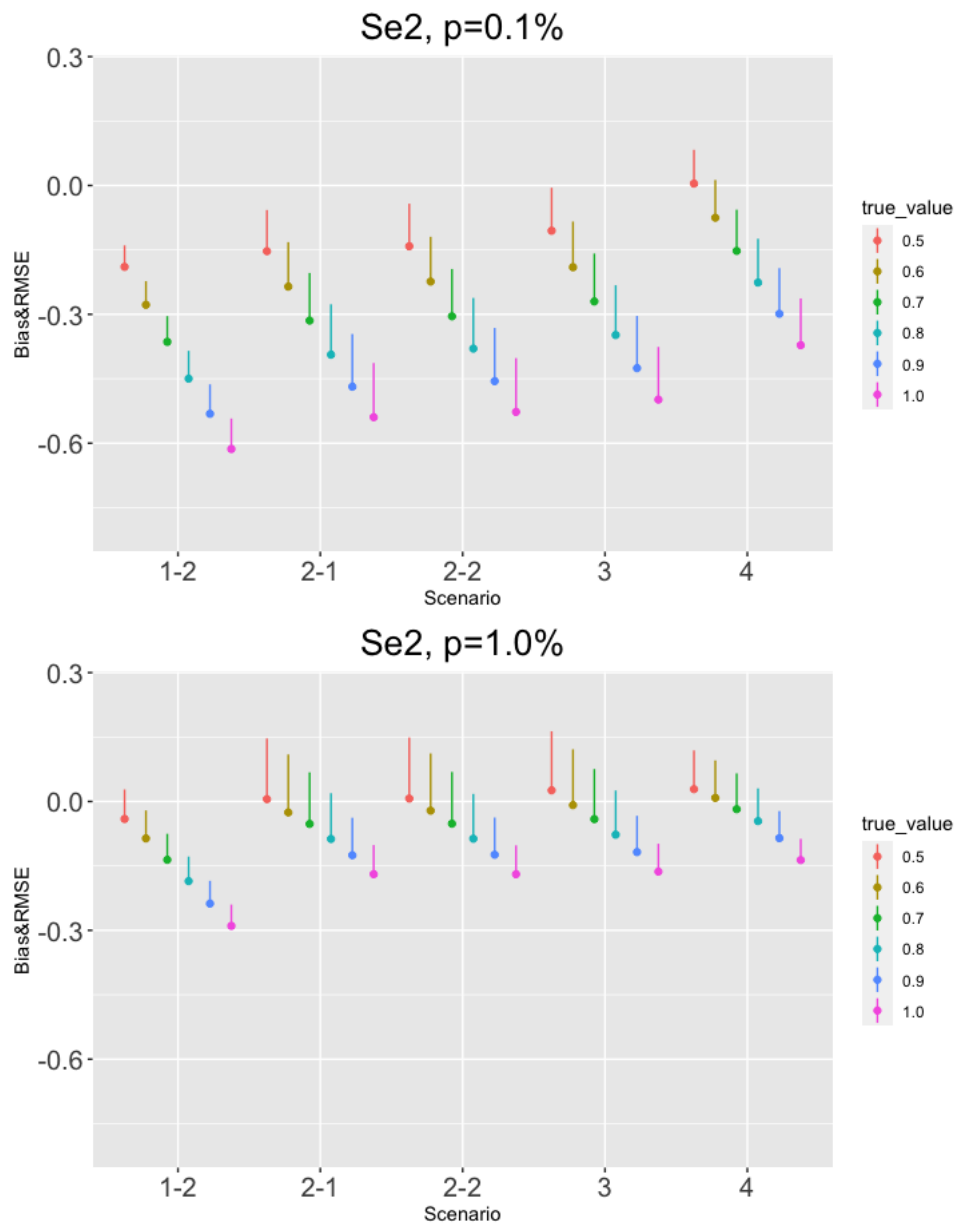


図 22: シミュレーション実験 2 における $Se2$ の結果. 有病率 p が 0.1% の場合 (上パネル) と 1.0% の場合 (下パネル) の結果を比較している. 点は $Se2$ のバイアスを表し, 実線の長さは $Se2$ の平均二乗誤差の平方根 (RMSE) を表す. 真値に応じて色付けしている.

11.4 比較手法の数理的背景

Hui and Walter (1980) で提案された推定方式 (以下, Hui-Walter モデルと記載) と Lu et al. (2018) で提案された推定方式 (以下, Lu モデルと記載) について紹介する. 2 集団 2 検査において, 集団 s の検査 j の偽陽性率を α_{sj} , 偽陰性率を β_{sj} , 有病率を π_s とおく. また, 集団 s の検査 1 における結果と検査 2 における結果がそれぞれ l, k になった人数を n_{slk} とおく. 条件付き独立性の仮定が成立するとき, 尤度 L は式 (1), (2) と同様に,

$$L = \prod_{s=1}^2 \{ (\pi_s(1-\beta_{s1})(1-\beta_{s2}) + (1-\pi_s)\alpha_{s1}\alpha_{s2})^{n_{s11}} \\ \times \{ \pi_s\beta_{s1}(1-\beta_{s2}) + (1-\pi_s)\alpha_{s1}\alpha_{s2} \}^{n_{s10}} \\ \times \{ \pi_s(1-\beta_{s1})\beta_{s2} + \pi_s(1-\alpha_{s1})\alpha_{s2} \}^{n_{s01}} \\ \times \{ \pi_s\beta_{s1}\beta_{s2} + \pi_s(1-\alpha_{s1})(1-\alpha_{s2}) \}^{n_{s00}} \}$$

となる. 推定するパラメータは $(\alpha_{sj}, \beta_{sj}, \pi_s)$ の合計 10 種類であるが, モデルの自由度は 6 のため一意に推定できない. Hui-Walter モデルでは, 集団間の感度・特異度はそれぞれ同値という仮定を入れることで, 自由度の問題を解決した. つまり,

$$\alpha_{sj} = \alpha_j, \quad \beta_{sj} = \beta_j$$

として, 最尤推定量を算出する方法を提案した.

一方, Lu et al. (2018) では条件付き独立性の仮定が現実的でない点について指摘し, 検査の取りうる結果の各確率を条件付き確率を利用して定式化する経験尤度の構成による推定方法を提案した. 検査 1 と検査 2 の結果を表す変数を T_1, T_2 とし, 確率 $P_{s,(u,v)}, p_{s(u|v)}, q_{s(u|v)}$ を次のように定義する.

$$P_{s,(u,v)} = P(T_1 = u, T_2 = v | s), \quad p_{s,(u|v)} = P(T_2 = u | T_1 = v, s), \quad q_{s,(u|v)} = P(T_1 = u | T_2 = v, s)$$

ただし, $u, v \in \{0, 1\}$ である. このとき, 集団 s における各検査結果の取りうる確率は

$$P_{s,(1,1)} = [\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1]p_{s(1|1)} \\ P_{s,(0,1)} = [\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1]p_{s(0|1)} \\ P_{s,(0,0)} = [\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1)]p_{s(0|0)} \\ P_{s,(1,0)} = [\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1)]p_{s(1|0)}$$

となる. 条件付き独立性の仮定をおいたモデルと比較すると, 括弧内の次数が 3 から 2 に減ったことが特徴の 1 つである. すると, 検査 1 のみまたは検査 2 のみでそれぞれ尤

度が記述できる.

$$\begin{aligned}
L_1 &= \prod_{s=1}^2 [(\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1)p_{s(1|1)}]^{n_{s11}} [(\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1))p_{s(1|0)}]^{n_{s01}} \\
&\quad \times [(\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1)p_{s(0|1)}]^{n_{s10}} [(\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1))p_{s(0|0)}]^{n_{s00}} \\
&\approx \prod_{s=1}^2 [(\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1)p_{s(1|1)}]^{n_{s11}+n_{s10}} [(\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1))p_{s(1|0)}]^{n_{s01}+n_{s00}}
\end{aligned}$$

$$\begin{aligned}
L_2 &= \prod_{s=1}^2 [(\pi_s(1-\beta_2) + (1-\pi_s)\alpha_2)q_{s(1|1)}]^{n_{s11}} [(\pi_s\beta_2 + (1-\pi_s)(1-\alpha_2))q_{s(1|0)}]^{n_{s01}} \\
&\quad \times [(\pi_s(1-\beta_2) + (1-\pi_s)\alpha_2)q_{s(0|1)}]^{n_{s10}} [(\pi_s\beta_2 + (1-\pi_s)(1-\alpha_2))q_{s(0|0)}]^{n_{s00}} \\
&\approx \prod_{s=1}^2 [(\pi_s(1-\beta_2) + (1-\pi_s)\alpha_2)q_{s(1|1)}]^{n_{s11}+n_{s10}} [(\pi_s\beta_2 + (1-\pi_s)(1-\alpha_2))q_{s(1|0)}]^{n_{s01}+n_{s00}}
\end{aligned}$$

両方の尤度において, 以下のような等式が成立する.

$$\begin{aligned}
P_{s,(1,1)} &= \{\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1\}p_{s(1|1)} = \{\pi_s(1-\beta_2) + (1-\pi_s)\alpha_2\}q_{s(1|1)} \\
P_{s,(1,0)} &= \{\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1)\}p_{s(1|0)} = \{\pi_s\beta_2 + (1-\pi_s)(1-\alpha_2)\}q_{s(1|0)} \\
P_{s,(0,1)} &= \{\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1\}p_{s(0|1)} = \{\pi_s(1-\beta_2) + (1-\pi_s)\alpha_2\}q_{s(0|1)} \\
P_{s,(0,0)} &= \{\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1)\}p_{s(0|0)} = \{\pi_s\beta_2 + (1-\pi_s)(1-\alpha_2)\}q_{s(0|0)}
\end{aligned}$$

この等式が成立する場合, 次のように $p_{s(\cdot|\cdot)}, q_{s(\cdot|\cdot)}$ を混合した表現で書くことができる.

$$\begin{aligned}
P_{s,(1,1)} &= \frac{1}{2} \left(\{\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1\}p_{s(1|1)} + \{\pi_s(1-\beta_2) + (1-\pi_s)\alpha_2\}q_{s(1|1)} \right) \\
P_{s,(1,0)} &= \frac{1}{2} \left(\{\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1)\}p_{s(1|0)} + \{\pi_s\beta_2 + (1-\pi_s)(1-\alpha_2)\}q_{s(1|0)} \right) \\
P_{s,(0,1)} &= \frac{1}{2} \left(\{\pi_s(1-\beta_1) + (1-\pi_s)\alpha_1\}p_{s(0|1)} + \{\pi_s(1-\beta_2) + (1-\pi_s)\alpha_2\}q_{s(0|1)} \right) \\
P_{s,(0,0)} &= \frac{1}{2} \left(\{\pi_s\beta_1 + (1-\pi_s)(1-\alpha_1)\}p_{s(0|0)} + \{\pi_s\beta_2 + (1-\pi_s)(1-\alpha_2)\}q_{s(0|0)} \right)
\end{aligned}$$

これにより, 新しい尤度として以下の経験尤度が定義でき, 感度・特異度の最尤推定量

を算出する方法を提案した.

$$\begin{aligned}
L &= \prod_{s=1}^2 P_{s,(1,1)} P_{s,(1,0)} P_{s,(0,1)} P_{s,(0,0)} \\
&\approx \prod_{s=1}^2 \left(\{ \pi_s(1-\beta_1) + (1-\pi_s)\alpha_1 \} p_{s(1|1)} + \{ \pi_s(1-\beta_2) + (1-\pi_s)\alpha_2 \} q_{s(1|1)} \right) \\
&\quad \times \left(\{ \pi_s\beta_1 + (1-\pi_s)(1-\alpha_1) \} p_{s(1|0)} + \{ \pi_s\beta_2 + (1-\pi_s)(1-\alpha_2) \} q_{s(1|0)} \right) \\
&\quad \times \left(\{ \pi_s(1-\beta_1) + (1-\pi_s)\alpha_1 \} p_{s(0|1)} + \{ \pi_s(1-\beta_2) + (1-\pi_s)\alpha_2 \} q_{s(0|1)} \right) \\
&\quad \times \left(\{ \pi_s\beta_1 + (1-\pi_s)(1-\alpha_1) \} p_{s(0|0)} + \{ \pi_s\beta_2 + (1-\pi_s)(1-\alpha_2) \} q_{s(0|0)} \right)
\end{aligned}$$

11.5 対数二項回帰と修正ポアソン回帰の性能比較の結果.

6.3 節のシミュレーション実験において, 共変量の数が 10 個の場合の結果と 15 個の場合の結果を表 10 と表 11 に示す.

11.6 動物性・植物性タンパク質の摂取量の組み合わせごとの結果

動物性タンパク質と植物性タンパク質の高摂取割合群と低接種割合群の組み合わせで群分けした場合のベースライン特性と有害事象を表 12 と表 13 に示す.

表 10: 対数二項回帰と修正ポアソン回帰の性能を比較した結果 (共変量が 10 個の場合).

			$r = 0$		$r = 0.1$		$r = 0.4$		$r = 0.8$	
RR	指標	p	mp	lb	mp	lb	mp	lb	mp	lb
0.7	Bias	0.1	1.18	53.39	1.00	49.27	1.76	47.23	4.27	47.52
		0.3	4.68	45.71	4.60	45.70	4.91	45.67	5.43	45.67
		0.5	9.15	45.87	9.50	45.87	9.30	45.87	9.56	45.87
	RMSE	0.1	0.17	0.88	0.17	0.81	0.19	0.75	0.31	0.76
		0.3	0.09	0.71	0.09	0.71	0.10	0.71	0.13	0.71
		0.5	0.11	0.71	0.12	0.71	0.11	0.71	0.13	0.71
	CP	0.1	1.00	0.74	1.00	0.68	1.00	0.66	1.00	0.65
		0.3	1.00	0.65	1.00	0.64	1.00	0.64	1.00	0.64
		0.5	1.00	0.65	1.00	0.65	1.00	0.65	1.00	0.65
1.0	Bias	0.1	2.60	22.43	3.23	17.66	4.05	18.25	9.33	18.43
		0.3	0.90	15.70	0.61	15.70	0.57	15.67	1.28	15.67
		0.5	0.35	15.79	0.36	15.87	0.30	15.87	1.09	15.87
	RMSE	0.1	0.26	0.70	0.26	0.60	0.28	0.61	0.48	0.63
		0.3	0.11	0.57	0.11	0.57	0.12	0.57	0.17	0.57
		0.5	0.07	0.57	0.07	0.57	0.08	0.57	0.11	0.57
	CP	0.1	1.00	0.71	1.00	0.67	1.00	0.66	1.00	0.65
		0.3	1.00	0.64	1.00	0.64	1.00	0.64	1.00	0.64
		0.5	1.00	0.65	1.00	0.65	1.00	0.65	1.00	0.65
2.0	Bias	0.1	7.89	-77.05	5.54	-80.61	5.66	-80.28	9.05	-78.10
		0.3	-30.35	-84.30	-30.56	-84.30	-30.62	-84.33	-30.14	-84.33
		0.5	-49.79	-84.13	-49.54	-84.13	-48.34	-84.13	-47.91	-84.13
	RMSE	0.1	0.59	1.02	0.54	1.00	0.66	1.00	0.80	1.01
		0.3	0.37	1.01	0.38	1.01	0.39	1.01	0.44	1.01
		0.5	0.52	1.00	0.52	1.00	0.51	1.00	0.53	1.00
	CP	0.1	1.00	0.69	1.00	0.67	1.00	0.65	1.00	0.65
		0.3	1.00	0.64	1.00	0.64	1.00	0.64	1.00	0.64
		0.5	1.00	0.65	1.00	0.65	1.00	0.65	1.00	0.65

¹ バイアスについては 100 倍した値を記載

² RR: リスク比の真値. CP: リスク比の推定値の収束割合. p : イベント発生率. r : 二変数間の相関の値.
mp: 修正ポアソン回帰, lb: 対数二項回帰.

表 11: 対数二項回帰と修正ポアソン回帰の性能を比較した結果 (共変量が 15 個の場合).

			$r = 0$		$r = 0.1$		$r = 0.4$		$r = 0.8$	
RR	指標	p	mp	lb	mp	lb	mp	lb	mp	lb
0.7	Bias	0.1	0.23	51.32	1.86	49.90	2.70	49.81	5.71	50.22
		0.3	6.70	45.67	6.78	45.67	6.34	45.67	7.19	45.67
		0.5	11.27	45.87	11.06	45.87	10.99	45.87	11.84	45.87
	RMSE	0.1	0.17	0.81	0.18	0.77	0.20	0.77	0.30	0.77
		0.3	0.11	0.71	0.11	0.71	0.11	0.71	0.15	0.71
		0.5	0.13	0.71	0.13	0.71	0.13	0.71	0.15	0.71
	CP	0.1	1.00	0.64	1.00	0.61	1.00	0.59	1.00	0.59
		0.3	1.00	0.64	1.00	0.64	1.00	0.64	1.00	0.64
		0.5	1.00	0.65	1.00	0.65	1.00	0.65	1.00	0.65
1.0	Bias	0.1	4.08	21.28	5.76	20.05	6.42	20.08	13.35	20.25
		0.3	0.28	15.67	0.49	15.67	0.37	15.67	1.44	15.67
		0.5	0.10	15.87	0.44	15.87	0.61	15.87	0.60	15.87
	RMSE	0.1	0.27	0.67	0.31	0.62	0.36	0.62	0.68	0.62
		0.3	0.10	0.57	0.11	0.57	0.11	0.57	0.17	0.57
		0.5	0.07	0.57	0.07	0.57	0.08	0.57	0.11	0.57
	CP	0.1	1.00	0.62	1.00	0.60	1.00	0.59	1.00	0.59
		0.3	1.00	0.64	1.00	0.64	1.00	0.64	1.00	0.64
		0.5	1.00	0.65	1.00	0.65	1.00	0.65	1.00	0.65
2.0	Bias	0.1	5.44	-79.99	6.63	-79.91	8.52	-79.92	14.84	-79.19
		0.3	-37.40	-84.33	-37.17	-84.33	-36.79	-84.33	-35.97	-84.33
		0.5	-54.15	-84.13	-54.24	-84.13	-54.14	-84.13	-52.21	-84.13
	RMSE	0.1	0.62	0.99	0.61	0.99	0.74	0.99	0.91	0.99
		0.3	0.43	1.01	0.43	1.01	0.43	1.01	0.47	1.01
		0.5	0.56	1.00	0.56	1.00	0.56	1.00	0.57	1.00
	CP	0.1	1.00	0.62	1.00	0.59	1.00	0.59	1.00	0.59
		0.3	1.00	0.64	1.00	0.64	1.00	0.64	1.00	0.64
		0.5	1.00	0.65	1.00	0.65	1.00	0.65	1.00	0.65

¹ バイアスについては 100 倍した値を記載

² RR: リスク比の真値. CP: リスク比の推定値の収束割合. p : イベント発生率. r : 二変数間の相関の値.
mp: 修正ポアソン回帰, lb: 対数二項回帰.

表 12: 動物性・植物性タンパク質の摂取量のペアに応じた慢性心不全患者のベースライン特性

特性	全体 (N=146)	低動物性/低植物性 (N=37)	低動物性/高植物性 (N=52)	高動物性/低植物性 (N=24)	高動物性/高植物性 (N=33)
年齢 (歳)	68 (60–77)	64 (53–72)	67 (61–75)	67 (61–80)	73 (63–80)
女性	46 (32)	6 (16)	16 (31)	11 (46)	13 (39)
BMI (kg/m ²)	22.9 (20.6–25.6)	23.2 (20.0–24.8)	24.1 (21.6–26.5)	22.9 (21.9–26.4)	21.9 (19.4–24.4)
LVEF (%)	45 (30–56)	49 (31–55)	45 (34–60)	48 (28–54)	41 (26–48)
NYHA class III	15 (10)	4 (11)	2 (4)	4 (17)	5 (15)
心不全の主な原因:					
虚血性心疾患	46 (32)	8 (22)	19 (37)	6 (25)	13 (39)
拡張型心筋症	46 (32)	14 (38)	15 (29)	9 (38)	8 (24)
その他の心疾患	54 (37)	15 (41)	18 (35)	9 (38)	12 (36)
高血圧	79 (54)	15 (41)	31 (60)	11 (46)	22 (67)
糖尿病	38 (26)	4 (11)	18 (35)	6 (25)	10 (30)
脂質異常症	103 (71)	27 (73)	40 (77)	13 (54)	23 (70)
検査値:					
ヘモグロビン (g/dl)	13.2 (11.9–14.3)	13.9 (12.8–15.1)	13.4 (11.9–14.3)	12.2 (11.5–13.5)	12.2 (11.6–14.0)
血清アルブミン (g/dl)	4.2 (3.9–4.4)	4.2 (4.0–4.3)	4.2 (3.9–4.4)	4.2 (3.9–4.4)	4.1 (4.0–4.2)
eGFR (ml/min/1.73m)	54.3 (38.4–67.6)	60.6 (45.7–73.4)	54.5 (38.7–68.8)	48.0 (40.9–65.8)	50.0 (35.2–59.4)
HbA1c (%)	5.8 (5.6–6.2)	5.7 (5.4–6.1)	5.9 (5.7–6.4)	5.8 (5.5–6.3)	5.8 (5.7–6.1)
hs-CRP (mg/dl)	484 (250–1,465)	473 (214–797)	542 (261–2,120)	651 (424–1,955)	446 (227–1,008)
血漿BNP (pg/ml)	156 (76–374)	117 (48–172)	154 (52–396)	170 (99–498)	220 (134–478)
6 分間歩行試験 (m)	432 (346–490)	450 (342–511)	432 (394–463)	375 (306–477)	426 (312–499)
MLHFQ	15 (5–27)	13 (5–20)	17 (4–27)	19 (8–27)	13 (4–28)
服用薬:					
ACE 阻害薬または ARB	111 (76)	25 (68)	39 (75)	19 (79)	28 (85)
β 遮断薬	129 (88)	35 (95)	45 (87)	18 (75)	31 (94)
MRA	85 (58)	20 (54)	28 (54)	17 (71)	20 (61)

中央値 (第 1 四分位数-第 3 四分位数) または n(%) で示す。コホート全体を総タンパク質 (閾値:16.0%), 動物性タンパク質 (閾値:10.3%), 植物性タンパク質 (閾値:6.0%) の高摂取群と低摂取群に分けた。すべての閾値は制限付き三次スプライン曲線の傾きの値に基づいて決定した。

表 13: 12 ヶ月の追跡調査期間中の有害事象

有害事象	全体 (N=146)	低動物性/低植物性 (N=37)	低動物性/高植物性 (N=52)	高動物性/低植物性 (N=24)	高動物性/高植物性 (N=33)
主要評価項目:					
全死亡および全入院	31 (21)	7 (19)	6 (12)	9 (38)	9 (27)
心血管死および心不全入院	18 (12)	3 (8)	4 (8)	6 (25)	5 (15)
副次評価項目:					
全死亡	6 (4)	0 (0)	1 (2)	4 (17)	1 (3)
心血管死	4 (3)	0 (0)	1 (2)	3 (12)	0 (0)
非心血管死	2 (1)	0 (0)	0 (0)	1 (4)	1 (3)
全入院	28 (19)	7 (19)	5 (10)	8 (33)	8 (24)
心不全入院	15 (10)	3 (8)	3 (6)	4 (17)	5 (15)
非心不全入院	13 (9)	4 (11)	2 (4)	4 (17)	3 (9)

中央値 (第 1 四分位数-第 3 四分位数) または n(%) で示す。コホート全体を総タンパク質 (閾値:16.0%), 動物性タンパク質 (閾値:10.3%), 植物性タンパク質 (閾値:6.0%) の高摂取群と低摂取群に分けた。すべての閾値は制限付き三スプライン曲線の傾きの値に基づいて決定した。各グループについては、“動物性/植物性” について記載。