

厚生労働科学研究費補助金(臨床研究等 ICT 基盤構築・人工知能実装研究事業)
総括研究報告書

機械学習を活用した診療情報の体系的な把握・分析に基づく、疾患との新たな関連性を発見するための研究

研究代表者 金谷 泰宏 国立保健医療科学院・健康危機管理研究部

研究要旨

稀少疾患は、症例が少ないが故に疾患概念を構築することが難しい。わが国においては平成 13 年度より全国規模で稀少疾患に関する患者情報を登録する特定疾患調査解析システムより症例の集積が行われてきた。本研究においては、これらデータベースを用いて人工知能による機械学習を試みることで、症例数が少なく、臨床所見、画像診断、遺伝子診断を総合的に組み合わせることで正確な診断が得られる多系統萎縮症(MSA)、脊髄小脳変性症(SCD)を取り上げ、人工知能による診断プロセスの妥当性について検証を試みた。また、腎臓疾患についても初診時における臨床所見と検査所見から重症化の可否との推計が人工知能で可能となるかについて検証を試みた。人工知能による MSA の診断結果に関する検証では、SND 及び OPCA はほぼ人工知能の診断結果と専門医の判定は一致することが認められた。一方で、SDS は SND と OPCA に含まれるものがそれぞれ 15%ずつ認められた。SCD に関しては、孤発性および痙性対麻痺でほぼ全数が専門医の診断と一致したが、常優性では 80%、他遺伝性および常劣性は全数で専門医の診断の一致が認められず、孤発、常優性及び痙性対麻痺に分類される等、診断プロセスにおける課題が示唆された。腎臓疾患については、データベースの構築を行い、慢性化の指標を明らかにしたところである。

分担研究者

市川 学	芝浦工業大学 システム理工学部 環境システム学科 准教授
佐々木秀直	北海道大学大学院 神経内科学教室 特任教授
鈴木祐介	順天堂大学医学部 大学院医学研究科 腎臓内科学講座 教授
眞野 訓	順天堂大学 革新的医療技術開発 研究センター 准教授
佐藤洋子	防衛医科大学校 防衛医学研究センター 医療工学研究部門

A . 研究目的

本研究は、厚生労働省が管理する難治性疾患データベースを活用し、人工知能を用いて診断基準の妥当性、診断基準との関連性が高い項目を明らかにする。特に、早期における診断が困難とされる神経疾患、腎臓疾患を取り上げ、専門医による診断と人工知能による診断との乖離を検証するものである。これらの検証を踏まえ、専門医以外により提供された診療情報から確実に対象となる疾患を絞り込めるプログラムを開発し、都道府県等での実装を目指す。

B . 研究方法

B.1 研究計画

初年度においては、厚生労働省における特定疾患治療研究事業によって収集されたデータベースを活用し、機械学習プログラムを利用して神経難病のうち MSA を、

腎臓疾患として顕微鏡的多発血管炎を対象として調査票における登録項目と確定診断との関係性について検討を行う(市川)。さらに、機械学習によって得られた各項目間の関連性について共分散構造分析を用いて検証を行う(金谷)。欧米の登録項目を参考に神経疾患及び腎臓疾患の病態把握に向けたミニマムデータセットを作成する(金谷、鈴木)。北海道大学病院においては、Hokkaido Rare diseases Consortium for MSA (HoRC-MSA)を用いて人工知能による診断プログラムの妥当性の検証を行う(佐々木)。また、ミニマムデータセットを用いた腎臓領域における症例データベースの構築を行う(眞野)。2年目においては、初年度における機械学習の結果を踏まえ、神経疾患(14疾患)及び腎臓疾患(3疾患)を対象とした診断プログラムの構築を行う(市川)。さらに、機械学習により登録項目の妥当性を検証し、神経疾患及び腎臓疾患に対するミニマムデータセットを構築する(金谷)。北海道大学病院及び順天堂大学革新的医療技術開発研究センターにおいて機械学習による診断アルゴリズムの検証を行う(佐々木、眞野)。3年目においては、2年次に設計した神経疾患及び腎臓疾患の診断プログラムの改修と検証を行う(市川)。2年次までの研究で明らかになった予後因子をアウトカムとして登録項目との関連性を共分散構造分析により明らかにする(金谷)。北海道大学病院及び順天堂大学革新的医療技術開発研究センターにおいてミニマムデータセット、診断プログラムの実装を行う(佐々木、眞野)。

B.2 研究資料

研究資料については、厚生労働省に登録された特定疾患治療研究事業・臨床調査個人票データベースを活用する。なお、当該データの利用については、厚生労働省健康局難病課より利用承認を得ている(健疾発 0708 第 1 号。平成 22 年 7 月 8

日)。また、地域における神経疾患のデータベースとして平成 27 年度より AMED の研究支援を受けて開始された Hokkaido Rare diseases Consortium for MSA (HoRC-MSA) との連携を図る。

(倫理面への配慮)

疫学研究に関する倫理指針を踏まえ、各機関において承認を得た。

C . 研究結果

C.1 神経難病データベースの構築

神経難病に関しては、2004~08 年度まで厚生労働省特定疾患調査解析システムに登録のあった脊髄小脳変性症(SCD)の新規登録症例 7,073 例、多系統萎縮症(MSA)の新規登録症例 4,949 例のデータクリーニングを行い解析用のデータベースを構築した。

C.2 人工知能エンジンの設計

機械学習用ライブラリ Chainer (<https://chainer.org>) を用いて 3 層構造からなるニューラルネットワークを構築した。

C.3 人工知能による神経難病の解析

ミニマムデータ・セットとして、神経分野に関して試行的に 初発症状、発病様式・経過、神経学的初見、画像所見、生活状況の各項目(NeuroI Med Chir, 2017)を選択し、機械学習による自動診断を試みた。多系統萎縮症(MSA)としてデータが揃っている 3577 例(SND904 例, SDS387 例, OPCA2116 例)を用いて、各 10 例をテスト用データとして残る 3347 例で機械学習を行った。機械学習と専門医との診断一致率は SND97%、SDS 71%、OPCA88%であった(表 1)。

表 1 MSA 機械学習結果

	Diagnoses by reviewing committee	Diagnostic probability estimated by AI		
		SND	SDS	OPCA
SND	n=10	0.97±0.03	0.02±0.02	0
SDS	n=10	0.15±0.06	0.71±0.08	0.14±0.06
OPCA	n=10	0.12±0.12	0.01±0.01	0.88±0.13

脊髄小脳変性症（SCD）に関しては、データが揃っている 7073 例を対象とした。内訳として、孤発型 SCD (sSCD) 2793 例、常染色体優性 SCD (AD_SCD) 1537 例、常染色体劣性 SCD (AR_SCD) 69 例、その他遺伝型 (Other_SCD) 60 例、痙性対麻痺 (SP) 373 例を用いた。結果として、sSCD 95%、AD_SCD 86%、AR_SCD 0%、Other SCD 0%、SP 83%と AR_SCD とその他遺伝型を除く SCD については 80% を超える正確性を持って専門医と診断結果が一致した。

表 2 SCA 機械学習結果

Diagnoses by reviewing committee		Diagnostic probability with AI				
		sSCD	AD_SCD	AR_SCD	Other_SCD	SP
sSCD	n=10	0.95±0.07	0.04±0.06	0.00	0.00	0.00
AD_SCD	n=10	0.09±0.26	0.86±0.30	0.01±0.03	0.02±0.01	0.00±0.04
AR_SCD	n=10	0.30±0.34	0.38±0.22	0.04±0.07	0.04±0.03	0.19±0.23
Other_SCD	n=10	0.34±0.39	0.60±0.35	0.03±0.03	0.03±0.02	0.00±0.06
SP	n=10	0.09±0.21	0.07±0.19	0.00±0.03	0.00±0.02	0.83±0.31

C.4 腎疾患データベースの構築

腎臓分野については、データセットとして、性別、登録時・発症時の年齢、合併症（感染症、消化器潰瘍、糖尿病、高血圧症、圧迫骨折、骨壊死、心筋梗塞、脳梗塞、悪性腫瘍、DIC）、検査所見（白血球減少、リンパ球減少、血小板減少、0.5g/日以上持続性タンパク尿、抗 Sm 抗体陽性、抗カルジオリピン抗体、ループス抗凝固因子、血清梅毒反応偽陽性）、治療状況（副腎皮質ステロイド、非ステロイド系抗炎症薬、免疫抑制剤、血漿交換療法）を選択した。

C.5 膠原病に伴う腎障害の予測

特定疾患調査解析システムより 2004-2008 年度に申請があった症例のうち、2 年以上連続更新し、初年度において血液透析未実施である 24,591 例を対象として、機械学習により患者毎の人工透析の必要性の是非について検証を試みた。しかしながら、人工透析を必要とした症例は全体の 0.76%にとどまり、機械学習により人工透析の必要の可否について示す

に至らなかった。

D. 考察

D.1 診断基準

診断基準とは、臨床所見、検査所見、画像所見、鑑別診断等の様々な客観的な視点から一定の共通する要素を持ち合わせた疾患の集団を捉えるために研究者の合意に基づいて設定される。このため、新たな検査法の出現に伴い、診断基準は見直される。とりわけ、数が少ない稀少疾患においては、数が少ないことから対象とする集団の全容を把握することは難しく、発症からの時間経過、何らかの介入により臨床病態は大きく修飾される。これが結果として確定診断を遅らせる原因ともなっている。そこで、本研究では、専門医が現行の診断基準で「ほぼ確定」以上の判断を下し、国としてデータ登録を行ってきた特定疾患治療研究事業対象疾患を対象に、登録時の臨床調査個人票データを用いて、機械的に過去の診断プロセスを学習させることで、神経難病の多系統萎縮症（MSA）及び脊髄小脳変性症（SCA）の診断が可能かについて検証を試みた。

D.2 人工知能の設計と課題

人工知能として Chainer を用いたが、当初 2 層で構築したものの回答を導くことが困難であり、最終的には 3 層構造とした。加えて、機械学習を行う上で、欠損の多いデータは除外するか、欠損した情報を補完する必要がある。また、今回の検証では、臨床調査個人票のデータを用いたが、北海道大学における MSA のデータベース（HoRC-MSA）に使用されている患者レコードの構造、使用されている各項目と必ずしも 1 対 1 で対応しておらず、事後、本システムの普及を図る上で、患者レコードの標準化が不可欠であると考えられた。さらに、今回使用したデータベースにおいて、臨床調査個人票において、開眼時立位能力の項目として、「支持なしで立位可能な場合」と「自力立位不

可能な場合」の双方を選択しているケースが相当数散見された。このようなケースをそのまま機械学習にかけた場合、誤った解釈を招くこととなるため、論理的な個人票の設計が求められる。

D.3 MSA 及び SCD の解析結果

MSA 及び SCD の解析に介して、解析に適さない症例が、MSA で 4949 例中 1372 例、SCD で 7073 例中 2241 例が認められる等、今後の人工知能の実装における課題である。今回の検証で、とりわけ欠損値が多い画像情報を外した場合、診断一致率は 70%にも満たない等、データによって大きく影響するもの、そうでないものの判別を進める必要がある。我々が設計した人工知能による MSA の診断結果に関する検証では、SND 及び OPCA はほぼ人工知能の診断結果と専門医の判定は一致することが認められた。一方で、SDS は SND と OPCA に含まれるものがそれぞれ 15%ずつ認められた。SCD に関しては、sSCD、AR_SCD、SP でほぼ全数が専門医の診断と一致したが、他遺伝性および常劣性は全数で専門医の診断の一致が認められず、多くの症例が、孤発、常優性に分類される等、診断プロセスにおける課題が示唆された。

D.4 腎臓領域への応用

SLE 症例を用いて、人工透析が将来的に必要とされるか否かについて、機械学習による予測の可能性を検証した。課題として、全症例 24,591 例のうち人工透析を必要とした症例は 186 例と機械学習に足りる症例数を得ることができず、予測には至らなかった。一方で、cox ハザード解析により、合併症として感染症、高血圧症、免疫抑制剤使用、血小板減少（10 万/mm³以下）、0.5g/日以上持続性蛋白尿が人工透析症例との関連性が高いことが示されることから、入力条件を絞ることで再度、人工知能による予後予測について再検討を試みる。

E . 結論

稀少疾患は、症例が少ないが故に疾患概

念を構築することが難しい。このため、わが国においては平成 13 年度より全国規模で稀少疾患に関する患者情報を登録する特定疾患調査解析システムより症例の集積が行われてきたところである。そこで、本研究においては、これらデータベースを用いて人工知能による機械学習を試みることで、症例数が少なく、臨床所見、画像診断、遺伝子診断を総合的に組み合わせることで正確な診断が得られる多系統萎縮症(MSA)、脊髄小脳変性症(SCA)を取り上げ、人工知能による診断プロセスの妥当性について検証を試みた。また、腎臓疾患についても初診時における臨床所見と検査所見から重症化の可否との推計が人工知能で可能となるかについて検証を試みたが、症例数が極めて少なく結果を得ることはできなかった。

F . 健康危険情報

なし

G . 研究発表

1. 論文発表

- 1) Eto A and Kanatani Y, Countering.Bioterrorism: Current Status and Challenges - A Focus on Pharmaceutical Products and Vaccines -.ADC Letter.2018;5(2): 50-52.
- 2) Uwatoko H, Hama Y, Iwata IT, Shirai S, Matsushima M, Yabe I, Utsumi J, Sasaki H. Identification of plasma microRNA expression changes in multiple system atrophy and Parkinson's disease. Mol Brain, 2019; 12(1), 49.
- 3) Shirai S, Yabe I, Takahashi-Iwata I, Matsushima M, Ito YM, Takakusaki K, Sasaki H. The Responsiveness of Triaxial Accelerometer Measurement of Gait Ataxia Is Higher than That of the Scale for the Assessment and Rating of Ataxia in the Early Stages of Spinocerebellar Degeneration.

- Cerebellum. 2019 Apr 16 [Epub ahead of print].
- 4) Suzuki H, Yasutake J, Makita Y, Tanbo Y, Yamasaki K, Sofue T, Kano T, Suzuki Y. IgA nephropathy and IgA vasculitis with nephritis have a shared feature involving galactose-deficient IgA1-oriented pathogenesis. *Kidney Int.* 93(3), 700-705, 2018.
 - 5) Asao R, Seki T, Takagi M, Yamada H, Kodama F, Hosoe-Nagai Y, Tanaka E, Trejo JAO, Yamamoto-Nonaka K, Sasaki Y, Hidaka T, Ueno T, Yanagita M, Suzuki Y, Tomino Y, Asanuma K. Rac1 in podocytes promotes glomerular repair and limits the formation of sclerosis. *8(1)*, 5061, 2018.
 - 6) Placzek WJ, Yanagawa H, Makita Y, Renfrow MB, Julian BA, Rizk DV, Suzuki Y, Novak J, Suzuki H. Serum galactose-deficient-IgA1 and IgG autoantibodies correlate in patients with IgA nephropathy. *PLoS One.* 13(1), e0190967, 2018.
 - 7) Suzuki Y. B Cells in IgAN and the Rationale for B Cell Directed Therapies in IgAN. *Kidney Dis.* 4, 145-194, 2018.
 - 8) Yamamoto R, Imai E, Maruyama S, Yokoyama H, Sugiyama H, Nitta K, Tsukamoto T, Uchida S, Takeda A, Sato T, Wada T, Hayashi H, Akai Y, Fukunaga M, Tsuruya K, Masutani K, Konta T, Shoji T, Hiramatsu T, Goto S, Tamai H, Nishio S, Shirasaki A, Nagai K, Yamagata K, Hasegawa H, Yasuda H, Ichida S, Naruse T, Fukami K, Nishino T, Sobajima H, Tanaka S, Akahori T, Ito T, Yoshio T, Katafuchi R, Fujimoto S, Okada H, Ishimura E, Kazama JJ, Hiromura K, Mimura T, Suzuki S, Saka Y, Sofue T, Suzuki Y, Shibagaki Y, Kitagawa K, Morozumi K, Fujita Y, Mizutani M, Shigematsu T, Kashiwara N, Sato H, Matsuo S, Narita I, Isaka Y. Regional variations in immunosuppressive therapy in patients with primary nephrotic syndrome: the Japan nephrotic syndrome cohort study. *Clin Exp Nephrol.* Clin Exp Nephrol. 22(6), 1266-1280, 2018.
 - 9) Kamei N, Yamashita M, Nishizaki Y, Yanagisawa N, Nojiri S, Tanaka K, Yamashita Y, Shibata T, Murakoshi M, Suzuki Y, Gohda T. Association between circulating tumor necrosis factor-related biomarkers and estimated glomerular filtration rate in type 2 diabetes. *Sci Rep.* 8(1), 15302. 2018.
 - 10) Inoue M, Yamamotoya T, Nakatsu Y, Ueda K, Inoue Y, Matsunaga Y, Sakoda H, Fujishiro M, Ono H, Morii K, Sasaki K, Masaki T, Suzuki Y, Asano T, Kushiyama K. The Xanthine Oxidase Inhibitor Febuxostat Suppresses the Progression of IgA Nephropathy, Possibly via Its Anti-Inflammatory and Anti-Fibrotic Effects in the gddY Mouse Model. *Int J Mol Sci.* 19(12), E3967, 2018.
 - 11) Suzuki H, Suzuki Y. Murine Models of Human IgA Nephropathy. *Semin Nephrol.* 38(5), 513-520, 2018.
 - 12) Yamasaki K, Suzuki H, Yasutake J, Yamasaki Y, Suzuki Y. Galactose-Deficient IgA1-Specific Antibody Recognizes GalNAc-Modified Unique Epitope on Hinge Region of IgA1. *Monoclon Antib Immunodiagn Immunother.* 37(6), 252-256, 2018.
 - 13) Gohda T, Suzuki H, Hidaka T, Ueda S, Suzuki Y. n Update of Pathogenesis and Treatment in Patients with Chronic Kidney Disease(CKD) and Cardio-Renal Syndrome. *Juntendo Medical Journal.* 64(3), 227-234, 2018.
 - 14) 金谷泰宏、市川学. 超スマート社会で医療ニーズに応え続けるためには何が

- 必要か-IoT, AI を活用した災害医療の研究・開発を主に. 新医療. 2018 June; No.522: pp18-21.
- 15) 禾千絵子、鈴木仁、鈴木祐介. IgA 腎症の発症機序から考える治療の妥当性 IgA 腎症の臨床 - Clinical medicine of IgA nephropathy - 編集:湯村和子. 東京医学社. 東京. 2018. 2-7.
 - 16) 鈴木仁、鈴木祐介. IgA 腎症の新規分子標的薬の可能性. IgA 腎症の臨床 - Clinical medicine of IgA nephropathy - 編集:湯村和子. 東京医学社. 東京. 2018, 187-189.
 - 17) 鈴木仁、鈴木祐介. 血管炎とアレルギー疾患 -内因性・外因性アジュバントの関わり -] Tonsil induced autoimmune/inflammatory syndrome (TIAS)としてのIgA腎症. アレルギー・免疫. 25(6), 802-809, 2018.
 - 18) 狩野俊樹、鈴木祐介. 【腸内細菌叢と腎泌尿器疾患】腎におよぼす腸管免疫の作用. 腎臓内科・泌尿器科. 7(6), 533-538, 2018.
 - 19) 狩野俊樹、鈴木祐介. 扁桃活性化機序からみたIgA腎症. 医学のあゆみ. 267(3), 236-238, 2018.
 - 20) 木原正夫、鈴木祐介. IgA 腎症治療の国際比較. 腎臓内科・泌尿器科. 7(1), 84-88, 2018.
 - 21) 狩野俊樹、鈴木祐介. 【腎疾患診療の未来 最新知見のエッセンシャル: 腎疾患の疫学・病因の研究の進歩】IgA 腎症の発症機構. 診断と治療. 106(4), 423-428, 2018.
 - 22) 出崎雄也、鈴木祐介. 【腎と透析ベッドサイド検査事典】(第11章)免疫血清学的検査 免疫グロブリン クリオグロブリン定性. 腎と透析. 84, 338-340, 2018.
 - 23) 二瓶義人、鈴木祐介. 【特集:腎臓病のエンドポイント】IgA 腎症の病態から考えるエンドポイント The endpoint based on the pathogenesis of IgA nephropathy. 日本腎臓学会誌. 60(5), 592-596, 2018.
 - 24) 禾千絵子、鈴木祐介. 【血管炎(第2版)-基礎と臨床のクロストーク-】小型血管炎 IgA 血管炎 疫学・病態. 日本臨床. 76, 3880392, 2018.
 - 25) 鈴木仁、鈴木祐介. Gd-IgA1(糖鎖異常IgA1)モノクローナル抗体. 腎・高血圧の最新治療. 7(3), 144-147, 2018.
 - 26) 佐藤大介、鈴木祐介. 【CKD(慢性腎臓病)慢性腎臓病のこれまでと今後:最新の実地診療のポイントの整理と活用】IgA 腎症の診断と治療. Medical Practice. 35(10), 1547-1552, 2018.
2. 学会発表
(発表誌名巻号・頁・発行年等も記入)
- 1) Sato Y and Kanatani Y, A comparative analysis of demographic information among 12 neural intractable diseases in a national registry of a rare disease in Japan, 5th Annual Congress on Rare Diseases and Orphan Drugs, 2018/8/30.
 - 2) Kanatani Y and Mizushima H. National registry of designated intractable diseases in Japan: Present status and future prospects, Japan-Korea Joint Session of The 92nd Annual Meeting of the Japanese Pharmacological Society, 2019/3/16, Osaka.
- H. 知的財産権の出願・登録状況**
(予定を含む。)
1. 特許取得
なし
 2. 実用新案登録
なし
 3. その他