

令和7年度厚生労働科学研究費補助金（障害者政策総合研究事業）
発達障害への地域支援に資する継続的な情報収集・活用方法・体制整備に向けた研究
分担研究報告書

臨床データ等を活用した情報システム構築に関する研究

研究分担者 浅尾 高行（信州大学学術研究・産学官連携推進機構 特任教授）

研究要旨

目的

発達障害関連疾患の診療、相談支援、教育現場等に存在するリアルワールドデータ（以下 RWD）の集積に関して持続可能なデータ集積方法の開発と実用性の検証を目的とした。

方法

市町村小児検診で扱う情報及びデータ処理の現状を調査し、使用されている問診票項目、データの集積法、報告書作成過程について調査した。専門外来の初診時間診票を参考に104項目からなるデータ集積用のチェックボックス形式の問診票を作成し、汎用スキャナーによる「手書き問診票集積システム」を試験実装し取り込み、精度の検証を行なった。問診情報と各種スコアの計算結果をデータベース化と同時に電子カルテへの入力を支援するシステムを考案し、業務 DX と RWD 集積を同時に実現する手法としての意義を検証した。

結果

手書き問診票集積システムの認識率は手書き数字の認識も含め100%を達成した。電子カルテ端末の PC のスペックでも実用レベルの速度で稼働した。全診療科の共通の外来、入院問診システムの導入を目指す協力病院にて電子カルテへの入力支援システムを試験導入し、業務改善の可能性と、テスト患者を用いたシミュレーション試験では問題なく稼働することが確かめられた。

結語

継続的疾患データベース構築において、電子カルテへの入力に加えて症例登録のための Web 入力の負担が解決すべき課題である。発達障害で扱われている情報は必要項目の数もデータ量も多く複雑で、現状の方法では相当な入力負荷が想定される。本プロジェクトで開発した新規システムは、専門外来で医師の労力を費やしている問診票の集計とスコア自動計算を行う業務支援と、RWD のデータ集積を同時に実現する手法であり、Web 入力の負担がなくデータベース構築できる有効な手法と考えられた。データ集積を業務改善 DX と関連付けて実現する取り組みは、病院以外にクリニック、自治体、各種施設などに重要な情報が散在している発達障害のレジストリー構築には特に有効性が高い。

A. 研究目的

発達障害関連疾患の診療、相談支援、教育現場等に存在する詳細情報は、確定診断となる検査診断手法が確立しておらず病因が多因子と推定される本疾患群において、疫学、病態究明の重要な情報となる可能性

がある。本研究は、臨床現場以外にも家族や社会など現場に存在する、発達障害に関するリアルワールドデータ（RWD）の集積に関して持続可能なデータベースの構築の検討と、実現のために必要なデータ集積方法を開発することを目的とした。

B. 研究方法

1) 市町村小児検診における調査

松本市、岡谷市において実施されている小児を対応として検診において、問診票項目、データの集積法、報告書作成過程について調査した。

2) 「手書き問診票集積システム」の検討

専門外来の初診時間診票を参考に 104 項目からなる OCR を用いたデータ集積用のチェックボックス形式の問診票を 1 年目にエクセルベースで作成したものをワードベースに作り直した。汎用スキャナーによる自動データ集積システム「手書き問診票集積システム」を試験実装し、下記要件を目指し改良を加え、データベース化の可能性を検証した。また、問診票のテンプレート作製方法をエクセルベースとワードベースの 2 つの方式を比較した。AI による手書き数字の画像認識システムモジュールは、1 年目に外注により調達したものを実装した。それ以外のプログラムは 3)以降のシステムを含めすべて研究分担者が自作し実装検証した。

システム開発における要件は以下の通りである。

<必要要件>

- ・ 質問項目と選択肢の表からチェックボックス形式の問診票テンプレートの半自動作成
- ・ 市販のドキュメントスキャナーでのデータ取り込み
- ・ 画像認識によりチェックボックスのチェックの有無をデータ化
- ・ AI 手書き数字認識システムによる取り込み

- ・ 問診票の読み取りから結果の出力がページ当たり 1 秒以内
- ・ ページ種類、ページ番号、患者識別番号の読み取りは認識率 100%
- ・ 手書き数字の読み取り精度 100%
- ・ スコア自動計算された各種スコアは電子カルテにクリップボード経由でペーస్తు
- ・ プログラミング言語 Python で実装し電子カルテ端末で稼働すること

3) 模擬患者データの読みとり精度試験

問診票のテストデータを作成しスキャンスナップ iX1600、RICOH 社製を用いて取り込み、精度の検証を行った。

4) 電子カルテへの入力支援システムの検討

「手書き問診票集積システム」で取得した問診情報と各種スコアの計算結果をデータベース化と同時に電子カルテへの記載を支援するプログラムを考案し、電子カルテ操作の省力化効果を検討した。また、患者ごとの問診票を印刷する操作の自動化を行った。システム開発要件は次の通り。

<必要要件>

- ・ スタッフが電子カルテに入力、出力している操作をエミュレート
- ・ スタッフがプログラムの知識がなくても変更できる汎用システム
- ・ 可能な限り現状で使用している文書の形態や運用方法を変える必要がない
- ・ メンテナンスが極力少なくて済むエコシステム
- ・ 現状で使用している端末 PC と共有フォルダで実用速度が保てる軽量システム

- ・ 端末 PC の管理情報（レジストリー）を書き換えずに Folder 内でプログラムが完結している
- ・ 電子カルテのほかのシステムと競合しないシステム
- ・ あくまで電子カルテの入力と出力を支援するシステム

C. 研究結果

1) 専門外来で扱う情報の特殊性

- ・ 発達障害の専門外来では問診を中心とした情報が診断や診療の決定に重要であり、その項目数は初診時の問診票だけでも 100 項目以上あり、登録データ入力の負担が大きいことが想定された。
- ・ 病理結果や採血の結果で診断が確定する疾患群と異なり、発達障害の診療は文章を中心とした非構造化データに依存していた。
- ・ 問診結果から 10 種類以上のスコアが利用されているが、計算機を用いた手計算により集計されており、診療開始時間前の医師の作業負担となっていた。
- ・ 病院ごとに使用している問診票は統一されておらず問診テンプレートは病院ごとのオーダーメイドが必要と考えられた。

2) 市町村小児検診での情報と利用

信州大学と研究協定を結んでいる 2 つの町村の高齢者向けの検診を含めて、検診で用いられている調査問診票は一定の要件は満たしているものの、統一の質問用紙使用されているわけではなく市町村独自に項目を設定し、それぞれの方法で実施されてい

た。質問用紙は手書き用紙で、チェックされた項目を職員が県指定のエクセルや Web 入力システムに項目ごとのデータを手作業で入力しており、現場担当者から業務の効率化を求める意見が聞かれた。

3) 「手書き問診票集積システム」の概要

開発テストを繰り返しながら必要要件を満たすシステム開発を実施した結果、最終的仕様は次のようになった。

- ・ 質問と選択肢、項目がチェック項目か自主項目かを記載したエクセルのマスタシートからワード型の問診票を半自動するシステムを採用した。理由は問診票のレイアウトを現状で使用中の質問票と一致させることが容易のためである。
- ・ データ管理に必要な問診票コード、ページ番号、患者番号についてバーコード、QR コード、数字の OCR 認識を比較し問診票コード、ページ番号については QR コードと OCR による文字認識の 2 重チェック機構を採用し、患者番号についてはスタッフが目で見ても認識可能な数字を OCR 認識する方法を採用した。
- ・ チェックボックスや手書き数字、自由入力欄の位置をスキャンデータが変わっても特定できるようにするための共通の原点として右上に配置した QR コードの左上角を採用した。
- ・ QR コードの右下角よりも上側と左側にある画像は以下の物体認識の対象から外すようにした。
- ・ 出来上がったワード問診票のチェックボックス 以外のすべての文字を白文字に置き換え、さらに を に置き換

える操作をワードの置換機能を用いて変換し、印刷した用紙をスキャンしてマッピング画像とした。この画像から物体認識ライブラリーを用いた位置座標取得のための問診票ごとページごとにマッピング画像から相対位置情報を取得して保存するプログラムを作成した。

- ・ 手書きで書かれた問診票をスキャンし問診票のコード、ページを QR コードで確定、その用紙から取り出すボックスの位置を取得、画像認識でチェックの有無、手書き数字の認識、自由記載の有無を判断するアルゴリズムを搭載した。
- ・ 手書き数字を認識する AI モジュールの実装
1 年目に調達した手書き文字認識 AI モジュールを実装した。手書き文字は公開数字文字データ 7 万件[1]を用いて AI 学習モデルを作成しアルゴリズムは Vision Transformer[2]を用いた。

4) 読みとり精度と速度の結果

「手書き問診票集積システム」で 1000 個のテストデータを用い検証での正解率は 100%であった。手書き数字の認識精度も同様に 100%であった。チェックボックスのみからなるテストデータを用いたスキャンから集計までの速度は電子カルテ端末 (CPU i5, 4GB メモリ) での処理で一枚当たりの処理時間は 1 秒で実用範囲内であった。

5) 電子カルテへの入力支援システムの開発結果

本開発システムを全診療科の外来、入院問診を本システムにて運用することを目指

している協力病院にシステムを試験導入し、実務スタッフと協議しながら開発を行った。システムを現場に合わせてスタッフが仕様変更、メンテナンスができるように Low code で運用するシステムとした。開発した電子カルテ支援システムの概略は次の通りである。テスト患者を用いたシステム運用のシミュレーション試験では問題なく稼働することが確かめられた。

- ・ スコアの自動計算アルゴリズム
- ・ テキストテンプレートをを用いたカルテ記載作成
- ・ カルテ記載をクリップボード経由で電子カルテに転記
- ・ ワードテンプレートをを用いたカルテ記載作成
- ・ 表、フォントなどの構造情報を維持して電子カルテ内のワード型文書に転記
- ・ 電子カルテの操作をエミュレートする機能、robotic process automation

D. 考察

これまでの登録中止となった疾患データベースプロジェクトの報告[1]には、臨床症例データ集積に関する共通する課題がまとめられている。データの集積が一時的に達成されたものの結果的に少数の特定の病院からの登録に偏り、さらに参加施設が増やせなかった原因として、入力の手間、維持コスト、有用な 2 次情報を得るのに時間とコストがかかるなどが報告されている。電子カルテへの入力に加えて、Web 入力の負担を考えると、診療現場の多忙化が加速している臨床の現状ではプロジェクトへの参加が増えないのも納得できる。

発達障害で扱われている情報は、これまでのデータベースプロジェクトに比較して、必要な項目もデータ量も多くかつ複雑で、しかも構造化しにくい自由入力テキストでカルテ記載されていた。一般的な Web 入力でこのような RWD を集めるためにはこれまでの Web 入力を基本とした方法では対応できないと判断された。実現のためには、これまでにはないデータ集積法が必要と考えられた。少なくとも入力が業務負荷になるような登録システムは継続的運用が困難なことが想定される。

一方で、発達障害の疾患情報の特性として、患者や家族の問診情報が重要であることが判明した。本プロジェクトでは、現状で多くの労力を費やしている問診票の集計とスコア計算を自動実行する業務支援と、RWDのデータ集積を同時に実現する新規システムの発想を得て、多忙な業務を軽減し、入力負担がなく、コストも安価なデータベース構築を目指して試験的開発と POC を検討した。

その結果、開発した「手書き問診票集積システム」は、業務の軽減を契機として参加者にとっての導入のメリットは大きく症例登録施設としての参加の契機になると考えられた。また問診票の内容をスコア化する前のリアルデータを集積できる点もこれまでのデータ集積では実現できなかった利点である。

本システムを病院に導入する際の注意点として、現場の特に受付業務のワークフローに変更が導入障害となりえる可能性が示唆された。本研究においても、今まで無記名の間診票を患者に渡すだけだった業務が、患者ごとに印刷して準備する業務に変

更することが容易ではなく実装できなかった事例を経験した。受付事務作業が病院職員でなく、外注派遣会社に所属し契約で業務が決められていることも一因と考えられた。

そこで導入をスムーズにする対策として、電子カルテを介さずに Python で問診票を直接印刷できるシステムも併せて開発した。電子カルテからの外来予約データを利用して一括印刷できることも可能で、協力病院では翌日の予約情報から患者 ID などデータの入った外来問診票の一括印刷が可能となっている。電子カルテへのスキャナー登録用のバーコードも含めて印刷されるため記入後の電子カルテへの取込みもワークフローの変更も伴わずに試行できる。このプログラムにより医師のみで問診票印刷からスキャン、電子カルテへのデータの取り込み、データ登録も完了する **One man operation** を可能とするエコシステムとしても運用も可能となった。今後の普及を考えると受付事務にとっても導入により「日常業務も楽になる」システムでなければ多くの病院での導入には難しい。データ集積システムを単独で導入するのではなく電子カルテの入力支援システムの一部として一括導入することが肝要と考えられた。

本プロジェクトで手掛けた、臨床研究のためのデータ集積を業務改善 DX と関連付けて実現する取り組みは 2025 年度の信州大学の第三期先鋭領域融合研究群[2]への採択を得て診療科や職種を超えた継続的取り組みとして多分野に発展することとなった。今後、各分野で実装を進め有効性を検証する必要があるが、本対象疾患のように病院以外にクリニック、自治体、各種施設

などに重要な情報が散在している疾患レジストーリー構築には、特にここで開発検証した手法が有効と思われた。

E. 結論

発達障害で扱われている情報には登録の必要なデータが多く、データ入力の負担が解決すべき課題である。データ集積を業務改善 DX と同時に解決する本法は、RWD の持続可能なデータ集積法として有望である。

F. 参考文献

[1] The MNIST database.

<https://huggingface.co/datasets/ylecun/mnist>

[2] An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby : Published: 13 Jan 2021, Last Modified: 12 Oct 2025ICLR 2021

[3] 日本リハビリテーション協議会
<https://square.umin.ac.jp/JARD/index.html>

[4] 信州大学 第三期先鋭領域融合研究群_標準化医療データ活用研究拠点
<https://shinshu-meddata-mars.org/>