

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)
分担研究報告書

人工知能を用いた診断アルゴリズムの設計に関する研究

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

研究要旨

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。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。SCAに関しては、孤発性および痙性対麻痺では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。腎疾患の予後予測について、全身性エリテマトーデス症例を用いて合併症としての人工透析に依存するリスクを予測できるか検証を試みた。

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。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。

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。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。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。

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)。(倫理面への配慮)

該当なし

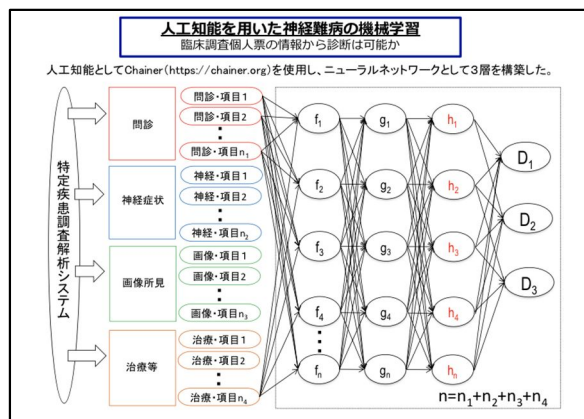
C . 研究結果

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4,957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 (Neurol Med Chir, 2017) を選択し、ニューラルネットワークによる解析を試みた。多系統萎縮症としてデータが揃っている 3577 例 (SND894 例, SDS377 例, OPCA2106 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。この中でデータ欠損値が多い画像診断の有無が診断一致率 (専門医の診断と AI の診断が一致) に与える影響を評価できた。画像診断無しの場合、一致率は SND64%、SDS 0%、OPCA85% であったが、有りの場合、SND と OPCA は 90%、SDS は 70% まで向上した。このように MSA では画像診断の有無が大きく診断結果に影響することが認められた。一方、SCD に関しては、画像診断の結果の有無にかかわらず、常劣型は一致率が 0% にとどまるなど、疾患毎で大きく異なる結果となっている。この背景として機械学習可能な症例数が少ないことがあげられる。

C.4 人工知能による予後予測に関する検証

2004~2008 年度に申請のあった全身性ループスエリテマトーデス (SLE) 患者のうち、2 年以上連続で更新され、かつ、初年度に人工透析未実施症例 24,591 例を用いて機械学習を試みた。なお、これらの症例のうち人工透析を必要とした症例は 186 例 (0.76%) と著しく少ないことから結果を得ることはできなかった。

D. 考察

D.1 人工知能の構築

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.2 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、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。

D.3 人工知能による予後予測

本研究においては、腎障害を伴う膠原病として SLE を対象としたが、機械学習すべき症例が 0.76% と極めて少なく、入力値から人工透析を必要とする症例であるとする考え方は否定された。

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(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。この中で、特記すべきは、常染色体優性 SCD については、機械学習によって高い専門医との診断一致率が認められ、これは、今日の診断基準が論理的であると言える。一方で、常染色体劣性とその他遺伝型は、症例数が少ないことも影響しているが、遺伝子型と表現型がマッチしていない可能性が示唆される。背景として、家族歴情報の欠損など、遺伝子型を決める重要な要素に欠損が多いことも課題であると考えられた。今後は、症例数を広げるとともに、2009 年度以降の新しい症例情報の活用を試みる

膠原病に伴う腎不全の可能性を予測可能かについて検証を試みたが、症例数が極めて少なく、人

工透析の導入の必要性を予測する手法、対象疾患の見直しが必要と考えられた。

F . 健康危険情報

なし

G . 研究発表

1. 論文発表

- 1) 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.

2. 学会発表

(発表誌名巻号・頁・発行年等も記入)

なし

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