

2024年4月4日

報道機関 各位

国立大学法人東北大学

高血圧になりにくい食事パターンが機械学習でわかった —コホート研究データ解析におけるAIの活用—

【発表のポイント】

- 集団の追跡調査であるコホート研究^(注1)（仙台卸商研究）^(注2)の簡易自記式食事歴質問票（BDHQ）^(注3)のデータを、教師なしの機械学習法で解析することにより、特徴的な食事パターンの検出に成功しました。
- 食事パターンはそれぞれの摂取頻度の高い食品から「海産物とアルコール飲料」「低タンパク・低食物繊維・高炭水化物」「乳製品・野菜」「肉」の4種類に分類されました。
- 2年間の追跡期間中の高血圧症^(注4)の発症は、「乳製品・野菜」および「肉」パターンはいずれも「海産物とアルコール飲料」に比べて高血圧の発症リスクが6割以上少ないことがわかりました。

【概要】

疾病リスクを軽減する上で食事は重要です。しかし、食事といっても調理方法や食行動など様々な因子が関係することから、食事の評価はその複雑さゆえに容易ではありませんでした。

東北大学大学院医工学研究科健康維持増進医工学分野の永富良一（ながとみりょういち）教授、大学院生李龍飛、大学院医学系研究科の稲田仁非常勤講師、門間陽樹准教授らの研究グループは、協同組合仙台卸商センター組合員男性447名を対象に2008年から2010年まで2年間追跡したコホート研究データから、AIによる食事パターンの分類を試み、高血圧発症リスクとの関連を検討しました。その結果、「乳製品・野菜」「肉」の二つの食事パターンは、「海産物とアルコール飲料」に比べて高血圧発症リスクが6割以上少ないことを明らかにしました。小規模なデータからも食事などの複雑な生活行動と健康との関連をAIによる解析により明らかにできることを示した研究であり、今後、さまざまな活用が期待されます。

本研究成果は、2024年2月25日に栄養科学の専門誌 European Journal of Nutrition のオンライン版に掲載されました。

【詳細な説明】

研究の背景

コホート研究における食事因子の疾病リスクへの関与は栄養素に重点がおかれ、例えば塩分の摂取が高血圧のリスクと関連していることが明らかにされてきました。しかし食事はさまざまな栄養素の組み合わせで構成されるだけでなく、調理方法や食事時間などの時間因子を含む複雑な生活行動です。これまで、心血管リスクの低下につながる地中海食や血圧をあげないDASH(Dietary Approaches to Stop Hypertension)食などは「食事」に注目することが疾病リスクの軽減につながることを示されてきました。しかし食事を評価することはその複雑さ故、容易ではありませんでした。従来は主成分分析や因子分析による食事の解析が行われてきましたが、調理方法や食行動などを反映させることは困難でした。そこで研究グループは、これらも含む簡易型自記式食事歴質問票(BDHQ: Brief-type self-administered diet history questionnaire) のデータ解析に人工知能 (AI) を用いました。質問票の回答から食事パターンの分類を行い、検出された食事パターンと 2 年間の高血圧発症との関連を検討しました。

今回の取り組み

協同組合仙台卸商センター組合員男性 447 名を 2008 年から 2010 年まで 2 年間追跡したコホート研究データから AI による食事パターンの分類を試み、高血圧発症リスクとの関連を検討しました。BDHQ に含まれる 58 の食品の一週間の摂取頻度、12 の食行動、9 つの調理方法のデータを次元削減^(注5)・可視化法の一つである UMAP(Uniform Manifold Approximation and Projection)で処理したのちに教師なし機械学習^(注6)による非階層的クラスタリング法^(注7)である k-means 法で分類を行いました。食事パターンはそれぞれの摂取頻度の高い食品から「海産物とアルコール飲料」「低タンパク・低食物繊維・高炭水化物」「乳製品・野菜」「肉」の 4 種類に分類されました。2 年間の追跡期間中の高血圧発症の発症は、「海産物とアルコール飲料」に比べて「乳製品・野菜」および「肉」パターンは、いずれも高血圧の発症リスク(オッズ比^(注8))が 6 割以上小さくなることがわかりました。

今後の展開

食事のような複雑な生活行動も AI による解析により健康との関連を明らかにできることを示すことができました。今後、さまざまな手段で集積されている健診や Personal Health Record (個人の健康・介護情報を集約した生涯型電子カルテ) のデータ解析への応用が期待されます。今回示されたパターンそれぞれの食品そのものが高血圧のリスクを低下させたり高めたりするものではありません。食事の組み合わせが重要です。

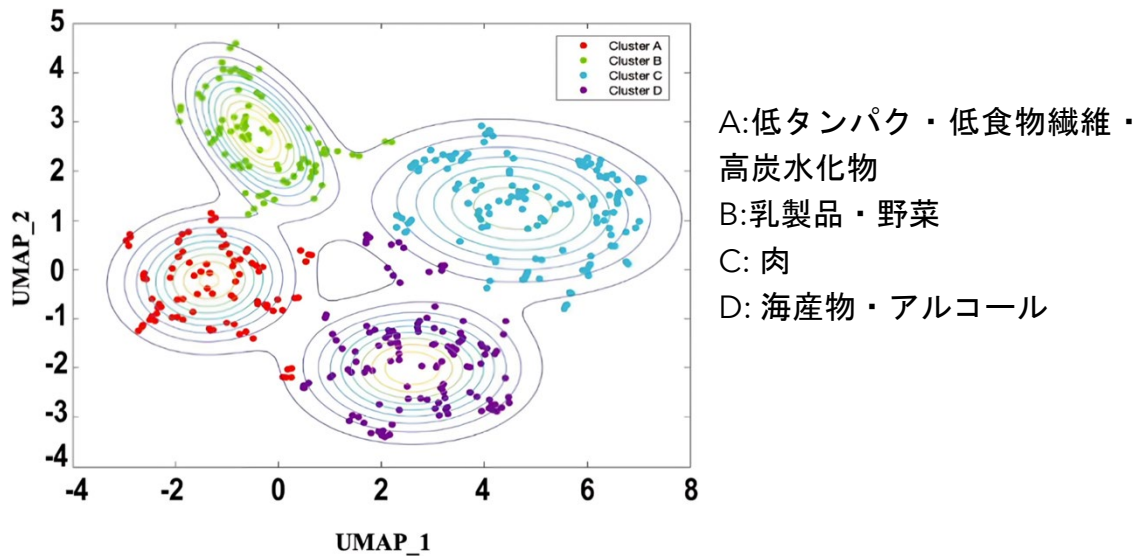


図 1. 非階層的クラスタリング法で得られた食事パターン

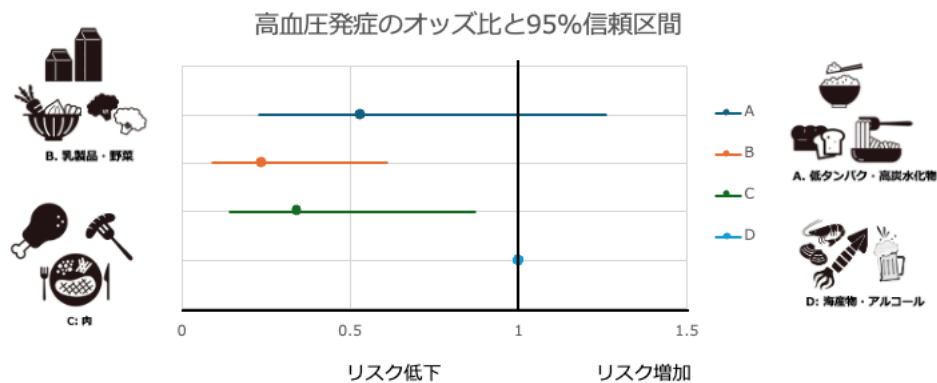


図 2. AI 食事パターンと高血圧発症のリスク(パターン D との比較)

【用語説明】

- 注1. コホート研究：人集団を一定期間追跡し、疾病等の発症に寄与すると考えられる要因を持つ集団と持たない集団との間で疾病の発症率などを比較する方法です。本研究では 2 年の追跡の間に高血圧を発症した人数を比較しています。
- 注2. 仙台卸商研究：文部科学省知的クラスター創成事業広域仙台先進予防型健康社会創成クラスター事業の一環として協同組合仙台卸商センター組合員を対象とし高血圧や動脈硬化危険因子だけではなく体力・抑うつ傾向に影響する生活習慣因子の探索を目的とした調査研究。2008 年にベースラ

イン調査、その後毎年の健診時に追跡データを2011年まで取得。東日本大震災に伴う健診手続きの変更に伴い追跡を終了した。国際誌に27の論文を発表。概要は運動疫学研究 2013;15(2):91-100, 保健の科学 2011;53(4):230-236 を参照。

- 注3. 簡易型自記式食事歴質問票(BDHQ: Brief-type self-administered diet history questionnaire) : 日本に住む人を対象に、通常の食事(サプリメント等を除く)から習慣的に摂取している栄養素量を、比較的簡便に、個人を対象として調べ、個人ごとの栄養素摂取量、食品摂取量、その他若干の定性的な食行動指標の情報を得るための質問票です。東京大学/国立健康栄養研究所の佐々木敏博士により開発されました。
- 注4. 高血圧 : 高血圧症。安静時収縮期血圧 140mmHg 以上、拡張期血圧 90mmHg 以上。脳血管疾患や循環器疾患の危険因子です。
- 注5. 次元削減 : 高次元のデータ(多数の変数)の中の類似点やまとまりをみつけ、できるだけ含まれる重要な情報を保持したまま低次元のデータ(少数の変数)に変換すること。このことにより変数間の関係性を明確化することが可能になります。
- 注6. 機械学習 : 人工知能技術(Artificial Intelligence AI)の一種。計算機に多量のデータを読み込ませ、データ内に潜む潜在的なパターンやデータ構造を学習させることにより、顕在化されていないルールを獲得することを可能にする解析方法です。
- 注7. 非階層的クラスタリング法 : 複雑なデータ構造を持つデータを予め決めた数のクラスタ(データのグループ)に自動的に分類する方法です。
- 注8. オッズ比 : 1 以上で発症リスク増加、1 未満は発症リスク低下。

【論文情報】

Title: Dietary patterns associated with the incidence of hypertension among adult Japanese males: application of machine learning to a cohort study

Authors: Longfei Li, Haruki Momma, Haili Chen, Saida Salima Nawrin, Yidan Xu, Hitoshi Inada*, Ryoichi Nagatomi*

タイトル : 高血圧発症と関連する日本人成人男性の食事パターン : コホート研究への機械学習の適用

著者 : 李龍飛, 門間陽樹, 陳ハイリ, Saida Nawrin, 徐芸丹, 稲田仁*, 永富良一* (*責任著者)

*責任著者 : 東北大学大学院医工学研究科 教授 永富良一

掲載誌 : European Journal of Nutrition

DOI : 10.1007/s00394-024-03342-w

URL: <http://www.sports.med.tohoku.ac.jp/index.html>

【問い合わせ先】

（研究に関すること）

東北大学大学院医工学研究科

健康維持増進医工学分野

教授 永富 良一

TEL: : 022-717-8588

Email : nagatomi*med.tohoku.ac.jp

（*を@に置き換えてください）

（報道に関すること）

東北大学大学院医工学研究科

TEL : 022-795-5826

Email : bme-pr*grp.tohoku.ac.jp

（*を@に置き換えてください）