

一般口演 | 電子カルテ・EHR

一般口演22

電子カルテ・EHR

2019年11月24日(日) 11:10 ~ 12:10 C会場 (国際会議場 2階国際会議室)

[4-C-2-01] 医療情報流通基盤としてのデータ集約型 EHR「千年カルテ」の構築

○ 糸直人¹、吉原博幸¹、黒田知宏²、荒木賢二³ (1. 京都大学大学院医学研究科EHR共同研究講座, 2. 京都大学医学部附属病院医療情報企画部, 3. 宮崎大学病院IR部)

キーワード: Electronic Health Records, Standardization, MML

従来、法的、慣習的制約の中で長く医療情報の医療機関外への保存は慎重に議論されてきたところである。近年、インターネット上のセキュリティ技術の向上を待って、バックアップや地域連携を目的とした医療情報の外部保存が認められ普及しつつあるといえる。一方で、昨今のコンピュータのデータ処理能力や機械学習を始めとする情報技術を医療情報の利活用に供するためには、そもそも大規模データを自在に取り扱う環境が必要である。多施設データを連携し大規模データとして取り扱う情報基盤は、日々新たに発生した医療情報を逐次、利活用を目的とする処理系に配布できる仕組みを備えることで、利活用に至るプロセスを大幅に短縮できると考えられる。

そこで、本研究は、医療情報流通基盤としてデータ集約型の Electronic Health Record (EHR) を構築することを目的とする。医療機関から直接データを出力しデータセンター (DC) に集約すると同時に、医療情報の流通と利用を分離したシステム構築をおこなう。また、データ集約後のデータベース構築コストを低減するためにデータ伝送には Medical Markup Language (MML) から派生した XMLフォーマットを用いる。

本研究では、EHRのフロントエンドにデータ流通基盤を構築し、病院・クリニックと DCを直結すると同時に、バックエンドに EHRサービスサーバを構築した。一次利用を目的としたサービスサーバをフロントエンドに配置した。主要な電子カルテベンダーが MML準拠の日次の XML出力に対応した。100以上の医療機関と DCを接続し、日次のデータ出力を行っている。今後、医療情報の利活用者へのデータ流通経路として情報基盤を提供する。

医療情報流通基盤としてのデータ集約型 EHR「千年カルテ」の構築

糸 直人^{*1}、吉原 博幸^{*1}、黒田 知宏^{*2}、荒木 賢二^{*3}

*1 京都大学 大学院医学研究科 EHR 共同研究講座、*2 京都大学医学部附属病院 医療情報企画部、
*3 宮崎大学医学部附属病院 病院 IR 部

Construction of A Centralized Database System of Electronic Health Record As A Medical Information Distribution Infrastructure

Naoto Kume^{*1}, Hiroyuki Yoshihara^{*1}, Tomohiro Kuroda^{*2}, Kenji Araki^{*3}

*1 Department of Electronic Health Record, Graduate School of Medicine, Kyoto University,

*2 Division of Medical Information Technology and Administration Planning, Kyoto University Hospital,

*3 Division of Institutional Research, University of Miyazaki Hospital

Abstract: Conventionally, Electronic Health Record (EHR) was sustained by subsidies, which is provided under limited period. Because of medical big-data needs, EHR is required much more sustainability to preserve patient records as long as possible. Therefore, EHR system should be managed by sustainable self-independent profit model instead of subsidies model. For this circumstance, new law for anonymized data use are enforced in 2018, in Japan. The authors propose EHR-centered medical information circulation model, which provides industries to use of named data and anonymized data, to achieve sustainable profit model. An EHR system is constructed as a centralized database, which consists with distribution infrastructure datacenter and EHR service datacenter. Also, EMR data retrieval is designed by XML data exchange, which is defined by Medical Markup Language as a structure, and a project-defined XML export protocol as operation management. In result, the proposed data export and upload system was installed over a hundred hospitals. The XML data from the connected hospitals were stored to the EHR service database in daily upload. The proposed distribution infrastructure datacenter preserves all daily XML files, and allow service developers to access the EHR service datacenter API to create new applications such as mobile apps. Thus, the developed datacenters would contribute as an information distribution infrastructure for secondary use under a proper manner of access control.

Keywords: Electronic Health Records, Standardization, MML.

1. はじめに

従来、地域医療連携システムは主に補助金によって構築され、継続的な補助金の獲得が常に課題となっている。あるいは、地域の中核病院の多大な負担の下に関連病院との連携をすすめられていることが一般的である。また、電子カルテの普及とともに蓄積された医療情報の二次利用に関するニーズの高まりとともに施設を超えたデータ連携が模索されつつあることは論をまたない。医療ビッグデータに対する期待とも相まって、例えば地域医療連携システムで交換されている医療情報をビッグデータとして利活用していくことは、今後の医療の質評価や Health Technology Assessment (HTA)の潮流に欠かせないものといえる。このようなニーズに応じた Electronic Health Record (EHR) の構築が求められる一方、昨今の医療費の国庫負担増加に対し補助金による EHR の構築・運用は容易に認められる情勢にない。そこで、医療ビッグデータの構築を目指した EHR は、データの利活用による自立的な運営を想定して設計する必要がある。本研究では、永続的な EHR の運用を可能とするスキームを含めたシステム構築を目的とする。

2. 背景

我が国では、政府の健康・医療戦略(平成 26 年 7 月 22 日閣議決定)の下、健康長寿社会の形成を目的として、新法「医療分野の研究開発に資するための匿名加工医療情報に関する法律(通称、次世代医療基盤法)」(平成 29 年法律第 28 号)が、平成 29 年 5 月 12 日に公布された¹⁾。本法により匿名加工した医療情報を広く利活用し、健康医療産業に供

する商業利用が認められた。

一方で、従来医療情報は個人情報保護法²⁾の下、病院施設内で管理することを前提としてきた歴史から、既存の地域医療連携システムにおいてもそれぞれの施設内データベースを相互参照する仕組みが主流であり、施設横断的な大規模データベースを構築している例はほぼない。

個人情報保護や病院施設による情報管理の視点からは、超高速な施設間データ参照により論理的な大規模データベースを構築することが望ましいと言えるが、一方で現在の技術的制約からそのような超高速データベースを構築することは困難であり、また各施設への投資に見合うパフォーマンスは期待できない。そこで、様々なデータ利活用のニーズに耐えうるデータベースを構築するためには、EHR をデータ集約型で設計することが望ましいといえる。

3. 手法

本研究では、データ利活用スキームの設計に合わせて、様々な利活用を想定したシステム設計を行う。データ利活用のパターンに応じて、動的にデータの収集と配信を制御することで、医療情報流通基盤としての振る舞いを実現すると同時に、基盤を維持する経済モデルが成立することを目的とする。

3.1 医療情報循環モデル

従来の健康医療情報の利活用には、個人情報保護法下における患者の同意に基づく利用が挙げられる。この場合、改正前個人情報保護法下における仮名化等の配慮は要するとしても原則記名データの利活用が中心となる。記名デー

タ利用においては、同意の取得方法が課題となるが、原則事業者責任となる。次世代医療基盤法下における利活用では、認定匿名加工医療情報作成事業者(以下、認定事業者)であれば、病院施設から記名データを収集し、匿名加工医療情報を作成した後、当該匿名加工医療情報を用いた事業が可能である。この場合、匿名加工品質については認定事業者が責任を負う。

本研究では、データ収集の入り口には EHR による地域医療連携サービスを設定し、記名および匿名のデータ利活用を出口としたシステム構築を行う。

図 1 に本研究で提案する医療情報循環モデルを示す。図下部は EHR 事業、図上部はデータ利活用事業を示す。EHR 事業では、医療従事者間の連携(B2B, 利用①)、および医療従事者一患者間の連携(B2C, 利用②)の 2 つを実施する。データ利活用事業では、認定事業者との連携による匿名加工医療情報の利活用、及び患者同意による記名情報の利活用を可能とする。ここで、EHR クラウドは NPO 法人による運営とし、BCP(データバックアップ)を主たる業務としつつ、B2B の運営支援をおこなう。ここで、BCP の対象となるのは、災害時に参照可能な診療情報であり、電子カルテ本体のバックアップではない。これは二次利用を前提とした EHR 設計において、電子カルテ本体のバックアップは実施負荷が大きい代わりに利活用用途が見いだせない事による。

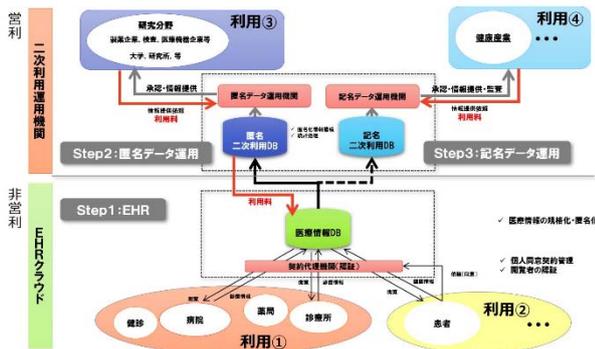


図 1 医療情報循環モデル

二次利用運用機関に関しては、認定事業者、ヘルステック事業者等、複数の企業体の複合によって成立することを想定している。これらの企業体が EHR クラウドを基盤として利用することでそれぞれのビジネスニーズに応じたデータ利用を実現することで、ひいては EHR 側の BCP, B2B, B2C といったサービスを持続的に運用することを想定する。

また、EHR の運用、研究開発等に関し NPO 法人単体での実施では困難であることから、NPO 法人と学術団体との連携体制を構築する。

3.2 システム構築

本研究では、EHR を中心とした医療情報循環モデルの実装を「千年カルテ」と称する。千年カルテは次の 3 つのシステムで構成される; 接続する医療施設の電子カルテ(EMR)からのデータ出力モジュール、出力データをデータ集約型データベースに流すデータアクセスポイント、EHR データベース。

電子カルテからの診療情報出力に関しては、病院ごとに採用している電子カルテシステムが異なることから、データ構造をある程度統一したフォーマットでのデータ収集が望ましい。

病院ごとに外部接続に関するシステムやポリシーが異なることから、接続方法の違いを吸収すると同時に、出力データの管理を行うアクセスポイントとなるデータセンターを設置する。また、収集したデータをデータベースに取り込み、B2B, B2C 等のサービスに供する EHR センターを設置する。

3.2.1 EMR データ出力

現在病院に導入されている EMR は殆どの場合、EMR ベンダーの提供するパッケージをカスタマイズした状態で供給されている。そのため、同一の EMR ベンダーを採用している病院間であっても、データ項目や利用しているデータベースカラム、あるいはマスター情報はそれぞれ異なっている。そのため、EMR にある情報を標準化して出力することが望ましい。しかし、一方で国際標準の規格を直接採用した場合、EMR ベンダーの開発工数がかかるうえ、EMR ベンダーごとの解釈によってデータが格納されて出力されるため、規格の準拠を仕様とするだけでは実質的には計算機処理に供するデータ収集は完結しない。そこで、本研究では、医療文書単位に着目したデータ出力をおこなう。この場合、個別の情報に関して利活用を目的とした場合に構造化の粒度が不十分な場合もあるが、一方で、構造化データ、非構造化データを同時に収容できるため、全体として収集可能な情報量が多くなる。特に、病院ごとに標準規格に対するデータの構造化レベルは異なるため、特定の病院では構造化されているデータが、他の病院では非構造化データのままとなっている事例がほとんどである。このような場合であっても、構造化、非構造化を問わずデータを収集し、収集後に必要があれば非構造化データをデータセンター側で構造化して利活用に提供するという選択肢を残す必要があると考えた。

以上より、本研究では医療文書単位でのモジュール化が可能な Medical Markup Language (MML)³⁾を採用し、XML によるデータ出力モジュールを各病院に導入する。

一方で、MML はデータの格納に関する規約ではあるが XML ファイル出力のルール、削除文書の表現、実施日時不明な場合の取り扱い、NULL 値の表現方法、MML 必須項目の型違いデータの出力等、実際のデータ出力に際して必要な運用ルールは規定されていない。そこで、本研究では千年カルテ用の MML 運用規約を別途規定し、MML で表現されていない実データ出力仕様を確定した。特に、テンプレートデータの出力等、病院個別に精細に構造化されたデータの取得に際しては、MML の拡張をおこない、JSON 等の別フォーマットでも格納可能とした。MML で定義されている、患者基本情報、保険情報、病名情報、生活習慣情報、経過記録情報、手術記録情報、臨床サマリー情報(退院時サマリー等)、検査結果情報、報告書情報、紹介状、バイタルサイン、体温表、処方、注射が出力対象文書である。ただし、EMR ベンダーごと、病院ごとに実際に出力可能な情報は異なる。

以上の仕様に基づいた EMR データを施設の営業が終了する時刻以降、深夜帯に日次バッチ処理で出力する。

3.2.2 アクセスポイント

本研究では、院内で EMR から出力されたデータを一旦院内のアップローダサーバに転送し、そこから外部接続ネットワークを通じてデータセンターのデータレイバサーバに再転送を行う。図 2 に院内とデータアップローダの関係図を示す。

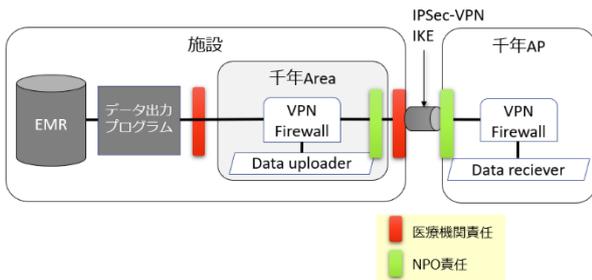


図2 データアップローダ

千年カルテに接続する施設は主に4つのネットワーク形態に分類できる；VPN直結型、専用線型、データセンター経由型、SSL自己証明書付インターネット接続型。これらを統合し、1つのデータリポジトリとして振る舞うデータセンターを設置し「千年AP」と称する。VPN直結型の場合、通常は千年APと施設間でそれぞれ固定IPを持って対向接続を行う。施設ポリシーによって固定IPを持たない場合はダイナミックグローバルIPに対してVPN接続を行う必要がある。またインターネットVPNがセキュリティ上問題になる施設では、施設が提供する専用線のノードを千年APに設置する。中小病院の場合、データバックアップやEMRサービスをEMRベンダーの管理するデータセンターで運用している場合がある。この場合には、EMRベンダーのデータセンターと千年AP間でVPN接続を行う。特に病院ごとのデータ管理が必要なため、複数病院でVPN経路は共有しつつ、データアップローダに関しては施設ごとに設置し、施設ごとにデータ受領サーバを千年AP内に構築する。特に県立病院や全国展開している私立病院、全国展開している中小病院EMRベンダーを中心とした病院群との接続がこの型に該当する。クリニックと接続する場合には、FirewallやData uploaderの設置が高コストになる一方出力されるデータ量が軽微であり、また多くの場合、基幹病院のデータ参照のみを目的として接続することから、SSL自己証明書付インターネット接続を採用する。

3.3.3 EHRセンター

千年APで受領したMMLファイルは、日次でEHRセンターにWebDAVで転送され、バッチ処理でHadoopデータベースに取り込む。Hadoopを採用することで収容するデータ量の増加によってデータベースを再構築する必要がない。取り込んだMMLデータは、Web画面での内容確認、アクセス権設定が可能で、また同時にB2C用のAPIを提供する。API経由で個人用のデータ取得が可能なAndroid、iOSアプリを提供し、患者自身がアカウントを用いてEHRにアクセスする。

図3にEHRセンターの構成概略を示す。Firewallは千年APとの直結用と、Web viewer、APIへのアクセス用に設定する。B2B連携におけるアクセス権制御は、MMLに設定される出力施設由来のアクセス権、Web viewerで設定される上書きアクセス権の2つを勘案して実施される⁴⁾。

図4に医師用のEHR画面を示す。左上に表示する文書種別を選択するボタン、左下に表示している文書一覧、右側に文書の内容が表示される。選択ボタンの文書種別はMMLのもつ文書種別モジュールに対応する。MMLには糖尿病モジュール等が存在するが、千年カルテでサポートしていない文書種別については表示されない。

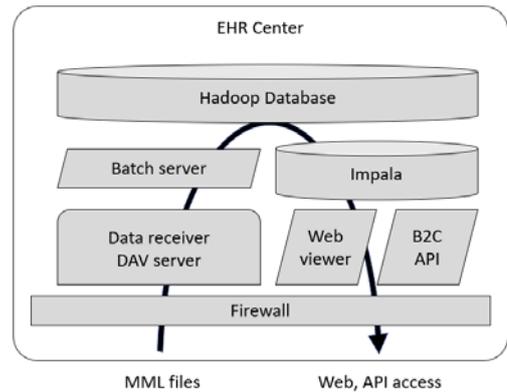


図3 EHRセンターの構成概略図

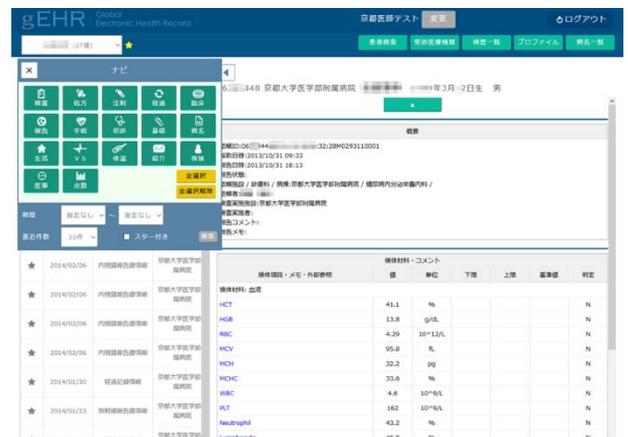


図4 EHR Web viewer画面（医師用）

4. 結果

2015年10月から、国立研究開発法人日本医療研究開発機構(AMED)の【臨床研究等ICT基盤構築研究事業】「全国共同利用型国際標準化健康・医療情報の収集及び利活用に関する研究（研究代表者・NPO法人日本医療ネットワーク協会 荒木賢二）」による支援をうけ、EMRデータ出力、アクセスポイント、EHRデータベースの実装をおこなった。

AMED研究事業の結果、2019年4月時点で108施設と接続し、それぞれの病院でのEMRデータ出力を順次開始している。

4.1 参画病院

2019年4月時点での千年カルテ参画施設は、108施設（約54,000床）で、延べ患者数7800万人/年、初年度180万人をカバーしている。表1に参画病院の属性と数を示す。内、DPC対象病院58、がん拠点病院32を含む。

また、参画病院に関しては、VPN直結型、データセンター経由型等を問わずVPNネットワーク接続を完了し、Data uploader設置を完了している。

表1 千年カルテ参画病院数

属性	病院数
国立大学病院	15
公立大学病院	3

私立大学病院	5
自治体立病院	33
その他公立病院	19
その他・団体	6
民間病院	27

4.2 EMR データ出力

千年カルテ運用規約に基づく MML 出力プログラム設置に関して、対象となった電子カルテベンダーと対象施設数を表 2 に示す。EMR データ出力対応済みベンダーに関しては、新規で千年カルテに接続する場合には出力プログラムの設置のみが必要で、新規の出力プログラム開発は不要となる。

表 2 千年カルテ運用規約準拠 MML 出力ベンダーと対応施設数(ベンダー名は通称表記)

EMR ベンダー名	病院数
富士通	41
日本電気	27
SSI	14
日本 IBM	5
CSI	5
コアクリエイト	5
キャノン	2
その他	9

4.3 アクセスポイント

ほとんどの施設は VPN 直結型で千年 AP に接続した。一方、県庁 LAN を有する県下の病院はセンター間接続型となった。リアルタイムで通信モニタリングを実施しており、アップリンクしていない施設を検知することができる。また、Data uploader のモニタリングも実施しており、アップロードが停止している施設、不安定・不定期的な施設等を確認することができる。とくにサービスイン直後は、ネットワークの問題と、データアップロード経路の問題が混在するため、切り分けが発生する。特に院内 Firewall の設定でアップローダとの通信が許可されていないケースが数例報告されている。

千年 AP の主たる役割はデータのアップロード制御であるが、同時に EHR センターで実施できない外部連携や、アプリ連携のポータルとしての役割を担っている。具体的には検査値の異常値アラートシステムや、健康系デバイスからのデータ受領等、追加機能を Web サービス連携で実装する場合の拠点サーバとして稼働する。この場合、診療データそのものは EHR センターの API 経由で受領するが、その上にかぶせるサービス群は千年 AP で実装される。以上の分離構造により、機能拡張を容易とするとともにデータの出入り口を管理することで診療情報の保全に寄与している。

5. 考察

本研究では、データ出力、データ流通経路、データ集約、データ可視化までの一連のシステムを実装した。一方で、データの二次利用に向けて千年 AP を中心とした二次利用機関(認定事業者等)との接続とデータ配置を今後検討していく必要がある。特に、千年 AP はファイルストレージとして機能しているが、一方でデータ要求は多岐に渡ることが予想されるため、データベース化が求められるといえる。EHR サービスに

関しては EHR センターが機能し、定常業務の中でモニタリングする情報(検査値)等は EHR センター側のルーチン処理で処理する余地がある。

今後、MML の出力結果に対して、利用者側のニーズに応じた EMR からのデータ出力可否が課題となる。ニーズに対するフィードバックを EMR ベンダーにどのように返していくか検討の余地がある。初回の出力プログラム作成までにはマニュアルでの MML データ確認を要し、機械的な MML 構造のチェックだけでは全く不十分であった。少なくとも既出のマニュアルチェック項目に関しては、機械的に出力データをチェックできる体制を取る必要がある。

本研究は、EHR の持続性を前提としてデータ利活用と収集の両輪を同時に実施する点に特徴がある。特に匿名加工医療情報の利活用を収益モデルの主軸においてシステム構築をおこなってきた。一方で利活用の要としていた次世代医療基盤法において平成 30 年 5 月 11 日の法律施行後、令和元年 8 月時点でも認定事業者が確定していない状況にある。このような社会的なリスク要因を鑑みると、記名データ利用に対して産業界との活発な連携が求められるものと考えられる。

6. おわりに

本研究では、持続可能な EHR システムの構築を目指し、データの利活用を前提とした EHR の設計をおこなった。データ集約型の EHR を基盤データセンターとサービスデータセンターに分離することで、データの流通とサービスの実装を別々のベンダーが実装できる体制を整えた。プロジェクトに参画する 108 病院に対して統一フォーマットでのデータ出力プログラムを実装し、データ集約が可能な基盤整備を行った。

参考文献

- 1) 医療分野の研究開発に資するための匿名加工医療情報に関する法律(次世代医療基盤法). 2018 年; 第 28 号: 第 2 節.
- 2) 個人情報の保護に関する法律. 2015 年; 第 57 号: 第 1 章第 16 条.
- 3) Medical eXtensible Markup Language Consortium. MML 4.1.2 規格書. (2019 年 6 月 1 日引用). URL: <http://medxml.net/>.
- 4) 桑直人, 吉原博幸. 地域医療連携システムにおける患者の受診歴を中心としたアクセス権制御. 遠隔医療学会雑誌, 2019-2-1, 2019.

謝辞

本研究は国立研究開発法人日本医療研究開発機構(AMED)の【臨床研究等 ICT 基盤構築研究事業 大規模健康・診療データ収集・利活用に関する研究】の支援によって行われた。

本研究は、AMED 研究事業に協力いただいた 108 施設の支援によって行われた。

本研究は、一部に EHR 共同研究講座における共同研究企業(エヌ・ティ・ティ・データ, グラクソ・スミスクライン, エス・アール・エル, みらか中央研究所, ラジエンスウエア)との研究成果を含む。