

一般口演

一般口演9

医療データ分析4（ビッグデータ・ゲノム情報・がん登録等）

2018年11月23日(金) 16:00 ～ 17:30 I会場 (福岡サンパレスH平安（中継末広）)

[2-I-3-5] 多施設臨床研究に適用可能なサンプル管理システムの開発

真鍋 史朗^{1, ○}, 武田 理宏¹, 白月 遼², 前田 大地³, 谷内田 真一⁴, 松村 泰志¹ (1.大阪大学大学院医学系研究科 医療情報学, 2.大阪大学医学部附属病院 がんゲノム医療センター, 3.大阪大学大学院医学系研究科 先端ゲノム医療学, 4.大阪大学大学院医学系研究科 がんゲノム情報学)

患者から臨床検体（サンプル）を取得し、解析を行う臨床研究では、患者 ID を匿名化した臨床サンプル ID を用いて研究を行う必要がある。中でも、ヒトゲノム・遺伝子解析研究は、より厳密な匿名化の管理が求められ、被験者 ID・臨床サンプル ID を二重匿名化した遺伝子サンプル ID を用いて研究を行う必要がある。我々は大阪府下の医療機関で、電子カルテと連携した CDCS という多施設臨床研究データ収集システムを開発・運用している。このシステムでは、患者 ID に対応する被検者 ID が発行され、臨床情報を管理している。被検者 ID と臨床サンプル ID、遺伝子サンプル ID を管理するシステムが存在すれば、研究者が患者 ID を知ることなく、臨床情報を取得することが可能となる。そこで、我々は医療機関の電子カルテシステムや CDCS と連携し、臨床サンプル ID、遺伝子サンプル ID を発番し、発番した ID をラベル印刷することができるサンプル情報管理システムを開発した。本システムは各医療機関でオーダ発行時や検体登録時に臨床サンプル ID を発番し、患者 ID との対応表を保持するサイトサーバと、研究施設で臨床サンプル ID から遺伝子サンプル ID を発番し、臨床サンプル ID と遺伝子サンプル ID の対応表を管理するセンターサーバで構築される。サイトサーバでは患者 ID と CDCS で発番される被検者 ID との対応表を取り込むことで、臨床サンプル ID と被検者 ID の対応表をセンターサーバに送ることを可能とした。センターサーバでは、遺伝子研究者が遺伝子サンプル ID を用いて臨床研究者に問い合わせを行うと、システムが被検者 ID に変換し臨床研究者に問い合わせを提示し、臨床研究者が被検者 ID を用いて遺伝子研究者に問い合わせを行った場合は、システムが遺伝子サンプル ID に変換し遺伝子研究者に問い合わせを提示することを可能とした。これらの機能により、臨床データとサンプルの解析データを、個人識別情報を見ることなく、紐づけることが可能になった。

多施設臨床研究に適用可能なサンプル管理システムの開発

真鍋史朗^{*1, *2}、武田理宏^{*1}、白月遼^{*2}、前田大地^{*3}、谷内田真一^{*4}、松村泰志^{*1}

*1 大阪大学大学院医学系研究科 医療情報学

*2 大阪大学医学部附属病院 がんゲノム医療センター

*3 大阪大学大学院医学系研究科 先端ゲノム医療学

*4 大阪大学大学院医学系研究科 がんゲノム情報学

Development of the Clinical Specimen Information Management System

Shirou Manabe ^{*1, *2}, Toshihiro Takeda ^{*1}, Ryo Shiotsuki ^{*2}, Daichi Maeda ^{*3}, Shinichi Yachida ^{*4}, Yasushi Matsumura ^{*1}

*1 Medical Informatics, Osaka University Graduate School of Medicine ,

*2 Cancer Genome Medical Center, Osaka University Hospital ,

*3 Clinical Genomics, Osaka University Graduate School of Medicine ,

*4 Cancer Genome Informatics, Osaka University Graduate School of Medicine

The clinical research for clinical genetic testing requires strict ID management. Therefore, we need to use gene sample ID which double-coded subject ID and clinical sample ID. We developed the clinical sample management system that can collaborate with the electric medical record system and electric data capture (EDC). This system consists of a site server that issues a clinical sample ID at the time of order issue or sample registration at each medical institution and holds a correspondence table with the patient ID and a gene server ID from the clinical sample ID at the research facility. It is constructed with a center server that manages a correspondence table between clinical sample ID and gene sample ID. The site server was set up at the Osaka University Hospital and the Osaka International Cancer Institute et al., and the center server was set up at the data center of Osaka University Hospital. When a clinical researcher queries a results of a clinical genetic testing to a gene researcher using a subject ID, the system converts the subject ID to the gene sample ID and presents the inquiry to gene researcher. With these functions, it became possible to use clinical research data and gene analysis data without looking at personal identification information.

Keywords: sample management system, multicenter clinical study, gene sample ID

1. はじめに

近年、日本において臨床研究をベースとした論文数が他国と比較して減少しているなどの問題が指摘されている¹⁾。その原因の一つとして、日本では医療施設が分散して存在している²⁾ため、1施設における特定の疾患に対する症例数を他国ほど集めることができない問題があると考えられる。そこで、ICTを活用して多施設からデータ収集を行い、臨床研究・治療を活性化することが求められている。我々は大阪府下の医療機関で、電子カルテと連携した Clinical Data Collecting System (CDCS) という多施設臨床研究データ収集システムを開発・運用している³⁾。このシステムでは、患者IDに対応する被験者IDが発行され、その対応表は施設毎に管理しており、臨床情報などの研究用データはセンターとなる施設で被験者IDをキーとして一元的に管理している。

多施設臨床研究において問題となるのが、検体検査や遺伝子解析を行う際の検体をセンターとなる施設に集めて処理を行う際、その検体に対する個人情報保護をどのように行うかということである。検体提供者のプライバシーの保護を目的として匿名化するだけでは、検体の管理を効率的に行うことができない。患者から血液サンプルや尿サンプルを取得する場合、病院での通常の血液検査、尿検査と同時にサンプルを取得することが多い。このため、電子カルテの検体検査オーダーで臨床サンプル取得をオーダーし、検体検査オーダーで発行される検体ラベルを読み取ることで臨床サンプルラベルを発行することができれば、患者にとっても、病院にとっても、研究者にとっても効率的にサンプル採取を行うことが可能となる。

患者から病理組織や血液などの体液を臨床検体(サン

プル)として取得し、特定の検査機関において一元的に解析を行う臨床研究では、サンプルが他施設に送られることになるため、患者IDを匿名化した臨床サンプルIDを用いる必要がある。中でも、ヒトゲノム・遺伝子解析研究は、その検体から得られる結果の重要性は高く、より厳密な匿名化の管理が求められ、被験者ID・臨床サンプルIDを二重匿名化した遺伝子サンプルIDを用いて研究を行う必要がある。さらに、その結果をどのようにして臨床情報と結びつけ、臨床研究者に見せるかを考えなくてはならない。

被験者IDと臨床サンプルID、遺伝子サンプルIDを管理するシステムが存在すれば、遺伝子研究者は患者IDを知ることなく臨床情報を取得することが可能となり、臨床研究者は遺伝子サンプルのことを知ることなく遺伝子情報を取得することが可能となる。

2. 目的

我々は医療機関の電子カルテシステムや多施設臨床研究用 CDMS (Clinical Data Management System) と連携し、臨床サンプルID、遺伝子サンプルIDを発番し、発番したIDをラベル印刷することができるサンプル情報管理システムを開発した。本システムの目的は、臨床データとサンプルの解析データを個人識別情報を見ることなく紐付けし、閲覧可能にすることである。

3. 方法

本システムは、多施設臨床研究支援システム、サンプル情報管理システム、遺伝子レポート秘匿情報交換システムからなる。

3.1 全体構成

本システムは各医療機関に設置するサンプル管理用サイトサーバーと、連携している病院のデータを集めるサンプル管理用センターサーバーにより構成される。サイトサーバーは電子カルテシステムと連携可能な場合は、電子カルテシステムと同じネットワークセグメントに設置する。サイトサーバーとセンターサーバーはアクセスコントロールを行うためのファイアウォールを介してIP-VPNで接続する。サイトサーバーが電子カルテシステムと連携できない場合は、サイトサーバーを1台のPCにインストールし、そのPCとセンターサーバー間をソフトウェアVPNで接続する。各種データはサイトサーバーからセンターサーバーへ送信するのみで、その逆の処理は行わない。また、データセンターにはセキュリティ確保のためIPS/IDSを設置し、各サーバーはアンチウィルスソフトウェアで保護を行う。全体構成図を図1に示す。

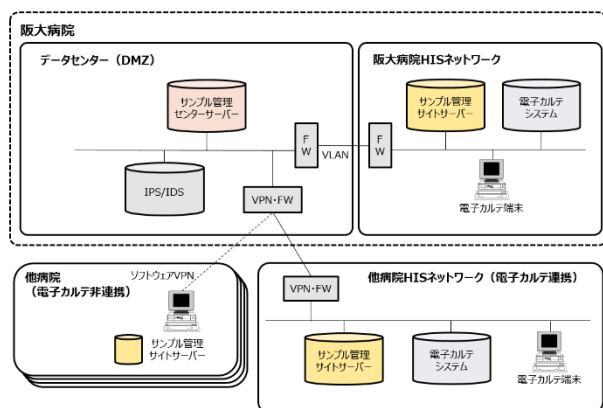


図1 システム構成図

電子カルテシステムと同じネットワークセグメントに各種サイトサーバーを設置し、病院内の情報を管理する。必要なデータのみをセンターサーバーに送信する。センターサーバーからサイトサーバーへの送信は行わない。

今回、センターサーバー機能を大阪大学医学部附属病院(阪大病院)データセンター内に設置した。また、サイトサーバー機能を阪大病院電子カルテ(HIS)ネットワーク内と、大阪国際がんセンターに設置した。これらの病院ではサイトサーバーがHISネットワーク内に設置されている。阪大病院内でのデータセンターとHISネットワーク内はVLAN/FWによる制御を、阪大病院内のデータセンターと大阪国際がんセンターはIP-VPNで接続している。

3.2 多施設臨床研究支援システム

臨床研究において、被験者のIDの発番・管理および臨床情報の入力・管理を行うシステムである。REDCap (vanderbilt univ.)やRave (Medidata)、Inform (Oracle)などがある。阪大病院ではCDCSを利用している。

3.3 サンプル情報管理システム

サンプル情報管理システムは、特定の臨床研究において中央で一括測定する臨床検査や遺伝子検査などで、複数の医療機関において患者から発生する臨床サンプルを中央で一括管理するシステムである。また、フリーザーなどに臨床サンプルを長期保存する際の保管場所の管理や、利用者の管

理・権限設定なども行う。

サンプル管理システムは、サンプル管理サイトサーバーとサンプル管理センターサーバー、サンプル識別用のラベルプリンター、情報確認用のバーコードリーダーから構成される。他システムから直接データを取り込み、処理を行うが、直接操作を行う場合、データの入力および管理はwebブラウザ上で行う。

本システムを利用した研究において臨床サンプルを収集する際は、研究(プロジェクト)毎にオーダ項目を作成し、「プロジェクトコード」と紐づけ、サブコードに「サンプル種別コード+材料種別コード」を紐づける。

3.4 遺伝子解析結果秘匿情報交換システム

多施設臨床研究において臨床サンプルを集めて遺伝子解析等の結果が得られた際に、臨床研究者や遺伝子研究者が、互いの担当者に個人識別情報を見ることなく、問い合わせを行い、結果を取得するシステムである。

3.5 IDについて

本システムで利用するIDについて、次のように定義する。IDの関係性を図2に示す。

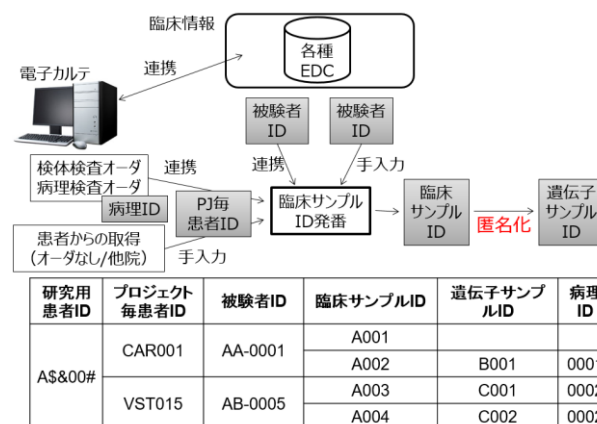


図2 IDの関係図

同一患者で、異なるプロジェクトに参加している場合は共通の研究用患者IDが設定される。各プロジェクトの参加者は意識しない。プロジェクト毎に被験者IDが発番され、遺伝子サンプルがある場合は遺伝子サンプルIDが発番される。

3.5.1 プロジェクト毎患者ID

プロジェクトコードをユニークとなるように管理し、プロジェクト内で被験者をユニークに識別するためのID。先頭2桁を施設コード、3桁を施設内の被験者の識別コードとし、5桁でプロジェクト内で被験者を識別し、プロジェクトコードを合わせた8桁により本システムで管理する範囲内で被験者を識別する。

3.5.2 臨床サンプルID

サイトサーバーで発行する。臨床サンプルか遺伝子サンプルかの識別区分1桁を先頭とし、8桁のプロジェクト毎患者ID、「術前」や「入院時」など、どのタイミングで取られたサンプルであるかを示すサンプル種別コードの2桁、「血液」や「病理組織」などサンプルの材料種を表すコード1桁、更に分注する場合の枝番1桁の計13桁で構成している。各コードは英数字を用いる。臨床サンプルIDのラベル書式を図3に示す。

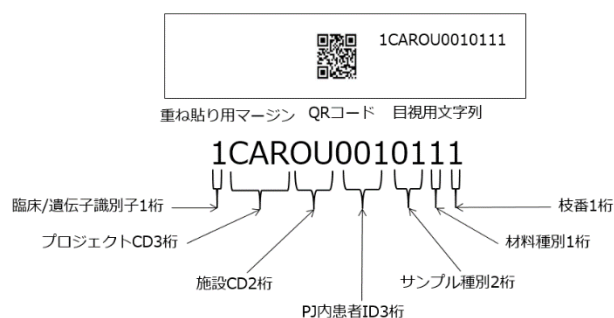


図3 臨床サンプル ID の例

13 桁のコードで表示される。バーコードは試験管や病理検体のブロックに貼ることを想定している。

3.5.3 遺伝子サンプル ID

センターサーバーで発行する。個々の遺伝子サンプルを管理するためのものである。患者情報から2重匿名化が行われているので、遺伝子サンプル ID だけでは個人を特定することができない。遺伝子サンプル ID は、保管場所、保管者、保管日、出庫者、出庫日、コメントを属性情報として持つ。

臨床サンプルか遺伝子サンプルかの識別区分1桁を先頭とし、3桁のプロジェクトコード、8桁の臨床サンプルIDを匿名化した匿名化ID、同一臨床サンプルを複数容器に分けて保存する場合に、それぞれを識別するために用いる複数のサンプルを区別するための枝番1桁の計13桁で構成されている。なお、各コードは英数字を用いる。遺伝子サンプルIDのラベル書式を図4に示す。

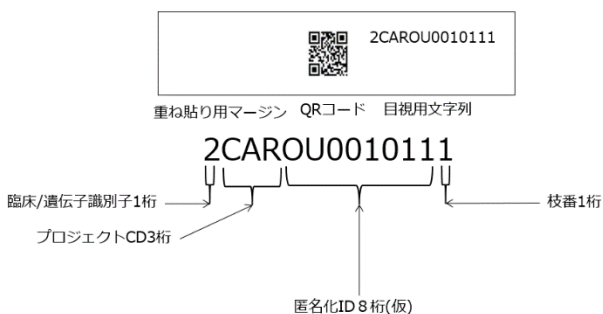


図4 遺伝子サンプル ID の例

13 桁のコードで表示される。バーコードはマイクロチューブの上面に貼ることを想定している。遺伝子サンプルIDを見ただけでは、どの施設の検体かわからない。

3.5.4 被験者 ID

臨床研究支援システムで発行する。CDCS を使う場合は、CRF レポーターで新規登録することにより発行され、本システムと連携が行われる。CDCS 以外の EDC を使う場合は、発番された被験者 ID をサイトサーバーに手入力する。被験者 ID は臨床研究を識別するコードと、臨床研究内 ID から構成されている。

3.5.5 研究用患者 ID

研究用患者 ID は、患者 ID を特定のルールにより別番号に変換し、施設コードと合わせることで構成している。システム内で臨床サンプルIDと紐付けられている。患者は複数のプロジェクトで登録されていることがあるため、プロジェクト間でサ

ンプルが同一患者のものであることを認識するために使用する。研究用患者 ID はシステム管理者以外は閲覧できない。

3-6. ユーザ管理

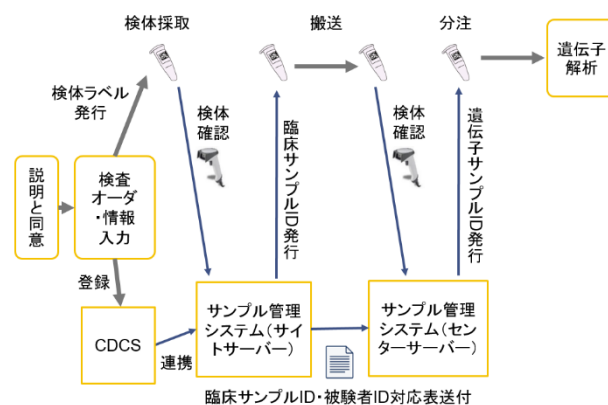
本システムの利用者は、主に臨床サンプル管理者、臨床研究者、遺伝子サンプル管理者、遺伝子研究者、システム管理者を想定している。それぞれ管理対象となる ID やマスタが異なり、システムによって閲覧制限を行う。

臨床サンプル管理者は採取後のサンプルを対象とし、患者 ID、プロジェクト毎患者 ID、被験者 ID を管理する。臨床研究者は臨床研究支援システムによって入力された自院のデータを対象とし、自院の患者 ID、プロジェクト毎患者 ID、被験者 ID、臨床研究データを管理する。遺伝子サンプル管理者は臨床サンプルから分注・薄切され、遺伝子検査を行うサンプルを対象とし、プロジェクト毎患者 ID、被験者 ID、遺伝子サンプル ID を管理する。遺伝子研究者は遺伝子サンプルから解析を行い、遺伝子サンプル ID、サンプルの材料、プロジェクト名、研究名を管理する。システム管理者は登録作業などを行うことはできないが、全ての ID を確認することが可能で、システム全体を管理する。

4. 結果

4.1 遺伝子解析を行う際の運用フロー

阪大病院を中心にシステム構築した、がんゲノム医療における運用フローを図5に示す。



サイトサーバーからセンターサーバーへ臨床サンプルID・被験者IDの対応表を送付する。これにより、匿名化状態で遺伝子検査の問い合わせが可能になる。

担当医から臨床研究に関する適格条件に合致した患者に対して説明と同意を行い、同意書を取得する。同意が得られた患者について、電子カルテ連携病院では、CRF レポーターを用いて新規症例登録を行う。(被験者 ID の発行)

次に、検査オーダーを行う。検査オーダー時、検体番号のみが検査システムと連携し、検体ラベルを発行する。検査オーダーのマスタは、プロジェクト毎に作成しているため、検査オーダーにより、何のプロジェクトに対する検体かの情報が取得される。(プロジェクト毎患者 ID の発行)

血液検査の場合は、検査オーダー後、患者が臨床検査室に行くと、検査室で当該試験用の検体ラベルが発行されているので通常の試験管にそのラベルを貼る。病理検体の場合は、病理部で検体を保管している袋に検体ラベルを貼る。その後、試験管や組織ブロックを施設の研究管理者(担当者)が、ラ

ベルをサンプル管理システムのバーコードリーダーで読み取る。CDCS では、患者 ID と被験者 ID の対応テーブルが作成されている。その対応テーブルをサイトサーバーが取り込み、研究用患者 ID と紐付けることにより、臨床サンプル ID を発行する。なお、CDCS 以外のシステムの場合は、被験者 ID をサイトサーバーに手入力する。(研究用患者 ID・臨床サンプル ID の発行)

読み取り後、その検体に対して新たにラベルが発行されると、中央管理施設へ送られる臨床サンプルに対してラベルの貼り替えを行う。その後は臨床サンプル ID をベースに臨床サンプルが管理される。

中央管理施設では臨床サンプル到着後、センターサーバーで臨床サンプル ID のラベルを読み取り、分注数を入力することにより、受け入れ処理を行い、遺伝子管理用のラベルを発行する。(遺伝子サンプル ID の発行)

臨床サンプルを測定用に分注・もしくは薄切を行う。それらに対して遺伝子サンプル ID を認識可能なラベルを貼り、個別のサンプルを保管・管理を行う。

遺伝子検査が終わると、阪大病院がんゲノム医療センターの担当者は、遺伝子解析を行った後のレポートを、センターサーバーへ登録する。

4-2. 遺伝子結果の問い合わせを行う際の運用フロー

臨床研究者は、研究責任者の承諾を取り、サンプル管理センターサーバーにログインして自身が関与している研究を選択し、目的の遺伝子解析を行った解析者を検索する。検索結果から、被験者のリストファイル(研究名+被験者 ID、無い場合はプロジェクト名+プロジェクト毎患者 ID で識別)をダウンロードし、知りたい解析項目(特定の遺伝子型など)を列に追記した表のファイルを作成し、解析者に対して送信する。

遺伝子解析者は、「研究名+被験者 ID(プロジェクト名+プロジェクト毎患者 ID)」が遺伝子サンプル ID に置き換わった表を受け取る。遺伝子解析者は各遺伝子サンプル ID について解析結果を受け取った表に記録し、依頼してきた臨床研究者に送信する。

臨床研究者は、リスト中の遺伝子サンプル ID が「研究名+被験者 ID(プロジェクト名+プロジェクト毎患者 ID)」に置き換わり、遺伝子解析の結果が記載された表を受け取る。

5. 考察

これらの機能により、臨床データとサンプルの解析データを、個人識別情報を見ることなく、紐づけることが可能になった。今までは臨床サンプル ID と遺伝子サンプル ID を管理するシステムはそれぞれ別個に存在しており、セキュリティ属性を高めるためオフラインでデータ照合を行っていた。それをオンラインで行うことにより利便性を高めている。セキュリティと利便性はトレードオフで、このようなシステムではセキュリティ的にオフライン運用を選択しがちであるが、セキュリティポリシーを明確にし、運用管理を行うことにより、オンラインでの運用を問題ないレベルにすることで対応可能と考える。

遺伝子検体管理システムとして、自施設の患者を対象にしたバイオバンクに対応したシステムを構築・運用している報告はある⁴⁾が、データの管理をどのように行っているか、詳細は不明である。本研究では、多施設に対して適用可能であること、電子カルテと連携し、さまざまな臨床情報に対応可能であること、などのメリットを持つ。さらに、患者 ID をキーとして臨床サンプル ID と被験者 ID の管理を同時に行うことにより、臨床

研究者と遺伝子解析者の間で患者の個人情報秘匿したまま、臨床データの問い合わせが可能となっている。

がんゲノム情報レポジトリへの対応について、現在、国立研究開発法人国立がん研究センターを中心として、がんゲノム医療中核拠点病院を手始めに臨床情報を含めたデータを収集することにより、レジストリ研究を行おうとしている⁵⁾。今後、中核拠点病院が連携病院の患者に対して遺伝子パネル検査による医療を連携して行うようになると、本システムのような仕組みにより、効率よい臨床サンプルの管理が必要になると思われる。

参考文献

- 1) 西村多寿子, 進正彦, 植岡健一, 他. 主要臨床医学雑誌の原著論文掲載数と研究内容の国際比較. 国際医療福祉大学学会誌 2017 ; 22 ; 37-48
- 2) 医療の国際数量比較. ニッセイ基礎研究所.
http://www.nli-research.co.jp/files/topics/52143_ext_18_0.pdf?site=nli (cited 2018-Sep-01)
- 3) Y Matsumura, A Hattori, S Manabe, et al. A strategy for reusing the data of electronic medical record systems for clinical research. Stud Health Technol Inform. 2016. 228:297-301
- 4) 増田史恵, 古谷由希, 板倉佳奈美, 他. 京都大学医学部附属病院バイオバンクの立ち上げから現在の運営報告. 医学検査 2018 ; 67 ; 84-89
- 5) 第1回がんゲノム医療推進コンソーシアム運営会議資料. 厚生労働省.
<https://www.mhlw.go.jp/content/10901000/000341604.pdf> (cited 2018-Sep-01)