

特集：臨床研究実施の現状と課題

第2部 医療データベースの疫学研究への活用：医療データベースの研究利用と社会還元

電子カルテデータの研究利用と標準化

東京大学大学院医学系研究科医療情報経済学分野

大江 和彦

はじめに

電子カルテシステム（以下では単に電子カルテ）やオーダリングシステムを導入する医療機関が漸増しており、特に大学病院や病床規模の多い病院での導入が着実に進んでいる。平成20年度の医療施設静態調査（厚生労働省）によれば、図1のように200床以上の病院におけるこれらシステムの導入率は20～90%である。オーダリングシステムでは処方、検査結果、病名データなどが電子化されているが、電子カルテではこれらに加えて日々の経過記載も電子化される。しかし、日々の経過記録などの電子カルテ特有のデータは文章で入力されることが多いので、このデータから研究利用可能な情報を得ることは技術的に困難である。そこで、本稿では、電子カルテデータのうち主として処方、検査結果、病名など、オーダリングシステムにおいても電子化されているデータを、研究利用という観点で扱うための課題と標準化について論じる。

1 データの種類、特性と取扱い

電子カルテのデータ種別は多岐にわたるが、研究利用の観点から主要なものについてデータの特性と取扱いの留意点を挙げる。一般的に電子カルテデータをそのまま研究に利用しようとする場合、データ取得目的が研究ではなく当該患者の診療であり、診療上の必要に応じて取得されるため、データ取得日時がデータ項目により一律でなく、欠損値が非常に多いことも問題となる。

1) 検体検査結果

データはほとんどが数値定量データか離散値をとる定性データで、計算機処理がデータ処理上は容易である。ただ、同一施設内でも分析装置、検査手法、試薬の変更により基準値や分布特性が変化することがあり、定性データの区分方法も診断基準や分類基準の変更により変化するため、経年データ処理ではこうした変更情報が検査結果データだけではわからないので注意が必要である。多施設データの統合ではこの問題が施設間で発生するため、定量データでは検査値ごとにその施設の基準値範囲を付与して収集する必要がある。この問題は次項の標準化でも述べる。

2) 処方、注射、点滴データ

多くのシステムで処方データは処方せんを作成し発行するためのシステムを運用する結果として蓄積される。また、外来患者では調剤薬局への調剤指示情報であり、入院患者では院内薬局に対して医薬品を調剤して病棟へ配送するための指示情報である。したがって、処方データは患者が服用したデータではない。特に入院患者では、処方オーダ情報を変更登録せずに、日々変化する病状に応じて処方薬の服用を中止したり新たな服用薬を臨時追加したりするため、電子カルテデータから得られる処方情報と実際の服用情報とのあいだには相当違いがあることに留意しなければならない。

注射や点滴のオーダシステムは、一般にオーダ入力が煩雑であるため導入していない病院も多い。ま

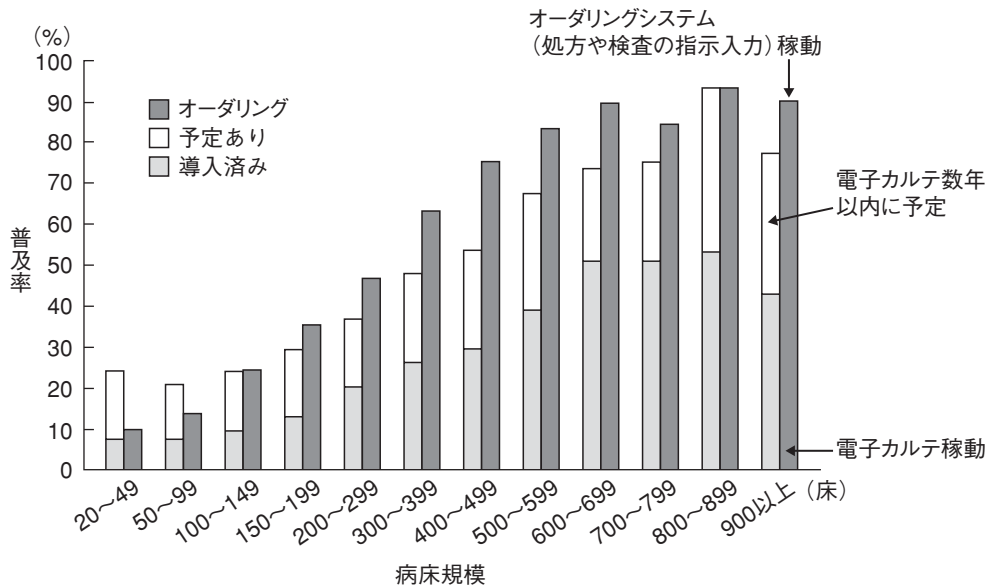


図 1 電子カルテ・オーダーリングシステム普及状況(平成 20 年度厚生労働省医療施設静態調査)

た、入力されている場合にも処方と同様に院内薬局から病棟への医薬品取寄せデータであり、実際の投薬データではないため研究利用には注意が必要である。特に病棟在庫の注射薬を使う場合には入力していない病院も多いので、このことにも留意する必要がある。

3) 病名、診断名情報

病名情報はレセプトの傷病名欄に反映させるために入力されており、検査を行うために必要な「疑い病名」などが多く含まれる。また、一度入力した病名を終了入力することを怠っている場合が多いため、実際にはすでに患者に存在していない病名がいつまでも付与されていることも多い。したがって、ある時点で実際に診療対象となっている病名を知ることが非常に難しい。DPC（急性期診療報酬包括評価制度）を導入している病院であれば、入院患者については DPC データの病名情報を使うほうが峻別された病名情報を得られる。これに対して、手術時診断名、病理診断名、内視鏡検査時の診断名など手術や検査時につけられた病名情報は、診療報酬請求目的でないため信頼性が高いが、これらの情報はレセプトへ反映させる必要がないため、多くのシステムでそれぞれ手術情報システムや各検査システムなどで個別に管理され、コード化されていないことが多いので計算機処理は簡単ではない。

4) 手術・麻酔情報、侵襲的検査・画像検査結果

手術や麻酔、侵襲的検査の実施時の情報については、多くのシステムである程度、区分コードデータや個々の項目データとして入力されているため、手術術式、麻酔方法、手術時間や麻酔時間、輸血量などは計算機処理可能な形でデータを得ることができる場合が多い。侵襲的検査結果や画像検査結果については、検査の生データの多くは標準化されたデジタルデータとして利用できるが、検査結果報告書に記載される評価結果情報の多くは自由文であったりシステムにより情報項目が異なったりするため、簡単な方法での利用は難しい。

5) 自覚症状、他覚所見、臨床経過など

電子カルテ特有のこれらの情報は、システムや医療機関の違いにより入力方法はさまざまで、テンプレートという入力雛型が用意されていて入力欄に個々の項目を入力するようになっている場合もあれば、自由文入力だけの場合もあり、両方が適宜使い分けられるシステムもある。いずれの場合にも、診療中の入力の手間を減らすために、どのような方法で入力するかは医師に任されていることがほとんどで、特定の項目情報、たとえば血圧測定値、体温、主たる訴えなどを個々に取り出して計算機処理することは難しい。

表 1 厚生労働省において保健医療情報分野の標準規格として認めるべき規格
「厚生労働省標準規格」

番号	標準名	説明
HS001	医薬品 HOT コードマスター	国内で認可されたすべての医薬品に標準コードを割り当てた標準コード表。7 桁または 9 桁。
HS005	ICD10 対応標準病名マスター	電子カルテおよびレセプトに電子記録する病名の文字列と病名コードの標準コード表。レセプトコードとも 1 対 1 対応。
HS007	患者診療情報提供書および電子診療データ提供書（患者への情報提供）	医療機関が患者に渡す診療情報提供書、処方、検査データを電子的に作成する場合の標準フォーマット規格。
HS008	診療情報提供書（電子紹介状）	医療機関が他医療機関に渡す診療情報提供書、処方、検査データを電子的に作成する場合の標準フォーマット規格。
HS009	IHE 統合プロファイル「可搬型医用画像」およびその運用指針	医療機関が他医療機関または患者に渡す医療画像データを CD や DVD に記録する標準フォーマット規格。
HS010	保健医療情報—医療波形フォーマット—第 92001 部：符号化規則	心電図、脳波などの生体検査波形データを電子的に記録し他システムと交換するための標準フォーマット規格。
HS011	医療におけるデジタル画像と通信 (DICOM)	医療画像検査のほぼすべてを網羅する標準データフォーマットと機器同士—システム間のデータ交換規格 DICOM。
HS012	JAHIS 臨床検査データ交換規約	検体検査データをシステム間、医療機関間などで交換するための標準データフォーマット規格。HL7v2.5 規格準拠。
HS013 (予定)	標準歯科病名マスター	HS005 の歯科用データ。
HS014 (予定)	臨床検査マスター	検体検査データをシステム間、医療機関間などで交換する際に使用する検査項目の標準コード表。日本臨床検査医学会の JLAC10 コード表準拠。

2 データと情報交換規格の標準化

計算機では、日本語文字列のままの生データや自由文データの処理は非常に困難である。したがって、前節で計算機処理できると述べたデータについても定量数値データ以外はコード化されている必要がある。もし医療機関ごとに別々の方法でコード化されていると、多施設データの統合化ができない。そこでこれまで多くの標準化努力がなされてきており、その成果物である標準コード表と標準コード化ルールは、標準コードマスターと呼ばれている。また、情報システムからデータを取り出す方法として、各システムや各機関がそれぞれ別々のデータフォーマットで出力したのでは、システムごとと機関ごととにその違いを識別してデータを集める必要があり非常に効率が悪いことから、データフォーマットの標準化も重要である。データフォーマットの標準化はシステムからのオンラインでの出し方も含めて、標準交換規格と呼ばれている。標準コードマスターと標準交換規格のうち主要なものは、表のように平成

22 年 3 月の厚生労働省通知（医政発第 0331 第 1 号）により厚生労働省標準規格として今後の各種医療情報システムで使用する事が強く推奨されている（表 1 には今後制定される予定の 2 規格を含む）。

今後、研究を指向する医療機関の電子カルテシステムが、研究用データを出力する場合には、少なくとも医薬品 HOT コードマスター、ICD10 対応標準病名マスター、JAHIS 臨床検査データ交換規約、そして臨床検査マスターに準拠することが必須である。これらの情報の詳細は HELICS 協議会のホームページ（<http://www.helics.jp>）から入手できる。

おわりに

電子カルテデータの研究利用にあたっては、データが診療上の必要により蓄積されたものであり研究目的で蓄積されたものでないことから、本稿で述べたように、さまざまな制約や留意点があることに注意したうえで、膨大な電子データを有効に活用することが必要である。そして、そのためには、研究用データに変換する過程で、標準コードと標準データ

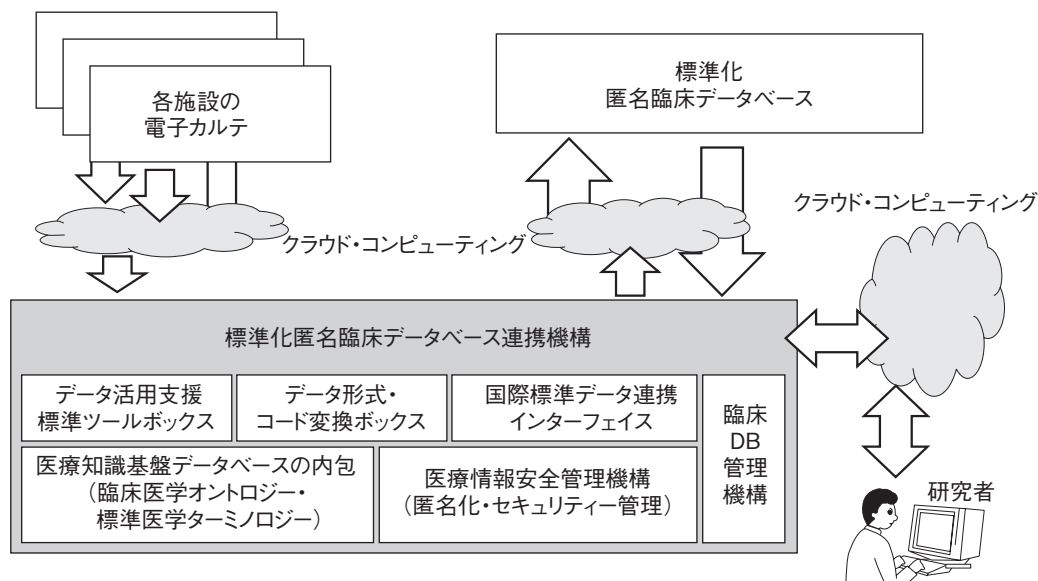


図 2 最先端研究開発支援プログラム「未解決のがんと心臓病を撲滅する最適医療開発」(平成 21～25 年・中心研究者：永井良三教授)
サブテーマ「臨床研究のための標準化臨床データベース連携機構の開発」(サブテーマリーダー：大江和彦)

フォーマットを採用することが重要である。なお、筆者らはこうした変換過程が多くの医療機関や研究機関で少ない労力で行えるように、国の最先端研究開発支援プログラム (中心研究者：永井良三東京大

学教授) において、クラウドコンピュータシステムを活用した臨床研究用標準 IT 基盤の開発研究と実用化を目指している (図 2) (<http://www.m.u-tokyo.ac.jp/medinfo/first/>)。