

令和7年度 厚生労働科学研究費補助金
(女性の健康の包括的支援政策研究事業)

課題名：健康寿命延伸に備えた女性の心身の健康支援のための普及啓発に向けた研究

分担研究報告書

研究名：特定健診データを利用した骨粗鬆症性骨折の予測モデルの検討

研究協力者 岩倉正浩 秋田大学大学院医学系研究科衛生学・公衆衛生学講座・助教
研究代表者 野村恭子 秋田大学大学院医学系研究科衛生学公衆衛生学講座・教授
研究分担者 長島健悟 慶應義塾大学病院臨床研究推進センター・特任准教授
研究分担者 谷原真一 久留米大学医学部公衆衛生学講座・教授
研究分担者 宮腰尚久 秋田大学大学院医学系研究科整形外科学講座・教授
研究分担者 寺内公一 東京科学大学大学院医歯学総合研究科
茨城県地域産科婦人科学講座・教授
研究協力者 木村竜太 秋田大学大学院医学系研究科整形外科学講座・助教

要旨 本邦における骨粗鬆症スクリーニングにあたる骨粗鬆症検診の受診率は2022年時点で5%程度と極めて低く、本邦の公衆衛生的課題である。しかし、65歳未満の女性において、骨粗鬆症スクリーニングの必要性を判断するリスクアセスメントは少ない。そこで、本研究では、保険者ベースの大規模保険請求・健診データベースを用い、40歳から60歳の女性を対象に特定健診にて取得される因子を用いて、将来の骨粗鬆症性骨折の発生を予測するモデルを開発し、その性能を内的検証にて評価することを目的とした。

本研究の対象者はデータベースに登録されていた40歳以上60歳未満の被保険者の女性とした。予測因子は特定健診の結果から、年齢、Body mass index (BMI)、腹囲、収縮期・拡張期血圧と標準的な質問票に含まれる項目を利用したモデル1、モデル1に血液検査結果を追加したモデル2、既存のリスクアセスメントであるOsteoporosis Self-Assessment Toolを改変したモデル3を開発した。アウトカムは、骨粗鬆症性骨折（大腿骨近位部骨折、椎体骨折、橈骨遠位端骨折を含む）とし、ICD-10コードに基づき定義した。欠測値を多重補完法で補完した後、死亡を競合リスクとしたFine & Grayモデルを当てはめた。モデルの性能は、4年間のアウトカム累積発生率を指標とし、識別、校正、臨床的有用性を評価した。内的検証はブートストラップ法を用い、optimismを考慮した性能を評価した。

追跡開始時点で40歳以上60歳未満であった2,226,596名の女性の内、975,381名の被保険者の女性が解析対象となった。年齢、BMIの中央値(Q1, Q3)はそれぞれ、46.5(41.6, 51.8)歳、21.4(19.5, 24.1) kg/m²であった。追跡期間の中央値(Q1, Q3)は35(16, 56)ヶ月であった。モデルの性能はモデル1とモデル2で概ね一致しており、モデル3は他の2つのモデルと比較し校正と識別、臨床的有用性が若干劣る結果であった。また、optimismを考慮した性能もほぼ同様の結果であり、過学習による各モデルの性能の過剰評価を示唆する結果は認められなかった。

モデル 1 と 2 の性能はほぼ同等であること、モデル 1 の方がモデル 2 と比較して予測因子が少ないこと、全ての予測因子は特定健診で一般に取得されている項目であることから、モデル 1 の実用性が高いと考えられた。今後は被扶養者の女性などにおける外的検証や、10 年間の累積発生割合を指標とした性能評価などが望まれる。

A. 研究の背景と目的

骨粗鬆症の有病率は全世界で約 20%と言われており、有病率は男性と比較して女性で 2.3 倍ほど高い¹。40 歳以上の日本人女性における有病率は 10～20%と推定されている²。さらに近年、骨粗鬆症の有病率は上昇している³。また、骨粗鬆症は脆弱性骨折の主要なリスク因子であり、障害調整生存年数 (disability-adjusted life-years [DALYs]) の損失につながる⁴。よって、骨粗鬆症を早期に発見し、治療を行うことは、骨粗鬆症性骨折予防の観点から重要である⁵。本邦でも健康増進法において骨粗鬆症の予防を目的に骨粗鬆症検診が実施されているが、本邦の骨粗鬆症検診の受診率は 2022 年時点でも約 5.5%と極めて低い⁶。さらに、世界的にも骨折のリスクがある女性の多くが必要な治療を受けておらず⁷、検診やスクリーニングの受診率と治療を受ける女性の割合を増加させることは重要な公衆衛生学的課題である⁸。

骨粗鬆症の早期発見と骨粗鬆症性骨折発生の予防を目指した様々なスクリーニング検査が存在する。代表的なスクリーニング検査としては、Dual-Energy X-ray Absorptiometry (DXA)を用いた大腿骨や腰椎の骨密度検査や、骨密度検査を伴うリスク評価の指標 (Fracture Risk Assessment Tool [FRAX] plus⁹や Fracture Risk Calculator [FRC]¹⁰ など)が広く用いられている。The US Preventive Services Task Force (USPSTF)からの最新のステートメント¹¹では、閉経後かつ 65 歳未満の女性で 1 つでも骨粗鬆症のリスク因子を有する人、または 65 歳以上の女性において、骨粗鬆症のスクリーニングは骨粗鬆症性骨折の予防に有効 (Grade B) であるとして、その実施が推奨されている。

特に、閉経後かつ 65 歳未満の女性で 1 つでも骨

粗鬆症のリスク因子を有する人においては、その中でどの女性がスクリーニング検査を受けるべきかを判断するためにリスクアセスメントツール (Risk Assessment tool) を使うことを提案している。リスクアセスメントツールには骨密度検査を伴わない FRAX や骨粗鬆症のリスクをみる the Simple Calculated Osteoporosis Risk Estimation (SCORE)¹²、Osteoporosis Risk Assessment Instrument (ORAI)¹³、Osteoporosis Index of Risk (OSIRIS)¹⁴、Osteoporosis Self-Assessment Tool (OST)^{15,16} などがある。FRAX¹⁷ は、その他のリスクアセスメントと比べて 10 年後の骨折リスクを評価できる点で有用であるが、糖質コルチコイドの服用歴や 2 次性骨粗鬆症の有無などが項目に含まれており、一般の人が自己報告として利用するにはハードルが高い。その他のリスクアセスメントツール¹²⁻¹⁵ は年齢や体重、その他数個の因子から構成されており簡便に実施可能あり、かつ骨粗鬆症の予測精度も中程度ではある。しかし、骨密度検査を伴わないリスクアセスメントが将来の骨折リスクを予測できるかについて検証すること、65 歳未満の女性を対象とした大規模な研究を行うことが、今後の課題であるとされている¹¹。また、これらのリスクアセスメントツール¹²⁻¹⁵ には、飲酒や喫煙、身体活動、バイオマーカーなど重要な骨折の予測因子は含まれていないことも限界点である。

これらのことから、65 歳未満の女性を対象に既存のリスクアセスメントツールの限界点を補いつつ、将来の骨折予測に有用なリスクアセスメントを作成することは重要である。そこで、本研究では株式会社 JMDC が保有する保険者ベースの保険請求・健診データベース (以下、JMDC Claims DB) を用い、40 歳以上 60 歳未満の女性を対象に、ルーティーンで実施される特定健診の結果から将来の骨粗鬆症

性骨折の発生を予測するモデルを開発し内的妥当性の検証をすることを目的とした。現行の特定健診で測定される検査項目から、将来の骨粗鬆症性骨折の発生を予測できれば、リスクが高い女性への骨粗鬆症検診受診の勧奨、治療が必要な女性への早期治療開始、ひいては骨粗鬆症性骨折の予防の一助となると考えられる。

B. 方法

本研究は JMDC Claims DB に登録されている 40 歳以上 60 歳未満の女性を対象に、ベースラインの特定健診データから将来の骨粗鬆症性骨折の発生リスクを予測するモデルを開発および検証するコホート研究である。本研究は匿名加工情報の提供を受けて行うため、対象者本人からの同意を得る必要はない。本研究は、ヘルシンキ宣言に則り、かつ秋田大学の研究倫理審査を受け、実施許可を得たうえで実施している（承認番号：3016）。本研究は、日常的に収集されるデータを用いた観察研究の報告ガイドラインである REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement¹⁸ および臨床予測モデルの報告ガイドラインである Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD)¹⁹ のアップデート版である TRIPOD+AI²⁰ に沿って報告した。

データソース

JMDC Claims DB は、日本全国の主に大規模事業所の保険者から基本情報、保険請求データ、および特定健診結果を収集し、標準化処理が行われている。JMDC Claims DB の詳細は先行研究にて報告されている^{21,22}。JMDC Claims DB は 2005 年 1 月からのデータを保有しており、累積母集団は 2021 年 8 月の時点で 1200 万人を超えており²²、データの悉皆性は国内の商用の保険請求データベースの中で最大規模である。本研究におけるデータの観察可能期間は 2005 年 1 月から 2023 年 6 月である。

対象

対象は、追跡開始時点の年齢が 40 歳以上 60 歳未満かつ追跡開始時点より少なくとも過去 2 年間は JMDC Claims DB に存在していた被保険者の女性とした。加えて、**Table 1** に示した除外基準を満たす場合は対象から除外することで、ベースラインの時点にて骨粗鬆症を有しておらずかつ、医師から既に骨の状態の評価や管理を受けている可能性が低い女性（本研究で開発する予測モデルのスコープに収まる女性）を解析対象とした。各除外基準の詳細は **Table S1–6** に示した。除外基準 4 における経口ステロイドのプレドニゾロン換算の成分量の計算は Liu D らのガイド²³ および National Center for Biotechnology Information の情報²⁴ を参考に行った。更に同ガイド²³ に基づき、骨密度の低下に影響するとされる 1 日のプレドニゾロン換算の処方量が 5 mg 以上の日数が累積で 90 日以上の場合に解析対象から除外した。なお、対象者の抽出後、被扶養者の女性では、特定健診の結果の欠測割合が 50% を越えていたため予測モデル開発・検証の目的に適さないと判断し、除外することとした。

アウトカムの定義と打ち切りへの対応

本研究では、先行研究^{25–27} を参考に ICD-10 コードと（財）医療情報システム開発センター（MEDIS-DC）が作成しているレセ電算コードおよび病名²⁸ に基づき、骨粗しょう性骨折を以下の様に操作的に定義した：骨粗鬆症性骨折（M80X）、大腿骨頸部骨折（S720X）、転子貫通骨折（S721X）、腰椎圧迫骨折（S32X）、橈骨遠位端骨折（S525X）。同じ ICD-10 コードの分類の中に外傷性骨折や多発骨折など高エネルギー性の骨折の可能性が高い標準傷病名が含まれたり、小児の骨端線に生じる骨折の標準傷病名が含まれたりしているため、これらの傷病名は除外することとした。また、上腕骨近位端骨折（S422X）は本研究の対象者である 40–59 歳の女性では稀であることを考慮し、アウトカムの定義からは除外した。アウトカムの定義に用いたコードの詳細は **Table S3** に示した。

本研究では、上記のいずれかに該当する確定診断

コードが1度でも付与された場合に、骨粗鬆症性骨折の発生「あり」と定義し、追跡開始月から診断月までを発生までの期間として採用した。なお、先行研究²⁹を参考に頭蓋/頸部、体幹/骨盤、上肢、下肢の4領域の内、2領域以上の骨折の診断コードまたは、多発骨折を示すT02の診断コード、交通外傷を示すV01～V99の診断コードのいずれかがアウトカム発生日の±7日以内に付与されていた場合には、高エネルギー性の骨折の可能性が高いと判断し、アウトカム非発生とみなし、追跡を続けた。コードリストの詳細はTable S7に示した。

本研究における打ち切りのタイミングは、以下のイベントの内いずれかが発生した最も早い月とした：①アウトカムの発生、②観察可能期間の終了（2023年6月）、③死亡による資格喪失のため追跡終了、④その他の原因による資格喪失のため追跡終了。これらの内、③死亡による資格喪失のため追跡終了は競合リスクとして扱った。よって、イベント状況として0：②観察可能期間の終了または④その他の原因による資格喪失のため追跡終了、1：①アウトカムの発生、2：③死亡による資格喪失のため追跡終了とした変数を作成した。

予測因子の候補

本研究では、特定健診にて一般的に取得されているデータを用いて予測モデルを開発することを目的としているため、予測因子の候補は年齢が40歳以上の人々において特定健診で実施される可能性が高い質問票の項目および血液検査結果から選定した。また、これらの項目の中で骨粗鬆症や骨粗鬆症性骨折、骨折のリスク因子として考えられる因子を候補とした。なお全てのデータは追跡開始時点から2年前までの期間中に存在する健診結果の内、最新のものを利用することとした。具体的には、年齢³⁰、body mass index (BMI)³⁰、腹囲³⁰、現在喫煙の有無³⁰、飲酒頻度³⁰、飲酒量³⁰、身体活動 (physical activity: PA)³¹、歩行速度^{32,33}、食習慣³¹、睡眠の状況^{34,35}、収縮期血圧 (systolic blood pressure: SBP)・拡張期血圧 (Diastolic blood pressure)^{36,37}、高血糖状態³¹、脂質異常^{38,39}、肝

機能異常^{40,41}、薬の利用状況（血圧を下げる薬、脂質を下げる薬、血糖値を下げる薬の利用の有無）、貧血に関連する指標（ヘマトクリット値、Hb⁴²、赤血球数）、既往歴（脳血管疾患^{43,44}、心疾患^{45,46}）、生活習慣の改善に関する行動変容ステージ（行動変容ステージ）を健診結果から取得した。詳細はTable S8に示した。

予測因子の選定と変数の型

前述の予測因子の候補の中から、年齢（歳）、BMI (kg/m²)、腹囲 (cm)、SBP (mmHg)、DBP (mmHg)、現在喫煙（はい/いいえ）、飲酒頻度（ほとんどない/時々/毎日）、自覚的な歩行速度が他者よりも速い（はい/いいえ）、中強度以上の身体活動の実施（はい/いいえ）、低強度の身体活動の実施（はい/いいえ）、週3回以上の朝食の欠食（はい/いいえ）、十分な睡眠（はい/いいえ）、血圧を下げる薬の利用（はい/いいえ）、脂質を下げる薬の利用（はい/いいえ）、血糖値を下げる薬の利用（はい/いいえ）、脳血管疾患の既往歴（はい/いいえ）、心臓病の既往歴（はい/いいえ）、行動変容ステージ（維持期/実行期/準備期/関心期/無関心期）、空腹時血糖 (fasting plasma glucose: FPG) (mg/dL)、hemoglobin A1c (HbA1c) (%) [National Glycohemoglobin Standardization Program: NGSP unit]、high-density lipoprotein (HDL) (mg/dL)、alanine transaminase (ALT) (U/L)、gamma-glutamyl transferase (GGT) (U/L)、hemoglobin (Hb) (g/dL)を予測因子として採用した。予測因子の候補のうち血液検査項目である随時血糖、中性脂肪、low-density lipoprotein (LDL)、aspartate aminotransferase (AST)、ヘマトクリット値、赤血球数は先行研究を踏まえつつ、著者内で協議のもと除外することとした。また、飲酒量は欠測割合が約35%とその他の変数と比べて欠測割合が15–20%ほど高く予測因子としては不適切と考えられること、飲酒頻度を利用可能であることから、飲酒量は除外した。

その他の情報

上記の予測因子に加え、対象者の基本特性として

生年（西暦）、CED（cohort entry date）の西暦を取得した。

統計解析

統計解析は R (Ver. 4.4)にて行った。対象者の予測因子およびアウトカムに関して、記述統計量として連続変数では中央値と 25 および 75 パーセンタイル値 (Q1, Q3) を、名義・カテゴリ変数では度数と割合を算出した。また全ての変数について、欠測値の度数と割合も算出した。

欠測は予測因子のみ生じており、予測因子の欠測状況は年齢や資格分類、資格取得日、その他予測因子と関連していた。よって、missing at random (MAR) と仮定し、{mice}パッケージ⁴⁷の{mice}関数を用いて多重補完法を行い 30 セットの補完済みデータセットを作成した。補完モデルには、全ての選定された予測因子とイベント状況（関心イベント：骨折、競合リスク：死亡、打ち切り）を含めた。さらに、競合リスクを考慮したアウトカム情報の要約として、関心イベント（骨折）の Fine-Gray サブディストリビューションハザードの基準累積ハザード推定値を含めた。なお、推定値は{survival}パッケージ^{48, 49}の{finegray}関数を用いて変換した共変量を含まない Fine-Gray 変換データを作成し、重み付き Cox モデルを{coxph}関数で当てはめ、{basehaz}関数により Breslow 型に推定した基準累積ハザードを、各個人の観測時間に基づいて評価を行った。予測因子の非線形性は restricted cubic splines (RCS) により表現した。多重補完では、スプライン基底を、欠測を含むオリジナルデータ上で事前に作成し、各基底項を Just Another Variable (JAV; transform-then-impute) として、他の変数と同様に補完モデルへ含めた⁵⁰。また、補完は変数の型に拠らず predicted mean matching を用いた。これは、predictive mean matching は 2 値または 3 値以上のカテゴリ変数にも用いることができること、計算速度が非常に速いことが理由である⁵¹。最後に、得られた結果は Rubin のルール⁵²に従って統合した。

予測モデルの開発と必要な最小サンプルサイズの推定

予測モデルの開発には、Fine & Gray モデルを用いた。追跡期間及びイベント状況を被説明変数、予測因子を説明変数とした予測モデルを作成した。モデルは運用の簡便さを考慮したモデル 1 (年齢と BMI、腹囲、SBP、DBP と質問票の項目の一部) と、モデル 1 に血液検査結果を予測因子として追加し、より高い精度を目指したモデル 2 を検証した。また、既存の骨折リスクの予測モデルとしてアジア地域での妥当性が検証されておりかつ、本研究で利用するデータベースの健診結果からほぼ同様のモデルが再現できる OST^{15,16,53,54} ([体重 kg - 年齢 歳] * 0.2) を参考に、Modified OST ([BMI - 年齢] * 0.2) を比較モデルとして採用した。

なお、予測因子のうち連続変数である年齢⁵⁵、BMI⁵⁶、腹囲⁵⁷、SBP・DBP⁵⁸、FPG⁵⁹、HbA1c⁶⁰、HDL⁶¹、ALT⁶²、γGTP⁶²、Hb⁶³ は骨折や骨粗鬆症と非線形の関連する可能性を示唆する先行研究があること、サンプルサイズとイベント数が十分であり、モデルの過学習の可能性は限定的と判断できることから、Knot 数 3 の固定マージン (開発データの各変数の 10、50、90 パーセンタイル値)の条件で{Hmisc}⁶⁴パッケージの{rcspline.eval}関数を用いて RCS 化した。また、連続変数は日本の標準的な健診・保健指導プログラム（令和 6 年度版）の「別紙 7-3 データ範囲のチェック」⁶⁵に記載されている範囲内に収まっていたため、極端な値をトリミングする、欠測とするなどの処理は行わなかった。モデルに投入した因子の詳細は以下の通りである。

Model 1: 年齢、BMI、腹囲、SBP、DBP、現在喫煙、飲酒頻度、歩行速度低下の自覚の有無、中強度以上の身体活動の実施、低強度の身体活動の実施、週 3 回以上の朝食の欠食、十分な睡眠、血圧を下げる薬の利用、脂質を下げる薬の利用、血糖値を下げる薬の利用、脳血管疾患の既往歴、心臓病の既往歴、行動変容ステージ

Model 2: Model 1 の予測因子 + FPG、HbA1c、HDL、ALT、GGT、Hb

Model 3: Control model; (BMI – 年齢) * 0.2

※Model 3 では既存モデルに沿って RCS 処理はせず、実数を用いた

モデルの当てはめには、{survival}パッケージ^{48,49}の{finegray}関数と{coxph}を用い、重み付き Cox モデルを用い、ハザード比 (HR) の点推定値と 95%CI、傾向の P 値を推定した。

本研究は既存データを用いた研究であるが、モデル開発に必要な最小のサンプルサイズを満たしているかどうかを確認するために、Richard DR らの方法⁶⁶に則り Time-to-event をアウトカムとした多変量予測モデルの開発に必要な最小サンプルサイズを以下の条件で計算した。なお、計算には{pmsampsize}パッケージ⁶⁶を用いた。計算はパラメーター数が最大となるモデル 2 で行ったところ、必要な最小サンプルサイズは最大で 34,905 名であった。

予測モデルの Apparent performance の評価

予測モデルの Apparent performance (“見かけの性能”: モデル開発に使ったオリジナルデータセット [以下、オリジナルデータ] で評価した予測モデルの性能)は、Calibration、Discrimination、臨床的有用性の観点から、以下の方法で評価した。

Calibration は 48 ヶ月時点までの骨折発生リスク (cumulative incident fraction: CIF) を用いて calibration plots で視覚的評価した。Calibration plots の作成は{riskRegression}パッケージ⁶⁷の{Score}関数で行った。本研究では先行研究を参考に、Calibration plots を 30 個の代入済みデータセットを縦型の stacked dataset に作り替え、全てのデータを用いて行った。また、量的な指標として calibration slope を算出した。

Discrimination は 48 ヶ月時点までの骨折発生リスクを用いて、Blanche ら⁶⁸の競合リスクを考慮した Time-dependent area under the curve (AUC) にて評価した。Time-dependent AUC の推定にも

{riskRegression}パッケージ⁶⁷の{Score}関数を用い、30 セットの補完済みデータにおいてそれぞれ推定を行い、結果を統合した。

予測モデルの臨床的有用性は決定曲線解析 (Decision curve analysis: DCA)⁶⁹で評価した。DCA は 4 年以内の骨粗鬆症性骨折を防ぐためにどのような人に骨粗鬆症スクリーニング (DXA など) をすべきかを予測モデルで決めた場合に得られるネット・ベネフィット (Net benefit) を、全員にスクリーニングを行うまたは誰にもスクリーニングを行わない場合の net benefit と比較して予測モデルの有用性を検証する手法である。Net benefit は、真の陽性 (True Positive: TP) の割合から偽陽性 (False Positive: FP) の割合に意思決定の閾値確率 (Probability Threshold: p_i) (4 年間で骨粗鬆症性骨折のリスクが $x\%$ 以上であればスクリーニングによって得られる益がスクリーニングによって生じる可能性がある害を上回ると判断する閾値) の重みを掛けたものを引いた値である ($\text{Net benefit} = \text{TP}/N - [\text{FP}/N] * [p_i / (1 - p_i)]$)。この閾値確率は患者や医師によって変動することが想定されるが、本研究では骨粗鬆症スクリーニングの p_i を 1-3% と設定した。この数値は、主要なガイドラインの 10 年の閾値確率 (骨粗鬆症スクリーニング: 約 3~10%⁷⁰) を期間中のハザードが一定であるとの仮定において、4 年時点における累積発生割合に近似した値として設定した。Net benefit の推定は、先行研究⁷¹⁻⁷³にない打ち切りを Inverse Probability of Censoring Weighting (IPCW) で補正しながら行い、30 セットの補完済みデータにおけるそれぞれの閾値確率ごとの net benefit の平均を算出した。

予測モデルの内的検証

本研究のサンプルサイズは 975,381 名であった。Collins GS らの予測モデルの開発と内的・外的検証に関するガイダンスでは、モデルの Apparent performance は、サンプルサイズが 10,000 を超える場合には overfitting によりモデルの性能が過剰評価されること、optimism に関しては内的検証は必須ではないとされている⁷⁴。しかし、optimism が生じてい

る可能性が完全には否定できないため、single imputation nested bootstrapping^{75,76} による内的検証を行い、optimism を調整したモデルの性能を評価した。Bootstrapping の回数は 200 回辺りから正確な optimism の推定が可能という先行研究⁷⁷の結果と、本研究のサンプルサイズが約 97 万例と非常に大きいことから 200 回に設定した。予測モデルの内的検証は、性能評価指標の内、Calibration slope と Time-dependent AUC、DCA について行った。

C. 結果

JMDC Claims database に登録されていた 20 歳以上の女性 6,809,127 名の内、Cohort entry date (CED) 時点で 40 歳以上 60 歳未満であった女性は 2,226,596 名であった。さらに、CED 時点から過去 2 年間の間に除外基準を満たした女性 71,661 名が除外され、2,154,935 名の女性が抽出された。この内、被扶養者では特定健診のデータの欠測割合が 50%を超えていたため、被保険者である 975,381 名を解析対象とした (Table 1)。

対象者の特性

解析対象者 975,381 名の内、追跡期間中に 5,296 件の骨粗鬆症性骨折が発生し、死亡の発生は 940 件であった。追跡期間の中央値 (Q1, Q3) は 35 (16, 56) ヶ月であった。解析対象者全体の CED 時点の年齢の中央値 (Q1, Q3) は 46.5 (41.6, 51.8) 歳であり、骨粗鬆症性骨折発生者の年齢が 52.8 (48.3, 56.2) 歳と、最も高い傾向にあった。解析対象者全体の BMI は 21.4 (19.5, 24.1) kg/m²、腹囲は 77 (71, 84) cm、現在喫煙者は 16%、アルコール摂取頻度はほとんどないが 50%であった。自覚的な歩行速度が他者よりも速いと回答した女性は 44%、中強度以上の身体活動を実施していたのは 14%、低強度の身体活動を実施していたのは 36%であった。朝食を週 3 回以上欠食している女性は 21%、睡眠で十分な休養を取れていると回答した女性は 54%であった。血圧を下げる薬、脂質を下げる薬、血糖値を下げる薬を利用している女性はそれぞれ 5.9%、1.5%、3.9%であった。脳卒中

の既往がある女性は 0.5%、狭心症か心筋梗塞の既往がある女性は 0.9%であった。行動変容ステージは、23%が改善の意思なし、42%が 6 ヶ月以内に改善の意思なしと回答した。その他、血圧や血液検査の結果の詳細は Table 2 に示した。

開発されたモデル

Table 3 に開発されたモデルのサマリーとして、各モデルに利用された予測変数の HR の点推定値と 95% CI、傾向の P 値を示した。RCS 処理をした連続変数については線形成分と非線形成分に分けて要約値を示した。また、RCS 処理を行う際のオリジナルデータにおけるノットの位置は Table S9 に示した。

Model 1 と 2 において、CED 時点の年齢と BMI の線型成分および非線形成分、腹囲と SBP の線型成分、飲酒頻度 (参照: Hardly ever) の内 Sometimes、週 3 回以上の朝食の欠食あり (参照: No)、睡眠による十分な休養なし (参照: Yes)、脳卒中の既往あり (参照: No)、狭心症または心筋梗塞の既往あり (参照: No)、行動変容ステージ (参照: Maintenance) の内 Preparation が骨粗鬆症性骨折と統計学的に有意な関連を示した。Model 2 において、血液検査結果のうち骨粗鬆症性骨折と統計学的に有意な関連を示したのは HbA1c と Hb の線型成分のみであった。

Model 3 では(BMI [kg/m²] * Age at CED [years]) * 0.2 は骨粗鬆症性骨折と統計学的に有意に関連していた。

モデルの Apparent performance

Discrimination の指標である Time-dependent AUC は Model 2 で最も高かった (0.759, 95% CI 0.758, 0.760)。Model 1 の Time-dependent AUC は Model 2 よりも僅かに低く (0.757, 95% CI 0.756, 0.758)、Model 3 で最も低かった (0.722, 95% CI 0.712, 0.723)。

Calibration の量的指標である、Calibration slope の点推定値は全てのモデルで 1.000 であり、信頼区間も概ね同等であった (Table 4)。各モデルの Calibration plots は Figure 1 に示した。3 つのモデルの 48 ヶ月時点での予測 CIF は概ね実測の CIF と一

致していた。しかし、Model 3 のみわずかではあるが 0-0.75%と 1%以降で実測とのずれが大きい傾向がみられた (Figure 1 (C))。

臨床的優位性を検証した DCA の結果は Figure 2(A)に示した。Model 1 と 2 は p_t 1-2.5%の範囲において、Test none と Test all の戦略よりも net benefit が高かった。Model 3 の net benefit は p_t 1-2.0%の範囲で Test none と Test all よりも高かった。また、Model 1 と 2 の net benefit は p_t 1-2.2%の範囲で Model 3 を上回っていた。

モデルの Optimism-corrected performance

Single imputation nested bootstrapping の結果、各モデルの optimism-corrected Time-dependent AUC は Model 1 で 0.755、Model 2 で 0.756、Model 3 で 0.721 であり、Apparent Time-dependent AUC とほぼ一致していた (Table 4)。Optimism-corrected calibration slope は、Model 1 で 0.992、Model 2 で 0.987、Model 3 で 1.001 であり、こちらも概ね Apparent performance と一致していた (Table 4)。Optimism-corrected DCA curves も同様に、全てのモデルにおいて Apparent DCA curves とほぼ一致していた (Figure 2(B))。

D. 考察

本研究では、全国規模の保険者ベースの診療報酬請求・特定健診データベースに登録された被保険者の女性を対象に、特定健診の結果を用いて骨粗鬆症性骨折を予測するモデルの開発と内的検証を行った。本研究の主要な結果は以下の通りである。①特定健診で測定される項目の内、標準的質問票に含まれる項目と年齢、BMI、腹囲、SBP、DBP のみを用いた Model 1、Model 1 に血液検査項目を加えた Model 2 は、比較モデルとして OST を改変した Model 3 と比べて calibration の程度は同等であった (calibration slope: Model 1 = 0.992; Model 2 = 0.987; Model 3 = 1.001)。また、calibration plot にて実測リスクとのずれが Model 3 でわずかに大きい傾向がみられた。②Model 1 と 2 の discrimination は、Model 3 と比較してわずかではあるが高かった(optimism-corrected Time-

dependent AUC: Model 1 = 0.755; Model 2 = 0.756; Model 3 = 0.721)。さらに、optimism 補正後の DCA においても、設定した p_t 1-3%の内 1-2.2%の範囲で Model 1 と 2 は Model 3 より有用であった。③Model 1 と 2 の性能はほぼ同程度であった。④最後に、内的検証で推定した optimism-corrected performance から、モデルの overfitting を示唆する明確なエビデンスは得られなかった。これらの結果から、本研究の対象となった女性と同等の集団において、本モデルを適用可能である可能性が示唆された。また、Model 1 と 2 の性能は同程度であること、Model 1 の方が予測変数の数が少ないこと、血液検査結果が測定方法によってばらつく可能性があることを考慮すると、Model 1 が実装上有用と考えられた。なお、年齢と BMI のみを利用した OST の改変版である Model 3 も、その他のモデルと比較して極端に性能が劣るというエビデンスは得られず、実装上有用と考えられた。ただし、本研究にて対象となった 40 以上 60 歳未満の女性では骨粗鬆症性骨折の累積発生率は約 0.6%と非常に低かったことを考慮すると、可能な限り性能が高いモデルを利用することが望ましいと考えられる。

本研究で開発したモデルの 4 年時点までの骨粗鬆症性骨折発生予測の性能について、discrimination の指標である optimism-corrected Time-dependent AUC は Model 3 が最小で 0.721、Model 2 が最大で 0.756 であった。Calibration の指標である optimism-corrected calibration slope は Model 1 で 0.992、Model 2 で 0.987、Model 3 で 1.001 であり、良好な calibration を示す 1 に非常に近い値であった。また、calibration plot から Model 3 は Model 1 と 2 に比べてリスクをわずかに過小推定する傾向がみられた。対象者の特性や追跡期間、アウトカムの定義、用いたモデルの違いなどから直接的な比較は困難であるが、日本の 40-74 歳の女性を対象に約 10 年間で骨粗鬆症骨折発生予測の性能を評価した先行研究⁷⁸では、骨密度の情報あり・なしの FRAX の AUC は 0.69 であった。アメリカ在住の 50-64 歳の女性を対象とした同様の先行研究⁷⁹においても、人種がアジアと回答した女性に

層別した時の AUC は骨密度の情報なしの FRAX が 0.65、OST が 0.62 であった。また、同研究内で人種がアジアである女性におけるリスクアセスメントの calibration slope は、骨密度の情報がない FRAX では 1.12、OST では 0.87 であった⁷⁹。ただし、上記の先行研究^{78,79}は既存のリスクアセスメントの外的検証の結果であり、本研究は内的検証の結果であることから、本研究で開発されたモデルの性能は外的検証時に低下⁷⁵することが予測される。しかし、本研究では約 97 万名の大規模データを用いていること、optimism-corrected performance と apparent performance はほぼ一致しておりモデルの overfitting の可能性を示す明確なエビデンスは確認できなかったことから、同等の年齢の女性を対象にモデルの外的検証を行う場合、顕著な性能の低下が起こる可能性は低いと考えられる。これらを踏まえると、本研究で開発された全てのモデルの性能は、既存のリスクアセスメントと比較して明らかに劣っているとは言えず、許容できる範囲である。

本研究では、比較モデルとして設定した Model 3 と比べて、特定健診で用いられる標準的質問票に含まれる項目と年齢、BMI、腹囲、SBP、DBP を用いた Model 1 と Model 1 に血液検査結果を加えた Model 2 は calibration と discrimination、臨床的有用性の観点からわずかに良好な性能を示した。また、興味深いことに、Model 1 に血液検査結果も加えた Model 2 と Model 1 の性能はほとんど差がなかった。初めに、Model 3 に比べて Model 1 と 2 の calibration と discrimination、臨床的有用性が良好だった理由としては、Model 3 では予測変数と骨粗鬆症性骨折の非線形の関連を考慮できていない一方で、Model 1 と 2 では連続変数を RCS 化しており予測変数と骨粗鬆症性骨折の非線形の関連を考慮⁸⁰できていることが考えられる。また、腹囲³⁰や飲酒頻度³⁰、朝食の欠食³¹、脳卒中の既往^{43,44}などをはじめとした骨粗鬆症性骨折のリスク因子を含んでいることも性能の向上に寄与しているものと考えられる。なお、変数の数が Model 3 と比較して Model 1 と 2 では多いことから、モデルの overfitting により Time-dependent AUC が

Model 1 と 2 で高く出た可能性も否定できない。しかし、本研究のサンプルサイズは 97 万名と非常に大きく最小の必要サンプルサイズ (34,905 名) を大きく上回っていること、少なくとも内的検証において optimism を考慮した Time-dependent AUC も apparent Time-dependent AUC と同様の結果であること、overfitting の可能性を示す calibration slope も全てのモデルでほぼ 1 に近い値であることから、その影響は小さいと考える。次に、Model 2 と比較して Model 3 の性能の改善は限定的であった。この理由の一つとして、Model 3 で追加となった血液検査結果は Model 2 に含まれている予測因子に関連しており、血液検査結果の追加による独立した予測情報の増分が限定的であった可能性が考えられる^{81,82}。例えば、糖尿病に関連する FPG や HbA1c は標準的質問票の血糖を下げる薬の利用に関する質問、HDL コレステロールは質問票の脂質を下げる薬の利用に関する質問や BMI、腹囲、Hb は BMI や質問票の週 3 回以上の朝食の欠食などがそれにあたると考えられる。これらから、実装上、性能が高くかつ予測変数の数が少ない Model 1 が適していると考えられた。

モデルの臨床的有用性に関して、optimism 補正後の DCA では設定した pt 1-3% の内 1-2.2% の範囲で Model 1 と 2 は Model 3 より net benefit が高かった。この結果は、主に骨粗鬆症検診や骨粗鬆症スクリーニング (DXA など) の実施を考慮する低い p_t 域における意思決定支援において Model 1 と 2 が Model 3 よりも優れていることを示唆する。すなわち、OST を改変した Model 3 では見逃され得る pt 近傍の患者の中で、骨粗鬆症検診や骨粗鬆症スクリーニングを行うべき女性を適切に選択できる可能性がある。一方で、約 2.2% 以降では Model 3 との net benefit の差はほぼ消失、約 2.5% 以降では各モデルの曲線はほぼ一致し、さらに treat none 戦略との差も小さくなったことから、この閾値域ではモデル使用による追加的な臨床的有用性は限定的であると考えられた。ただし、これは本モデルの臨床的有用性を否定するものではない。むしろ、その有用性は検診・スクリーニングの文脈でこれらの実施を比較的低い閾値で検

討する場面に集中していると解釈される。本研究の対象は、脆弱性骨折歴や長期グルココルチコイド使用などの明らかな高リスク因子を有する女性を除いた40-59歳女性である。したがって本研究の結果は、高リスク患者全体ではなく、明白な高リスク所見を欠く検診・スクリーニング対象の女性に最も関連する。USPSTF¹¹ は65歳未満の閉経後女性を十分に含む新規リスク評価ツールの開発・検証を研究課題として挙げている。本研究はその対象集団と完全には一致しないものの、二次予防的な骨粗鬆症検診・スクリーニングの文脈における低閾値領域におけるさらなる検査の紹介・実施の決定の改善可能性を示す点で、この課題に関連する知見を提供するものと考えられる。

本研究の強みは以下の7点である：①全国規模の保険者ベースの診療報酬・特定健診データベースを用いたことにより悉皆性が高くサンプリングバイアスと over fitting のリスクは低いと考えられていること、②保険者ベースのデータベースであるためアウトカムである骨粗鬆症骨折の発生および競合リスクである死亡による追跡終了時点を比較的正確に測定することができていること、③予測因子に特定健診でルーティーンに取得される項目のみを用いており追加の測定コストが無いこと、④特定健診で取得される情報は全国で標準化が行われていること、⑤競合リスクを考慮した Fine-Gray モデルを用いて予測モデルを開発したこと、⑥リスクアセスメントの開発が望まれる比較的年齢が低い女性を対象としたこと、⑦モデル1で利用している予測因子は年齢、BMI、腹囲、血圧と質問票の項目のみであるため、日本のように特定健診がルーティーンで行われていない国においても容易に実装可能であることである。

一方、本研究では以下の限界点が存在するため、慎重な結果の解釈が必要である。1つ目は、本研究で利用した診療報酬・特定健診データの提供元である保険者は大企業が中心であることである。よって、中小企業に勤める女性を中心とした集団におけるモデル性能の外的検証が必要であると考えられる。2つ目は、本研究では予測因子の欠測割合が高かった被

扶養者の女性が除外されている点であり、被扶養者の女性における外的検証も望まれる。3つ目は、骨粗鬆症骨折の定義が診断コードのみに基づいている点であり、アウトカムの誤分類によるバイアスが生じている可能性が否定できない。この点については、先行研究²⁵⁻²⁷を参考に整形外科専門医との協議の元、高エネルギー骨折である可能性が高い傷病名はアウトカムの定義から除外する、複数部位の骨折の診断コードや交通外傷コードなどが付与されている場合はアウトカムありと判定しないなど、誤分類の可能性を最小限に抑える対応をした。4つ目は、予測期間が一般的なリスクアセスメントの10年ではなく4年であったため、既存の10年リスク指標との直接比較可能性は限定される点である。ただし、これは本モデルの臨床的有用性を直ちに損なうものではない。USPSTF¹¹ は10年リスク値をスクリーニング実施の機械的閾値として用いることを意図しておらず、リスクアセスメントはあくまで追加スクリーニングの要否の判断を補助するものと位置づけている。したがって、本モデルは10年間の骨折リスクを予測するツールの代替としてではなく、検診の受診勧奨やスクリーニング実施の要否の判定を支援する短期的なリスクアセスメントとして活用するのが適切であると考えられる。

これらを踏まえると、今後は中小企業に勤める女性や被扶養者の女性を対象に外的検証を行うこと、リスクアセスメントの性能は人種によって異なる可能性があるため⁷⁹、国外や異なる人種の女性を対象に外的検証を行うこと、診療報酬請求データから骨粗鬆症性骨折を定義する妥当なアルゴリズムを作成すること、10年間での骨粗鬆症性骨折発生リスクや骨粗鬆症の判定に対する性能などを検証すること、モデルの実装が骨粗鬆症検診の受診率、骨粗鬆症の早期発見・治療、ひいては骨粗鬆症性骨折の予防に寄与するかを検証することなどが求められる。

E. 結論

本研究では全国規模の診療報酬請求・特定健診データベースに被保険者として登録されていた40-59

歳の女性を対象に、骨粗鬆症性骨折のリスクアセスメントとして、特定健診で取得される情報を用いて4年間の骨粗鬆症性骨折発生リスクを予測するモデルの開発と内的検証を行った。年齢、BMI、腹囲、血圧、標準的質問票に含まれる生活習慣に関連する質問を利用したモデルは内的検証においても良好な性能を示し、特に骨粗鬆症検診やスクリーニングの実施の必要性を判断するうえで、OSTの改変版モデルや全員検査・全員無検査の戦略よりも有用である可能性が示唆された。

引用文献

1. Xiao PL, Cui AY, Hsu CJ, et al. Global, regional prevalence, and risk factors of osteoporosis according to the World Health Organization diagnostic criteria: a systematic review and meta-analysis. *Osteoporos Int.* 2022;33(10):2137-2153. doi:10.1007/s00198-022-06454-3
2. Chandran M, Brind'Amour K, Fujiwara S, et al. Prevalence of osteoporosis and incidence of related fractures in developed economies in the Asia Pacific region: a systematic review. *Osteoporos Int.* 2023;34(6):1037-1053. doi:10.1007/s00198-022-06657-8
3. Naso CM, Lin SY, Song G, Xue H. Time trend analysis of osteoporosis prevalence among adults 50 years of age and older in the USA, 2005-2018. *Osteoporos Int.* 2025;36(3):547-554. doi:10.1007/s00198-025-07395-3
4. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos Int.* 2006;17(12):1726-1733. doi:10.1007/s00198-006-0172-4
5. LeBoff MS, Greenspan SL, Insogna KL, et al. The clinician's guide to prevention and treatment of osteoporosis. *Osteoporos Int.* 2022;33(10):2049-2102. doi:10.1007/s00198-021-05900-y
6. Japan Osteoporosis Foundation. [Osteoporosis screening rate]. Accessed April 22, 2025. <https://www.jpof.or.jp/research/kenshin.html>
7. McCloskey E, Rath J, Heijmans S, et al. The osteoporosis treatment gap in patients at risk of fracture in European primary care: a multi-country cross-sectional observational study. *Osteoporos Int.* 2021;32(2):251-259. doi:10.1007/s00198-020-05557-z
8. Chotiyarnwong P, McCloskey EV, Harvey NC, et al. Is it time to consider population screening for fracture risk in postmenopausal women? A position paper from the International Osteoporosis Foundation Epidemiology/Quality of Life Working Group. *Arch Osteoporos.* 2022;17(1):87. doi:10.1007/s11657-022-01117-6
9. FRAX plus | FRAXplus®. Accessed April 2, 2025. <https://fraxplus.org/frax-plus>
10. Lo JC, Pressman AR, Chandra M, Ettinger B. Fracture risk tool validation in an integrated healthcare delivery system. *Am J Manag Care.* 2011;17(3):188-194.
11. US Preventive Services Task Force, Nicholson WK, Silverstein M, et al. Screening for Osteoporosis to Prevent Fractures: US Preventive Services Task Force Recommendation Statement. *JAMA.* 2025;333(6):498-508. doi:10.1001/jama.2024.27154
12. Lydick E, Cook K, Turpin J, Melton M, Stine R, Byrnes C. Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density. *Am J Manag Care.* 1998;4(1):37-48.
13. Cadarette SM, Jaglal SB, Kreiger N, McIsaac WJ, Darlington GA, Tu JV. Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. *CMAJ.* 2000;162(9):1289-1294.
14. Sedrine WB, Chevallier T, Zegels B, et al. Development and assessment of the Osteoporosis Index of Risk (OSIRIS) to facilitate selection of women for bone densitometry. *Gynecol Endocrinol.*

- 2002;16(3):245-250.
15. Koh LK, Sedrine WB, Torralba TP, et al. A simple tool to identify asian women at increased risk of osteoporosis. *Osteoporos Int*. 2001;12(8):699-705. doi:10.1007/s001980170070
16. Chen SJ, Chen YJ, Cheng CH, Hwang HF, Chen CY, Lin MR. Comparisons of Different Screening Tools for Identifying Fracture/Osteoporosis Risk Among Community-Dwelling Older People. *Medicine (Baltimore)*. 2016;95(20):e3415. doi:10.1097/MD.0000000000003415
17. FRAX Fracture Risk Assessment Tool. Charts to download. Accessed April 2, 2025. <https://frax.shef.ac.uk/FRAX/charts.aspx>
18. Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Medicine*. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885
19. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-73. doi:10.7326/M14-0698
20. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378
21. Nagai K, Tanaka T, Kodaira N, Kimura S, Takahashi Y, Nakayama T. Data resource profile: JMDC claims database sourced from health insurance societies. *J Gen Fam Med*. 2021;22(3):118-127. doi:10.1002/jgf2.422
22. Kumamaru H, Togo K, Kimura T, et al. Inventory of real-world data sources in Japan: Annual survey conducted by the Japanese Society for Pharmacoepidemiology Task Force. *Pharmacoepidemiol Drug Saf*. 2024;33(1):e5680. doi:10.1002/pds.5680
23. Liu D, Ahmet A, Ward L, et al. A practical guide to the monitoring and management of the complications of systemic corticosteroid therapy. *Allergy Asthma Clin Immunol*. 2013;9(1):30. doi:10.1186/1710-1492-9-30
24. National Center for Biotechnology Information. PubChem Compound Summary for CID 5875, Paramethasone. Accessed August 4, 2025. <https://pubchem.ncbi.nlm.nih.gov/compound/Paramethasone>
25. Lix LM, Azimaee M, Osman BA, et al. Osteoporosis-related fracture case definitions for population-based administrative data. *BMC Public Health*. 2012;12:301. doi:10.1186/1471-2458-12-301
26. Fujiwara S, Buchanan-Hughes A, Ng A, Page J, Adachi K, Li H. Real-world evaluation of osteoporotic fractures using the Japan Medical Data Vision database. *Osteoporos Int*. 2022;33(10):2205-2216. doi:10.1007/s00198-022-06472-1
27. Konstantelos N, Rzepka AM, Burden AM, et al. Fracture definitions in observational osteoporosis drug effects studies that leverage healthcare administrative (claims) data: a scoping review. *Osteoporos Int*. 2022;33(9):1837-1844. doi:10.1007/s00198-022-06395-x
28. 財団法人医療情報システム開発センター. ICD10 対応標準病名マスター Ver. 5.12. June 1, 2023. Accessed May 7, 2025. <https://www2.medis.or.jp/stdcd/byomei/download2019.html>
29. Lane NE, Saag K, O'Neill TJ, et al. Real-world bone turnover marker use: impact on treatment decisions and fracture. *Osteoporos Int*. 2021;32(5):831-840. doi:10.1007/s00198-020-05734-0
30. Nuti R, Brandi ML, Checchia G, et al. Guidelines for the management of osteoporosis and fragility fractures. *Intern Emerg Med*. 2019;14(1):85-102. doi:10.1007/s11739-018-1874-2
31. National Osteoporosis Guideline Group UK. Clinical

- guideline for the prevention and treatment of osteoporosis NOGG 2024. December 2024. Accessed May 1, 2025. <https://www.nogg.org.uk/full-guideline>
32. Coupland CA, Cliffe SJ, Bassey EJ, Grainge MJ, Hosking DJ, Chilvers CE. Habitual physical activity and bone mineral density in postmenopausal women in England. *Int J Epidemiol*. 1999;28(2):241-246. doi:10.1093/ije/28.2.241
 33. Tomita Y, Arima K, Mizukami S, et al. Association between self-reported walking speed and calcaneal stiffness index in postmenopausal Japanese women. *BMC Geriatr*. 2020;20(1):466. doi:10.1186/s12877-020-01858-4
 34. Zhao H, Jia H, Jiang Y, et al. Associations of sleep behaviors and genetic risk with risk of incident osteoporosis: A prospective cohort study of 293,164 participants. *Bone*. 2024;186:117168. doi:10.1016/j.bone.2024.117168
 35. Zhu C, Sun J, Huang Y, Lian Z. Sleep and risk of hip fracture and falls among middle-aged and older Chinese. *Sci Rep*. 2024;14(1):23273. doi:10.1038/s41598-024-74581-4
 36. Ye Z, Lu H, Liu P. Association between essential hypertension and bone mineral density: a systematic review and meta-analysis. *Oncotarget*. 2017;8(40):68916-68927. doi:10.18632/oncotarget.20325
 37. Du XP, Zheng ML, Yang XC, Zheng ML. High blood pressure is associated with increased risk of future fracture, but not vice versa. *Sci Rep*. 2024;14(1):8005. doi:10.1038/s41598-024-58691-7
 38. Cao L, Wu W, Deng X, Guo H, Pu F, Shao Z. Association between total cholesterol and total bone mineral density in US adults: National Health and Nutrition Examination Survey (NHANES), 2011-2018. *J Orthop Surg Res*. 2023;18(1):40. doi:10.1186/s13018-022-03485-8
 39. Chen HM, Hsu CY, Pan BL, et al. Association of Decreased Bone Density and Hyperlipidemia in a Taiwanese Older Adult Population. *J Endocr Soc*. 2024;8(5):bvae035. doi:10.1210/jendso/bvae035
 40. Kim KJ, Hong N, Yu MH, et al. Elevated gamma-glutamyl transpeptidase level is associated with an increased risk of hip fracture in postmenopausal women. *Sci Rep*. 2022;12(1):13947. doi:10.1038/s41598-022-18453-9
 41. Chen G, Chen L, Huang B, et al. Is alanine aminotransferase associated with osteopenia in middle-aged and elderly Chinese? *Endocr Pract*. 2014;20(8):775-784. doi:10.4158/EP13166.OR
 42. Lee EA, Shin DW, Yoo JH, Ko HY, Jeong SM. Anemia and Risk of Fractures in Older Korean Adults: A Nationwide Population-Based Study. *J Bone Miner Res*. 2019;34(6):1049-1057. doi:10.1002/jbmr.3675
 43. Lee D, Cho IY, Chang WH, et al. Fracture Risk Among Stroke Survivors According to Poststroke Disability Status and Stroke Type. *Stroke*. 2024;55(6):1498-1506. doi:10.1161/STROKEAHA.123.044953
 44. Zhang N, Guo L, Yu Y, et al. New-onset stroke on the risk of hip fracture: the Kailuan cohort study in China. *BMC Public Health*. 2023;23(1):925. doi:10.1186/s12889-023-15787-5
 45. Liu XP, Jian XY, Liang DL, et al. The association between heart failure and risk of fractures: Pool analysis comprising 260,410 participants. *Front Cardiovasc Med*. 2022;9:977082. doi:10.3389/fcvm.2022.977082
 46. Broussard DL, Magnus JH. Coronary heart disease risk and bone mineral density among U.S. women and men. *J Womens Health (Larchmt)*. 2008;17(3):479-490. doi:10.1089/jwh.2007.0593
 47. Bartlett JW, Morris TP. Multiple Imputation of Covariates by Substantive-model Compatible Fully Conditional Specification. *The Stata Journal*. 2015;15(2):437-456. doi:10.1177/1536867X1501500206
 48. Therneau TM, Grambsch PM. *Modeling Survival*

- Data: Extending the Cox Model*. Springer; 2000. doi:10.1007/978-1-4757-3294-8
49. Therneau TM. A Package for Survival Analysis in R. R package version 3.8-3. 2024. Accessed December 18, 2025. <https://CRAN.R-project.org/package=survival>
 50. Seaman SR, Bartlett JW, White IR. Multiple imputation of missing covariates with non-linear effects and interactions: an evaluation of statistical methods. *BMC Med Res Methodol*. 2012;12:46. doi:10.1186/1471-2288-12-46
 51. Austin PC, van Buuren S. Imputation of incomplete ordinal and nominal data by predictive mean matching. *Stat Methods Med Res*. 2025;34(11):9622802251362642. doi:10.1177/09622802251362642
 52. Toutenburg H. Rubin, D.B.: Multiple imputation for nonresponse in surveys. *Statistical Papers*. 1990;31(1):180-180. doi:10.1007/BF02924688
 53. Sheng YH, Wu TY, Liaw CK, Hsiao SH, Kuo KL, Tsai CY. Real world fracture prediction of fracture risk assessment tool (FRAX), osteoporosis self-assessment tool for Asians (OSTA) and one-minute osteoporosis risk test: An 11-year longitudinal study. *Bone Reports*. 2024;20:101742. doi:10.1016/j.bonr.2024.101742
 54. Yang Y, Wang B, Fei Q, et al. Validation of an osteoporosis self-assessment tool to identify primary osteoporosis and new osteoporotic vertebral fractures in postmenopausal Chinese women in Beijing. *BMC Musculoskeletal Disorders*. 2013;14(1):271. doi:10.1186/1471-2474-14-271
 55. Ensrud KE. Epidemiology of Fracture Risk With Advancing Age. *J Gerontol A Biol Sci Med Sci*. 2013;68(10):1236-1242. doi:10.1093/gerona/glt092
 56. Nishikura T, Kitamura K, Watanabe Y, et al. Body mass index, height, and osteoporotic fracture risk in community-dwelling Japanese people aged 40–74 years. *J Bone Miner Metab*. 2024;42(1):47-59. doi:10.1007/s00774-023-01478-z
 57. Turcotte AF, Jean S, Morin SN, Mac-Way F, Gagnon C. Added value of waist circumference to body mass index for predicting fracture risk in obesity: a prospective study from the CARTaGENE cohort. *Arch Osteoporos*. 2023;18(1):92. doi:10.1007/s11657-023-01302-1
 58. Jin H, Zhao H, Jin S, et al. Menopause modified the association of blood pressure with osteoporosis among gender: a large-scale cross-sectional study. *Front Public Health*. 2024;12. doi:10.3389/fpubh.2024.1383349
 59. Shieh A, Greendale GA, Cauley JA, Karvonen-Gutierrez CA, Karlamangla AS. Prediabetes and Fracture Risk Among Midlife Women in the Study of Women's Health Across the Nation. *JAMA Netw Open*. 2023;6(5):e2314835. doi:10.1001/jamanetworkopen.2023.14835
 60. Gao L, Liu Y, Li M, Wang Y, Zhang W. Based on HbA1c Analysis: Bone Mineral Density and Osteoporosis Risk in Postmenopausal Female with T2DM. *Journal of Clinical Densitometry*. 2024;27(1):101442. doi:10.1016/j.jocd.2023.101442
 61. Tang Y, Wang S, Yi Q, Xia Y, Geng B. High-density Lipoprotein Cholesterol Is Negatively Correlated with Bone Mineral Density and Has Potential Predictive Value for Bone Loss. *Lipids in Health and Disease*. 2021;20(1):75. doi:10.1186/s12944-021-01497-7
 62. Breitling LP. Liver Enzymes and Bone Mineral Density in the General Population. *J Clin Endocrinol Metab*. 2015;100(10):3832-3840. doi:10.1210/jc.2015-2016
 63. Jaiswal R, Johansson H, Axelsson KF, et al. Hemoglobin Levels Improve Fracture Risk Prediction in Addition to FRAX Clinical Risk Factors and Bone Mineral Density. *J Clin Endocrinol Metab*. 2023;108(12):e1479-e1488. doi:10.1210/clinem/dgad399
 64. Harrell Jr F. Hmisc: Harrell Miscellaneous. Published

- online October 5, 2025. Accessed January 7, 2026. <https://cran.r-project.org/web/packages/Hmisc/index.html>
65. 厚生労働省. 標準的な健診・保健指導プログラム (令和6年度版). April 2024. Accessed November 5, 2025. https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000194155_00004.html
66. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med*. 2019;38(7):1276-1296. doi:10.1002/sim.7992
67. Ozenne B, Sørensen AL, Scheike T, Torp-Pedersen C, Gerds TA. riskRegression: Predicting the Risk of an Event using Cox Regression Models. *The R Journal*. 2017;9(2):440-460. doi:10.32614/RJ-2017-062
68. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. 2013;32(30):5381-5397. doi:10.1002/sim.5958
69. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol*. 2018;74(6):796-804. doi:10.1016/j.eururo.2018.08.038
70. Section 4: Intervention thresholds and strategy. Accessed October 29, 2025. <https://www.nogg.org.uk/full-guideline/section-4-intervention-thresholds-and-strategy>
71. van Geloven N, Giardiello D, Bonneville EF, et al. Validation of prediction models in the presence of competing risks: a guide through modern methods. *BMJ*. 2022;377:e069249. doi:10.1136/bmj-2021-069249
72. Spieker AJ, Illenberger N, Roy JA, Mitra N. Net benefit separation and the determination curve: a probabilistic framework for cost-effectiveness estimation. *Stat Methods Med Res*. 2021;30(5):1306-1319. doi:10.1177/0962280221995972
73. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak*. 2008;8(1):53. doi:10.1186/1472-6947-8-53
74. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ*. 2024;384:e074819. doi:10.1136/bmj-2023-074819
75. Steyerberg EW. Validation of Prediction Models. In: Steyerberg EW, ed. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer International Publishing; 2019:329-344. doi:10.1007/978-3-030-16399-0_17
76. Brand J, van Buuren S, le Cessie S, van den Hout W. Combining multiple imputation and bootstrap in the analysis of cost-effectiveness trial data. *Stat Med*. 2019;38(2):210-220. doi:10.1002/sim.7956
77. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KGM. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol*. 2003;56(5):441-447. doi:10.1016/s0895-4356(03)00047-7
78. Tamaki J, Iki M, Kadowaki E, et al. Fracture risk prediction using FRAX®: a 10-year follow-up survey of the Japanese Population-Based Osteoporosis (JPOS) Cohort Study. *Osteoporos Int*. 2011;22(12):3037-3045. doi:10.1007/s00198-011-1537-x
79. Crandall CJ, Larson JC, Schousboe JT, et al. Race and Ethnicity and Fracture Prediction Among Younger Postmenopausal Women in the Women's Health Initiative Study. *JAMA Intern Med*. 2023;183(7):696-704. doi:10.1001/jamainternmed.2023.1253
80. Harrell FE. General Aspects of Fitting Regression Models. In: Harrell Jr Frank E, ed. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and*

Survival Analysis. Springer International Publishing; 2015:13-44. doi:10.1007/978-3-319-19425-7_2

81. Cook NR. Quantifying the added value of new biomarkers: how and how not. *Diagn Progn Res*. 2018;2(1):14. doi:10.1186/s41512-018-0037-2
82. Hlatky MA, Greenland P, Arnett DK, et al. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation*. 2009;119(17):2408-2416. doi:10.1161/CIRCULATIONAHA.109.192278

F. 研究発表

なし。論文執筆中であり、今後査読付きの学術雑誌に投稿予定。

G. 知的所有権の取得状況

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

Table 1. 対象者の選択の流れ

Process	対象者数	除外された人数
20 歳以上の女性	6,809,127	
追跡開始時に 40 歳以上かつ 60 歳未満の女性	2,226,596	4,582,531
除外基準 (追跡開始日から 2 年前までの期間における情報)		
1: 骨粗鬆症の診断コードを付与されていた女性 (ICD-10: M80 to M82)	2,186,787	39,809
2: 骨粗鬆症性骨折の診断コードを付与されていた女性 (アウトカムと同じ定義)	2,183,240	3,547
3: 骨粗鬆症の治療薬の処方コードが付与されていた女性	2,168,950	14,290
4: 経口グルココルチコイドの処方コードを付与されていた女性 (プレドニゾロン換算で投与量が 5 mg/日の累積処方日数が 90 日以上)	2,163,002	5,948
5: 骨腫瘍か転移の診断コードを付与されていた女性 (ICD10: C40, C41, C79.5, D160, D480)	2,160,092	2,910
6: 骨粗鬆症または骨粗鬆症性骨折のリスクとなる疾患の診断コードを付与されていた女性*	2,159,433	659
7: アロマターゼ阻害薬の処方コードを付与されていた女性	2,154,935	4,498
追加の除外基準		
被扶養者の女性 (理由: 健診データが 50%以上の女性で欠測していたため)	975,381	1,179,554

*: 骨粗鬆症または骨粗鬆症性骨折のリスクとなる疾患は以下のものを含む: osteomalacia (ICD-10: M83); Chronic kidney disease, stage 4 (severe) (ICD-10: N18.4); Chronic kidney disease, stage 5 / End-stage renal disease (ICD-10: N18.5); Dependence on renal dialysis (hemodialysis/peritoneal dialysis) (ICD-10: Z99.2); Transplanted organ status (kidney, liver, heart, lung, etc.) (ICD-10: Z94); and Failure or rejection of transplanted organ (ICD-10: T86).

略語: ICD-10, International Classification of Diseases, 10th Revision.

Table 2. 対象者のベースラインの特性

Characteristic	全体
N	975,381
追跡開始時の年齢, years	46.5 (41.6, 51.8)
生年, n (%)	
1940s	489 (<0.1%)
1950s	33,346 (3.4%)
1960s	322,153 (33%)
1970s	492,611 (51%)
1980s	126,782 (13%)
追跡開始年, n (%)	
2007–2009	4,767 (0.5%)
2010–2014	62,178 (6.4%)
2015–2019	434,037 (44%)
2020–2023	474,399 (49%)
BMI, kg/m ²	21.4 (19.5, 24.1)
Missing	109,851 (11%)
腹囲, cm	77 (71, 84)
Missing	118,572 (12%)
現在喫煙; yes, n (%)	137,809 (16%)
Missing	117,580 (12%)
飲酒頻度, n (%)	
ほとんどない	404,940 (50%)
時々	269,904 (33%)
毎日	133,030 (16%)
Missing	167,507 (17%)
同年代と比較して歩行速度が速い; はい, n (%)	346,369 (44%)
Missing	190,236 (20%)
中強度以上の身体活動の実施; はい, n (%)	111,062 (14%)
Missing	182,347 (19%)
低強度の身体活動の実施; はい, n (%)	286,500 (36%)
Missing	183,933 (19%)
週 3 回以上の朝食の欠食; はい, n (%)	170,045 (21%)

Characteristic	全体
Missing	181,202 (19%)
睡眠で十分な休息をとれている; はい, n (%)	422,812 (54%)
Missing	189,724 (19%)
血圧を下げる薬の利用; はい, n (%)	50,589 (5.9%)
Missing	121,986 (13%)
血糖を下げる薬またはインスリンの利用; はい, n (%)	11,908 (1.5%)
Missing	163,313 (17%)
脂質を下げる薬の利用; はい, n (%)	31,815 (3.9%)
Missing	163,320 (17%)
脳血管疾患の既往 (脳卒中); はい, n (%)	3,918 (0.5%)
Missing	190,528 (20%)
心血管疾患の既往 (狭心症か心筋梗塞); はい, n (%)	7,323 (0.9%)
Missing	190,462 (20%)
行動変容ステージ, n (%)	
維持期	71,283 (9.1%)
実行期	75,414 (9.7%)
準備期	128,323 (16%)
関心期	327,837 (42%)
無関心期	177,581 (23%)
Missing	194,943 (20%)
SBP, mmHg	114 (104, 126)
Missing	110,154 (11%)
DBP, mmHg	70 (63, 79)
Missing	110,154 (11%)
FPG, mg/dL	90 (85, 96)
Missing	207,569 (21%)
HbA1c, % (NGSP)	5.40 (5.20, 5.60)
Missing	182,229 (19%)
HDL cholesterol, mg/dL	71 (60, 82)
Missing	110,989 (11%)
ALT, U/L	14 (11, 19)

Characteristic	全体
Missing	112,082 (11%)
GGT, U/L	17 (13, 25)
Missing	112,304 (12%)
Hb, g/dL	13.2 (12.4, 13.9)
Missing	289,041 (30%)
追跡期間, months	35 (16, 56)

データは中央値 (Q1, Q3) または n (%) で表示した。欠測のある変数については、欠測数 (%) も表示した。

略語: ALT, alanine transaminase; BMI, body mass index; DBP, diastolic blood pressure; FPG, fasting plasma glucose; GGT, gamma-glutamyl transferase; Hb, hemoglobin; HbA1c, hemoglobin A1c % [National Glycohemoglobin Standardization Program: NGSP unit]; HDL, high-density cholesterol; SBP, systolic blood pressure.

Table 3. 予測モデルのサマリー

Predictors	Model 1		Model 2		Model 3	
	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
追跡開始時の年齢, years	1.194 (1.171, 1.217)	<0.001	1.193 (1.170, 1.216)	<0.001	N.A.	N.A.
追跡開始時の年齢, nonlinear component (RCS; 3 knots at 10/50/90%)	0.965 (0.944, 0.986)	0.001	0.968 (0.947, 0.989)	0.004	N.A.	N.A.
BMI, kg/m ²	0.867 (0.842, 0.893)	<0.001	0.872 (0.846, 0.899)	<0.001	N.A.	N.A.
BMI, nonlinear component (RCS; 3 knots at 10/50/90%)	1.141 (1.093, 1.191)	<0.001	1.133 (1.085, 1.183)	<0.001	N.A.	N.A.
腹囲, cm	1.028 (1.017, 1.039)	<0.001	1.028 (1.017, 1.040)	<0.001	N.A.	N.A.
腹囲, nonlinear component (RCS; 3 knots at 10/50/90%)	0.987 (0.973, 1.002)	0.084	0.987 (0.973, 1.001)	0.076	N.A.	N.A.
SBP, mmHg	0.992 (0.986, 0.998)	0.007	0.991 (0.985, 0.997)	0.003	N.A.	N.A.
SBP, nonlinear component (RCS; 3 knots at 10/50/90%)	1.005 (0.999, 1.012)	0.115	1.005 (0.999, 1.012)	0.093	N.A.	N.A.
DBP, mmHg	1.005 (0.996, 1.013)	0.276	1.006 (0.998, 1.015)	0.159	N.A.	N.A.
DBP, nonlinear component (RCS; 3 knots at 10/50/90%)	1.002 (0.992, 1.011)	0.726	1.001 (0.992, 1.011)	0.778	N.A.	N.A.
現在喫煙, はい (ref. いいえ)	0.989 (0.910, 1.074)	0.787	0.999 (0.918, 1.088)	0.983	N.A.	N.A.
飲酒 (ref. ほとんどない)						
時々	1.127 (1.064, 1.193)	<0.001	1.071 (1.007, 1.140)	0.029	N.A.	N.A.
毎日	1.023 (0.970, 1.078)	0.406	1.007 (0.955, 1.062)	0.800	N.A.	N.A.
同年代と比較して歩行速度が速い; いいえ (ref. はい)	0.967 (0.909, 1.029)	0.285	0.963 (0.905, 1.025)	0.234	N.A.	N.A.
中強度以上の身体活動の実施, いいえ (ref. はい)	0.932 (0.851, 1.022)	0.133	0.937 (0.855, 1.027)	0.162	N.A.	N.A.
低強度の身体活動の実施, いいえ (ref. はい)	1.004 (0.940, 1.073)	0.899	1.008 (0.944, 1.077)	0.803	N.A.	N.A.
週3回以上の朝食の欠食, はい (ref. いいえ)	1.209 (1.117, 1.309)	<0.001	1.200 (1.108, 1.300)	<0.001	N.A.	N.A.
睡眠で十分な休息をとれている, いいえ (ref. はい)	1.091 (1.026, 1.160)	0.006	1.088 (1.023, 1.157)	0.007	N.A.	N.A.
血圧を下げる薬の利用, いいえ (ref. はい)	0.961 (0.861, 1.071)	0.469	0.972 (0.871, 1.084)	0.609	N.A.	N.A.
血糖を下げる薬またはインスリンの利用, いいえ (ref. はい)	0.888 (0.714, 1.104)	0.284	0.888 (0.689, 1.145)	0.359	N.A.	N.A.
脂質を下げる薬の利用, いいえ (ref. はい)	0.995 (0.865, 1.144)	0.944	0.992 (0.863, 1.139)	0.904	N.A.	N.A.
脳血管疾患の既往 (脳卒中), はい (ref. いいえ)	1.732 (1.284, 2.335)	<0.001	1.726 (1.279, 2.329)	<0.001	N.A.	N.A.
心血管疾患の既往 (狭心症か心筋梗塞), はい (ref. いいえ)	1.082 (0.818, 1.433)	0.579	1.080 (0.816, 1.429)	0.589	N.A.	N.A.
行動変容ステージ (ref. 維持期)						
実行期	0.985 (0.899, 1.078)	0.736	0.980 (0.895, 1.073)	0.662	N.A.	N.A.
準備期	0.911 (0.841, 0.986)	0.021	0.910 (0.841, 0.986)	0.021	N.A.	N.A.

関心期	1.017 (0.946, 1.094)	0.644	1.018 (0.947, 1.095)	0.623	N.A.	N.A.
無関心期	0.992 (0.921, 1.068)	0.827	0.991 (0.920, 1.068)	0.817	N.A.	N.A.
FPG, mg/dL	N.A.	N.A.	0.998 (0.991, 1.006)	0.636	N.A.	N.A.
FPG, nonlinear component (RCS; 3 knots at 10/50/90%)	N.A.	N.A.	1.003 (0.996, 1.010)	0.378	N.A.	N.A.
HbA1c, %	N.A.	N.A.	0.784 (0.632, 0.973)	0.027	N.A.	N.A.
HbA1c, nonlinear component (RCS; 3 knots at 10/50/90%)	N.A.	N.A.	1.139 (0.946, 1.373)	0.169	N.A.	N.A.
HDL cholesterol, mg/dL	N.A.	N.A.	0.997 (0.992, 1.001)	0.161	N.A.	N.A.
HDL cholesterol, nonlinear component (RCS; 3 knots at 10/50/90%)	N.A.	N.A.	1.006 (1.001, 1.011)	0.011	N.A.	N.A.
ALT, U/L	N.A.	N.A.	1.012 (0.997, 1.026)	0.107	N.A.	N.A.
ALT, nonlinear component (RCS; 3 knots at 10/50/90%)	N.A.	N.A.	0.987 (0.969, 1.005)	0.149	N.A.	N.A.
GGT, U/L	N.A.	N.A.	1.005 (0.997, 1.014)	0.243	N.A.	N.A.
GGT, nonlinear component (RCS; 3 knots at 10/50/90%)	N.A.	N.A.	0.993 (0.978, 1.009)	0.395	N.A.	N.A.
Hb, g/dL	N.A.	N.A.	0.951 (0.912, 0.992)	0.018	N.A.	N.A.
Hb, nonlinear component (RCS; 3 knots at 10/50/90%)	N.A.	N.A.	0.995 (0.951, 1.042)	0.836	N.A.	N.A.
(BMI [kg/m ²] * 追跡開始時の年齢 [years]) * 0.2	N.A.	N.A.	N.A.	N.A.	0.560 (0.548, 0.573)	<0.001

Hmisc::rcspline.eval を用いて RCS 化（ノット数：3 knots，ノット位置：the 10th, 50th, and 90th percentiles）を行った。変数ごとのノットの値は Table S13 に示した。新たな個人で予測を行う場合、RCS 化処理を行う必要がある。

以下の数値と数式を用いて計算を行うことで、48 ヶ月時点での絶対リスクの予測を行うことが可能である。本研究では、モデルごとのベースラインの累積サブディストリビューションハザード（`baschaz(..., centered=TRUE)`）を推定した。 $\text{eta_raw} = \sum(\text{beta_j} * \text{x_j})$ とする。オリジナルデータセットにおける、ベースラインの累積サブディストリビューションハザードと線形予測子の平均をそれぞれ $H0_48$ ， lp_mean とすると、以下の計算式で骨折リスクを計算できる：

$\text{pred_risk_48} = 1 - \exp(-H0_48 * \exp(\text{eta_raw} - lp_mean))$ 。

なお本研究では、多重補完を行っているため、 $H0_48$ と lp_mean は補完後の各データセットの平均値を報告している。

Model 1: $H0_48 = 0.00476388858359615$, $lp_mean = 6.81441039245127$ 。

Model 2: $H0_48 = 0.00470394168810565$, $lp_mean = 4.96422683263494$ 。

Model 3: $H0_48 = 0.00446502522938077$, $lp_mean = 2.92318397903768$ 。

略語: ALT, alanine transaminase; AST, aspartate aminotransferase; BMI, body mass index; CI, confidence interval; DBP, diastolic blood pressure; FPG, fasting plasma glucose; GGT, gamma-glutamyl transferase; Hb, hemoglobin; HbA1c, hemoglobin A1c; HR, hazard ratio; HDL, high-density lipoprotein; LDL, low-density lipoprotein; N.A., not applicable; NGSP, National Glycohemoglobin Standardization Program; RCS, restricted cubic splines; SBP, systolic blood pressure.

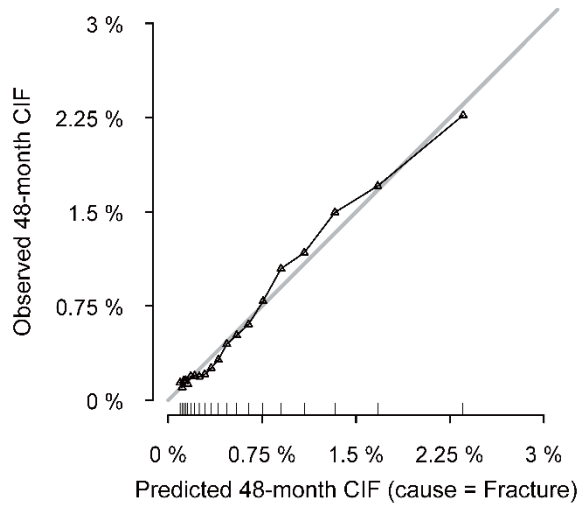
Table 4. 予測モデルの見せかけの性能および Optimism を考慮した性能

モデル	識別: Time-dependent AUC		較正: Calibration slope	
	見かけの性能 (95% CI)	Optimism 補正後の性能	見かけの性能 (95% CI)	Optimism 補正後の性能
モデル 1	0.757 (0.756, 0.758)	0.755	1.000 (0.968, 1.032)	0.992
モデル 2	0.759 (0.758, 0.760)	0.756	1.000 (0.969, 1.031)	0.987
モデル 3	0.722 (0.712, 0.723)	0.721	1.000 (0.963, 1.037)	1.001

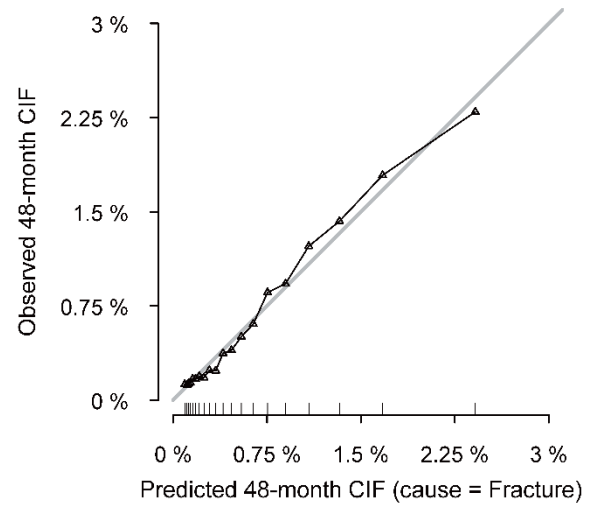
すべての性能指標は、競合リスクを考慮した上で、対象イベント（原因＝脆弱性骨折）について 48 か月時点で評価した。Time-dependent AUC は、IPCW 法に基づいて推定した。Calibration slope は、線形予測子を共変量とした Weighted Fine-Gray calibration model により推定した（slope = 1 は理想的な校正を示す）。見かけの性能は、元のコホートにおいて多重代入（m=30）を用い、Rubin のルールを適用して推定した。Optimism は、各ブートストラップサンプル内で単一代入（M=1）を行うネステッド・ブートストラップ法（B=200）により推定した。各ブートストラップ反復において、ブートストラップサンプルで見かけの性能を算出し、元のコホートでテスト性能を算出した。Optimism は、ブートストラップサンプルでの見かけの性能とテスト性能の差として定義した。Optimism 補正後の性能は、元の見かけの推定値から Optimism の平均を差し引くことで求めた。

略語: AUC, area under the curve; CI, confidence interval; IPCW, inverse probability of censoring weighting.

(A) Model 1



(B) Model 2



(C) Model 3

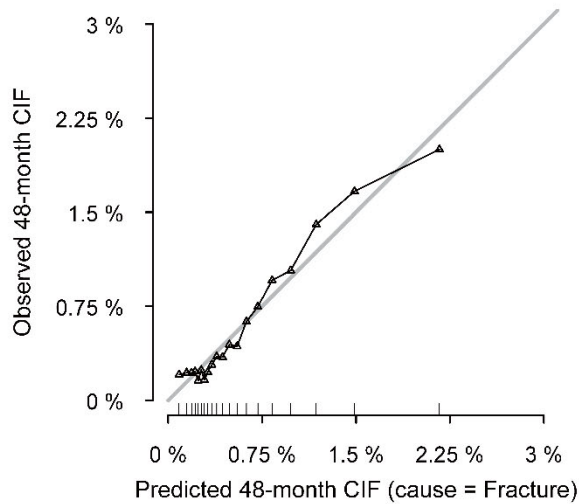
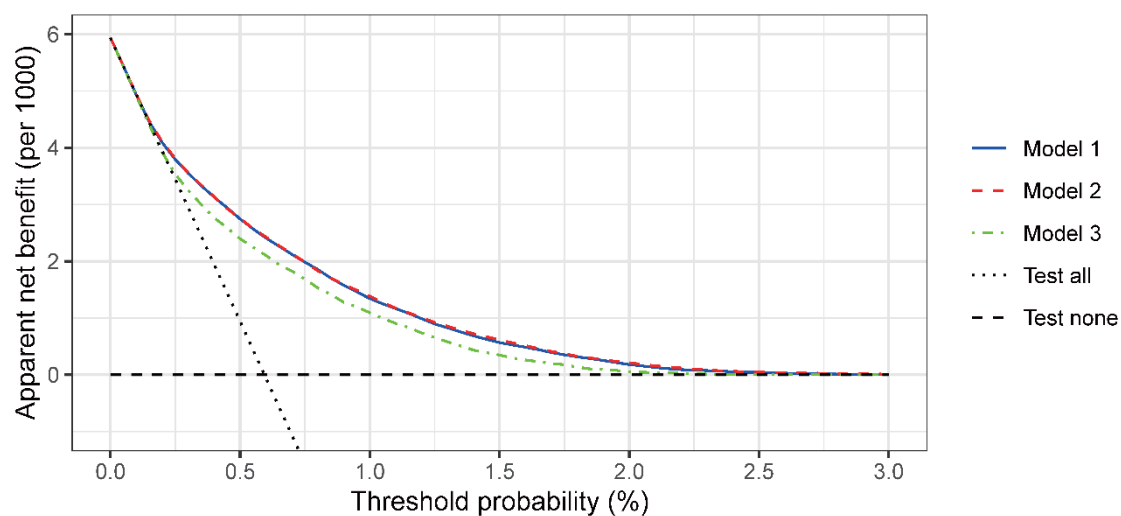


Figure 1. 開発したモデルで4年間の骨粗鬆症性骨折リスクを予測した際の calibration plots

Calibration plots は多重補完後の 30 セットのオリジナルデータを縦型に連結し、1 つのデータセットとして作成した。横軸は予測された4年間の骨粗鬆症性骨折のリスクを示し、縦軸は実際のリスクを示す。斜め45度の線は予測リスクと実測リスクが完全に一致している状態を示し、この線から上部に逸脱すると予測モデルはリスクを過小評価していること、下部に逸脱するとリスクを過大評価していることを示す。横軸上部にある縦線は、予測リスクを5パーセンタイル刻みでビン分けした分位点を示している。

(A) Apparent performance



(B) Corrected performance

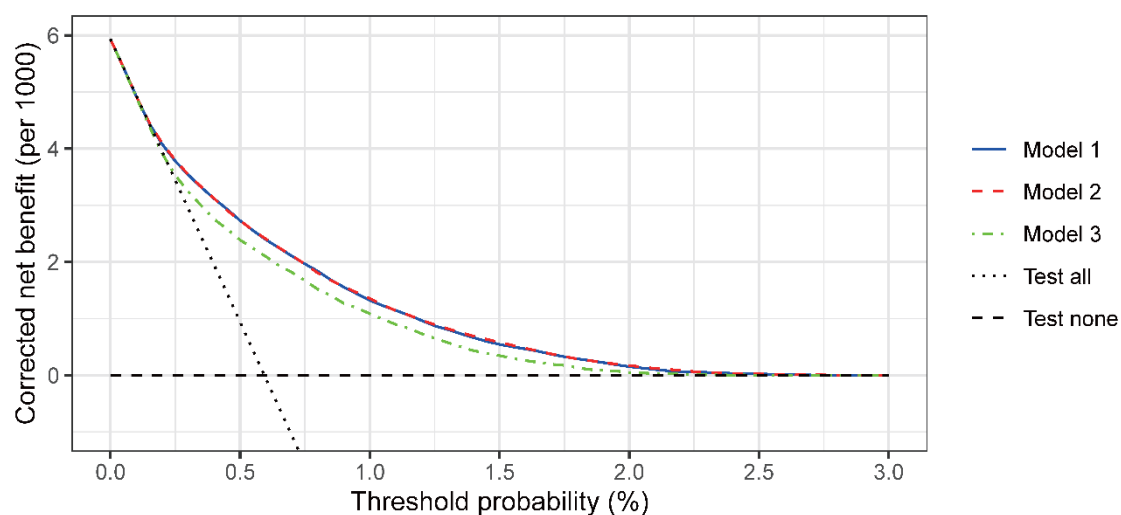


Figure 2. 開発したモデルを用いて 4 年間の骨粗鬆症性骨折発生リスクを予測した際の臨床的有用性

横軸は閾値確率, 縦軸はパネル A が見かけのネット・ベネフィットを, パネル B は Optimism を補正したネット・ベネフィットを示す. 青線がモデル 1、赤線がモデル 2、緑線がモデル 3、黒の点線が Test all (全員検査)、黒の点線が Test none (全員検査をしない) 戦略を示す. 本研究では整形外科医との協議の元, 事前に閾値確率を 1–3%と定めている. 当該閾値内で, ある曲線が上部に位置する場合, その曲線にあたる戦略が他の戦略よりもネット・ベネフィットが高い (臨床的有用性が高い) と判断する.

※日本の標準的な健診・保健指導プログラム (令和 6 年度版) の「別紙 7-3 データ範囲のチェック」⁶⁸より引用