

共同企画

共同企画13

包括的・重層的症例データベース J-CKD-DBにより可能となった臨床研究と今後の発展への期待

2021年11月21日(日) 15:10 ～ 17:10 H会場 (2号館3階234)

[4-H-3-03] AI・ICTを活用した CKD患者管理システムの開発 Devdevelopment of CKD therapy sytem based on AI and ICT

*神田 英一郎¹ (1. 川崎医科大学)

*Eiichiro Kanda¹ (1. Kawasaki Medical School)

キーワード : Database, Machine learning, CKD, Dialysis

慢性腎臓病(CKD)・透析患者は死亡・心血管疾患の高リスク患者である。特に高齢患者では、原疾患、合併症、低栄養などが合併しやすいため、早期発見と管理が重要である。しかしながら、臨床現場では全ての患者に対して複数の危険因子を同時にコントロールすることは難しい。そこで講演者らは、患者データベースに AI・ICT 技術を活用し、高リスク患者をスクリーニングするための機械学習モデルおよび治療管理システムを開発した。

CKD患者データを川崎医科大学附属病院電子カルテデータベースより抽出し、透析・死亡リスクを予測する26の機械学習モデルを開発した。8つの機械学習モデルは既存マーカーよりも高い精度を示した。また、機械学習による予測確率は腎・生命予後に強い関係性を示した。サブグループでも同様の結果であった。さらに臨床で活用するため、ユーザーが患者情報を入力すると予後を予測する WEBシステムを開発した。

機械学習モデルは従来の統計モデルと異なり、その解析結果を人間が理解することが困難である。そこで講演者らは解析内容をある程度理解可能にするため、日本透析医学会統計調査データベースを用いて機械学習のアンサンブルモデルを開発した。k-means法で解析したところ、人間が理解できる特徴で患者は5つのグループに分類された。最も生命予後の悪かったグループは、高齢かつ protein energy wasting の患者であった。本システムの使用によりこのような高リスク患者の早期発見が可能になることが示唆された。また、講演者らは、食事療法管理システムや検査値・治療薬等を自動的に解析する治療管理システムを開発した。これらシステムではディープラーニングが搭載されており高い精度を達成した。

患者データベースに AI・ICT を活用したシステムを臨床応用することにより、高リスク患者の発見から治療内容の管理まで、患者を中心とした医療システムの構築が可能になる。また、CKD・透析患者の予後改善が期待される。

AI・ICT を活用した CKD 患者管理システムの開発

神田 英一郎^{*1}

^{*1} 川崎医科大学 医学部

Development of CKD therapy system based on AI and ICT

Eiichiro Kanda^{*1} ^{*1} Medical Science, Kawasaki Medical School

Abstract: The number of chronic kidney (CKD) patients in Japan has been increasing because of the aging of its population. Accurate risk prediction of CKD progression is an important tool for treating CKD. Recently, with the development of artificial intelligence (AI) and information and communication technology (ICT), big databases of patients have been constructed and used for clinical research and therapy. For example, analyses of big databases of CKD patients enable the accurate prediction of CKD progression. Using the new technology, we have obtained new information on CKD, which cannot be clarified by standard statistical methods. It is expected that the utilization of AI and ICT will lead to the development of new CKD therapeutic strategies and the improvement of life and renal prognosis of CKD patients.

Keywords: database, chronic kidney disease, dialysis, artificial intelligence, information and communication technology.

緒論

加齢や生活習慣病とともに腎機能は低下するため、CKD 患者は増加傾向にある。CKD の進行および発症をいかに予防し進行を抑制するかが、患者および医療費の観点から重要な課題である。近年人工知能(artificial intelligence, AI)や情報通信技術(information and communication technology, ICT)が発達し、CKD の病態解明や治療にこれらの技術が取り入れられるようになってきている。そこで本稿では、筆者らが行ってきた、患者データベースに AI・ICT を活用した CKD に関する研究を中心に概説する。

データベースの構築と活用

データベース技術の発展に伴って、従来は使用できなかったビッグデータの解析が可能になってきている。日本腎臓学会と日本医療情報学会は包括的 CKD データベース(J-CKD-DB)を構築した(1)。このデータベースには、大学病院の電子カルテ情報から 20 万人以上の CKD 患者の全情報(検査値、治療内容、診断名など)が、中央のデータセンターに集められている。データがコンピューターにより自動抽出されるため入力負荷が軽減される。また、他のデータベースとの突合が可能であるため、電子カルテ情報以外の腎生検などのデータと合わせて解析することが可能である。このデータベースを用いて、CKD の治療パターンの解析が進められている(2, 3)。

日本透析医学会は長年にわたり統計調査(Japanese Society for Dialysis Therapy Renal Data Registry, JRDR)を 1966 年から継続しており、これをもとにデータベースを拡張してきた(4)。JRDR の 2019 年末時点における年次調査は 4,487 施設を対象に実施され、98.3%の施設データを収集している。このような全国約 34 万人の透析患者を網羅したデータベースは世界でも類がない。

さらに、健康診断や医療保険のデータベースを使用した研究も可能になった。医療保険データベースには、患者の検査データだけでなく処方薬や医療費のデータも含まれている。筆者らは高カリウム血症の医療費に対する影響を報告した(5)。

患者データベースはコホート研究だけでなく、患者レジストリとして捉えることによりランダム化比較試験などへの活用も考えられる。さらに電子カルテ情報だけでなく様々なデータベースとの統合により、従来の疫学的研究では明らかにでき

なかった CKD の多様性が解明される可能性がある。

CKD 患者の予後予測

患者の予後を正確に予測できれば、予後の悪い患者をスクリーニングすることができ、検査・治療を集中することができる。また治療法の選択による予後の変化を予測できるため、ベストな治療法の選択が可能になる。すると薬剤の組み合わせの検討や透析導入の時期などの決定に参考になる。

これまで CKD 患者や透析患者の予後予測指標は、統計学的モデルを用いて開発されてきたが、患者ビッグデータを用いて予後を予測する機械学習モデルの報告は数少ない。そこで CKD 患者の予後を予測する AI システムを開発し、実用化を試みた。

開発用データでリスクを予測するモデルを作成し検証用データで確認するという基本的な研究の流れは、統計学的モデルと機械学習モデルのいずれでも同様である。まず、電子カルテデータベースより CKD 患者データ(n=66,981)を抽出した。平均 eGFR は 54.2 ml/min/1.73m²であった。

次に CKD 患者の末期腎不全・死亡リスクを予測する 26 の機械学習モデルを開発し、予測精度を既存のバイオマーカーと 1,000 回のシミュレーションで比較した。8 つの機械学習モデルは既存マーカーよりも高い C 統計量を示し(p<0.05)、特にランダムフォレストは 0.916 (95% CI 0.88, 0.952)であった。また Cox 比例ハザードモデルは機械学習による予測確率と腎・生命予後に強い関係性を示した(p<0.0001)。サブグループ解析でも同様の結果であった。

さらに開発した機械学習モデルを搭載した WEB システムをユーザーフレンドリーな仕様で開発することに成功した。このシステム開発により AI による予後予測システムの臨床での活用の可能性が示唆された。

健康診断参加者における早期 CKD への進行の予測

腎硬化症は、慢性糸球体腎炎や糖尿病性腎症と異なり、尿蛋白が検出されない場合が多く、早期の発見が困難である。このような潜在的リスクを有する患者の多くは健康診断では GFR の低下が認められないため数年間放置され、血清クレアチニン値が上昇した後医療機関を受診することが多い。このように進行する早期 CKD の患者を発見する臨床研究が長年

必要とされてきたが、早期 CKD 患者を対象とした臨床研究ではエンドポイントの発生数が非常に少なく、解析困難になることが多かった。そこで、筆者らは 7,465 人を 7 年間追跡した山形県の健康診断データに機械学習を用いて、各対象の CKD への進行リスクを予測する研究を行った(6)。

腎機能の低下には血圧、体重、糖尿病の有無など様々な背景因子が複雑に関与している。またアウトカム発生率が非常に低く統計学的有意差が得られにくい傾向があった。そこで本研究では、頻度論的統計学による解析ではなく、ベイジアンネットワークを用いて危険因子を探索し、その選択された因子を用いてアウトカムを正確に機械学習で予測できるか試みた。

選択された危険因子を用いた support vector machine (SVM)を学習させた後、腎機能が将来的に低下する患者分布の予測を行った。結果をヒートマップで示す(図 1)。

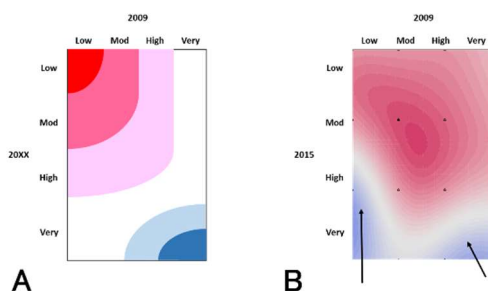


図 1 CKD 重症度分類の経年的変化に基づいた腎機能低下が予想される患者の分布

ヒートマップは SVM によって予測された腎機能が将来低下する患者分布を示し、青色は高分布、赤色は低分布を表す。
A 通常患者分布予測: 横軸は 2009 年の CKD 重症度分類を表し、縦軸は 20xx 年の CKD 重症度分類を示す。
B SVM による患者分布予測: 横軸は 2009 年の CKD 重症度分類を表し、縦軸は 2015 年の CKD 重症度分類を示す。腎機能低下が予測された患者分布が SVM により示された。2009 年に low risk であっても、その後高リスクとなる症例群が出現する(図 B→)。

研究開始時に low risk であれば、その後も腎機能低下の確率が低いと予想される(図 1A)が、経年的に high risk や very high risk の対象が徐々に出現し(図 1B)、それらの腎機能低下の確率が高いことが分かった。また、そのような傾向が明らかになるまで、少なくとも 3 年以上の経過観察が必要であった。つまり健康診断では 1 回の検査で low risk であったとしても決して安心することはできず、3 年以上経過を観察しなければ予後予測が困難であることが示唆された。

本研究のように一般的な統計による解析が困難な場合であっても、ビッグデータに機械学習を用いて解析することが可能になる。

透析患者の予後予測

透析患者の低栄養の特徴の一つは筋肉や脂肪などの蛋白質やエネルギー貯蔵が減少している状態であり、protein energy wasting (PEW)と呼ばれている。この状態では、低栄養と炎症が合併することにより、生命予後に影響を及ぼす。PEW の予防と治療には、透析患者の栄養スクリーニングを定期的に行い、低栄養のある患者には詳しい診察を行う必要がある。また、高リスク患者を早期に発見し合併症の改善と治療の最適化は、リスクを減少させ予後を改善する。しかしながら、

臨床現場では全ての患者に対して、多数のリスクファクターを同時にコントロールすることは難しい。そこで、JRDR を解析し、高リスク患者を発見する機械学習モデルを開発した(7)。

機械学習モデルはデータの分類や予測の精度が非常に高い。しかし、従来の統計モデルと異なり、その解析結果を人間が理解することが困難である(ブラックボックス)。そこで、解析内容をある程度理解可能にするため、複数の機械学習モデルを組み合わせて構築するアンサンブルモデルを作成した。

データベースをモデル開発用と検証用に分けた(各 n=39,990)。まず、開発用データセットを用いて、k-means 法にて患者を 5 つのクラスターに分類した。次に各クラスターに対して、SVM を学習させ、プライマリーアウトカムである 1 年後の死亡を予測させるモデルを作成した。K-means 法と SVM を統合することで、アンサンブルモデルが得られた(図 2)。

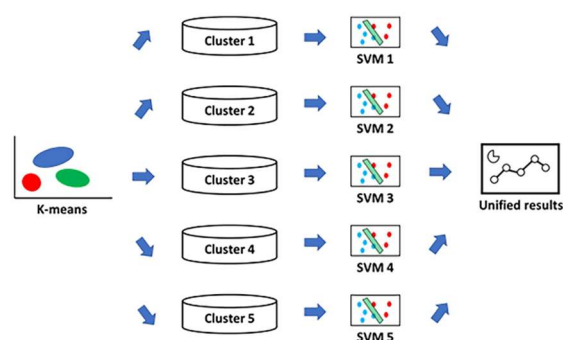


図 2 アンサンブルモデル

本研究で作成したアンサンブルモデルは 5 つのクラスターとそれぞれに対する SVM で構成されている。患者は k-means 法にて 5 つのクラスターに分類される。次に各クラスターの患者データを SVM で解析する。その結果を統合し、患者の予後を予測する。

次に検証用データセットを用いて、モデルの性能を評価した。各クラスターの患者背景には、クラスター1 は若い男性で原疾患は糸球体腎炎であった。クラスター2 は女性・糸球体腎炎、クラスター3 は糖尿病性腎症、クラスター4 は高齢者・腎硬化症、クラスター5 は高齢・protein energy wasting であった。Cox 比例ハザードモデルを用いて 1 年死亡リスクとクラスターの関係を調べたところ、クラスター1 から 5 の順に生命予後の悪い傾向が認められ、クラスター1 と比べてクラスター5 のハザード比は 8.86 (95% CI 7.68, 10.21)であった(図 3)。

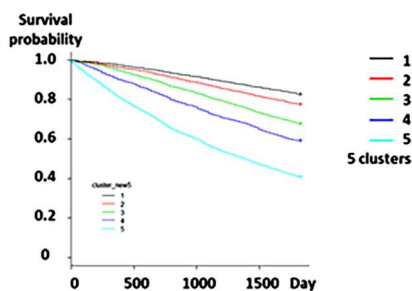


図 3 クラスターと 5 年生存率の関係

クラスター1 が最も予後が良く、クラスター5 が最も悪かった。

このモデルの予測正確度は 0.948 であり、ロジスティック回帰

モデル(0.938)、SVM(0.937)や深層学習(0.936)よりも高い精度であった。

本研究では高い予測パフォーマンスを示すモデルの開発を通して機械学習の解析の途中経過が明らかになった。本モデルのようなリスク予測システムは臨床現場へ適用できるため、高リスク患者の早期発見・治療を可能にすると考えられる。

AIを使った透析患者の食事療法管理システム

栄養摂取量の正確な推定には、食事記録が欠かせない。しかしながら、高齢者が食事記録を継続することは困難である。そこで筆者らは食事の写真をスマートフォンで撮影し、その写真を解析することで栄養摂取量を記録できるAIを開発し、それを活用した食事療法管理システムを構築した(図4)(8)。

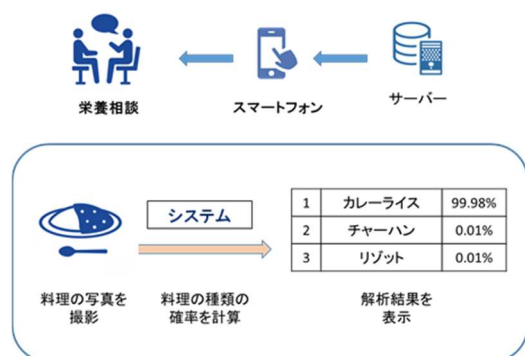


図4 食事療法管理システム

スマートフォンで撮影された料理のデータはディープラーニングを用いたシステムで解析される。写真の料理がどの種類なのか確率を計算し判別結果を表示する。

このシステムでは、ユーザーが簡単に利用できるよう、スマートフォンアプリケーションを使用している。このシステムには食事の写真を認識することが出来るディープラーニングが組み込んであり、食事の写真を撮影することでメニューを判別し、標準的な栄養摂取量を計算する。より正確な推定のため、食事摂取量の入力も可能である。記録された食事記録のデータはサーバーに保管された後自動的にグラフ化されるため、栄養指導で活用することができる。またこのデータは患者本人の他、主治医、看護師、栄養士などの医療スタッフや、家族も参照することが出来る。

結論

医療データベースと機械学習といった最新のビッグデータの構築・解析手法の開発が進んでおり、これまでの解析手法では明らかにできなかったエビデンスが発見されてきている。AI・ICTを活用した新規のアプローチにより、診療に活用される新しい治療法が開発され、患者の予後が改善されることが期待される。

参考文献

1. 柏原, 桑原, 長州, 岡田. 包括的慢性腎臓病データベース(J-CKD-DB). 日腎会誌. 2017;59(7):1034-41.
2. Nakagawa N, Sofue T, Kanda E, Nagasu H, Matsushita K, Nangaku M, et al. J-CKD-DB: a nationwide multicentre electronic health record-based chronic kidney disease database in Japan. Sci Rep. 2020;10(1):7351.
3. Sofue T, Nakagawa N, Kanda E, Nagasu H, Matsushita K,

Nangaku M, et al. Prevalence of anemia in patients with chronic kidney disease in Japan: A nationwide, cross-sectional cohort study using data from the Japan Chronic Kidney Disease Database (J-CKD-DB). PLoS One. 2020;15(7):e0236132.

4. 日本透析医学会. わが国の慢性透析療法の現況 2019 年末. 2020.

5. Kanda E, Kashihara N, Kohsaka S, Okami S, Yajima T. Clinical and Economic Burden of Hyperkalemia: A Nationwide Hospital-Based Cohort Study in Japan. Kidney Med. 2020;2(6):742-52.e1.

6. Kanda E, Kanno Y, Katsukawa F. Identifying progressive CKD from healthy population using Bayesian network and artificial intelligence: A worksite-based cohort study. Sci Rep. 2019;9(1):5082.

7. Kanda E, Epureanu BI, Adachi T, Tsuruta Y, Kikuchi K, Kashihara N, et al. Application of explainable ensemble artificial intelligence model to categorization of hemodialysis-patient and treatment using nationwide-real-world data in Japan. PLoS One. 2020;15(5):e0233491.

8. 神田. AI・ICTを用いた透析管理 —AIを用いたCKD患者の食事療法管理—. 臨床透析. 2020;36(5):471-6.