

口演 18 特定健康診査の質問票を使用したフレイル状態の判別

ひらみつよしみち

○平光良充（名古屋市衛生研究所）、小嶋雅代（名古屋市健康福祉局）

【要旨】

特定健康診査の質問票を使用して、フレイル該当者を判別するためのモデルの作成を試みた。その結果、実際の業務で使用できるレベルの判別能力を有するモデルを作成することができた。

【背景と目的】

“フレイル”とは、高齢期になり心身の活力が低下し、介護を必要とするリスクが高まった状態をいう。2020年度に本市保険年金課が実施した調査によれば、本市後期高齢者では38.5%がフレイルに該当し、前期高齢者でも12.5%がフレイルの状態であった。したがって、前期高齢者の段階からフレイル対策を行うことが必要である。わが国において、一般集団におけるフレイル評価には、基本チェックリスト（KCL）の質問票が広く用いられている。本市では、毎年多くの前期高齢者が特定健康診査（以下、特定健診）を受診している。特定健診とは、生活習慣病予防を目的として40～74歳を対象に実施される健診のことで、俗に“メタボ健診”とも呼ばれる。特定健診では、主に生活習慣病のリスク評価を目的とした質問票が使用されている。本研究では、この特定健診質問票を使用し、KCLで評価したフレイルを判別できるか検討することを目的とした。

【方法】

1. 使用したデータ

本市保険年金課が2021年度に実施した『フレイル予防に関する調査』の回答データ及び2021年度特定健診のデータを個人単位で突合して分析に使用した。『フレイル予防に関する調査』の質問項目にはKCLが含まれていた。本研究では、両データが突合でき、かつ分析に使用する変数に欠損がなかった1,697人を分析対象者とした。

2. 使用した変数

従属変数はKCL得点に基づくフレイル判定結果とした。フレイル判定は、KCL得点が8点以上を「フレイル」、0～7点を「非フレイル」と定義した。また、独立変数は、性別、年齢、特定健診質問票（22項目）、及びフレイルとの関連が予測される体重減少幅及びBMIとした。

3. 統計解析

判別モデルを単一のデータで作成すると、そのデータだけに“過剰適合”したモデルが作成される危険がある。“過剰適合”したモデルは、作成時に使用したデータに対しては高い判別能力を発揮できるが、他のデータに対しては判別能力が極端に低下するため、実用的でないモデルとなる。そのため、一般的に機械学習の分野では、“過剰適合”への対策として、作成したモデルを他のデータにも当てはめ、モデルの判別能力を確認する作業が行われる。そこで、本研究でも以下の手順で判別モデルを作成した。①まず、分析対象者1,697人を70：30の比率でTraining data（判別モデルを作成するためのデータ）とTest data（作成したモデルの判別能力を確認するためのデータ）に無作為分割した。②次いで、Training dataを使用して、KCL得点に基づくフレイル判定結果を従属変数とした二項ロジスティック回帰（変数増減法）により判別モデルを作成した。最適カットオフ値にはYouden Indexを使用した。③最後に、作成した判別モデルにTest dataを当てはめ、判別モデルが“過剰適合”していないか確認した。

モデルの判別能力の指標には、感度、特異度、ROC曲線下面積（AUC）を使用した。なお、感度は「実際に問題がある人を“問題あり”と正しく判別できる確率」、特異度は「実際には問題のない人を誤って“問題あり”と判別しない確率」、AUCは「全般的な判別能力（モデルの正確さ）」を示す指標である。感度、特異度及びAUCは、いずれも0～1の範囲の値をとり、値が1に近いほどモデルの判別能力が高いことを示す。一般に、AUCが0.7以上の場合は“使用可能なレベルのモデル”と解釈される。

4. 倫理的配慮

本研究は、名古屋市衛生研究所等疫学倫理審査委員会の承認を得て行われた（受付番号46）。

【結果】

1. Training dataによる判別モデルの作成

二項ロジスティック回帰分析（変数増減法）を行ったところ、「ほとんど噛めない」「同年齢と比べて歩行速度が遅い」「睡眠で休養が十分でない」「体重前年比3kg以上減少」など13項目の変数で構成される判別モデルが作成された。判別モデルのAUCは0.774であり、Youden Indexに基づく最適カットオフ値は0.177であった。0.177をカットオフ値とした場合は、感度0.733、特異度0.679であった。

2. Test dataに判別モデルを適応した結果

判別モデルにTest dataを適用した場合のAUCは0.790であった。また、前述の最適カットオフ値（0.177）で判別した場合は、感度0.782、特異度0.676であった。

【考察】

Training dataとTest dataとで感度、特異度、AUCに大きな差はみられなかった。したがって、今回作成した判別モデルは、“過剰適合”が起きておらず、異なるデータ（今回使用したデータとは異なる年度の特定健診結果など）を適用した場合でも概ね同水準の判別能力を発揮できると考えられる。また、Training data、Test dataともにAUCが0.7以上であったことから、今回作成した判別モデルは実際の業務で使用できるレベルの判別能力を有していると考えられる。