

(令和) 6 年度厚生労働行政推進調査事業費補助金
(医薬品・医療機器等レギュラトリーサイエンス政策研究事業)
総括研究年度終了報告書

適切な医薬品安全性評価のための国際整合化を考慮した医療情報データベースの
品質管理・標準化に関する研究

研究代表者 横井英人 国立大学法人香川大学医学部附属病院 医療情報部 教授

研究要旨

本研究は、日本における医薬品の安全性評価に資する医療情報データベースの品質管理手法を最適化するための“あるべき姿”を明確にすることを目標に、国際的な品質管理手法の比較、MID-NET®の情報提供医療機関における課題の分析、および電子カルテシステムの仕様調査の各分析を実施することで電子カルテシステムを情報源とする医薬品の安全性評価に資する医療情報データベースの品質を維持、向上させるための具体的な検討課題を浮き彫りにした。これにより、我が国における医薬品の安全性評価の信頼性を高め、医療現場での実用性を向上させるための電子カルテシステムを情報源とする医療情報データベースの国際的整合性を踏まえた品質管理における留意点等の整理を図る。

中島 直樹

国立大学法人九州大学大学院 医学研究院
医療情報学講座 教授

武田 理宏

国立大学法人大阪大学大学院 情報統合講座
医療情報学 教授

石川 慎一郎

国立大学法人佐賀大学 医学部附属病院 医
療情報部 准教授

原田 紗世子

独立行政法人医薬品医療機器総合機構 医療
情報科学部 調整役

■ 国際整合化

日本の医療情報データベースを国際的な基準に整合させ、グローバルな研究や評価に対応できるようにする。

■ 効率的な品質管理手法の確立

データ収集、管理、チェックの各プロセスにおいて効率的かつ効果的な品質管理手法を確立し、運用コストを削減する。

B. 研究方法

以下の方法で国際的整合性を踏まえた信頼性の高い医療情報データベースを構築するための検討課題の抽出とその対応策を検討した。

1. GAP 分析

我が国のMID-NET®の品質管理方法をベースにした米国のSentinel Initiativeと英国のCPRDとの品質管理方法との違いについて比較分析し、“医療情報データベースの品質管理のあるべき姿”を検討した。なお、この検討には、「効率的な品質管理手法の確立」という点から“工数”を考慮した。

そして、①医療情報データベースの利用目的、②データ情報源、③データ集積方法、④データチェック、⑤データベース管理の5つの評価軸

A. 研究目的

本研究の目的は、医薬品の安全性評価に資する医療情報データベースの品質管理手法を最適化し、国際的整合性を踏まえた信頼性の高いデータベースを構築するための検討課題について下記の3つの視点を考慮して整理することを目指した。

■ データの信頼性向上

医療情報データベースにおけるデータの正確性、妥当性、完全性を確保し、医薬品の安全性評価における信頼性を向上させる。

を設定し、各評価軸に対して具体的な選択肢を提示し、それぞれのメリット・デメリットを分析した。

2. 実地調査

医療情報データベースにデータ提供する際の英国および米国の品質管理の状況について現地のデータ提供側の関係者へのインタビューも含めて情報収集し、特に、当該国における診断名登録の品質保証やコーディング、データクリーニングの手法について、我が国との違いと我が国の医療情報データベースの品質管理としての導入の是非について分析した。

3. 電子カルテシステムの仕様調査

主要ベンダー7社8製品に対してヒアリング調査を実施し、電子カルテシステムのデータ管理方法、標準化対応状況、セキュリティ機能などを調査した。“医薬品の安全性評価に資する医療情報データベースの情報源としての電子カルテシステム”という観点から、傷病名や医薬品副作用情報のコーディングを含む管理方法に焦点を当て、情報収集を行った。

4. MID-NET®のデータ提供医療機関における課題検討

MID-NET®の構築時と運用時に情報提供医療機関から提供された電子カルテデータのMID-NET®の統合データベースへの移行に際して見られた問題事例を分析し、医療情報データベースの情報源となる電子カルテデータの品質管理における課題について検討した。

C. 研究結果

以下の主要な知見が得られた。

1. GAP分析（別添資料①参照）

“医療情報データベースの品質管理のあるべき姿”について下記のとおり各評価軸で論点（課題）を整理した。

- ① 利用目的を「医薬品の安全性の評価に限定して利用する」のか、あるいは「医薬品の安全性に限らない幅広い情報分析に利用する」かでデータの収集範囲（種類等）が異なる

ことになるため、医療情報データベースの品質管理は、その集積データの特性に応じた対応が課題となる。例えば、データ収集の目的を医薬品の安全性評価に限定することで、収集すべきデータの範囲を狭め、品質管理が容易になる一方、幅広い研究を対象とする場合はデータの多様性が向上するが、品質管理の工数が増加することなど医療情報データベースの品質管理は集積データの特性に影響を受けることなどを整理した。

- ② 医療情報データベースの情報源（医療機関の種類や規模、データの種類や量など）の特性に応じた対応が課題となる。
- ③ 情報源である医療機関等において既に標準化されたデータ（標準的なコーディングが行われたデータを含む）あるいは統一されたフォーマットでのデータを集積するのか、標準化されていないデータを集積するのかで情報源から医療情報データベースへのデータの ETL(Extract、Transform および Load)における品質管理が課題となる。
- ④ 情報源におけるデータの品質管理（データクリーニングの実施の有無やその方法、データの妥当性や正確性のチェック等）の程度や手法に応じた対応が課題となる。
- ⑤ 情報源である医療機関等からデータが正しく医療情報データベースに取り込まれ、そのデータの品質が医療情報データベースにおいて適正に維持するための対応が課題となる。

2. 実地調査（別添資料②、③参照）

医療情報データベースにデータ提供する際の英国（UK）や米国（US）における診断名登録の品質保証やデータクリーニングの重要性が確認された。

■ UKにおける診断名登録

1) 診療情報登録

- 医師（GP）が診療内容を電子カルテに直接

- 登録するか、ソフトウェアが自動的に記録
- 2) 品質保証
 - GP が診療後、コードの正確性を確認（診療 8 時間に対し 2 時間程度を要する）
 - Compliance Software が地域や疾患領域などの特性が類似したデータと比較し、不整合を確認（約 40%の GP が当ツールを使用）
 - 医療コーダーが電子カルテにアクセスし、適切な診断名や介入方法を選択し、不整合がある場合は医師に問い合わせることでデータの正確性を向上
 - 3) コーディング
 - ICD-10／SNOMED-CT などの標準コードを利用して、各医療情報データベースに情報を登録

UK の特徴として、GP が診療情報の正確性を確認するだけでなく、電子カルテシステムの Compliance Software によって他の類似した（地域等）データと比較し、診断名が正しいかを判断するようになっていることであった。そして、数多くの医療コーダーによる不整合のチェックなど徹底された登録情報の正確性確認が行われていることであった。また、診療内容を自動的に記述するソフトウェアが多くの GP で利用されており、記述の精度も向上していること、GP を対象に年一度程度、データ品質のトレーニングやドキュメンテーションのチェックが実施され、NHS の Health Informatics Services が監査を実施し、その監査は Clinical Audit Platform などのプラットフォームにデータをアップロードして行われていることなどが分った。

■ US における診断名登録

- 1) 診療情報登録
 - 医師が診療内容を電子カルテに直接登録
- 2) 品質保証
 - 医療コーダーが医師から診療情報を収集・確認し、疑義があれば医師に連絡
 - 医療コーダーを対象として、四半期ごとや年に 2 回ほど病院が監査を実施
- 3) コーディング
 - 医師の記録を ICD-10／CPT®コードブックを使用して標準化された医療コードに変換

- ICD-10/CPT コードマッピングリストの最新化を頻繁に（年二回など）実施
- コード化されたデータをデータベースに入力後、請求システムに反映させ、医療請求担当者に送付

US の特徴として、会計システムに対応できるような患者に係る多数の情報を電子カルテに登録する必要があるが、監査の際には 90-95%の正確性が求められ、会計側の承認が得られないと病院の収入が確保できないため、正確性に関する病院のモチベーションが高いことが分った。特に保険会社等の支払い側の権限が強く、特定の病院等が統計データと比較し疑わしいデータが確認された際は第三者に監査の依頼をすることがある。このような背景から医療コーダーのところで不正確な内容が高確率で修正される体制となっており、その監査は UK では GP を対象としていたが、US では医療コーダーを対象に行われていた。医師にはそこまでの責任や正確性は求められないが、医療コーダーにはかなりその正確性が求められ、年に 1 度のトレーニングが医療コーダーに対して義務付けられ、知識のアップデートが求められていることなどが分った。

3. 電子カルテシステムの仕様調査（別添資料④参照）

日本の電子カルテシステムは、“医薬品の安全性評価に資する医療情報データベースの情報源”という観点から下記のとおり「傷病名」や「医薬品副作用情報の管理」および「データの標準化」において改善の余地があると考えられた。

- 1) 傷病名・日付の取り扱いについて
 - 傷病名については、診療報酬請求のためにレセプト記載に用いる傷病名（レセプト病名）が記録されている（ICD-10 とのマッピングはできている）
 - レセプト病名とは異なる、患者の診療に重要な位置を占める「主要な傷病名」が別途付される場合、プロブレムリスト・既往歴として任意で管理できる形となっている
 - 「主要な傷病名」の ICD 等へのマッピング

の可否はベンダーごとに異なっており、統一がなされていない

- 開始日・終了日については、レセプト病名の開始日（初診日）・終了日（転帰確定日）が記録されている
- レセプト病名の開始日・終了日とは異なる、疾患自体の発症日・治癒日が別途付される場合、上記同様プロブレムリスト・既往歴として任意で管理できる形となっている
- 疾患の発症日・治癒日のデータの持ち方はベンダーごとに異なっており、統一がなされていない
- 傷病名・開始日・終了日とも、妥当性チェックはレセプト記載事項のみとなっており、「主要な傷病名や発症日・治癒日」の一元的な情報収集・妥当性チェックはなされていない⇒医学的な傷病名等の正確性が担保されていない
- 「レセプト情報」と、「主要な傷病名や発症日・治癒日」との整合に関する確認もなされていない（多少実態と異なるグレーゾーンの運用ができるようになっている）ものと思われる⇒レセプト病名等の正確性も十分担保されていない可能性がある
- 加えて上記の通り、レセプトごとの情報管理となっていることから、「主要な傷病名や発症日・治癒日」に関する構造化データを、電子カルテシステムから直接取得することは難しい⇒主要な傷病名等に関する時系列での情報取得は困難となっている

2) 医薬品副作用情報の管理

- 医薬品副作用情報については、患者プロフィール等に任意で記載する形となっている
- 従って、日々の経過記録と関連した時系列に入力する仕組みにはなっていない
- 薬剤標準マスターとのマッピング可否はベンダーごとに異なっており、統一がなされていない⇒医薬品副作用情報に関する構造化データの取得は困難となっている

3) データの標準化

- 傷病名（ICD-10）・薬剤（YJコード等）を除き、標準コードの普及は進んでいない
- 特に検査（JLAC10）については、コーディン

グが出来ない、あるいは外付けの変換マスターが必要な製品もみられた

- 4) 薬事を想定した病名・症候管理の観点から見た電子カルテのデータ構成についての考察（別添資料⑤参照）
 - MID-NET®のアウトカム定義バリデーションの結果、本邦では、傷病名单独での陽性的中率（PPV）は、欧米の類似したデータベース（DB）に比較して、著しく低いことが明らかになりつつある
 - 国内の電子カルテを調査した結果、レセプト・DPCにデータ連携する「病名登録システム」以外にあるシステム内の病名のコード化・構造化は遅れていることが判明
 - 医療 DX として今後 FHIR 化される「診療情報提供書」や「退院時サマリ」に記載される種々の病名もコード化が推奨されると思われる。これらの2文書で扱う病名には、「既往歴」や「プロブレムリスト」も含まれる。それらのコード化を行うことは状況を好転させるきっかけとなりうるが、コード化の範囲は大きく拡大することから、システム側および医療機関の対応が可能かについては懸念がある
 - 臨床家の多くは、データのコード化・標準化に注意を払っていない（後からコーダーや AI が対応すればよいと思っている）。疾患には、病態発生の把握が容易な急性期病変（感冒など）と、発生の把握が困難で病態の解釈が医師によっても分かれうる疾患（間質性肺炎）も存在
 - 「医師の働き方改革」による医師の業務時間短縮や種々の医療従事者の待遇改善が進むことで、これまでと同コストで、上記のデータを生成できるとは言えないと思われる
 - このような現状に於いて、本邦の医療 DB の病名データの信頼性については改めて検討する必要があると考える

4. MID-NET®のデータ提供医療機関における課題検討

情報源としての電子カルテデータの医療情報データベースへの移行（電子カルテデータの

抽出：Extract、SS-MIX2 規約に基づくデータ変換：Transform、MID-NET®統合データベースへの変換データの読み込み：Load）において見られた 12 件の問題事例は下表のとおりであった。

1	処方 院内の指示のために利用しているデータ（ex. 持参薬報告）や処方として必要なデータ（ex. 持参薬処方）を電子カルテで管理していない場合があるが、それによって不要なデータ（ex. 持参薬報告）の送信が発生したり、必要なデータ（ex. 持参薬処方）が得られていない状況が発生した。
2	検体検査 見出し（ex. 血液脂質）のみのデータが読み込まれた。
3	外来受診 外来受診の一部のデータが送信されていないや外来受診を示す情報ではないデータが読み込まれた。
4	退院時サマリ 部門システム側での様式変更に伴って日付に関する XML タグが変更され、日付が送信されなかった。
5	細菌検査 部門システムでデータを版数更新し、その情報を電子カルテ側にも共有しているが、処理シーケンスに考慮不足があり、削除→更新ではなく、更新→削除で送信したため、データが最後に削除され、報告書の版数の順番に不整合が発生した。
6	処方 電子カルテでは、医薬品のオーダーをする時の投薬量は指示量という形で出されることからそれが SS-MIX2 規約に基づく変換がされた時にミスが起こり、投薬量（1 日量、1 回量、総投与量）が正しく設定されていなかった。
7	診療科 入院患者の外来受診の場合に、外来受診した診療科ではなく、入院時の診療科が送信された。
8	検体検査 検査結果値の取得先を誤っていた。

9	注射情報 電子カルテでは、投与経路をローカルコード・名称で運用しているが、一部の投与経路の変換マスターに誤った設定があった。
10	注射情報 電子カルテでは、投薬量単位をローカルコード・名称で運用している。一部の投薬量単位の変換マスターに誤った設定があった。
11	細菌検査 薬剤感受性検査の単位は一律「μg/ml」であるため、データベース上では管理されていなかったため、単位に相当するデータが全て空欄であった。
12	検体検査 測定法が変わり、対応する JLAB10 も変わったが、同じローカルコードを使用していたため、変換テーブルの世代管理ができなかったことで、時期によって違う JLAB10 に変換することができなかった。

MID-NET®は施設間差を最小限にするべく行ったデータの標準化に基づき、構築された各協力医療機関に配置される「統合データソース」を、共通的な検索式を用いて検索することで、複数の医療機関のデータベースを統合したのと同様の使用感で、データ解析を行うことができる。その形態はいわゆるデータウェアハウスといえることができ、その構築・データ集積の手法は、データウェアハウスにおけるデータ統合の理念と同一である。即ち、

- データソースから各施設の運用を考慮して、目的とするデータを特定し、抽出(E)する
- 上記のデータに対して、標準化の手法を適用しての変換(T)を行う
- 変換されたデータをしかるべきデータ項目の保存場所に格納(L)する

という 3 つの工程に対して、これまで問題になった点はどのように関わっているかを考察することとする。

MID-NET®における ETL は、電子カルテから SS-MIX 2 標準化ストレージへの ETL と、同ストレージから統合データソースへの ETL の 2 工程に大別できる（別添資料⑥参照）が、後半の ETL

ではエラー発生頻度が少ないと考えられることから、今回は前半の ETL に限定して議論を行う。

1) エラーの検知手法

MID-NET®の品質管理には、データの日次モニタリング・月次モニタリングを通して、システムの運用方法が変わったなどの変化がないかを監視することと、電子カルテの抽出ツール等を用いて、オリジナルデータとの一致性確認を定期的に行うことなどが行われている。

モニタリングにより、何らかのシステムエラーが起きてデータ送信が滞っているなどの問題に即時的に対応でき、またオリジナルデータとの一致性確認により、ETL ロジックや設定テーブルのミスを的確に指摘できており、有効に機能していると考えられる。

2) 指摘された課題についての考察

原田分担研究者により上表のとおり 12 件の課題が提起され、これを ETL のうち、どの部分に関する問題かを検討したところ、主として抽出作業ミスに関連するものが 8 件、変換作業ミスに関連するものが 4 件、格納作業ミスに関連するものが 1 件、いずれにも関連しないと判断されたものが 1 件となった。抽出作業ミスは、システムの考慮不足やプログラムミスの他、設定者の誤謬（データの解釈間違い）によって起こった可能性が示唆された。No. 11 は常に同じ単位として同じ「 $\mu\text{g}/\text{ml}$ 」が用いられていたとのことで、データベース上にはそのデータがなく、送信元としてのデータが存在しない状況であり、その結果 1 つのデータも送信されず、当該項目は全て空欄となっていた事例であった。

変換作業ミスは人為的なミスが 2 件あったが、投薬量（1 日量、1 回量、総投与量）が正しく設定されていなかったとする処方に関する課題が 1 件あり、また残り 1 件（No. 12：後述）については電子カルテの運用方針自体の妥当性も問われることとなった。格納作業ミスとしては、通常、データ更新をする場合に、同じキー ID のデータに対して、先に「削除」処理をし、その後「新規（登録）」処理をすることで更新後のデータが残ると考えられるが、その逆に「新規（登録）」、「削除」の順にデータが送信され、結果的に

データが削除された、という事例であった。トラブルが起きた工程が格納作業であったので、まずそちらで 1 件と数えたが、データ抽出の時点で、時系列通りに抽出をしていないことから、抽出作業のミスとも考えられ、2 工程のミスとカウントした。

3) 電子カルテ側で考えるべき問題

• No. 12 についての検討

この件は、変換テーブルの世代管理が可能で、期間ごとに異なるデータ変換が可能であれば解決できる、という意見があり、それも有効ではあるが、その場合どの検査項目がいつどのような変更が加えられたのかという、いわゆるメタデータの存在が必須となる。そのような「データを説明するためのデータ」は通常、データベース自体に記録されていることはないので、関連文書として適切な文書管理の下で、データ閲覧者に提供されなくてはならない。特に検体検査などは、試薬・検査機器・検査に関する運用により、データの解釈を変える必要がある可能性があり、特に厳密にメタデータを管理する必要がある。そのようなデータがあれば、データの扱いが期間によって違うとき、それに合わせて世代管理された変換ロジックをかけることにより、期間ごとに分けたデータの扱いができることになり、その点は MID-NET®システムの今後の課題と考えられた。

しかし、上記のような解決を行う場合、各医療機関のメタデータの情報（いつ、どのように検査法が変わったかなど）を元に複数世代の変換テーブルを準備する必要がある。一方、測定法が変わり、測定値も大きく変化するような変更に対しては、明示的に違う検査コードを付与して区別すべきという意見もあり、同じローカルコードで運用していた電子カルテ側の運用方法の妥当性も検討するべきとの結論に至った。

• データ発生の経緯に関する課題

電子カルテは、一人の患者に関しても、短期間に大量のデータのやり取りが行われ、その結果がデータベースに記録される。ただし、それぞれの記録が患者にとってどのような意味を持つのかは、病院ごとに違う患者の動線、つまりは運用によって変わりうる。No. 3 はそのよう

な背景を考慮しなければならない。大学病院などでは、非常に細かい業務レパートリーと細分化された部門システムが緊密に連携し合っていることから、その運用を単純化することは難しく、またその経緯を説明すること自体が難しい面があり、臨床現場のステータス管理データを元にしてのデータ処理には課題が残っていると考えられる。

4) 全体を通しての課題

分担研究者の何人かからは、十分な検証・テストの仕様を確立し、それを元に ETL システムを発注することが重要である、との指摘があった。現実的には MID-NET[®]などの DB がどのように利用されるかを想定しないと、妥当な品質管理手法を立案できない点もあると思われるが、少なくとも単純な人為的ミスを排除するようなテスト仕様の検討も必要であろうと考える。

これらの問題事例を今回分析し、その原因について本研究の代表研究者および分担研究者が所属する各医療機関の実情を合わせて議論し、電子カルテシステムを情報源とする医療情報データベースの品質を担保するためには、下記3点の対応が重要であることで意見が一致した。

- 電子カルテベンダーの送信プログラムの標準パッケージでの検証（バリデーション）のみでは不十分で医療機関ごとに実データを用いた確認が必要
- 医療情報データベース構築時だけでなく、送信プログラムに影響を及ぼす可能性がある変更がされていないかの継続的な確認が必要
- 電子カルテのリプレイス等のたびに、医療情報データベース構築段階と同じような実データを用いたバリデーションが必要

D. まとめ

本研究では、医薬品の安全性評価に資する医療情報データベースの品質管理手法について、我が国の MID-NET[®]の品質管理をベースにした英国・米国での医療情報データベースの品質管理の体制や手法との比較（GAP 分析）を通じて、“医療情報データベースの品質管理のあるべき姿”を検討した。その結果、その品質管理は

医療情報データベースの情報源となるデータの特性に大きく影響されるとの結論に至った。そこで、本研究が対象とする「医薬品の安全性評価に資する医療情報データベース」の利用において、その情報源となる電子カルテシステムにおける各種診療情報データに焦点を絞った検討を行うこととした。そして、電子カルテデータの中で医薬品の副作用のアウトカムと成り得る傷病名データに着目し、その電子カルテシステム上での品質管理を中心に各種分析を行った。これにより、医薬品の安全性評価に資する国際的整合性を踏まえた日本の医療データベースの品質管理のあるべき姿を示し、その“あるべき姿”のための課題や留意点などを整理した。今後は、これらの知見を元に『医薬品の安全性評価に資する医療情報データベースの品質確保のための留意点』（仮）という内容のガイダンス的な文書作成を行い、医療情報データベースを利用した医薬品の安全性評価の信頼性と有効性をさらに向上させるための一助となる情報提供に繋げる研究に発展させていく。

E. 健康危険情報

該当なし。

F. 研究発表

該当なし。

G. 知的財産権の出願・登録状況

該当なし。