

2024/1/24

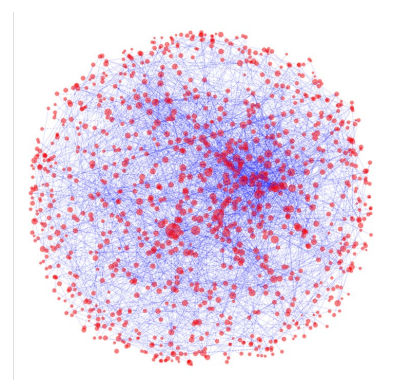
報道関係 各位

【研究発表】言葉の力で慢性腎臓病(CKD)の進行を予測する！
ビッグデータ+AIで構築した医学用語のバーチャル空間モデル

CKD 患者を文章のように扱う AI を開発 CKD 進行を正確かつ簡単に予測！ (川崎医科大学 医学部 神田英一郎特任教授ら研究グループ)

【概要】

本学医学部 神田英一郎学長付特任教授(データサイエンス・腎臓内科学)らのグループは、本学高齢者医療センター 柏原直樹病院長とともに、**医学用語のバーチャル空間を AI によって構築しました**。この空間内の**医学用語のネットワーク**を詳細に検証したことにより、**慢性腎臓病(CKD)患者を文章のように扱うことができることを世界で初めて医学的かつ数学的に証明し、CKD の進行を正確かつ容易に予測できるシステムと予後指標を開発**しました。本研究は、Scientific Reports (<https://www.nature.com/articles/s41598-024-52235-9>)で発表されました。



医学用語の空間のイメージ

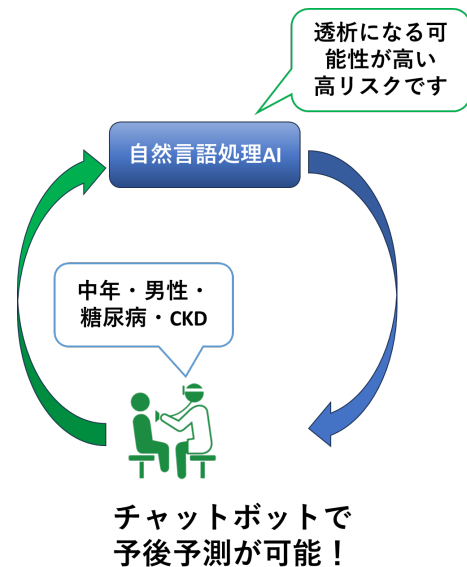
- ・ 赤点, 医学用語の分布 (サイズは重要度を表す)
- ・ 青線, 用語間の関係性を示す

ポイント① 自然言語処理(NLP) AI の解析が、現実の疾患病態を反映していることを、**ビッグデータ**と最新数学(圏論)で証明した。**NLP AI の医学的信頼性**が示されました。

ポイント② 患者の情報を、実際の数値が無くても、**文章のように扱うことが可能**であることが示されました。

活用例① CKD の進行予測

高齢者、男性、糖尿病、降圧剤 A を使用、尿蛋白ありなどと、入力することで、透析などが将来どれくらいの確率で生じるか、**予後を正確に予想**できます。この技術は、近年話題になっている **NLP AI で容易に実用化**することができます。



活用例② CKD の進行の新しい指標

患者データを医学用語バーチャル空間内で解析することで、正確に予後を予測することが可能になりました。さらに、**患者ベクトルと透析ベクトルの関係(内積)**が**新しい予後指標**であることが示されました。

ポイント③ 医学用語のネットワークが、CKD の病態を反映していることを示しました。

活用例③ 新規治療法のシーズ探索

信頼性の高い研究シーズ探索エンジンの基盤となります。これにより、**新規治療法の開発や疾患の病態解明が促進**されます。

【研究の背景】

医学研究では、大量の論文を読むことが不可欠です。しかし、CKD に関する論文でも年間 1 万本以上発表されており、全ての論文から、新しい研究シーズを見つけ出すことは非常に困難でした。そこで、医学論文を解析するシステム「**Doctor K**」を開発したことをきっかけに、本研究を開始しました。

論文が科学的に正しいことを前提に研究は進められます。また、その医学用語の関係性は、当然医学概念を反映し、ネットワークを構築していると考えられます。しかし、その正しさは概念的なものであり、どの程度正しいのか、**現実の患者データとマッチしているのか**について検討された研究はありませんでした。

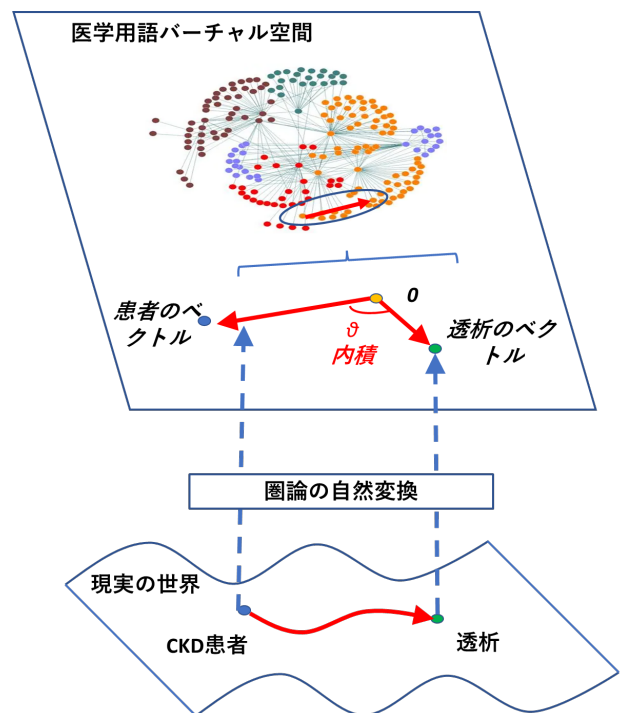
近年、**ChatGPT や Bard** のような NLP AI の開発が進められ、話題になっています。もともと Word2Vec というモデルによって、単語をベクトルとして計算できるだけでなく、言葉の意味も保たれていることが示されたことが、今日の隆盛の始まりとなっています。そこで、私どもは、**Word2Vec** を用いて **165,271 本の医学論文を解析し、医学用語のネットワークを構築**しました。そして、そのネットワーク内の**医学用語の関係性を手作業で確認**し、AI の解析が、医学的な意義とベクトルとしての計算の両方を担保していることを明らかにしました。この作業は、あまりに骨が折れ大変すぎるため、これまでだれも行ったことがありませんでした。

次に、このネットワークと現実の **CKD 患者データ(26,433 人)** の関係を、**最新数学の圏論** を基盤とした解析方法で確認しました。圏論の活用は、物事の関係性を解析する分野で、データサイエンスや認知科学で注目されていますが、**臨床研究系分野では世界初の試み**です。これにより、**論文の医学概念と実際の患者データが、高い精度で一致**していることが示されました。

患者の予後予測を行うには、誤診を少なくし、高い正確度が要求されます。また、どのように AI が解析しているか明らかにならなければ、信頼することができません。本研究により、**NLP AI が正確に現実の疾患病態を反映し、信頼できる**ことが分かりました。

【今後の展開】

最新の NLP AI に、音声などで患者情報を入力することで、予後を予測するモデルの開発を行っています。また、新規の研究シーズを発見する検索システムの開発も行っています。



医学用語バーチャル空間と現実の世界の関係性

- 医学用語のネットワークと現実の患者の関係が、圏論の自然変換で示された
- 医学用語バーチャル空間で患者データを解析することで、容易に予後を予測することができるようになった
- 新規予後指標(内積)が開発された

【論文情報】

タイトル: New marker for chronic kidney disease progression and mortality in medical-word virtual space

タイトル和文: 医学用語バーチャル空間における慢性腎臓病進行の新規指標

著者名:

神田英一郎; 川崎医科大学 医学部

Bogdan I Epureanu; 米国ミシガン大学 機械工学部

安達泰治; 京都大学 医生物学研究所 生命システム研究部門 バイオメカニクス

佐々木環; 川崎医科大学 腎臓・高血圧内科学

柏原直樹; 川崎医科大学 高齢者医療センター

掲載日: 2024年1月18日

掲載誌: Scientific Reports 14, Article number: 1661 (2024)

DOI: 10.1038/s41598-024-52235-9

<問合せ先> 広報について
川崎医科大学 事務部
庶務課 課長 浅沼 淳
Tel 086-462-1111(代)

<問合せ先> 研究について
川崎医科大学 医学部
学長付特任教授 神田 英一郎
Tel 086-462-1111(代)