

<研究課題> 高齢者に対する人工知能を用いた運動、栄養の健康効果の検証とヘルスケアの個別化の実現化

代表研究者 東京大学大学院医学系研究科 糖尿病・代謝内科 届出研究員、英国キングス
カレッジロンドン Institute for Population Health 客員教授 山田 朋英
共同研究者 東京大学 糖尿病・代謝内科 特任准教授 庄嶋 伸浩
東京大学 糖尿病・代謝内科 教授 山内 敏正
台湾軍人病院 准教授 Chia-lin Lee
King's College London 教授 渋谷 健司

【まとめ】

精度の高い生活習慣病合併症リスクエンジンの開発により、コホートにある栄養や運動などの問診などの個別化情報からの自動で診断し、科学的根拠に基づいた精密な予防や治療方針決定の実現化を目標とします。

具体的には、1) 個別化治療戦略アプローチ、2) 標準化治療戦略アプローチ、3) 個別化および標準化治療戦略アプローチに当てはまらないその場の“空気”による戦略アプローチです。人工知能の解析統合と体系化を行います。

1. 研究の目的

1-1 個別化治療戦略アプローチとしては、食・生活習慣による人工知能を用いた従来の統計学的手法に比べ、精度の高い長期的死亡、高齢化に伴う疾患、QOL 低下、メンタルヘルス、メタボリックシンドロームを予測するリスクエンジンの開発を行います。

食・生活習慣がなく、死亡リスクが高い特定の人々のうち、ある特定の食・生活習慣の介入（例：牛乳を飲む、減塩を促す、たん白摂取を促すなど）を促すことで、将来の死亡率を下げることができる集団を人工知能で特定します。

1-2 集団の治療戦略アプローチとしては、“生活習慣”に関するリアルタイムシステムティックレビュー・メタアナリシスの実現化を目的としました。

過去に発表された、食生活に関する膨大な医学情報内に記載されたテキストの検索/閲覧の省力化をテキストマイニングの手法を用い可能とします。

人工知能によるニューラルネットワークを介したディープラーニングには一般に膨大な教師データが必要となる、論文検索/閲覧

支援プログラムの開発により、ディープラーニングのための教師データの作成の省力化の実現が期待されます。

過去の食生活と生活習慣病に関連した膨大な臨床試験や疫学試験などの科学的根拠を統合解析し、臨床的社会的な位置づけを明確にして、効果的な予防に活用します。

2. 研究方法と経過

2-1 個別化治療戦略アプローチ

研究対象：日米大規模コホートデータを利用。米国民健康栄養調査（NHANES）は米国民の健康状態、栄養状態を定期的に調査し、面接による人口動態学的データ、社会経済学的状況、食事、健康状態の聞き取りのほか、医学的諸検査、身体測定を行っている。

これらのデータは、主要疾患、疾患発症危険因子の有病率、栄養状態の把握とその健康増進、疾病予防との関連の検討に供される。

参加者の食品アンケート記録詳細、生活習慣病発症、死亡率に関連するデータを用いる。日本のコホートに関しては魚沼コホートのデータを用いました。

統計解析：データは必要に応じて、トレーニングデータ、検証データ、およびテストデータに分割します。

トレーニング検証データを使用して、モデル構造のハイパーパラメーターを調整します。テストデータはモデルテストと一般性テストに使用し Up sampling、SMOTE matching を用います。

ディープラーニングによってアルゴリズムを作成します。

検証データでパラメーターを調整し、トレーニング/検証データでロジスティック回帰と深層学習を比較します。

ディープラーニングまたはロジスティック

回帰の最終決定バージョンアルゴリズムがテストデータに適用されます。テストデータでアルゴリズムをテストし、再度データをテストする際のロジスティック回帰と深層学習の精度/AUCの比較を行います。最終的にはオンライン上などでの実装を目標とし、誰もが自由に無料で利用できることが望まれます。

2-2 ヘルスケアの集団的アプローチ

ニューラルネットワークを介した人工知能では、文章を読んだ人が重要だと判断する暗黙の基準を学習(ディープラーニング)し、多数の文書からその基準に沿ったものを抽出できます。

教師データを学習する際に、教師データに含まれる単語ごとに教師データとの関連性と単語同士のつながりに関して伝達情報量を通して数値化し、ニューラルネットワークを介し、特徴量として利用しています。

教師データを学習した後、新たに評価用に投入されたデータにも同様の処理がなされ、教師データによって付与された特徴量により、それぞれの文章はスコア化され、そのスコア(=正解である予測確率)を基に教師データとの関連性を、食事療法に関する臨床ガイドラインに採用されたシステムティックレビュー・メタアナリシス論文を評価しました。

2-3 標準化治療戦略アプローチ

国内外の診療ガイドラインで実際に引用されたシステムティックレビュー・メタアナリシスを用いた検証を行い、人工知能アルゴリズムの実用化に向けた改良を行いました。食事療法介入研究試験(DPP)のメタアナリシス論文で実際にシステムティックレビューに用いた論文を、正解論文(最終的に系統的レビューの対象となった論文)、非正解論文(系統的レビュー対象とならなかった論文)に分けました。

人工知能が、教師データをもとに機械学習し、評価データにスコア付けをしました。

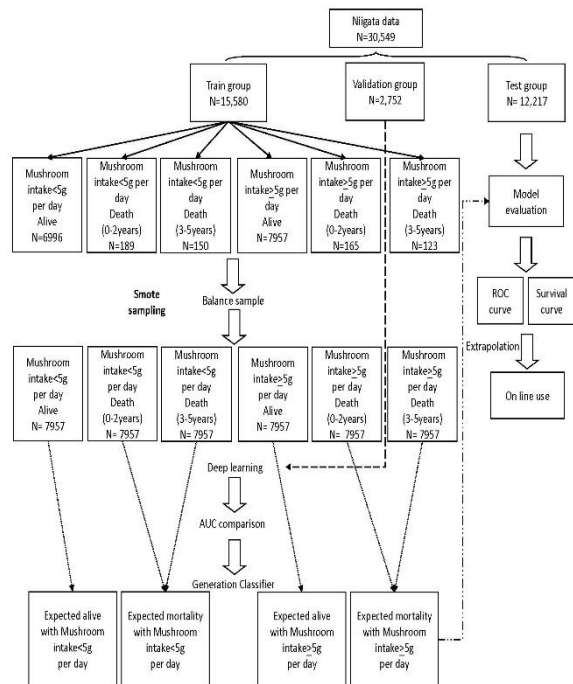
3. 研究の成果

3-1 ヘルスケアの個別化アプローチ

参加者を Train group と Validation group, Test group に分割しました。

Train group を、ある種の食事(図の例ではきのこを例とした)摂取と、5年追跡期間中での死亡の有無の4グループに分割しました。

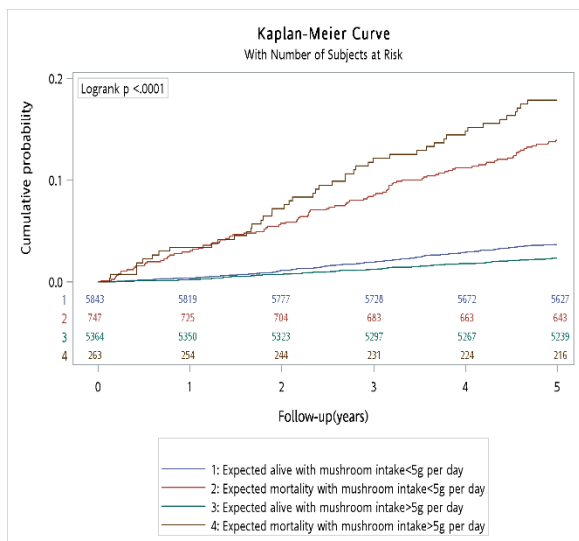
ADASYN Sampling し、仮想データ構築と Matching を行い、バランスをとりました。Deep learning し、これに Validation group を用い、Hyper parameter を調整しました。その後 Generation Classifying し、Test group で Validation を行いました。Logistic model と AUC について比較を行いました。



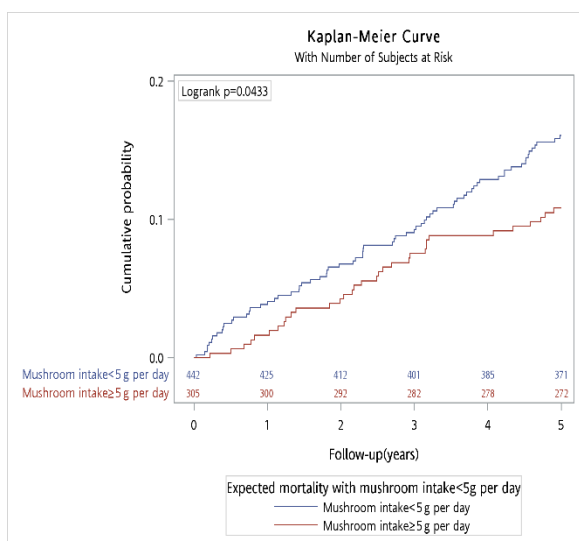
この食事摂取に関するAIアルゴリズムを用いることで、個人は以下の4つのグループのいずれかに分類されました。

この上でAIアルゴリズムを用いた死亡リスクに関するシミュレーション解析結果を示します。

この4グループは、もし彼らが食生活を変えなかった場合の生存予測シミュレーションのグループです。



これらの4グループにおいて、彼らがある食べ物(ここではキノコを例とする)を食べる習慣を変更(続けるか、あるいはやめるかで、さらに8つのグループに分類しました。もし個人が、下記のグループに分類された場合、キノコを食べる新しい習慣を始めると、それ以下の量しか食べない場合に比べて、5年後の死亡リスクの有意な低下が認められることが推定されました。(統計学的有意差あり $p=0.04$)



しかし一方で、食生活の変更をしても死亡リスクに変化のない集団グループも分類できました。

3-2 標準化治療戦略アプローチ

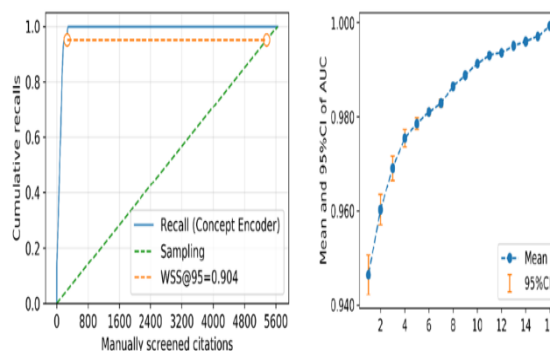
ディープニューラルネットワークベースの機械学習により、系統的レビューのスクリーニング作業負荷が軽減されたことを報告しました。アメリカ糖尿病学会 (ADA)、アメリカ心臓病学会 (AHA)、アメリカ心臓協会 (ACC)、

およびアメリカ脳卒中協会(ASA)からの最近の臨床ガイドラインで引用された介入ランダム化比較試験のすべての系統的レビューとメタアナリシスが評価されました。結果2報の論文をランダムに抽出し、機械学習を行うことで、専門家の暗黙知を理解しました。

その後AIが独自に付けた高スコア順に論文を閲覧した場合、全候補論文の10%未満の閲覧で、100%の正解論文をもれなく抽出することができ90%の論文スクリーニングの省力化が得られました。

ディープラーニングにより食事生活習慣介入研究試験(DPP)のシステムティックレビューにおける、90%の論文スクリーニングの省力化が得られた。

Yamada et al. J Med Internet Res. 2020 Dec 30;22(12):e22422. doi: 10.2196/22422



Diabetes Prevention Program Research Groupによる検証

3-3 個別化治療戦略アプローチである精密栄養の応用としては、一般的な変数を使用して人工知能モデルを構築することが挙げられます。

今後は様々な食事や運動習慣での検討を、遺伝データやリアルワールドデータなどと組み合わせることで、より精度の高いAIアルゴリズムを構築できると考えます。

健康診断施行施設などと協力することができ、健康診断の後、人工知能は死亡リスクを自動的に計算して予測することが可能となります。

すなわち、人工知能による予測により、リスクの高い人々に必要な生活習慣を提案することができるようになると思います。

また、標準化治療戦略アプローチとしては、文献検索省力化に関しより精密に検討を進めます。

この他我々は、論文の質的評価の高速化を目的としたソフト、PICORON-EBM(PICO and Risk Of bias Navigator for Evidence Based Medicine)の開発を行いました。システムティックレビュー・メタアナリシスにおいて、多くの医学臨床論文の質(バイアスなど)に関して適切に評価することが求められます。

見落としを防ぐためのテキストマイニングソフトウェアPICORON(PICO and Risk of Bias Navigator)を開発しました。

PICORON に英文の臨床試験の論文を読み込ませることで、質的評価に必要な単語が自動でハイライトされ、論文の質的評価時間は3分の1まで短縮することができました。

また、医療情報からのデータ抽出ソフト T-Library の開発を行いました。

医療情報(退院サマリ)の情報を項目ごとに自動で分割抽出し、エクセルデータ上の対応項目に分割され転記することができました。

T-Libraryにより退院サマリデータの転記にかかる時間は5分の1に省力化できました。

このソフトにより数千数万のテキストベースの退院サマリなどのビッグデータを解析することが容易となり、自動更新される疾患リスクエンジンや人工知能を用いた解析に利用することが可能となります。

4. 今後の課題

煩雑な作業を、専門家の暗黙知を学んだ人工知能が代行することで、人間はより創造的で生産的な仕事に従事することが可能となります。

今後も人工知能による臨床エビデンスの統合と体系化を実現するため、各個人の生活情報や他のバイオマーカーなどの膨大な個人情報から、人工知能を用いて統合解析を進めてまいります。

精度の高い生活習慣病合併症リスクエンジンの開発により、コホートにある栄養や運動などについての問診などの個別化情報からの自動で診断し、科学的根拠に基づいた精密な予防や治療方針決定の実現化を今後も目標とします。

5. 研究成果の公表方法

【参考文献】

1)Yamada T, Yoneoka D, Hiraike Y, et al, Deep Neural Network for Reducing the Screening Workload in Systematic Reviews for Clinical Guidelines: Algorithm Validation Study.

J Med Internet Res 2020 Dec 30;22(12):e22422.

2)Mudaliar U, Zabetian A, Goodman M, Echouffo-Tcheugui JB, Albright AL, Gregg EW, et al, Cardiometabolic risk factor changes observed in diabetes prevention programs in US settings: a systematic review and meta-analysis.

PLoS Med 2016;Jul;13(7):e1002095

3)Yamada T, Kondo Y, Momosaki R, PICO and Risk Of bias Navigator for Evidence Based Medicine. PeerJ Preprints 2019;7:e27684v1 DOI 10.7287/peerj.preprints.27684v1

4)Yamada T, Kondo Y, Momosaki R, Automated data extraction software for medical summary using text mining (T-Library) PeerJ Preprints 2019;7:e27685v1 DOI 10.7287 / peerj.preprints.2768

【公表方法】

Lancet,NEJM,BMJ 誌などへの投稿です。

以上

(Personalized healthcare for the elderly using artificial intelligence)

Primary Researcher: (Tomohide Yamada)
(Visiting Professor), (King's College London, London, UK)
(Registered Researcher), (The University of Tokyo, Tokyo, Japan)

Co-researchers: (Nobuhiro Shojima)
(Specially Appointed Associate Professor), (The University of Tokyo, Tokyo, Japan)
(Toshimasa Yamauchi)
(Professor), (The University of Tokyo, Tokyo, Japan)
(Chia-lin Lee)
(Associate Professor), (Taichung Veterans Hospital, Taichung, Taiwan)
(Hisashi Noma)
(Professor), (Institute of Mathematical Statistics, Tokyo, Japan)

Evidence-based medicine (EBM) involves determining treatment that matches the needs of each patient by integrating the best and latest available "scientific evidence" and "clinical skills". Systematic review and meta-analysis refers to the process of searching databases and performing statistical analysis to integrate the results of multiple independent studies conducted in the past. These results obtained provide the highest quality evidence, which has become the foundation of various clinical guidelines. The objectives of this project are to reduce the time required to perform a systematic review by employing artificial intelligence (AI) and to improve the precision of the method. We also aimed to identify particular people those who have benefited from lifestyle intervention at high risk of 5-year mortality by Artificial Intelligence. This approach takes into account the role played by differences in genetics, epigenetics, the microbiome and other environmental factors including but not limited to other nutrients, environmental toxicants, pharmaceuticals, disease status, and physical activity. This could provide significant benefit by facilitating the acquisition of new evidence for clinical guidelines in all disease fields as well as lifestyle related diseases in the elderly.

Key words: Artificial intelligence, Systematic Review, Life-related diseases