

# M1-24 疫学データを用いた主成分分析に基づく骨折リスクの要因抽出 ～骨折リスクマップによる視覚化と簡易診断手法の提案～

知能システム制御研究室 河越 祐太

## 1. はじめに

近年、医学の進歩と保健衛生環境の充実により、日本の平均寿命は伸びつつある。その一方で、自立した生活を過ごせる期間である健康寿命との差が広がっている。厚生労働省の調べによると、高齢者が要支援者になる原因として、骨折・転倒は 14.2%、関節疾患は 18.9% を占める。一方、要支援者が要介護者になる原因として、骨折・転倒は 12.0%、関節疾患は 6.9% を占める [1]。また、骨折が要介護の直接的な原因でないとしても、骨折の治療中に運動機能が低下することから、多くの場合で要介護になる。そのため、高齢者が骨折リスクの上昇を早期に気づき、保健行動につなげて予防することが重要である。

現在、骨折を日常生活の中で簡易的に診断する方法の 1 つに FRAX<sup>®</sup> (Fracture Risk Assessment Tool) が存在するが、数値で骨折リスクを評価しているため、高齢者に対して骨折リスクの重要度が伝わりにくく、骨折に対する意識改善に繋がっていない [2]。そのため、骨折に対するリスク意識向上を伝える新たな指標が必要である。

そこで本研究では、高齢者が骨折リスクの上昇を早期に自覚し、保健行動につなげられるようにするため、健康診断などの疫学データに基づき骨折リスクを予測し、分布マップ (予測マップ) を構築する。予測マップは、骨折リスクを可視化できるという利点がある。

## 2. 疫学データの収集条件と前処理

実験協力者は 136 名 (女性)、年齢は 50~74 歳 ( $65.94 \pm 6.79$ )、収集した疫学データの項目数は 397 項目である。なお、136 名の実験協力者の内 28 名は疫学データ収集後 1 年以内に転倒した事実がある。今回は、骨折の危険因子である転倒を骨折の分類基準とする。また、「項目内の分散が 0 となる項目」、「重複する項目」、「1 項目内で 100 名以上の値に欠損がある項目」、「アンケート (主観) に該当する項目」については分析精度を向上させるため除外した。

## 3. PCA を用いた骨折リスクマップ

### 3.1 主成分に基づく要因抽出

まず、主成分分析 (以下、PCA) を行い、多次元の疫学データを合成変数として縮約し新たな意味づけを有した軸 (以下、主成分) を複数作る。PCA を行うことで疫学データの可視化ができる。疫学データ

を転倒事実の有無に分け、各々 PCA を行った結果、転倒事実の有無のいずれにおいても第 1~3 主成分の寄与率合計は 50% 以上と過半数を超えていた。そのため、疫学データ全体の変化を第 3 主成分までで概ね表現可能である。

つぎに、主成分負荷量 (主成分と疫学データとの相関関係) に着目して、骨折リスクの要因を抽出を行った。主成分負荷量の絶対値が 0.7 以上の項目に着目すると、転倒事実がある実験協力者において、第 1 主成分では脂肪に関する項目、第 2 主成分では握力や筋肉に関する項目、第 3 主成分では歩行速度 [m/s] や歩行周期時間 [s] といった歩行能力に関する項目が見られた。同様に、転倒事実がない実験協力者において、第 1 主成分では脂肪や筋肉に関する項目、第 3 主成分では両脚支持時間 [s] や荷重応答時間 [s] といった歩行能力に関する項目が見られた。第 2 主成分には主成分負荷量絶対 0.7 以上となる疫学データは見られなかった。

さらに、各主成分ごとに抽出した疫学データに対し、転倒事実の有無の相互で共通する項目を除外する。共通する項目を除外すると、転倒事実がある実験協力者の疫学データは第 1 主成分で 4 項目 (全身筋肉量スコア、左足筋肉量スコア、右足筋肉量スコア、体水分率)、第 3 主成分で 3 項目 (歩行速度 [m/s]、Cadence[step/s]、歩行周期 [s]) であった。一方で、転倒事実がない実験協力者群の疫学データは第 1 主成分で 6 項目 (除脂肪量、筋肉量、推定骨量、脚点、左足筋肉量、左足除脂肪量)、第 3 主成分で 6 項目 (荷重応答時間 [s][%]、両脚支持時間 [s][%]、つま先離地時間 [s]、遊脚時間 [%]) であった。ここで時間 [%] とは、歩行 1 周期にかかる時間の割合である。また、第 2 主成分に関しては転倒事実なしにて主成分負荷絶対値量 0.7 以上となる項目が存在しないことから、第 2 主成分は議論の対象外とする。

### 3.2 主成分得点に基づく骨折リスクマップ

3.1 で抽出された疫学データから、主成分得点を算出して分布マップを作成する。まず、第  $i$  主成分の主成分得点  $z_i$  を (1) 式に示す。

$$z_i = \sum a_p u_{np} \quad (1)$$

ここで、添え字  $p$  は疫学データの項目番号、 $n$  は実験協力者番号を表している。また、 $a_p$  は疫学データにおける主成分負荷量、 $u_{np}$  は実験協力者  $n$  にお

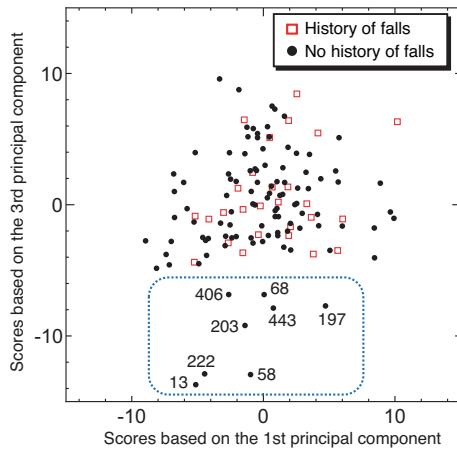


Fig. 1 骨折リスクマップ

る疫学データ  $p$  の標準化された値である。なお、 $a_p$  については同じ疫学データ  $p$  であっても転倒事実の有無によって主成分負荷量  $a_p$  は異なる。

(1) 式に基づいて、各主成分ごとに主成分得点を算出し、2次元マップ上に表示したものを Fig. 1 に示す。Fig. 1 の横軸は第1主成分に起因して算出した主成分得点、縦軸は第3主成分に起因して算出した主成分得点であり、四角(□)は転倒事実がある実験協力者、丸(●)は転倒事実がない実験協力者を表している。ここで、Fig. 1 の破線四角枠に着目すると、転倒事実がない実験協力者8名(実験協力者ID: 13, 58, 68, 197, 203, 222, 406, 443)がまとまって存在していることが分かる。これは、第3主成分に起因する疫学データ(歩行能力)に基づいて転倒事実の有無を切り分けできることを意味している。

#### 4. 骨折リスクマップの解析

転倒事実なしを特徴付ける第3主成分の6項目について転倒との関係を解析する。6項目の内、荷重応答時間[s][%]、両脚支持時間[s][%]、つま先離地時間[s]の5項目は主成分負荷量の値が正であり、遊脚時間[%]は負の値となっている。そこで、疫学データ値について、主成分負荷量の値が正の場合は昇順に、負の場合は降順に並び替えを行う。その結果、各々の項目において、転倒リスクの低い実験協力者上位20名に着目したところ、Fig. 1 の破線四角枠に含まれる転倒歴のない実験協力者8名が必ず含まれていた。つまり、上記6項目を指標として実験協力者を評価する際、各指標の上位20名(いずれも転倒リスクが低い順)に入っている数が多いほど転倒リスクが低いことが予想される。そこで、上記6項目を満たす(上位20名に入っている)数をカウントし、Fig. 1 を改めて図にしたのが Fig. 2 である。Fig. 2 をみると、第3主成分が負の方向ほど6項目を満たす数が増えており、転倒リスクが低くなる事を示唆している。

つぎに、「転倒事実のある実験協力者群(28名)」と

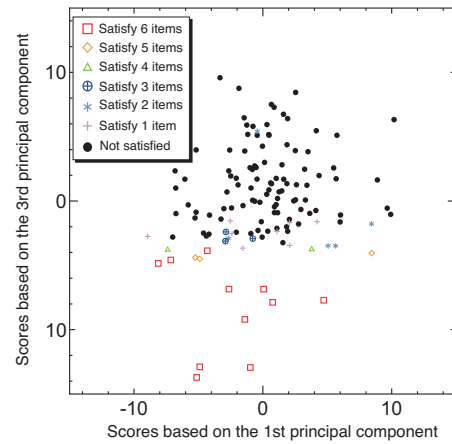


Fig. 2 6つの疫学データと骨折リスクマップの関係

「転倒事実のない実験協力者のカテゴリ分けされた各群」に対し、2標本t検定(片側検定)によって有意差を調べる。カテゴリ分けは、疫学データ6項目にて上位20名以上を満たす個数別に6項目, 5項目以上, 4項目以上, 3項目以上, 2項目以上, 1項目以上, 0項目以上の7つのカテゴリ群とする(各々をカテゴリ6~0とする)。Fig. 2 の縦軸に示される第3主成分得点に起因する得点に着目し、帰無仮説を「転倒事実のある実験協力者と転倒事実のない実験協力者(7つのカテゴリごと)において得点の平均値は等しい」、対立仮説を「転倒事実のある実験協力者と転倒事実のない実験協力者(7つのカテゴリごと)の得点の平均値に差がある(転倒事実あり > 転倒事実なし)」とし、有意水準を2.5%と定めてp値を算出した。結果として、カテゴリ6~カテゴリ1では $p < 0.025$ となり(カテゴリ6:  $t(37) = 7.014, p = 5.577 \times 10^{-7}$ , カテゴリ5:  $t(39) = 6.859, p = 2.147 \times 10^{-7}$ , カテゴリ4:  $t(40) = 6.771, p = 1.425 \times 10^{-7}$ , カテゴリ3:  $t(43) = 6.332, p = 1.601 \times 10^{-7}$ , カテゴリ2:  $t(46) = 5.090, p = 5.708 \times 10^{-6}$ , カテゴリ1:  $t(52) = 5.086, p = 2.756 \times 10^{-6}$ , カテゴリ0では $p > 0.025$ となった( $t(134) = 0.884, p = 0.191$ )。これらより、各6項目を満たす項目が1つ以上ある実験協力者は、転倒事実がある実験協力者と比べて第3主成分に起因する得点に有意な差があった。

#### 5. おわりに

鳥取県日野郡日野町で取得された疫学データに対し、PCAを行うことで骨折に関連する6つの疫学データを抽出し、骨折リスクマップを構築した。また、各6項目の内を満たす項目が1つ以上ある実験協力者は転倒事実のある実験協力者とは有意な差があることが分かった。

#### 参考文献

- [1] 細井俊明, 橋本千春, 小池康浩, 中内健治: 国民生活基礎調査の概況, 厚生労働省, p. 24, 2019
- [2] 松本浩実, 荻野浩: 一般在宅住民における骨折リスク値に関する意識調査, 第14回日本骨粗鬆症学会抄録, 2012