

厚生労働行政推進調査事業費補助金
(難治性疾患等政策研究事業(難治性疾患政策研究事業))
分担研究報告書

難病患者データの活用方策の検討 - 1 -

研究分担者 松山 晃文 (国立研究開発法人医薬基盤・健康・栄養研究所 創薬資源部 部長)
秋丸 裕司 (国立研究開発法人医薬基盤・健康・栄養研究所
難治性疾患治療開発・支援室 研究調整専門員)
羽鳥 裕 (公益社団法人日本医師会 常任理事)
王子野 麻代 (日本医師会総合政策研究機構 主任研究員)
研究協力者 澤 倫太郎 (日本医師会総合政策研究機構 研究部長)

(研究要旨)

平成 29 年後半より難病データ登録システムへの患者データの入力スタートする。今後、300 万件以上の難病患者のデータベースの蓄積がなされる中、現時点で考えられるデータ活用方策及び個人情報保護法改正による利用上の課題を検討した。

- (1) 指定医が記入する臨床調査個人票データの精度向上のため、診断基準等の解説テキスト作成を行った。
- (2) 難病データベースの活用方策と課題を提案した。
- (3) 難病データベース単体の活用以外に、小児特定慢性疾患データベースや Clinical Innovation Network(CIN)などとの連携及び利活用における個人情報保護法改正に伴う同意取得のあり方や倫理審査の必要性を議論した。
- (4) 難病データベースの利用者が企業であり、且つ疾病研究以外の創薬研究に利活用する場合の課題を挙げた。

A . 研究の目的

平成 26 年から始まった指定難病の患者データ登録システム構築から 3 年半を経て平成 29 年後半に患者データのデータベースシステムへの入力がいよいよスタートする。その間、データ登録システムを取り巻く状況が大きく変わった。すなわち、入力プラットフォームは難病指定医が使用する web アプリから光学文字認識方式(OCR)の臨床調査個人票による文字認識入力への変更、入力業務フローも自治体での認定審査済みの臨床調査個人票の写しを疾病登録センター(仮称)に送付してセンターで一括入力、平成 29 年 4 月より第 3 次指定難

病 27 疾病(既存疾病に追加される 3 疾病を含む)の医療費助成が開始されるため疾病追加の準備がなされた。(詳細は研究課題「難病データ登録システムの開発」を参照)

データ登録システムに入力される患者データは平成 29 年 4 月より使用される OCR 臨床調査個人票(以下、OCR 臨個票)とそれ以前の臨個票(平成 27 年 1 月から平成 29 年 3 月までの旧臨個票。ただし、平成 26 年 12 月までの特定疾患治療研究事業 56 疾患の臨個票(旧々臨個票)での申請も可能な経過措置が取られている)にそれぞれ記載されて各都道府県に保管されている。そのデータ数は、平成 27 年度の認定者数約 94

万3千人及び不認定者と毎年の新規(毎年5%ずつの増加と想定)及び更新申請から推定して、平成29年後半までに約300万件以上の臨個票がデータ登録システムへの入力対象と考えられる。

データ登録システムに収納されたデータベースの活用の際には、様々な課題が横たわっているが、本研究ではデータ登録システムにおける前後の課題を下記のように挙げて検討を進めた。

1. データベースに収納する臨個票データ精度向上のため、難病指定医用の疾病テキスト作成
2. 小児慢性特定疾病データベース(以下、小慢DB)やClinical Innovation Network(CIN)などとの連携

データベース連携とは両者のサーバをインターネットで物理的に繋げる(ハードの連携)だけでなく、データ規格の互換性(場合によってはデータベースフォーマットの統一)、個人情報保護への対応、同意取得、データの匿名化といった倫理的課題などソフトの連携について課題検討が必要である。

具体的には、難病DBと小慢DBとの連携、難病DBと個別疾病の患者レジストリ(例えば、我々が構築したR Square)及び国立高度専門医療研究センター(National Center, NC)が独自に持つデータベースとの連携による臨床開発インフラの整備構想CINの運用に係る課題を検討した。

B. 研究方法

1. 難病指定医向けのテキスト作成

未収集のテキストを収集するため、保健医療科学院の金谷部長及び成育医療センターの掛江室長らに継続して協力をいただきテキストの収集にあたった。

2. 難病DBの活用方策

研究課題「難病データ登録システムの開発」にて構築を進めているデータベースの登録項目や入力業務フローなどを鑑みて、難病DBの活用方策を検討した。

企業数社に対して難病DB利活用の要望についてヒアリングを実施した。

3. 他のデータベースとの連携

(1) 小慢データベース

小慢の神経・筋疾患群に分類される65疾病の中で指定難病と移行連携が可能な17疾病を選択し、該当する指定難病の臨個票の「診断基準に関する項目」が小慢の医療意見書の項目とどの程度一致しているか表1に示す比較のcriteriaに従って調査を行い、疾病毎の項目Match率を算出した。

また、データ移行連携のための課題をそれぞれのデータベースの特徴を検証して抽出した。

(2) CINなどとの連携

CINなどで難病DBと他のデータベースを連携及び利活用する場合の課題を個人情報保護法改正に則して検討した。

(倫理面への配慮)

個人情報の取り扱い等、倫理規定に関連する事項はない。

C. 研究結果

1. 難病指定医用の疾病テキスト作成

疾病関連学会を通じて110疾病のうち、91疾病でテキストが収集できた。

2. 難病DBの活用方策

難病DB自体の活用により把握できる情報は以下に列記した項目が挙げられる。

(1) 疾病に関する情報

疾病毎及び都道府県毎の認定者数・不認定者数

疾病毎の認定・不認定の年齢区分

(2)患者に関する情報

疾病患者の自治体・医療機関レベルの位置情報

他の都道府県に跨って受診する患者数
疾病患者の臨床所見や重症度などの経年変化

軽症者の実態

データ登録だけを行うボランティア患者数

(3)指定医に関する情報

指定医毎の認定・不認定率

(4)臨個票に関する情報

臨個票の記入エラー頻発項目の有無

(5)自治体での活用

現行のデータ登録システム運用の業務フローでは見送られることになったが、データベースに付随する自動診断システムによる1次診断結果の自治体審査会での使用は認定審査の大幅な効率化の改善をもたらす。

3. 他のデータベースとの連携

(1)小慢データベース

小慢DBから難病DBへ移行連携を考える上では、臨個票と医療意見書レベルの検討が必要であることから以下の調査を行った。

臨個票と医療意見書の一致項目

小児慢性特定疾病制度では14疾患群704疾病(平成29年4月より18疾病が追加され、722疾病が対象)が医療費助成の対象である。診断項目の面からデータベース連携(この場合、データベース間のデータ移行)が可能であるかを検証するため、難病臨個票と小慢医療意見書の診断基準に必要なそれぞれの項目を抽出して両者の照合を行い、一致項目のMatch率を算出した。神経・筋疾患に属する65小慢疾病のうち、17疾病の

臨個票の診断項目数と項目Match率を表2に纏めた。17疾病の臨個票と医療意見書で一致する診断項目数は数個~数十個と幅があるものの、項目数の多少に関係なく平均項目Match率は約65%であることが分かった。

データ規格

指定難病のデータ規格は「独自CSV」である。

データベース管理システム

指定難病と小慢のデータベースシステムはそれぞれ「PostgreSQL」と「マルチバリュー(NoSQL)」と最初の構築段階から統合不一致な形式で開発が進められている。データベースの統合を考えるならば、どちらかに統一するためデータベースの抜本的な改修が必要となる。

両制度の受給者番号の連携

小慢疾病制度と難病制度で医療費受給認定された場合、それぞれの精度に固有の「受給者番号」が付与される。現時点では、両者の紐付がデータベース間で出来る状況にはない。

(2)CINとの連携

データベース連携及び利活用の同意取得

難病DBは法律に基づく行政事業のために利活用することが前提で患者データの収集と利活用の法的枠組みがあるため(「行政利用型データベース」に分類)、「人を対象とする医学系研究に関する倫理指針」(以下、医学系指針)の対象外なので、データベース構築にあたっては倫理審査は不要である。また、臨個票に記載された難病データの疾病研究(学術研究)への利用に関する患者の同意は医療費申請時の支給認定申請書に記載されている以下の同意文で取得されている。

「厚生労働省では、難病の研究を推進するため、本申請書に添付した診断書(臨床調

査個人票)をデータベースに登録し、厚生労働省の研究事業等の基礎資料として使用することとしています。」

しかし、上記の限定した研究利用以外の二次利用の場合、例えば、臨床調査研究班が実施する臨床研究等へ協力を求める場合や難病データベースを他のデータベースと利用連携及び利活用する場合は、患者の明示的な同意取得や倫理審査のあり方を検討する必要がある。

注意が必要なのは、個人情報保護法改正により、「病歴その他の診療情報」が「要配慮個人情報」となりオプトアウトでの利用は不可となったことである。行政事業のデータベースであっても、目的外利用で二次利用される場合や他の事業で構築された(患者レジストリによる)データベースとの連携の際には、同意取得の手続きのあり方を検討する必要がある。

さらに、企業の創薬研究は疾病研究(学術研究)に該当しないため、個人情報保護法の適応除外対象とはならないので、オプトアウトでのデータベース連携及び利活用は、医学系指針に則して医薬品医療機器等法により行われる研究活動であれば可能であるが、そうでない場合は、データベースの利用は困難である。

難病 DB 二次利用のデータ保護

将来的に難病データの二次利用が可能になることを想定して、二次利用にあたっては個人情報保護と第三者利用の防止の観点から下記の方策が有効と考える。

- (a)解析データの暗号化
- (b)解析データのコピー制限
- (c)利用者側 PC 端末の MAC アドレス認識による PC 端末の限定利用
- (d)データ複合化の Password コード発行

(3)ゲノムデータベース

臨個票の遺伝学的検査には、変異があっ

た遺伝子名を選択、或いは新規変異遺伝子名を記載するのみで、具体的な遺伝子変異の塩基ポジションや変異の様態(SNV, deletion, insertion, translocation, inversion など)の情報を記載する様式ではない(数疾病を除いて)。複数企業から難病 DB の二次利用に関するヒアリングを行ったところ、希少疾病の発症に關与する原因遺伝子群の塩基変異のゲノム情報を難病 DB から利用したい要望は非常に高かったが、現時点では難病 DB からは変異の詳細情報を利用することはできない。

診断基準に関わる遺伝子検査は保険収載され、その大部分がかずさ DNA 研究所や難病研究班で実施された遺伝子データベース(これらは「診療情報型データベース」)があり、また、AMED のゲノム医療推進のための臨床ゲノム情報統合データベース(「研究利用型データベース」に位置付け)が構築されつつある。しかし、患者のゲノム解析を行うと reference sequence に対して数千～数万の変異(single nucleotide variation)が存在するため、これらの変異が自然変異(spontaneous mutation)か、疾病の原因遺伝子か支持的遺伝子かを見分けることは塩基配列の比較だけでは不可能である。表現型となる臨床所見や検査所見との比較分析が必須で、そのためには遺伝子検査等による塩基変異情報と難病 DB の症状・検査等情報との連携は互いに補完的かつ相乗的な医療情報をもたらすことは明白であるが、連携の課題も多い。

前述の 3 型のデータベースはそれぞれの利用目的のために同意取得されて構築及び利用されているものであるため、相互に連携を進めるためには、同意取得及び倫理審査、必要によっては法的な枠組みを考慮しなければならない。個人情報改正により医学系指針ガイドラインにあるように「ゲノムデータ」(全ゲノムならびに全エクソームシーク

エンス、SNP 解析などから個人を認証できるもの)は「個人識別符号」となったため、同意取得には医学系指針に加えて「ヒトゲノム・遺伝子解析研究に関する倫理指針」(以下、ゲノム指針)も考慮する必要がある。

D . 考察・結論

1 . 臨個票記入の手引き

指定医の疾病理解のため、関連学会に疾病テキストの作成依頼を行ったが、難病情報センターには医療従事者向けに疾病の概要・診断基準・重症度分類が厚生労働省 HP に掲載の局長通知とほぼ同様な内容でアクセス可能である。疾病テキスト作成によって情報過多になるよりは、疾病毎の診断及び臨個票記入の手引きの作成が臨個票記入データ精度の向上には有効であると思われる。難病対策課及び難病普及班(「指定難病制度の普及・啓発状況の把握および普及・啓発のための方法論の開発」班)との連携により手引き作成を検討することを提案する。

また、今後、追加疾病を検討する際には、政策研究班(及び関連学会)に局長通知や臨個票の内容確認だけでなく、記入の手引き(臨床所見、検査所見、遺伝学的検査、鑑別診断、重症度分類への記入に特化した手引き)作成を依頼することが効率的と考えられる。それらの記入の手引きを難病情報センターから発信することも検討されたい。

2 . 難病 DB の活用

330 疾病(病型、類縁疾患を含めると約 1100 疾病)に対して 300 万件を超える悉皆性のある難病データベースの分析データの二次利用は、難病の発症機構、診断及び治療法の研究推進のみならず、難病指定医の臨個票記入の精度向上、認定の公平性確保、自治体審査会の受給判定の客観的正確性の担保など難病政策あり方の改善に向けて正のスパイラルの起点となり得る可能性を有していると考えられる。

難病 DB の二次利用に加え、難病データ登録システムの機能である自動診断システムから判定される一次診断結果の自治体審査会での利用は、自治体の負担軽減には非常に有効である。旧制度の特定疾患調査解析システム WISH で自動診断の正確性に問題があったが、難病 DB の自動診断システムは精度が大幅に改善されているため、システム運用の業務フローの見直しを行い、自治体審査会で自動診断を利用できるよう行政サイドで検討することを提案する。

3 . 小慢データベースとの連携

両データベースとも診断に必要な項目数は限定されるものの、データ移行が十分可能な項目一致(Match 率)を疾病毎に有していることが明らかになった。両データベースのフォーマットが異なることが統合に向けての最大のハードルであるが、診断基準項目のデータ移行だけを論点にするならば、両者のデータ規格を揃えるデータ形式変換ソフト(Mapper)を利用すればデータ共有は可能と考えられる。

また、小慢疾病制度と指定難病制度で付与されるそれぞれの「受給者番号」の紐付は、医療意見表と臨個票の様式の違い及び両データベースの項目の違いにより現状は不可能であるが、将来的に臨個票に小慢疾病の「受給者番号」記載欄を設け、難病 DB に収納できるようにすれば、データ移行連携や将来的なデータベース統合の際には、両 DB の患者データ紐付マスターとして有効利用が期待できる。

4 . 他のデータベースとの連携

難病 DB は悉皆性があり患者の自然歴を追跡できる医療情報を持つ一方、限られた調査項目しか診断しないため、同意が取られた他の利用目的のデータベース、例えば、「研究利用型データベース」で我々が

構築している患者と研究者を繋ぐレジストリ R-Square(Rare Disease patients x Researchers)、個別の患者レジストリ、学会レジストリや診療情報型データベースとの連携が難病研究を加速させる。

これらの連携における課題を個情報改正も踏まえて、アカデミアのみならず企業をも含めた利活用を想定して種々の側面から下記に上げた。

< 倫理・法的面 >

明示的な同意取得

- a. 同意文のあり方
- b. 企業へデータ提供の場合、利用目的の患者への説明の仕方
- c. オプトインによる同意
- d. オプトアウトを可能とするならば、その方策

倫理審査

法的基盤の見直し、整備

連携の基準、規範となる Good Registry Practice のガイドライン策定

利用審査

データ情報公開のあり方

- a. 個人情報保護
- b. 第3者利用の禁止

< 技術面 >

データベース形式の統一

データクリアランス

統合データベースの構築費用

データベース運用管理費用

E . 健康危険情報

なし

F . 研究発表

1. 論文発表

なし

2. 学会発表

なし

G . 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

表 1 臨個票と医療意見書の項目比較の基準

No.	要件	比較内容
1	比較の対象	指定難病の下記に関する箇所を対象とし、それに合致する小慢の項目を検出する。 指定難病：診断基準に関する項目（診断のカテゴリーについては、記載はするが項目としての比較はしない） 小慢：上記指定難病の項目に対応する項目（基本は、1.臨床所見、2.検査所見、3.その他の所見とするが、1～3に記載がなくても基礎情報（身長や体重など）等での合致があれば記載する）
2	リンケージの候補選択	リンケージの候補選択は、マニュアルで実施する。 （優先1）指定難病で抽出された項目を小慢の医療意見書において検索し、合致した項目を選択。 （優先2）Web等にて同義の語を調査し、表記は異なっても同義であると類推された項目を選択。 その際、単位や回答方法が異なっても、上記方法にて合致と考えられる場合は合致とする。
3	候補の比較	合致候補あり(match)と合致候補なし(non-match)の2択とする。 上記2において、完全一致ではなく部分合致であっても候補が記載されている場合にmatchとする。
4	match率の定義	【当該小慢の医療意見書でmatchした指定難病の項目数】/【当該指定難病の臨個票から抽出した全項目数】×100とする。 指定難病の1項目が小慢の複数項目で合致している場合、当該指定難病の1項目がmatchとする。 指定難病の複数項目が小慢の1項目で合致している場合、当該指定難病の複数項目それぞれでmatchとする。
5	項目の抽出記載について	項目（質問）・回答の空白スペースは、臨個票通りではなく適宜短縮する。 項目（質問）の番号は、群内での共通性を確認するため一部削除する。

表 2 神経・筋疾患群における指定難病 17 疾病の臨個票と医療意見書の Match 率

項目 Match 率 = [意見書と臨個票の一致項目数] / [臨個票の診断基準項目数] x 100

No.	臨個票番号	指定難病名	臨個票の診断基準項目数	意見書と臨個票の一致項目数	項目 Match 率 (%)
1	113	筋ジストロフィー	24	16	66.7
2	118	脊髄髄膜瘤	18	10	55.6
3	131	アレキサンダー病	22	14	63.6
4	134	中隔視神経形成異常症/ドモルシア症候群	5	4	80.0
5	138	神経細胞移動異常症	27	18	66.7
6	177	有馬症候群	16	13	81.3
7	156	レット症候群	18	8	44.4
8	158	結節性硬化症	18	14	77.8
9	191	ウェルナー症候群	15	12	80.0
10	192	コケイン症候群	26	16	61.5
11	139-1	先天性大脳白質形成不全症 (ペリツェウス・メルツバッハ病)	8	5	62.5
12	139-2	先天性大脳白質形成不全症 (ペリツェウス・メルツバッハ様病1)	7	4	57.1
13	182	アペール症候群	10	6	60.0
14	181	クルーゾン症候群	9	6	66.7
15	183	ファイファー症候群	10	6	60.0
16	184	アントレー・ビクスラー症候群	14	7	50.0
17	22	もやもや病	16	12	75.0
				Av.	65.2