

TRIPOD+AI Expanded チェックリスト 有志翻訳による日本語版 version.1.1

2025年3月19日作成

原文：Collins G S, Moons K G M, Dhiman P, Riley R D, Beam A L, Van Calster B 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

訳者：佐藤俊太郎¹，中島誉也²，後藤隆之介³，後藤匡啓^{4,5}

所属：

1. 長崎大学病院臨床研究センター
2. 長崎大学大学院医歯薬学総合研究科麻酔集中治療医学
3. Department of Biomedical Data Science, Stanford University School of Medicine
4. 横浜市立大学学術院医学群データサイエンス研究科ヘルスデータサイエンス専攻
5. TXP Medical 株式会社

著作権：CC BY 4.0

本文内の「適用」におけるD/Eは下記参照

D: Development、モデル開発

E: Evaluation、モデル検証

より適切な翻訳があれば、佐藤（shuntarosato@nagasaki-u.ac.jp）までご連絡ください。

TITLE (タイトル)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|---|---|-----|
| 1 | D;E | 研究が多変数(多変量)を用いた予測モデルの開発もしくは性能評価の研究であることを明確にし、対象となる集団と予測対象となるアウトカムを明示すること。 | <ul style="list-style-type: none"> - 情報量の多いタイトルは潜在的な読者や系統的レビューをおこなう研究者にも役立つ。 - 対象集団と予測されるアウトカムについて、重要な情報を含む分かりやすいタイトルにする。 | |

ABSTRACT (要旨)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|--------------------------------|--|-----|
| 2 | D;E | TRIPOD+AI のための要旨チェックリスト参照すること。 | <ul style="list-style-type: none"> - TRIPOD+AI のための要旨チェックリストの各項目に対応する要旨を報告する。 | |

INTRODUCTION (緒言)

Background (背景)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|--|--|-----|
| 3a | D;E | 既存のモデルを参照しながら、医療に関する背景（診断か予後予測かどうかを含む）と予測モデルを開発または評価する根拠・必要性を説明すること。 | <ul style="list-style-type: none"> - 予測モデルが使用される、または必要とされる医療環境について説明する。 - 既存の予測モデルが利用可能な場合は、新しいモデルを開発する明確な根拠を示す。 - 既存のモデルを評価する研究では、評価をおこなう根拠を示し、評価対象となるすべてのモデルの元論文を参考文献として示す。 | |

| | | | | |
|----|-----|---|--|--|
| 3b | D;E | 予測モデルが診療プロセスのどの場面で使われるのかを説明し、対象集団、目的、想定されるモデルの使用者（医療専門家、患者、一般市民など）を明確にすること。 | <ul style="list-style-type: none"> - 開発または評価されたモデルの対象集団について説明する。たとえば、どのような年齢の人々か、どの国の人々か、または特定の疾病を持つ人々か。 - モデルが意図している目的を説明する。モデルがどのような臨床的な判断や指針を補助することを意図しているか（例：追加検査への紹介や入院、トリアージ、治療の開始や生活環境の変更）、および診療プロセスのどの時点でモデルを使用するか。 - 想定している使用者が誰なのか、そしてそのモデルが医療専門家、患者、一般市民、あるいはその他の人向けなのか説明する。 | |
| 3c | D;E | 社会人口統計学的グループ間の既知の健康格差を明示すること。 | <ul style="list-style-type: none"> - モデルが使用されると想定している医療環境において、対象となる集団における社会人口統計学的グループ間の既知の健康格差を明示する（その健康格差を裏付ける引用もつけること）。 | |

Objectives (目的)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|---|--|-----|
| 4 | D;E | 研究の目的を具体的に述べること。これには、研究が予測モデルの開発なのか検証なのか（あるいはその両方なのか）を含む。 | <ul style="list-style-type: none"> - 研究のすべての目的を明示的に述べ、その研究が予測モデルの開発なのか、予測モデルの性能評価なのか、あるいはその両方なのか説明する。 | |

| METHODS (方法) | | | | |
|--------------|-----|---|---|-----|
| Data (データ) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 5a | D;E | 開発用と評価用のデータセットについて、それぞれのデータソース（例：無作為化試験、コホート、日常診療データ、レジストリデータ）、およびそれらのデータを使う理由と代表性について説明すること。 | <ul style="list-style-type: none"> - モデル開発とモデル性能評価に使用されたデータソースについて説明すること。これには無作為化試験、コホート研究、電子的な日常診療記録など、どのようなソースからデータが得られたかも含む。 - 研究が既存のデータを使用しているのか、予測モデル研究の目的のために新たにデータを前向きに収集しているのかを明記する。 - 既存のデータを使用する場合（元々は異なる目的でデータが集められていた場合）、それらのデータを使用する根拠を示し、そのデータが適切な理由（特にデータが想定される対象集団と異なる状況や国から得られた場合）、および想定される対象集団とその集団の背景に関するデータの代表性についてコメントする。 - データソースの説明は、すべてのデータセットについて、さらに開発用と評価用それぞれに対しておこなう。 - 人工合成データが使われた場合はその理由を説明し、人工合成データがどのように作成されたか（コードについては項目 18f を参照）、研究でどのように使用されたか説明する。 | |
| 5b | D;E | 参加者データの対象期間となる日時を明らかにすること。これには、参加者の募集開始日と終了日、該当する場合は追跡期間の終了日も含む。 | <ul style="list-style-type: none"> - 参加者あるいは使われたデータが選択された期間の開始日と終了日を明らかにする。 - 予後予測モデルの場合は追跡期間が重要であるため、追跡調査の終了日を報告する。 | |

| Participants (参加者) | | | | |
|---------------------------|-----|--|---|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 6a | D;E | 研究の設定における重要な要素（例：プライマリケア、セカンダリケア、一般集団）を明らかにすること。これには、施設の数や場所も含む。 | <ul style="list-style-type: none"> 研究が実施された医療環境（例：一次医療、三次医療、外来、入院、あるいは施設の特徴や医療機関の種類）を説明し、研究の参加者がどこから募集されたか記載する。 研究がおこなわれた地理的な位置（少なくとも国）と施設（施設数を含む）を報告する。 | |
| 6b | D;E | 研究参加者の適格基準を説明すること。 | <ul style="list-style-type: none"> 参加者の適格基準は、予測モデルの適用可能性や一般化可能性を理解するために重要。 適格基準には、連続変数の範囲についての報告も含む。たとえば、対象となる参加者の適格性を定義するために使用した年齢範囲など。 | |
| Data preparation (データの準備) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 7 | D;E | データの前処理と品質チェックについて説明すること。これには、関連する社会人口統計学的グループ間で同様の処理がおこなわれているかどうかも含む。 | <ul style="list-style-type: none"> 特徴量エンジニアリング、生データの変換、特徴量の削減、およびデータの品質チェックなどを含むデータクリーニングの手順について説明する。データクリーニングに用いたすべてのコードは公開する（項目 18f 参照）。 複数のソースのデータを使用する分析の場合（例：異なる研究データセット、コホート、レジストリ）、アウトカムや予測因子の標準化、統一化について説明する（訳註：同じ因子か、定義が統一されているか）。 すべてのデータ前処理/データクリーニングの手順が主な社会人口統計学的グループ間で同様であったかどうかを確認する（該当する場合）。 データの前処理/データクリーニングの手順が広範囲に及ぶ場合は補足資料で報告する。 | |

| Outcome (アウトカム) | | | | |
|-----------------|-----|--|--|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 8a | D;E | 予測対象となるアウトカムと予測の時間的範囲を明確に定義すること。これには、アウトカムの評価方法と時期、このアウトカムを選択した根拠、およびアウトカムの評価方法が社会人口統計学的グループ間で一貫しているかどうかを含む。 | <ul style="list-style-type: none"> - 診断予測モデルの場合、アウトカムを明確に定義する。アウトカムの有無を判定する際に、（広く受け入れられている）参照基準（ground truth）を使用したかどうかも含む。 - 予後予測モデル（将来のアウトカムを予測するモデル）の場合、著者はアウトカム予測の時間的範囲を報告する。たとえば、心臓胸部手術後 28 日以内の死亡リスクの予測や、骨粗鬆症患者における 10 年間の骨折リスクの予測などである。また、追跡期間中のアウトカム評価の頻度も報告する。 - ICD コードなどの一般的な定義を使用する場合は、それを明記し、参考文献を示す。 - 社会人口統計学的グループ間でアウトカム評価に差異がある場合は報告する。 - 場合によっては、予測因子がアウトカムの定義に使用されていないこと、またはアウトカムの代替指標となっていないことを確認する。 | |
| 8b | D;E | アウトカムの評価に主観的な解釈が必要な場合は、評価者の資格と人口統計学的特徴について説明すること。 | <ul style="list-style-type: none"> - 主観的な解釈を必要とするアウトカム（例：画像検査の結果の解釈）については、評価者の人数および資格、人口統計学的特徴について説明する。 - アウトカムの測定と解釈に（追加の）トレーニングや特別な指示が必要な場合は、それらについて報告する。 - 内容が広範な場合は補足資料で報告する。 | |
| 8c | D;E | 予測するアウトカムの評価における、盲検化について報告すること。 | <ul style="list-style-type: none"> - 予測されるアウトカムは、予測因子に関する情報を知らない状態で評価されるべきである。特に主観的な解釈を必要とするアウトカムにおいて重要で、データ（を識別するための情報）の漏洩（リーク）を防ぐためである。 - 著者はアウトカム評価者がどのような情報にアクセスできたかを説明し、アウトカム評価の盲検化のために取られた具体的な対策について報告する（該当する場合）。 | |

| Predictors (予測因子) | | | | |
|-------------------|-----|---|--|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 9a | D | はじめに候補として考えた予測因子の選択方法（例：文献、過去のモデル、利用可能なすべての予測因子）と、モデル構築前の予測因子の事前選択について説明すること。 | <ul style="list-style-type: none"> - はじめに候補とした予測因子をどのように決めたかを詳述する。それらの予測因子が文献の（系統的）レビュー、臨床的な観点（分野の専門家）に基づいて選択されたのか、あるいは、単に利用可能なデータ内のすべての予測因子を使用したのかを説明する。 - モデル構築前に予測因子の事前選択をおこなった場合は、その方法を詳述する。たとえば、欠測データが多いため予測因子から除外したのか、あるいは予測対象のアウトカムとの（臨床的な）関連性が妥当でないため除外したのかなどを説明する。 - はじめに候補とした予測因子のリストが広範な場合、補足資料での報告を推奨する。 | |
| 9b | D;E | すべての予測因子を明確に定義し、測定方法や時期（およびアウトカムや他の予測因子に対してある予測因子の評価をどのように盲検化したか）を含めて説明すること。 | <ul style="list-style-type: none"> - モデリングに含まれるすべての予測因子は、測定単位およびカテゴリカルな予測因子のすべてのカテゴリとともに明確に定義する。これにより読者や他の研究者がモデルを再現、実装、または性能を評価することができる。 - 予測因子の値がどのように、いつ測定されたのか詳述する。予測因子は、モデルの使用が意図される時点またはそれ以前に測定されるべきであることに注意すること。 - 主観的な解釈を必要とする予測因子については、モデリングで考慮される他の予測因子の値に対して盲検的に解釈することが重要な場合がある（データの漏洩を回避するなど）。著者は、予測因子の測定評価を他の予測因子に対して盲検化するための対策について報告する必要がある。 - 特に診断モデルの場合、予測因子とアウトカムとの関連性が人為的に強化される可能性があるため、個人のアウトカムを知ることなく予測因子の測定をする必要がある。著者は、予測因子の測定評価をアウトカムの値に対して盲検化するための措置について報告する必要がある。 - 予測因子の数が多すぎる場合、すべてを論文本文に記載すると煩雑になるため、補足資料で報告することを検討する。この際においても、予測因子を全て明確に定義する重要性は変わらない。 | |

| | | | | |
|----|-----|---|--|--|
| 9c | D;E | 予測因子の測定に主観的な解釈が必要な場合は、予測因子の評価者の資格と人口統計学的特徴について説明すること。 | <ul style="list-style-type: none"> - 主観的な解釈を必要とする予測因子（例：画像検査の結果の解釈）については、予測因子の評価者の資格と人口統計学的特徴を報告する必要がある。 - 測定と解釈に（追加の）トレーニングや特別な説明が必要な場合は、それらを報告する必要がある。これらは補足資料で報告することができる。 | |
|----|-----|---|--|--|

Sample size (サンプルサイズ)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|--|---|-----|
| 10 | D;E | 研究の規模（サンプルサイズや施設数など）がどのように決定されたかを（開発と評価とでわけて）説明し、その規模が研究疑問に答えるのに十分であることを示す。サンプルサイズ計算の詳細を含めること。 | <ul style="list-style-type: none"> - サンプルサイズの決定方法を説明する。モデル開発に必要なサンプルサイズと、モデルの性能評価に必要なサンプルサイズについて別々に説明する必要がある。ここで、データが前向きに集まっているか、既存のデータを使用しているかは関係ない。 - サンプルサイズ計算に使用されたすべての推定値と詳細を示す。 - 統計学的なサンプルサイズ計算をおこなわなかった場合（例：利用可能なすべてのデータを使用した場合）、データの規模が研究疑問に答えるのに十分であったかどうかの妥当性を説明する。 | |

Missing data (欠測データ)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|--|--|-----|
| 11 | D;E | 欠測データをどのように対処したのか説明すること。データを除外した場合は、その理由を説明すること。 | <ul style="list-style-type: none"> - 欠測データはあらゆる所でみられる問題である。著者は、モデルへ組み込むことを検討している各予測因子について、欠測の数を報告する必要がある。 - 欠測値の取り扱いについて、欠測の理由に関する仮定を含めて報告する必要がある。 - 欠測値により個人（または予測因子）が除外された場合は、その旨を報告し、理由を説明する必要がある。 - 欠測値が補完された場合は、欠測値を補完するための方法の詳細をすべて報告する必要がある。 - 欠測値を補完した場合は、訓練データとテストデータで別々におこなわれたことを確認すること。データ漏洩を避けるためである。 | |

| Analytical methods (分析方法) | | | | |
|---------------------------|----|---|--|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 12a | D | データが分析でどのように使用されたか（例：モデルの開発とモデルの性能評価のため）を説明すること。必要なサンプルサイズを考慮しながら、データの分割について説明すること。 | <ul style="list-style-type: none"> - モデルの開発と性能評価のために、利用可能なデータがどのように使用されたかを説明する。これには、データを分割したかどうか、分割方法、およびデータを分割した理由を含む（例：モデル開発、ハイパーパラメータの調整、モデル性能の評価、内部-外部交差検証）。 - データを分割した場合、分割時にサンプルサイズの要件（項目 10 参照）を考慮したかどうか、および分割後のデータサイズが分析を実行し、研究疑問に答えるのに十分であるか報告する。 - データを訓練データ（ハイパーパラメータの調整を含む）とテストデータに分割した場合、データ漏洩が発生していないことを確認すること。 - データに同一個人からの複数のレコードやサンプルが含まれ、データを訓練データ（ハイパーパラメータの調整を含む）とテストデータに分割した場合は、分割されたデータ間で個人の漏洩が生じていないことを確認すること。もし漏洩がある場合は、分析でどのように対処したか説明する（項目 12c 参照）。 | |
| 12b | D | モデルの種類に応じて、予測因子が分析でどのように扱われたかを説明すること（関数形、再スケーリング、変換、または標準化）。 | <ul style="list-style-type: none"> - 分析時に変換された予測因子（再スケーリングや標準化）については、その方法を説明する。 - カテゴリカルな予測因子において、カテゴリーの統合をおこなった場合（例：サンプルサイズが小さい/アウトカムのイベントが少なすぎる）は、その詳細と理由を示す。 | |

| | | | | |
|-----|---|---|--|--|
| 12c | D | <p>モデルの種類、根拠[すべてのモデル構築アプローチで分けて]、ハイパーパラメータ調整を含むすべてのモデル構築手順、および内部検証の方法を明記すること。</p> | <ul style="list-style-type: none"> - 開発しているモデルの種類（ロジスティック回帰、Cox 回帰、ランダムフォレスト、ニューラルネットワークなど）を明確にし、それぞれのモデル構築方法を使う根拠を示す。予測されるアウトカムの種類と、予測モデルが実際にどのように実装されるかを考慮すること。 - 各モデルについて、ハイパーパラメータ調整を含むモデル構築のすべての手順を明確に説明し、どのハイパーパラメータが調整され、どのようにおこなわれたかを記述する。多くのモデル構築アプローチが適用され、文字数制限で完全な説明ができない場合は、補足資料を使用して詳細を示すこと。 - 複数のモデルを開発する研究（異なるモデル構築方法を使用するなど）の場合、（もしあれば）どのモデルを提案するかを選択する基準を明確に説明する（項目 12e および項目 23 のモデルの性能評価を参照）。 - モデル開発時の内部検証アプローチ（モデル性能を評価するため）を明確に説明する。たとえば、k 分割交差検証またはブートストラップが使用されたかどうかなどである。内部評価の方法時にすべてのモデル構築手順（ハイパーパラメータ調整を含む）が再現されたかどうかを明確にすること。 - モデルの安定性（予測因子の選択、予測性能、個別の予測の観点から）を調べるために使用された方法（ブートストラップなど）を明確に説明すること（Riley & Collins, <i>Biom J</i> 2023; 65: 2200302 [DOI: 10.1002/bimj.202200302]）。 - データに同一個人からの複数のレコードまたはサンプルが含まれている場合、モデル構築と内部検証でこれらがどのように扱われたかを説明する（例：k 分割交差検証が使用された場合、個人のすべてのレコード/サンプルが同じ分割に含まれたことを確認する（データ漏洩を避けるなど））。 | |
|-----|---|---|--|--|

| | | | | |
|-----|-----|--|---|--|
| 12d | D;E | <p>クラスター（例：病院、国）間でのモデルパラメータ値とモデル性能の推定値の不均一性が考慮され、定量化された場合、その方法を説明すること。追加の検討事項は TRIPOD-Cluster を参照せよ。</p> | <ul style="list-style-type: none"> - モデル開発時、またはモデル性能の評価時にデータのクラスター（例：複数の研究からの個別参加者データの組み合わせ、医療施設/病院別、または国別にクラスター化されたデータ）を考慮した分析をおこなった場合、クラスターを考慮した根拠と方法を明確に説明する。 - クラスターとモデルパラメータ値および性能の不均一性を考慮した予測モデル研究の具体的な報告推奨事項については、TRIPOD-Cluster チェックリストを参照すべきである（Debray et al, BMJ 2023; 380: e071018 [DOI: 10.1136/bmj-2022-071018]）。 | |
| 12e | D;E | <p>モデル性能（例：識別能、較正、臨床的有用性）を評価し、複数のモデルを比較するために使用したすべての指標と図を（その根拠とともに）明記すること。</p> | <ul style="list-style-type: none"> - モデル性能を評価するために使用したすべての指標を報告する。最低限、モデルの識別能と較正（較正プロットを含む）を提示することが一般的に期待される。 - 予測モデルが生存時間（time-to-event）アウトカムを予測する場合、生存時間アウトカムの特徴（つまり、打ち切り）を考慮するために使用された指標と方法を明確に説明する。同様に、競合リスクの取り扱いも記述すべきである（該当する場合）。 - 予後モデルの場合、モデルの予測性能が評価されたすべての時点で報告する。 - 較正プロット（滑らかな較正曲線を含む）や意思決定曲線など、モデル性能の可視化に使用した方法を報告する。 - 複数のモデルを比較する場合、つまり既存のモデルと比較したり複数のモデリングアプローチを比較したりする場合、これらのモデルを比較するために使用した方法と、どのような基準で優劣を判断したのかを明確にする。 | |
| 12f | E | <p>モデル評価から生じたモデルの更新（例：再較正）について、全体的なものか、特定の社会人口統計学的グループまたは特定の環境に関するものかを記述すること。</p> | <ul style="list-style-type: none"> - 検証後にモデルが更新された場合（再調整や再度のあてはめなど）、それが全コホートに対してか、特定の社会人口統計学的グループに対してかにかかわらず、モデルの更新に使用した方法の詳細を示す。 | |

| | | | | |
|-----|---|--|--|--|
| 12g | E | <p>モデル評価において、モデルの予測がどのように計算されたか（例：数式、コード、オブジェクト、API）を説明すること。</p> | <ul style="list-style-type: none"> - 別のデータセットで既存のモデルを評価する研究（つまり、外部検証の研究）の場合、モデルから個々の予測値がどのように算出されたか詳細を示す。モデルが自由に/公的に利用できない場合は、予測値がどのように得られたかを説明する。 - 回帰モデルの数式を評価する場合は、その回帰式の詳細を示す（例：回帰式を提示する。式を開発した元の研究の引用を提示する）。 - 式のない予測モデル（例：ニューラルネットワーク、ランダムフォレスト）を評価する研究の場合、予測がどのようにおこなわれたかの詳細を示す。たとえば、コード、ソフトウェアオブジェクト、API など、およびそれらがどこで見つけられるか（つまり URL[uniform resource locator の略]や DOI[digital object identifier の略]）を示す。 - モデルからの個人予測が新たなリスクグループや分類の作成に使用された場合（モデル開発時に特定されていなかったもの）、どのように分類をおこなったか、その理由と詳細を報告する必要がある（項目 15 を参照）。 | |
|-----|---|--|--|--|

Class imbalance (クラス不均衡)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|---|--|-----|
| 13 | D;E | <p>クラス不均衡に対処する方法を使用した場合は、その理由と方法、続けて、モデルやモデル予測を再較正するために使用した方法を説明すること。</p> | <ul style="list-style-type: none"> - クラス不均衡に対処する方法（例：アンダー/オーバーサンプリング、SMOTE[synthetic minority oversampling technique の略]）を使用した場合は、その使用した根拠と実施方法を示す。その際、サンプルサイズへの影響（例：アンダーサンプリング法の場合）を考慮する。 - 不均衡の補正はモデルの較正に影響を与え（van den Goorbergh et al, J Am Med Inform Assoc 2022; 29: 1525--1534 [DOI: 10.1093/jamia/ocac093]）、確率推定値を過大にする傾向にある（これはリスクグループの定義にも影響を与える）。モデルやモデル予測を再較正するために使用した方法について説明する。 | |

| Fairness (公平性) | | | | |
|-----------------------|-----|---|---|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 14 | D;E | モデルの公平性に対処するために使用したアプローチとその根拠を説明すること。 | <ul style="list-style-type: none"> - 公平性とは、予測モデルが人種、性別、年齢などの個人的属性に基づいて個人やグループに対する差別をしないことを保証することを指す。公平性に対処するために使用されたすべてのアプローチは、その根拠とともに明確に説明する必要がある。 - モデルの開発とその性能の評価において、データが（標的集団の）代表的なグループを含んでいることを確認することが重要であり、研究者はこれを実証するよう努めるべきである。 - 予測モデルが十分に代表されていないグループのデータを使用して開発された場合、または特定のグループが含まれていない場合、これらのグループにおけるモデルを評価するためには、代表的なデータでのこれらのグループにおける評価が必要である。これは開発および評価データに含まれる個人のグループを超えて、より多くのグループへの一般化可能性を高めるためである。 | |
| Model output (モデルの出力) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 15 | D | 予測モデルの出力（例：確率、分類）を特定する。あらゆる分類について、またそれらの閾値がどのように決めたかについての詳細を示すこと。 | <ul style="list-style-type: none"> - ほとんどのモデルは個人に対する確率の推定値を出力するが、一部のモデルは出力を分類（例：低リスクまたは高リスクグループ）に変換する。これは明確に述べるべきである。分類やリスクグループが作成された場合、治療経路の文脈においてそれをおこなう根拠と、これらのリスクグループが臨床判断にどのように影響するかを説明する必要がある。 - 分類やリスクグループを出力するモデルについては、これを明確に報告し、グループを定義する推定確率の範囲などの閾値を指定する必要がある（これらが文献、臨床ガイドライン、統計的検討事項、またはアドホックに基づいているかどうかにかかわらず）。 - 個々の予測モデルの出力に対して不確実性の区間が提示された場合は、その実装方法の詳細を示す（例：パラメータ推定値の分散共分散行列または conformal prediction を使用）。 | |

| Development versus evaluation (モデルの開発と評価) | | | | |
|---|-----|---|---|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 16 | D;E | 医療環境、適格基準、アウトカム、および予測因子において、開発データと評価データの間にある違いを特定すること。 | <ul style="list-style-type: none"> ある環境、施設、または国で開発された予測モデルは、必ずしも異なる環境、施設、または国で有用とは限らない。異なるソースのデータ間では、適格基準、アウトカムおよび予測因子の定義が（意図的に）異なる可能性がある。開発データとモデル性能の評価に使用されたデータとの違いを説明することは、元のモデル開発データの文脈におけるモデルの性能と一般化可能性を理解し解釈するのに役立つ。 | |
| Ethical approval (倫理承認) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 17 | D;E | 研究を承認した治験審査委員会または倫理委員会の名称を示し、参加者のインフォームドコンセントまたは、倫理委員会によるインフォームドコンセントの免除について説明すること。 | <ul style="list-style-type: none"> 研究に治験審査委員会または倫理委員会の承認がない場合は、承認がないこととその理由を明確に述べる。 | |
| OPEN SCIENCE (オープンサイエンス) | | | | |
| Funding (研究資金) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 18a | D;E | 研究資金の出所と資金提供者の研究における役割を示すこと。 | <ul style="list-style-type: none"> 研究が資金提供を受けたかどうかの詳細と、研究における資金提供者の役割に関する詳細を記載する。 すべての著者に対する追加的な資金源を記載する。 | |

| Conflicts of interest (利益相反) | | | | |
|------------------------------|-----|---|---|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 18b | D;E | すべての著者の利益相反と財務状況を開示すること。 | <ul style="list-style-type: none"> 著者の人間関係や活動のうち、読者が研究に関連すると考える可能性があるもの、または研究の設計、実施、解釈、報告に影響を与えた可能性があるものを開示する。 | |
| Protocol (プロトコル) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 18c | D;E | 研究のプロトコルを入手できる場所を示すか、プロトコルが準備されていない場合はその旨を述べること。 | <ul style="list-style-type: none"> プロトコルの入手可能性に関するすべての詳細を記載する。これには、プロトコルの保管場所（例：出版情報、補足資料、Open Science Framework などのリポジトリ）、URL または DOI を含む。 研究のプロトコルが作成されていない、または公開されていない場合は、（その理由とともに）明確に述べる。 研究のプロトコルで明記したものからの重大な逸脱がある場合は、その概要と理由を記載する。 | |
| Registration (登録情報) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 18d | D;E | 研究の登録情報（登録名と登録番号を含む）を示すか、研究が登録されていない場合はその旨を述べること。 | <ul style="list-style-type: none"> 研究が登録されている場合（例：clinicaltrials.gov、Open Science Framework）、登録番号、登録先の名前、登録先へのリンク（DOI を含む）に関する詳細を記載する。 研究が登録されていない場合は、その旨を明記する。 | |

| Data sharing (データの共有) | | | | |
|-----------------------|-----|---|---|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 18e | D;E | 研究データの利用可能性に関する詳細を示すこと。 | <ul style="list-style-type: none"> - 研究データの利用可能性に関する詳細を示す。これには、データの保管場所（例：公開リポジトリ、URL、DOI）、データの取得方法、データの取得と使用に関する条件や制限事項が含まれる。共有されるデータには、データ仕様書を添付する必要がある。 - データを共有できない場合は、その理由を記載する。 - 「正当な要求があれば利用可能」などの決まり文句を避ける。何が正当な要求であるかの条件を指定せずに、この文言は使用しないこと。 | |
| Code sharing (コードの共有) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 18f | D;E | 分析コードの利用可能性に関する詳細を示すこと[分析コードには、データクリーニング、特徴量エンジニアリング、モデル開発、モデル評価を含む]。 | <ul style="list-style-type: none"> - 分析に用いたコード（およびコードの実行方法に関する資料）の利用可能性に関するすべての詳細を示す。これには、コードの場所（例：コードリポジトリ、DOI、リンク）、コードの取得方法、コードの取得と使用に関する条件やライセンスを報告する必要がある（バージョンも含む）。 - 分析コードとは、研究の報告された結果と知見をすべて（原則として）再現するために必要なすべてのコードである（データクリーニングのためのコードを含む）。研究の知見を（原則として）再現するために必要なソフトウェアとパッケージは、（バージョン番号を含めて）報告される必要がある。場合によっては、コンピューティング環境に関するより詳細な情報を報告する必要がある場合がある（例：ハードウェア、オペレーティングシステム、CPU、RAM）。 | |

| PATIENT & PUBLIC INVOLVEMENT (患者と市民参画) | | | | |
|--|-----|---|--|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 19 | D;E | 研究の設計、実施、報告、解釈、または普及において患者や市民がどのように関与したかを詳述するか、関与がなかった場合はその旨を述べること。 | <ul style="list-style-type: none"> - 患者や市民が研究の計画、設計、実施、報告、または知見の普及にどのように関与したかを説明する。 - 研究の知見は患者や市民に提示されたかどうかを確認する。 - 研究における患者と市民の関与を報告するために GRIPP2 声明を使用することを検討する (Staniszewska et al, BMJ 2017; j3453 [DOI: 10.1136/bmj.j3453]) 。 - 研究のいかなる側面においても患者や市民が関与しなかった場合は、その旨を明確に述べる。 | |
| RESULTS (結果) | | | | |
| Participants (研究参加者) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 20a | D;E | 研究の流れを説明し、アウトカムイベントの有無ごとの参加者数や、該当する場合は追跡期間の要約を示す (図を用いるとわかりやすい) 。 | <ul style="list-style-type: none"> - フロー図は研究を通じての参加者の流れを説明するのに役立つ。フロー図の始点となるのは参加者の募集元 (病院、コホートなどの集団) で、その後の流れは適格基準、追跡調査 (該当する場合)、データの利用可能性に関連する。 - フロー図に表示するその他の有用な情報には、欠測値のある参加者の数、およびアウトカムイベントの数が含まれる。 - 予後研究や診断研究において、参照基準の評価が遅れておこなわれる場合、追跡期間の要約を報告する必要がある (例: 中央値、範囲) 。 | |

| | | | | |
|-----|-----|--|--|--|
| 20b | D;E | 表を用いるなどして、全体的な特徴と、該当する場合は各データソースまたは設定ごとの特徴を報告すること。これには、重要な日付、重要な予測因子（人口統計学的特性を含む）、受けた治療の内容、サンプルサイズ、アウトカムイベントの数、追跡期間、および欠測データの数を含む。主な人口統計学的グループ間の違いを報告すること。 | <ul style="list-style-type: none"> - 使用されたすべてのデータセットの要約を、アウトカム、予測因子（例：平均/中央値、標準偏差/四分位範囲、頻度）の分布、受けた治療、サンプルサイズ（およびアウトカムイベントの数）、追跡期間の要約、および各予測因子の欠測値の数と割合を含めて、表を使用するなどして報告する。 - 必要に応じて、主な人口統計学的グループ間での違いを報告する。 | |
| 20c | E | モデルの評価において、開発データとの比較を示し、重要な予測因子の分布（人口統計学的特性、予測因子、アウトカム）を比較する。 | <ul style="list-style-type: none"> - 既存のモデルの性能を評価する研究（モデル開発研究の一部としておこなわれるものも含む）では、重要とされる変数の分布（平均/中央値、標準偏差/四分位範囲、頻度など）について比較を示す。これには人口統計学的特性、モデル内の予測因子、アウトカム、そして欠測値の割合も含む。これらの情報は表形式で示すのが最も良く、また「疾患あり/なし」などのアウトカムの状態別に分けて報告するのが望ましい。 | |

Model development (モデルの開発)

| 項目 | 適用 | 内容 | 具体的説明 | ページ |
|----|-----|---|--|-----|
| 21 | D;E | 各分析（例：モデル開発、ハイパーパラメータチューニング、モデル評価）における参加者とアウトカムイベントの数を明確に述べること。 | <ul style="list-style-type: none"> - サンプルサイズ（アウトカムイベントの数を含む）は、各分析（つまり、各モデル開発、各モデル評価）について報告されるべきである。予測モデル研究（例：データ分割後、モデルハイパーパラメータチューニング後）のさまざまな分析間でしばしば異なる可能性があり、特に欠測データがある場合はそうなる可能性が高い。 - データに同一個人からの複数のサンプルまたはレコードが含まれている場合は、個人の人数も報告する。 | |

| Model specification (モデルの仕様) | | | | |
|------------------------------|----|--|---|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 22 | D | <p>新たな個人での予測を可能にし、第三者による評価と実装を可能にするために、完全な予測モデル（例：予測式、コード、オブジェクト、API）の詳細を記載すること。アクセスや再利用に対する制限（例：自由に利用可能、所有権）も含めること。</p> | <ul style="list-style-type: none"> - 予測モデル開発研究の「成果物」は予測モデルである。したがって、モデルの詳細と、新たな個人の予測を可能にする使用方法に関する詳細を示すことが重要である。たとえば、回帰モデルの式を示すか、モデルが式として「書き下せない」方法を使用して開発された場合は、コード、ソフトウェアオブジェクト、またはAPIの利用可能性に関する詳細を示し、他者が自身のデータでこのモデルを評価したり、日常診療で実装したりできるようにする。 - モデルの使用方法を説明して、他者が自身のデータでこのモデルを評価したり、日常診療で実装したりできるようにする。 - 複数のモデルが開発された場合は、すべてのモデルの利用可能性に関する詳細を示す。 - モデルの使用方法を説明して他者が新たな個人に対する予測を行えるようにする。 - 第三者のテスト、実装、モニタリングを可能にするためのハードウェア要件、ソフトウェア（およびパッケージ）の詳細を示す。 - モデルを公開できない場合（例：商業的理由のため）、これを明確に報告し、新たな個人の予測計算を可能にし、第三者による評価を可能にするためのモデルへのアクセス条件を報告する必要がある。 | |

| Model performance (モデルの性能) | | | | |
|----------------------------|-----|--|--|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 23a | D;E | <p>主要なサブグループ（例：社会・人口統計学的特性）を含む、モデルの性能の推定値（信頼区間含む）を報告すること。図表での可視化を検討すること。</p> | <ul style="list-style-type: none"> - 項目 12e で説明されたすべてのモデル性能指標の推定値は、信頼区間とともに提示されるべきである。 - 全体の集団、および関心（例：公平性チェックの一環としての関心）のある主要なグループ（例：性別、民族）に対するモデル性能の推定値を信頼区間とともに報告する。 - 可視化することでモデルの評価を補助する。たとえば、較正プロット（滑らかな較正曲線と予測値の分布を含む）や意思決定曲線などである。 - 実施されたすべての評価（例：開発データでの評価、評価データでの評価、内部検証プロセスからの評価など）の性能の推定値を報告し、検討された各時点での評価（予後モデルの場合）も含める。 - モデルの安定性に関する検討結果を報告する。たとえば、ブートストラップサンプルで開発されたモデル間での性能の推定値や、個々の予測の変動性などについてである（Riley & Collins, Biom J 2023; 65: 2200302 [DOI: 10.1002/bimj.202200302]）。 - 性能の推定値を提示するために使用されたデータを明確に示す。 | |
| 23b | D;E | <p>クラスター間のモデル性能の不均一性を評価した場合は、その結果を報告すること。追加の詳細については TRIPOD Cluster を参照。</p> | <ul style="list-style-type: none"> - モデル性能の評価がデータのクラスター（例：複数の研究からの個別参加者データの組み合わせ、または施設/病院、または国によってクラスター化されたデータ）を考慮した場合、信頼区間とともに結果を報告する必要がある（項目 23a 参照）。 - クラスターとモデル性能の不均一性を考慮した予測モデル研究の具体的な報告推奨事項については、TRIPOD-Cluster チェックリストを参照すべきである（Debray et al, BMJ 2023; 380: e071018 [DOI: 10.1136/bmj-2022-071018]）。 | |

| Model updating (モデルの更新) | | | | |
|-------------------------|-----|--|--|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 24 | E | モデル更新の結果を、更新されたモデルとその後の性能を含めて報告すること。 | <ul style="list-style-type: none"> - 検証後に予測モデルが更新された場合（例：再較正、再度のあてはめ）、第三者による評価と実装を可能にするために、更新された予測モデルの詳細を報告する必要がある（アクセスや再利用に対する制限を含む）（項目 22 参照）。 - 更新されたモデルの性能は報告する必要がある（項目 23a、必要あれば 23b を参照）。 | |
| DISCUSSION (考察) | | | | |
| Interpretation (解釈) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 25 | D;E | 主な結果の全体的な解釈を提示すること。これには目的と先行研究の文脈における公平性の問題を含める。 | <ul style="list-style-type: none"> - 研究結果を解釈する際は、それを他の研究結果と比較して意味づけする必要がある。既存のモデルがある場合は、それら先行研究の文脈において知見を議論する。 - 既存の予測モデルの性能を評価する研究では、もし他の研究がすでにそのモデルの性能を評価している場合、それらの知見を議論・要約し、研究の文脈の中に位置づけることが重要である。 - モデルの過大解釈や「スピン (spin)」を防ぐため、知見の解釈がモデルの開発と評価から報告された知見を越えないことを確認すること。 - 読者が、評価データでのモデルの性能が他の評価研究とどのように比較されるかを理解できるようにする。結果が異なる場合、モデル性能の差異の考えられる理由について考察する。 | |

| Limitations (研究の限界) | | | | |
|---|-----|---|--|-----|
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 26 | D;E | 研究のあらゆる限界（例：代表的でない標本、サンプルサイズ、過剰適合、欠測データ）とそれらがバイアス、統計的不確実性、一般化可能性に与える影響について議論すること。 | <ul style="list-style-type: none"> 研究限界を認識することは科学論文の重要な側面であり、研究デザイン、実施、分析のあらゆる側面に言及できる。分析に用いたデータの代表性、サンプルサイズ、過剰適合、欠測データ/データ品質に関する懸念を考慮した、研究の限界についての意義のある議論をおこなう。 | |
| Usability of the model in the context of current care (日常診療におけるモデルの有用性) | | | | |
| 項目 | 適用 | 内容 | 具体的説明 | ページ |
| 27a | D | 日常診療におけるモデルの有用性：予測モデルを実装する際に、品質の低い入力データや利用できない入力データをどのように評価し処理すべきかを説明すること。 | <ul style="list-style-type: none"> モデルが日常診療の一部として使用される際に、利用できない予測因子の値をどのように扱うべきかについて言及する。モデルが実際に使用される時点で欠測値を補完するための戦略も評価する（したがって、論文中の方法と結果の節でも言及する）。 同様に、実装の時点における低品質な入力データ（例：画像解像度、データ形式）の処理について議論する（該当する場合）。 | |
| 27b | D | 入力データの取り扱いやモデルの使用において、利用者が関わる必要があるかどうか、また利用者にとどの程度の専門知識が求められるかについて議論すること。 | <ul style="list-style-type: none"> 予測モデルが意図したとおりに使用されるために、利用者がどのようにモデルと関わるかが期待または求められているかについて詳細を示す。たとえば、入力データの処理に関する考慮事項など。 モデルの使用、入力データの処理や収集に専門知識やトレーニングが必要とされるか、必要とされる場合は、その詳細を示す。 | |
| 27c | D;E | モデルの適用可能性と一般化可能性に特に焦点を当てた、将来の研究のための次のステップについて議論すること。 | <ul style="list-style-type: none"> モデルのさらなる評価が必要か（例：異なる集団やサブグループでの評価）、あるいはモデルは臨床試験での評価、または診療プロセスの一部としての実装の準備ができているか。 | |