

厚生労働科学研究研究費補助金

医薬品・医療機器等レギュラトリーサイエンス総合研究事業

**医薬品等の市販後における
有効性、安全性の評価方法に関する研究**

平成17年度 総括・分担研究報告書

主任研究者 竹内正弘

平成18(2006)年4月

目 次

I. 総括研究報告書

医薬品等の市販後における有効性、安全性の評価方法に関する研究	1
竹内正弘	
(資料1) 班会議議事録	
(資料1-1) 第1回議事録	21
(資料1-2) 第2回議事録	25
(資料2) パイロット試験実施計画書(最終版)	29
(資料3) 統計解析結果報告書	
Analysis of Total Cholesterol Data Random Mixed Effect Model	46
(資料4) ITツールの開発	
(資料4-1) データ統合作業実施における技術的課題と検討に関する報告	55
(資料4-2) データ統合の作業	66

II. 分担研究報告書

1. 北里大学病院における電子的患者個人情報の管理方法と多施設共有化の可能性に関する研究(北里大学病院)	97
矢後 和夫	
2. 電子化された患者情報を、複数の医療機関から収集、集積して解析する方法に関する研究(国立国際医療センター)	99
吉野 信次	
3. 電子化された患者情報を複数の医療機関から収集、集積して解析する方法に関する研究実施に伴う国立循環器病センター高度先駆的医療・研究専門委員会の申請および審査結果と電子化データダウンロードにおける問題点の評価、検討について(国立循環器病センター)	103
森下 秀樹	
4. 電子化された患者情報を、複数の医療機関から収集、集積して解析する方法に関する研究(国立病院機構東京医療センター)	113
青木 誠	
5. 電子化された患者情報を、複数の医療機関から収集、集積して解析する方法に関する研究(徳島大学病院)	115
水口 和生、森口 博基	
6. 広島大学病院における医療情報抽出についての検討(広島大学病院)	118
木平 健治	

III. 研究成果の刊行に関する一覧表

なし

IV. 研究成果の刊行物・別刷

なし

厚生労働科学研究費補助金
(医薬品・医療機器等レギュラトリーサイエンス総合研究事業)
平成 17 年度 総括研究報告書

医薬品等の市販後における有効性、安全性の評価方法に関する研究

主任研究者： 竹内正弘（北里大学薬学部臨床統計部門 教授）

分担研究者： 矢後和夫（北里大学病院 薬剤部長 教授）

青木 誠（国立病院機構東京医療センター 診療部部長）

森下秀樹（国立循環器病センター 薬剤部長）

吉野信次（国立国際医療センター 薬剤部長）

水口和生（徳島大学医学部附属病院 薬剤部長 教授）

森口博基（徳島大学医学部附属病院 医療情報部長 教授）

木平健治（広島大学医学部附属病院 薬剤部長 教授）

研究要旨

本研究は、医療機関で電子化されている医療情報を有効利用して、速やかに安全性情報を収集・解析し、適正使用情報を得ることを目指している。この目的に沿って、医療情報の電子化状況を把握するためのアンケート調査、問題点を把握するための多施設市販後調査及びこれに適した IT ツールの開発を行った。

初年度は、全国の病院（施設）を対象に、医療情報の電子化状況を把握することを目的として、アンケート調査を実施した。また、多施設共同調査を実施する 6 つの医療機関について、電子化状況に関する詳細なアンケートを実施した。

第二年度は、HMG-CoA 還元酵素阻害剤の有効性と安全性に関する多施設共同市販後調査（以下、パイロット試験という）の実施計画書の作成、並びにこのパイロット試験を実施するために必要な IT ツールの試作・検討を行った。

最終年度は、パイロット試験について各施設の倫理委員会で承認を取得し実施した。各施設で抽出したデータは、研究班で開発した試作 IT ツールを用いて施設毎にデータ変換及び統合を行った。統合したデータを用い主要評価項目である総コレステロール値について解析を実施した。

統一した実施計画書を用いて、各施設の倫理委員会審査を受けたが、疫学調査に係る倫理指針が公表されているにもかかわらず、施設毎で対応が異なった。今後、倫理委員会の対応が統一されるためには、一般市民に対する医薬品の市販後調査の重要性を啓発していく必要性を感じた。

各施設のデータは、日常臨床を通じて医療機関で蓄積された電子化データであり、なん

らかの調査を目的として入力されたものではない。データ統合を実施するには、項目毎にデータクリーニングが必要であった。また、データ解析には、欠側値や1点の測定値しか測定されていない患者のデータが含まれており、それらの状況に対応可能な統計解析モデルが必要である。今回は、総コレステロール値（TC）について線形混合効果モデルを応用し、解析を実施した。日常臨床のデータを使用したため、不定期にTC値の測定が行われていたが、このモデルを用いて有効性の解析が可能であることが示唆された。ただし、欠側値については、missing completely at random, または、missing at random の条件下で解析可能となった。

A. 研究目的

パイロット試験の実施を通じて、電子化されたデータを利用した市販後の多施設共同調査の問題点を把握する。

1) 各施設での倫理委員会への諮問

各施設の倫理委員会で指摘された問題点を整理し、市販後の多施設共同疫学試験を実施するために必要な事項を整理する。また、研究の実施にあたり、各施設の倫理規定に関して必要な情報を収集する。

2) 各施設でのデータの抽出及び匿名化

各施設で実際のデータ抽出及び匿名化の実施を行い、問題点を把握する。

3) 各施設から提出されたデータの統合及びITツールの検討

提出された多施設のデータの統合を行い、各施設のデータ構造の違いが統合・解析に与える影響など、問題点を把握する。

試作したITツールを使用して、各施設のデータの抽出、多施設のデータの統合に必要なITツールについて検討を行う。

4) データ解析方法の検討

提出されたデータを用いて解析を行い、医療機関の日常臨床のデータを用いた解析の問題点を把握する。

B. 研究方法

各施設で、パイロット試験の実施計画書に基づいて、その実施を倫理委員会に諮り承認を得る。承認取得後、各施設でデータの抽出と匿名化を行い、主任研究者（北里大学薬学部臨床統計部門、竹内教授）に提出する。北里大学では提出されたデータを統合して解析を行い、統合・解析に向けた問題点を検討する。あわせて、市販後の多施設調査に必要なITツールの開発に向けた検証を行う。

1) 各施設での倫理委員会への諮問

各施設で倫理委員会への諮問は以下のように行われた。

① 北里大学附属病院

「HMG-CoA リグラーゼ阻害剤の使用成績調査」の研究実施計画書に基づき、当院B倫理委員会への審査申請を行った。

② 東京医療センター

当院倫理審査委員会への申請書の作成にあたっては、「臨床研究に関する倫理指針」(厚生労働省、医政発1228002号、平成16年12月28日)、「疫学研究に関する倫理指針」(文部科学省、厚生労働省、平成14年6月17日[平成16年12月28日全部改正][平成17年6月29日一部改正])及び「個人情報の保護に関する法律」を参考とし、当院で定められた様式に従って作成した。

③ 徳島大学附属病院

当院倫理指針、「疫学研究に関する倫理指針」(文部科学省、厚生労働省、平成14年6月17日[平成16年12月28日全部改正][平成17年6月29日一部改正])、「臨床研究に関する倫理指針」(厚生労働省、医政発1228002号、平成16年12月28日)及び「個人情報の保護に関する法律」を参考とし、作成した共通実施計画書と申請書を当院臨床研究倫理審査委員会に提出、審議依頼を行った。

④ 広島大学附属病院

病院医療情報管理システムから抽出した診療情報を院外施設に提供することに関して、昨年度の調査結果に基づき臨床試験申請書および臨床試験審議依頼書を作成し、共通実施計画書を添えてIRBに提出し審査を受けた。

⑤ 国立医療センター

本研究はヘルシンキ宣言および「疫学研究に関する倫理指針」(文部科学省、厚生労働省、平成14年6月17日)を遵守して実施する。

データの抽出・加工の第一段階は「疫学研究に関する倫理指針」の「特定の患者の治療を前提とせずに、ある疾病の治療方法等を検討するため、研究者等が所属する医療機関内の当該疾病を有する患者の診療録等診療情報を収集・集計し、院内又は院外に結果を報告する行為。」に該当することから、倫理指針の対象外と考えられる。一方、データの抽出・加工の第一段階は「ある疾病の患者数等を検討するため、複数の医療機関に依頼し、当該疾病の患者の診療情報を収集・集計し、解析して新たな知見を得たり、治療法等を調べる行為。」に該当することから、倫理指針の対象と考えられる。

本研究において、対象者からインフォームドコンセントを受けることが非常に困難であることから、倫理審査委員会の承認を得た後に施設長の許可を受け、インフォームドコンセントの取得を簡略化するものとする。

⑥ 国立循環器病センター

国立循環器病センターの規定に伴い、医薬品等の市販後における有効性、安全性の評価方法に関する研究を進展させるため、共通プロトコル、疫学研究に関する

倫理指針、臨床研究に関する倫理指針及び個人情報の保護に関する法律を勘案し、審査申請書を作成した。今回、研究対象者個別にインフォームドコンセント取得は不可能と考えられたため、疫学研究に関する倫理指針第3の7の規程のインフォームドコンセントの簡略化等に関する細則に従い、研究対象者が含まれる集団に対し、資料の収集・利用の目的及び内容を院内に掲示し、その方法も含めて広報するため、申請用院内広報を作成した。

2) 各施設でのデータの抽出及び匿名化

① 北里大学附属病院

予定症例数：2003年4月1日～2004年9月30日の間に当院でスタチン系薬剤の処方を受けた全ての患者。年齢：全ての年齢を対象とする。性別：男性・女性いずれも対象とする。疾患：高脂血症（基礎疾患は問わない）。また、本研究は既存資料の収集・解析によるものであり、被験者に対し直接的に物理的な侵襲を与えるものではないが、個人情報の取り扱いに関し、個人の権利利益を保護しなければならない。取り扱われる電子情報は当大学病院規定に基づき暗号化を行い連結は施設内のみとし、施設外へ提出される情報には個人を特定できるような情報を一切含まないものとする。

② 東京医療センター

第一段階として、東京医療センターオーダリングシステム上から、上記の診療情報をダウンロードし、CSV形式のデータ構造に変換する。抽出された情報は分担施設研究責任者が管理する。なお、施設内において患者情報は連結可能匿名化し、連結のための対応情報は当院の臨床研究センター長が厳重に管理する（保管期間は、本研究が終了するまでとする）。

第二段階として、施設識別番号と患者シーケンシャル番号で構成されるキー情報を、収集されたデータに付加し、代表研究者の管理責任のもとに、北里大学薬学部臨床薬学研究センターに物理的輸送する。

③ 徳島大学附属病院

当院臨床研究倫理審査委員会の承認後、当院において、分散されている患者情報、処方箋情報、一般検査情報等の医療情報を、施設内で電子的に連結する方法について検証し、施設内の情報統合にかかる問題点と解決策を整理した。

④ 広島大学附属病院

IRBの承認が得られた後、EMIRを用いて診療情報の抽出を行う。抽出結果はCSVで保存する。抽出した診療情報は、MS Excelを利用して匿名化を含む加工を施した後、FDに書き出し、Internetを介してMS Excel fileで主任研究者に送付する。EMIR専用端末内には、匿名化コード連結情報のみを残し、抽出した診療情報は匿名化した情報も併せて全て削除する。診療情報の抽出および匿名化は、全て医療情報部に配置されているEMIR専用端末内で医療情報担当者監督の下で行う。

⑤ 国立医療センター

電子カルテシステムからスタチンを服用している患者情報、処方歴（投薬歴）、検査値を抽出する。抽出したデータについては、患者識別コードをシーケンシャル No. に置換し、情報を匿名化する。

⑥ 国立循環器病センター

高度先駆的医療・研究専門委員会から条件付承認とされ、ダミーデータの使用を条件とされたため、パイロット研究への参加は断念した。

3) 各施設から提出されたデータの統合及び IT ツールの検討

各施設から提出されたデータを統合した。

臨床試験や市販後調査のデータマネジメントシステム（HITCANDIS/DM：日立製作所の製品）構築の経験を活かし、医療機関における医薬品の適正な使用実態を把握することを狙いとした特定の薬効群や患者群で大規模にデータを収集可能な基盤（情報収集拠点医療ネットワーク）を構築するためのシステム化の方式検討をベースにプロトタイプにて、検証すべき技術的課題を整理し評価を行った。

5つの実施医療機関よりダウンロードされた CSV 形式のデータから XML 形式のデータに変換することでデータの統合を行う。

- ① 各医療機関において、収集対象となるデータ項目を医療情報システムからダウンロードする。その際、匿名化対象の情報に関しては、データの提供前に各医療機関にて匿名化等を実施する。データを特定できるようにデータ提供前に施設を識別できるコードと患者のシーケンシャル No. で構成される KEY 情報を付加する。
- ② 提供される CSV 形式などのファイルを XML 形式に変換する際に、CSV などデータ型が存在しないファイル形式では変換時にデータ型やサイズなどを設定する必要がある。そこで、CSV 形式のファイルに関する添付資料として提出された各カラムの属性情報を参照して設定を行う。設定した内容をスキーマとして登録することで、同様のデータフォーマットを取り込む際に再利用することができる。

提供された医療機関からの XML データ毎にデータ構造変換スクリプトを指定することで異なるデータ構造の XML データから特定のデータ構造の XML データに変換する。

4) データ解析方法の検討

3087 名が含まれていたが、220 名の総コレステロール値は測定されていなかったため（欠側値）、統計解析からは、除外された。各自、総コレステロール値の測定時間、測定総数が相違しているため、線形混合効果モデルを応用し、総コレステロール値の経時的な動向、総コレステロール値に影響を与える重要な因子を考察することにした。

（倫理面への配慮）

パイロット試験の実施にあたり、各施設の倫理委員会の承認を得て実施した。匿名化は各施設で実施され、院外へのデータ提供は各施設の倫理委員会の指示に従って実施さ

れた。

C. 研究結果

1) 各施設での倫理委員会への諮問

各施設の諮問結果は以下であった。

① 北里大学附属病院

電子媒体による個人情報の入手には、倫理委員会の許可によって、また当該データの院外への提供に関しては、倫理委員会の許可証並びに当該担当部門の個人情報保護管理者の承諾を受け、更に病院長決裁によって許可された。情報の抽出にあたっては、同一検査における検査項目コードが変更されていたこと、基準値の変更などがあり、連続した経時的データとしての抽出に関して課題が残された。

② 東京医療センター

本研究は、日常診療での患者診療録を基にした“後ろ向きコホート研究”であり、侵襲的項目はなく、また、遺伝子分析を行うものでない。従って、前記の疫学研究に関する倫理指針（第3章7(2)-②-イ）における「新たに資料を採取しない（既存資料等のみ用いる）観察研究」に該当し、倫理的には、研究対象者からインフォームド・コンセント（IC）を受けることを必ずしも必要としないが、ICの簡略化等に関する細則で求められている下記要件を満たしているかが問題となる。①本研究は、研究対象者に対して最小限の危険を超える危険を含むものではない、②当該方法によることが、研究対象者の不利益とならない、③当該方法によらなければ、實際上研究を実施しえず、又は当該研究の価値を著しく損ねる、④適切な場合には、常に、次のいずれかの措置が講じられること。ア. 研究対象者が含まれる集団に対し、資料の収集・利用の目的及び内容を、その方法を含めて広報すること、イ. できるだけ早い時期に、研究対象者に事後的説明（集団に対するものも可）を与えること、ウ. 長期間にわたって継続的に資料が収集又は利用される場合には社会に、その実情を、資料の収集又は利用の目的及び方法も含めて広報し、社会へ周知される努力を払うこと、⑤当該疫学研究が社会的に重要性が高いと認められるものであること。これらの点について本研究との整合性について検討した。

①、③、⑤は、本研究と相反する事項はないと判断した。②については、過去の日常診療に伴う情報を収集することから、その不適切な利用、開示、漏洩から情報を保護する十分な対策がなされているかが問題となると考えた。④については、研究対象者が含まれる集団に対し、資料の収集・利用の目的及び内容を、その方法を含めて広報するにはどのようにしたら良いのかが問題になると考えた。

②については、患者特性の抽出項目を生年・月、性別、身長、体重の限定し、これらに投薬状況、検査データなどを加えたデータを収集する。施設内では、患者情報は連結匿名化し、連結のための対応情報は、当院の情報管理責任者のもとで厳重に管理

する。収集されたデータは、主任研究者の管理責任の下、主任研究者の施設に直接電子媒体を介さず搬送する。主任研究者の施設では、当院のデータは、主任研究者のもとに厳重に管理される。データの保存期間は、本研究が終了するまでとする。

④については、アをどのように行うかが問題であるが、院内に別添1のお知らせを掲示し、匿名情報であったとしても自分のデータを利用することに同意しない対象者については、指定の部署まで申し出てもらうようにした。また研究内容について別添2を掲示すること、また東京医療センター倫理委員会ホームページに掲載することとした。

以上を、当院で定められた様式に従って作成した申請書として倫理委員会に提出した。情報収集、解析を担当する北里大学の倫理委員会で承認がえられた時点で研究を開始することで承認された。

③ 徳島大学附属病院

当院臨床研究倫理審査委員会における審議では、対象となる患者の同意について説明を求められた。しかし、今回の検討対象となる資料は既に連結不可能匿名化されている情報のみを用いる研究であること、また当該情報を医療情報部において適切に管理することを説明した結果、承認を得ることが出来た。

④ 広島大学附属病院

診療情報の院外への提供については、IRBで審議された。その結果、患者への連結が不可能な処理を行った後、院外施設に提供されること。また、当該研究では、追跡や照合を行なうために院外施設で個人を特定する必要がないと考えられることから、患者個人情報の保護について倫理的問題はないと判断され、IRBの承認が得られた。

⑤ 国立医療センター

倫理委員会からの質問に回答し、承認を得た。

⑥ 国立循環器病センター

高度先駆的医療・研究専門委員会に審査申請したところ、審議の結果は次の3か条の条件付き承認であった。

1. 本研究計画では、患者の処方箋情報等を他施設に提供する場合に、患者の同意取得を原則としていないが、患者の同意を不要とする根拠が明確でないとの意見が委員の中からあった。
2. 本研究デザイン上は、必ずしも患者の実データを必要としていないことから、薬剤の有効性及び安全性に関する情報提供に関しては再検討を行うこと。
3. 他施設に対して提供する患者データについては、ダミーデータを使用すること。

以上の審査結果を受け、対象者個別のインフォームドコンセントを取得することは困難であり、またダミーデータを提供することは、研究成果つまり解析結果に影響を与えることとなるため実施できないと判断し、申請を取り下げることとした。

2) 各施設でのデータの抽出及び匿名化

各施設でのデータ抽出及び匿名化の結果は以下であった。

(ア) 北里大学附属病院

(ア) 電子情報として抽出可能な項目

診療科、性別、生年月日、入院・外来区分、各種薬剤の投薬状況、血液生化学検査

(イ) 電子情報抽出にあたっての制限

- ・ ホストコンピュータへの負荷に関する問題（抽出に要する時間：1～数時間迄、抽出を行う時間帯：夜間帯が望ましい）
- ・ 抽出するデータ量：診療上、ネットワークに影響せず上記で処理できる範囲
- ・ 作業人員：ダウンロードプログラム作成に要する人員ならびにダウンロード操作人員（夜間帯）

(イ) 東京医療センター

診療情報データのダウンロード量が莫大であり、項目によっては個別に収集する必要があり、莫大な時間を要した。抽出作業は、サーバーへの負担を軽減するため夜間に行った。

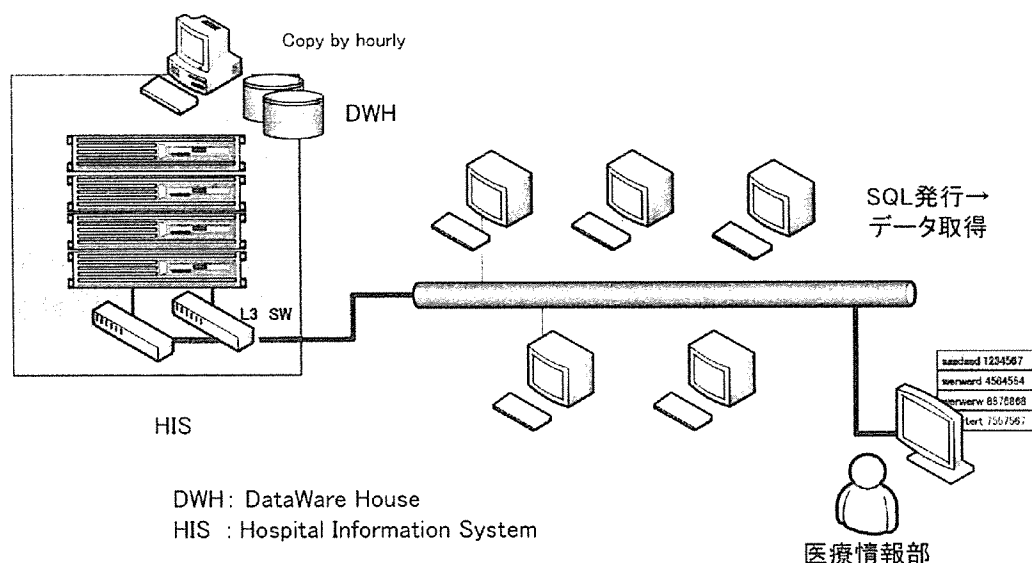
(ウ) 徳島大学附属病院

データ結合に関する検証を行った。以下に手順を示した。

1. 徳島大学病院では、病院情報システムの処方・注射等の診療に関連するデータから1時間ごとに差分データを取り込み、時系列的にデータが参照できるデータベースであるDWH(Dataware House)が稼働している。
2. Microsoft Access を用いてDWHにSQL(Structured Query Language)文の発行を行い、必要なデータを出力した。
3. 出力したデータの中で、患者IDは個人情報保護のため別のユニークな番号に変換した。

以上の結果より、徳島大学病院では患者個人データの収集と匿名化に特に大きな問題点は認められなかった。HIS と DWH の概略図を以下に示す。

厚労省班研究報告書 HISとDWHの概略図



(エ) 広島大学附属病院

病院診療情報システムからの診療情報の抽出は、特に患者個人情報の保護に留意した。診療情報の抽出に利用したEMIRは病院診療情報システムから患者背景、投薬状況、及び検査情報をそれぞれ個別のファイルにコンバートし、診療情報の検索環境を提供するシステムであり、データ抽出はテーブル毎に行わなければならない。データ抽出時に各テーブルで設定できる最大抽出項目数は20項目である。さらに抽出データの出力ファイルも4~5MBに制限されており、これを超過した場合の情報の正確性は保証されない等の制約がある。以上、EMIRの機能制限を踏まえて、今回は抽出作業を、薬品（成分名）ごとに患者背景と病名情報は各1回、投薬情報4回、検査情報3回に分けて行った。プラバスタチンの場合、採用品目である3剤形（メバロチン5mg錠、同10mg錠、同細粒）を同時に検索・抽出したところ、抽出対象期間とした2003年3月1日~2005年3月31日で1,275名の患者が検索され、検索および抽出に要した時間は合計約3時間であった。抽出した情報はCSVで書き出した。分割抽出した情報の結合及び匿名化は、抽出した情報各々に患者IDと日時を含めておいたので、これらをキーにしてMS-Excelを用いて約2時間で完了した。すなわち1薬品（成分名）の診療情報抽出に約5時間、全5薬品の抽出に約25時間を要したことになるが、これは、情報端末の処理時間を単純に累計したものであり、実際には1台しかない情報端末を使用するための待ち時間や、情報端末から抽出データをFDで書き出すためにデータを分割し、さらに外部で再結合するための時間が必要であり、その結果、情報抽出に費やした時間は延べ2週間程度になる。XMLへの変換は既存の汎用ソフト

では対応できなかったので、主任研究者へは Internet を介して MS-Excel file で送付した。EMIR 専用端末内には、匿名化コード連結情報のみを残し、抽出した診療情報は匿名化した情報も併せて全て削除した。

(オ) 国立医療センター

データ収集項目をダウンロードするため、患者基本情報、外来患者履歴、患者入院履歴、処方データ、注射データ、病名データ、検査結果データ、血圧データ、以上の 9 データベースから必要なデータをダウンロードした。ダウンロードした項目は、診療科、性別、生年月、身長 (cm)、体重 (kg)、患者区分、診断名、合併症、HMG-CoA リダクターゼ阻害剤の投薬状況、フィブラート系薬剤の投薬状況、併用薬の投薬状況、血液生化学検査、血液検査、肝機能検査、腎機能検査、尿検査、血清脂質検査、血圧、以上の 18 項目であった。

9 データベースからダウンロードした 9 ファイルについて、患者 IDNo をシーケシャル No に変換した。変換した 9 ファイルのデータがシーケシャル No をキーに連結可能であることを検証した。

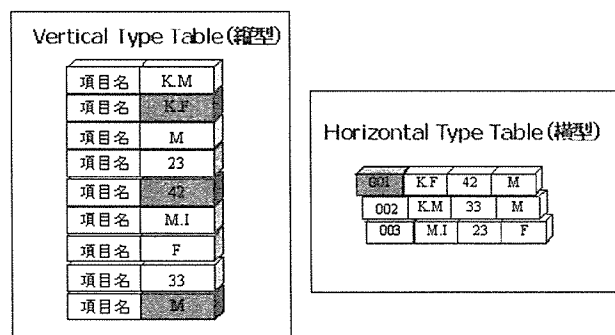
3) 各施設から提出されたデータの統合及び IT ツールの検討

各医療機関から提供された CSV ファイルとファイルのフォーマットを記述したファイル仕様書をもとに各医療機関側の XML データを生成し、予め定義しておいた本研究側の統合データベース側の構造 (Schema) を生成し、各医療機関側の XML データと統合データベース側の構造 (Schema) との Mapping 作業を行いデータ変換することでデータ統合を実現した。

① ファイルフォーマット

各医療機関から提供されたダウンロードファイルの内容を分析すると大きく 2 種類のタイプに分類することができる。

下図に提供されたファイルの概要を示す。



患者ごとに複数の項目が横一列に並んでいるタイプのファイルで Horizontal Type Table(横型)形式と 1 項目ごとに縦一列に積み上げたタイプのファイルで Vertical Type Table(縦型)形式に大別できる。

② データ量

各医療機関から提供されたデータ量を下表に示した。

(単位：人)

患者数	北里大 学病院	徳島大 学病院	広島大 学病院	東京医 療セン タ	国立循 環器病 センタ	国立国 際医療 センタ
全体	5577	1310	2437	2407	—	3087
男性	2668	506	1008	1040	—	1362
女性	2909	804	1429	1429	—	1725
外来の み	2390	677	1232	—	—	1653
入院の み	1204	264	174	—	—	116
外 来 / 入院	1320	369	915	—	—	1318

③ 統合側のデータフォーマット

統合側のデータフォーマットとしては、Vertical Type Table(縦型)形式が相応しいと考えられる。

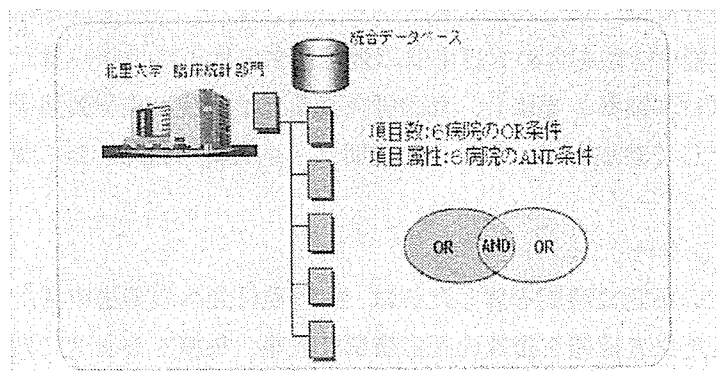


図 統合データベースの定義

上図に示したように、統合データベースの定義に関しては、項目の数は参加病院の OR 条件で設定し、項目属性は参加病院の AND 条件で設定した。

4) データ解析方法の検討

解析方法を検討するために、国立医療センターのデータを用いて検討した。患者数は、全体で 3087 名であったが、220 名の患者では、解析の主要評価項目である総コレステロール値が測定されていなかったため、解析からは除外した。そのため、2867 名が解析対

象者となった。患者毎に、総コレステロールの測定時間、測定数も異なっており、解析対象者の全てのデータを統計解析に使用するため、経時データ解析の代表的なモデル、線形混合効果モデルを適応した。総コレステロール値は、平均して、100 日毎に 1.67mg/DL 減少、女性患者は男性患者と比べると、平均して 15.54mg/DL 高く、10 年毎に 5.7mg/DL 減少しているという解析結果が得られた。本研究では、副作用発生頻度についても解析する予定であった。しかしながら、副作用情報は電子化されていなかったため、電子的に統合できず、解析はできなかった。

D. 考察

1) 各施設での倫理委員会への諮問

① 北里大学附属病院

本研究は「疫学研究に関する倫理指針」（文部科学省、厚生労働省）第3の7の規定によりインフォームドコンセントの取得は簡略化できるものと判断されたが、1)医療機関の診療情報の取り扱いであるにも関わらず本研究に製薬協がオブザーバーとして参加している理由、2) 個人情報管理者には研究代表者、研究協力者以外の人を指名すること、について返答を求められた。1)については、本研究が承認後の再審査、再評価の際に市販後情報を収集することを想定しており、その中には承認条件など法的枠組みの中で企業により実施される市販後調査が含まれ、実際に市販後調査を実施する当事者である企業の意見を聴取することが必要であること、また、今回実施されるパイロット試験について、実施可能性を含めて日米欧の製薬企業の代表に批評や意見を戴く観点から、製薬協関係者（日本製薬工業協会、米国研究製薬工業協会、欧州製薬業団体連合会）がオブザーバーとして参加していることを説明し了承を得られた。2)に関しては、指摘に従った。

② 東京医療センター

倫理委員会での承認に時間を要したのは、倫理委員会への審議申請が多数あったためと、電子化された患者情報を複数の医療機関で収集、集積する研究に関しては、初めての審議であったためである。本研究に関しては倫理委員会委員より、医薬品の有効性と安全性の情報が得られると一定の評価と大きな期待がされた。

③ 徳島大学附属病院

「患者特定が行えないような情報」について適当な範囲が示された指針などは存在しないが、個人情報の匿名化、すなわち当該個人情報から、当該情報に含まれる氏名、生年月日、住所等、個人を識別する情報を取り除く事で、特定の個人の識別はできないと考えられ、また当院臨床研究倫理審査委員会においても承認を得ることができた。

④ 広島大学附属病院

診療情報を院外施設で利用することに関して、患者個人への連結が不可能な匿名化処理により倫理面の問題はクリアされると判断された。しかしながら、実作業では患者個人情報を抜くことから、匿名化コード連結情報の取り扱い含めて、診療情報の取り扱い方法に関する具体的なマニュアルが必要であると思われる。

⑤ 国立医療センター

倫理委員会からの質問に対し文書による回答を行い、本研究に対する理解を得ることができた。

⑥ 国立循環器病センター

高度先駆的医療・研究専門委員会の審査結果が条件付承認となったが、いずれの条件も満たすことが困難であった。第1の患者同意を不要とする根拠が明確でないとの意見に対しては、疫学研究に関する倫理指針に基づく申請が理解されていないことが明らかであった。また第2の本研究デザイン上は、必ずしも患者の実データを必要としないとの意見に対しては、この研究の趣旨が電子化された患者情報を複数の医療機関から収集、集積して解析する方法の検証と捉えられていたため、薬剤の効果と安全性の signal detection としての解析の重要性が認識されていなかった。そのため、第3のダミーデータの提供指示がなされたと考えられる。以上のことから、データの提供、収集等が必須となる多施設共同研究において高度先駆的医療・研究専門委員会の審査申請は、規定や倫理指針を遵守した審査申請書を作成提出するだけでなく、各委員に対して研究デザインについての理解を得ることが重要であると思われる。

5) 各施設でのデータの抽出

① 北里大学附属病院

情報の抽出にあたっては、同一検査における検査項目コードが変更されていたこと、基準値の変更などがあり、連続した経時的データとしての抽出に関して課題が残された。

② 東京医療センター

今回の抽出にあたっては、サーバーへの負担を軽減するため、データのダウンロード作業は夜間に行った。そのため、収集には時間と労力が必要であった。

当院では、2006年1月より電子カルテが稼動したため、将来的には二次利用の活用が可能であるが、今後はいかにして各施設共通の情報を抽出できるかが課題である。

③ 徳島大学附属病院

実際の検証では患者 ID は個人情報保護のため別のユニークな番号に変換が可能であった。また必要な患者個人データの抽出も可能であると考えられた。

④ 広島大学附属病院

EMIR を利用した診療情報の抽出は、同検索ツールの種々の制限のため非常に時

間を要した。EMIR のような簡易検索システムを、本研究のような大規模な疫学調査を目的に利用することは不適切であると思われる。

今回、当院で実施した診療情報の抽出および院外施設への提供において、その情報の信頼性および匿名性に関する信頼性は作業者に大きく依存している。本研究の目的である複数施設からの診療情報の収集を実現するためには、各施設に合った情報抽出方法の確立と、匿名化や本研究で共通フォーマットとした XML データへの変換を自動化できるソフトウェアの開発が必要であると思われる。

⑤ 国立医療センター

今後は、データの二次利用を考慮したシステムの開発が必要であり、院内でデータ処理できる部署の設立も検討する必要があると思われる。

電子カルテシステム自体の問題として、当院では一部に手書き処方箋・注射箋がシステム化されずに運用されている。また、システム構造の問題として、データの解析に必要な有害事象、患者のアレルギー歴、副作用歴、嗜好歴などが記録できるシステムではないことも挙げられる。今後、電子カルテの二次利用を行ううえで、完全にペーパーレスとして運用できる完成度の高いシステムの導入が必要である。

6) 各施設から提出されたデータの統合及び IT ツールの検討

今回実施した実証実験の結果を当初設定した技術的検証項目に沿って、統合作業中に発生した課題と対応について考察を行う。

① データ統合について

機械的にデータの統合作業は進んでいくが、統合されたデータの内容を詳細に見ていくと統合後にデータ処理を行う際に問題を発生させることになる状況が幾つか発生した。ここでは、その代表例を紹介する。

表 6：重複データの存在

患者 ID	採取日時	検査項目名	検査値
同一患者	2004/09/16 8:54	ASL (GOT)	22
	2004/09/16 8:54	ALT (GPT)	22
	2004/09/16 8:54	ALT (GPT)	22
	2004/09/16 8:54	LDH	199
	2004/09/16 8:54	ALP	150
	2004/09/16 8:54	γ GTP	13
	2004/09/16 8:54	γ GTP	12

表に示すように、同一患者で採決日時が同じ検査データが存在し、検査結果が同じものや違うものまで存在する。

表：履歴データの存在

患者 ID	実施日	実施時間	血圧	結果
同・患者	20031226	144248	血圧（上）	174MH mmHG
	20031226	144248	血圧（下）	174MH mmHG
	20031226	144248	血圧（上）	174MH mmHG
	20031226	144248	血圧（下）	76MH mmHG

表に示すように、血圧の上下に同じデータを入力してしまったため、再度正しいデータを入力し直している形跡が見られる。このようなケースは、直接データの内容を見ないと判断できないことが分かる。

表：明らかに間違ったデータの存在

患者 ID	実施日	実施時間	血圧	結果
同・患者	20030805	104425	血圧（上）	0MH mmHG
	20030805	104425	血圧（下）	130MH mmHG

表に示すように、血圧の上下を逆に入力されているケースが存在した。これも、直接データの内容を見ないと判断できない。

②メッセージ交換について

今回作成した医療情報統合ツール(XML Transform)により、基本的には解決することが出来た。XML Transform により、各医療機関側から提供されるデータ構造 (Schema) とセンタ側の統合データベースのデータ構造 (Schema) との対応付けを行うことでデータのメッセージ交換は実現した。

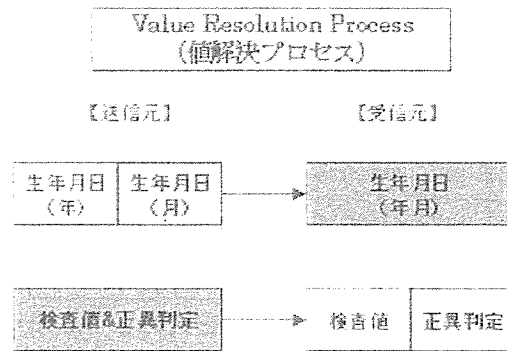
参加 6 病院の定義作業が必要なため、今後医療機関側での標準化が進めば変換作業のための定義が 1 つに集約することが可能になる。

医療サイドの国際的な標準仕様である HL7 Ver3.x に準拠した電子カルテシステムの導入が医療機関において普及してくると各医療機関からの患者の診療情報の電子的な提供を HL7 CDA Release2 に基づいて入手できる基盤が出来上がると考えられる。

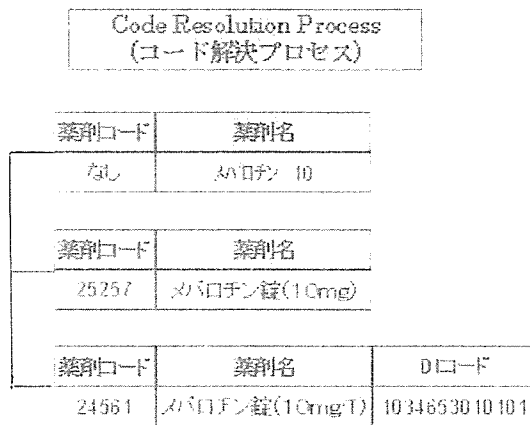
③データ変換について

医療情報統合ツール(XML Transform)の紹介の【データ変換】で説明した方法により、統合されたデータが統一したデータ体系へ変換が行われ、データ処理において大きな効果が得られた。

データ変換の方法は、2つのアプローチが存在し、データを結合・分割などして変換する方法と変換テーブルを参照するコード変換する方法がある。



図：値解決プロセス



図：コード解決プロセス

④タイムスライスについて

統合したデータを統計解析する場合、複数の患者データを投与開始から同じ相対日時のタイムフレームに揃えるという前処理が必要になる。

表：検査間隔の違い

検査回数	患者 A		患者 B	
	リピートル 10		リピートル 5	
	検査日	検査間隔	検査日	検査間隔
1 回目	2003.6.7		2004.6.28	
2 回目	2003.6.9	2 日月	2004.8.27	60 日月
3 回目	2003.6.10	1 日月	2004.12.27	122 日月
4 回目	2003.6.11	1 日月	2005.1.17	21 日月
5 回目	2003.6.13	2 日月	2005.2.28	42 日月
6 回目	2003.6.16	3 日月		
7 回目	2003.6.17	1 日月		
8 回目	2003.6.18	1 日月		
9 回目	2003.6.19	1 日月		

(同一薬剤/同じ用量)

検査回数	患者 C		患者 D	
	メバロチン 10		メバロチン 10	
	検査日	検査間隔	検査日	検査間隔
1 回目	2003.12.15		2003.6.10	
2 回目	2004.1.19	35 日月	2003.7.8	28 日月
3 回目	2004.1.22	3 日月	2003.8.5	28 日月
4 回目	2004.1.23	1 日月	2003.9.2	28 日月
5 回目	2004.1.24	1 日月	2003.9.30	28 日月
6 回目	2004.1.29	5 日月	2003.11.11	42 日月
7 回目	2004.2.2	4 日月	2003.12.16	35 日月
8 回目	2004.2.6	4 日月	2004.1.20	35 日月
9 回目	2004.2.9	3 日月	2004.3.2	42 日月

実際に統合したデータの検査日と検査間隔の状況の事例を表に示した。

投与開始ポイントの決定は、今回の実証実験における対象治療期間である 2003 年 4 月 1 日から遡って 1 ヶ月前のプレ情報を収集し、このプレ期間にデータの発生がなく、それ以降初めて発生したデータを初診データと位置付けることにした。つまり、プレ期間にデータが発生している患者は、それ以前から治療が行われていると判断した。

⑤重複投与について

特に、外来患者においては、来院の際に薬剤の処方を実施されるため、患者の来院の都合により、処方量の重複が予想される。

表：重複処方の状況

処方回数	患者A			患者B		
	メバロチン 10			メバロチン 10		
	処方日	処方日数	重複日数	処方日	処方日数	重複日数
1回目	2003.4.15	6日		2003.4.17	56日	
2回目	2003.4.21	7日	0日	2003.6.12	49日	0日
3回目	2003.4.28	14日	0日	2003.8.14	63日	-14日
4回目	2003.5.1	21日	11日	2003.10.16	49日	0日
5回目	2003.5.21	30日	1日	2003.12.4	56日	0日
6回目	2003.6.25	30日	-5日	2004.1.22	30日	7日
7回目	2003.7.24	30日	1日	2004.2.19	42日	2日
8回目	2003.8.20	30日	3日	2004.4.15	55日	-14日
9回目	2003.9.17	30日	2日	2004.6.9	56日	0日

(同一薬剤/異なる用量)

実際に外来患者の処方状況を調べて見ると表8に示すような重複処方の状況が確認された。処方の前後において、処方量が重複するケースと処方量の不足が発生するケースが存在した。

解析を実施する際に、それぞれのケースをどのように取り扱うかがポイントとなる。さらに、処方量と実際の服薬量との違いがあることも重要である。

また、今回の参加医療機関においては、倍量処方は行われていなかった。

⑥欠測値(Missing Data)について

各医療機関から提供されたデータを統合し、そのデータの内容を参照して見ると、当初の予想通り、非常に多くの欠測値の存在が確認できた。

図は実際にデータ参照を行った際の欠測値の分布状況です。