

HL7バージョン3の病院情報システムへの適用

星本弘之¹⁾ 増田剛²⁾ 坂本憲広¹⁾

神戸大学医学部附属病院 医療情報部¹⁾

(財)先端医療振興財団 臨床研究情報センター 遺伝子データベース研究部²⁾

Adoptation of HL7 version 3 for Laboratory Information System integration with Hospital Information System.

Hiroyuki HOSHIMOTO¹⁾ Gou MASUDA²⁾ Norihiro SAKAMOTO¹⁾

Department of Medical Informatics, Kobe University Hospital¹⁾

Laboratory for Clinical Genome Informatics, Foundation for Biomedical Research and Innovation²⁾

Abstract: This paper describes about the adaptation of HL7 version 3 to Hospital Information System integration in Kobe University Hospital. HL7 v3 is employed as a technical basis for healthcare information network in UK and US. In Japan, HL7v3 is a designated standard for healthcare information exchange by Ministry of Labor, Health and Welfare. For system implementation, we replaced 11 message interactions by HL7 v3 message interactions, 9 are from HL7 v3 Ballot 6 document and rest are originally developed for our system. By our experience, defining and mapping the element to the model is actually easy. However, we found some difficulties when we implement the model to actual messages. We are going to extend the system development by HL7 v3, and will publish the guideline for HL7 v3 implementation, to share the experiences.

Keywords: HL7, Standardization, Hospital Information System, Laboratory Information System

1. はじめに

HL7v3¹⁾は、厚生労働省の「保健医療分野の情報化にむけてのグランドデザイン」²⁾において、HL7v.2.4以降やDICOMとともに使用すべき規格されている。海外においては、英国の国家レベルの保健医療情報ネットワーク(National Programme for IT)³⁾で基盤に採用されているほか、米国CDCの感染症情報サーベイランスシステム⁴⁾の標準メッセージ規格として採用されるなど、欧米諸国を中心にひろく採用されており、保健医療情報分野において現在もっとも注目すべき規格である。しかし、国内では九州大学と福岡市医師会との共同研究による「糖尿病電子カルテ共有システム」⁵⁾、東京大学医学部附属病院での入院患者向け臨床情報提供システム⁶⁾等の事例はあるが、病院情報システムに本格的に導入された例はなく、実装経験の蓄積が求められている。本研究ではHL7v3の実装について検討し、神戸大学病院のシステムに適用した結果を報告する。

2. 対象と方法

我々は神戸大学病院の検査部門システム(LIS)更新に際し、病院情報システム(HIS)との接続にHL7v3を適用した。このシステム間接続は、独自に定義した電文によって接続され、「検査依頼」「検査結果」「検体到着確認」「患者基本情報」「処方実施」「注射実施」「入退院」「転科・転棟」「物品発注・棚卸し」など、合計11のインタラクションがあったが、それぞれ送信項目を整理し、HL7v3投票用パッケージ第6版(Ballot6)に規定されているメッセージインタラクションとの対応を決定し、HL7v3

メッセージ情報モデルの要素に項目をマッピングした。

対応項目、メッセージインタラクションの決定に当たり、HL7バージョン2シリーズの規格書などを参考に、HL7v3における項目・属性マッピングを決定した。

3. 結果

物流関係以外のインタラクションは、Ballot6中に定義されているメッセージインタラクションで対応可能だった。しかし、従来の接続でのインタラクションとHL7v3のBallot6に定義されているメッセージインタラクションの間では、1対1の対応が可能なものもあるが、多くの場合、そのインタラクションによるシステムの状態の変化を考慮してHL7v3のメッセージインタラクションを決定する必要があった。典型的なものとして、従来接続の「検体到着確認」は、検査室に検体が到着した際にLISからHISに送信され、このメッセージ以後、検査オーダーが実施中の状態の変化することから、Ballot6で定義されている“Filler Order Confirmation”に相当するものとしてマッピングを決定した。また、従来は異なる電文で送信していた細菌検査依頼と検体検査依頼の電文は、検査項目コードで識別することにより、単一のメッセージインタラクションに統合された。特に細菌検査依頼メッセージは、旧電文では「真菌」「好酸菌」「一般細菌」など、検査依頼や結果通知において必要となる項目の微妙な違いから異なる電文として定義されていたが、HL7v3においては、検体検査も含めて単一のメッセージモデルで対応が可能であった。Ballot6には利用可能なメッセージモデルとインタラクションが定義されていなかったの

は、物流関係の「物品発注」と「棚卸し」であった。これらについては、処方領域のメッセージ情報モデルに薬剤配送に関する情報を記述するモデルが定義されているため、それを適用し、インタラクションIDに日本独自であることを示す Realm コード“JP”とバージョン番号“01”を付与した。

メッセージ中の要素の数は、整理・統合の結果、例えば検査依頼では92から68へと大幅に減少した。これは旧電文では異なっていた項目も、可能な場合はHL7v3では同一のクラス・属性に割当て、コードで識別するようにしたほか、旧電文ではすでに使用していなかった項目を削除したことによる。

検査項目コードなどのボキャブラリは、神戸大病院で使用中の旧システム用のマスタに、新規項目を追加して作成した。追加に当たってはHL7v3との整合性を考慮し、可能な限りHL7v3関連のものを使用した。

メッセージモデルの定義は比較的容易に完了したが、実装に入る際に、モデルの情報を現実にもどるようにメッセージで表現するのかについて、詳細に定義した実装仕様書が必要であった。これは、同一のメッセージモデルから、複数のメッセージ形式が導出可能であるため、ユースケースごとに現実には検討し、メッセージの実装を定義した。

4. 考察

検査部門システムと病院情報管理システムの接続のHL7v3への更新において、一部メッセージインタラクションを定義する必要などがあったが、特に問題はみられなかった。また、HL7v3への置き換えにより、旧電文と比べて項目数・メッセージ種類とも減少し、メンテナンス性が向上した。

今回のHL7v3によるシステム接続の更新においては、HL7v3メッセージ情報モデルとその項目への対応づけ以前に、旧電文の項目の意味・内容について整理することに大きな労力をとられた。しかし、HL7v3などの標準化規格による接続では、このよう

な作業はほぼ不要となり、蓄積した情報の継続性を確保することができると考えている。HL7v3は、現在も開発中の規格であるが、今回の経験から、その定義済みの情報モデルや新規のモデル記述力は、病院内の各種情報システム間の接続を全てカバーすることが可能であると考えている。

しかし、今回の我々の経験でも、従来の全てのシステム間接続のメッセージが容易にはHL7v3で定義されているものに対応付けられないものも有り、また、必要とするメッセージインタラクションがそもそも定義されていない場合もあった。そのような場合での対応について、現状で不足するメッセージについて、医療機関がそれぞれ独自のメッセージモデルを定義したり、メッセージインタラクションの対応を独自に決定したりしていくのでは、メッセージの相互運用性の確保という標準化規格の目的から逸脱し、大きな問題である。このような事態を避けるためにも、日本の保険医療機関において必要となるメッセージを定義し、その標準的な実装方法のガイドラインを作成・公開する必要があると考えている。

我々は、今後は他のシステム接続についてもHL7v3対応を推進し、新規に開発したメッセージや実装上の問題点について国際的に還元し、国内外の医療機関とも実装の際に得られたノウハウなどについて共有していく予定である。

参考文献

- [1] <http://www.hl7.org/>
- [2] <http://www.mhlw.go.jp/shingi/0112/s1226-1a.html>
- [3] <http://www.nhsia.nhs.uk/>
- [4] <http://www.cdc.gov/phn/>
- [5] 中島直樹, 坂本憲広, ほか: “公開鍵基盤を利用した広域分散型糖尿病電子カルテネットワークシステムの実証実験” 医療情報学 22 (1), 2002: 11-18
- [6] 星本弘之, 渡辺宏樹, ほか: “入院患者向け検体検査結果閲覧サービスの構築と評価” 医療情報学 23 (Suppl.), 2003: 381-384

HL7バージョン3処方オーダーメッセージの開発

増田剛¹⁾ 坂本憲広²⁾

財団法人先端医療振興財団臨床研究情報センター遺伝子データベース研究部¹⁾
神戸大学医学部附属病院医療情報部²⁾

Design and Implementation of HL7 Version 3 Pharmacy Order Messages

Gou Masuda¹⁾ Norihiro Sakamoto²⁾

Laboratory for Clinical Genome Informatics, Foundation for Biomedical Research and Innovation¹⁾
Department of Medical Informatics, Kobe University Hospital²⁾

Abstract: HL7 Version 3 (HL7V3) is one of the standard protocols for healthcare information exchange. In applying HL7V3 messages to practical applications, a great deal of familiarity with HL7 information models such as RIM is required. For this reason, providing not only message definitions, but also some referenced implementation guides that take into account of practical usecases in Japan is useful. In this paper, we design and implement a set of HL7 Version 3 pharmacy order messages. First, we investigate 18 usecases for pharmacy order such as "Activate Inpatient Order" and "Abort Order". Finally, we define 17 message interactions and 4 message types. We are implementing these messages and evaluating them in Kobe University Hospital as a model case. We hope these message definitions are used as a basis for discussing a referenced HL7 V3 implementation guide for Japan.

Keywords: HL7 Version 3, Pharmacy Order Entry System

1. はじめに

保健医療分野における電子的情報交換のための標準化規格の一つであるHL7では、保健医療分野のより複雑な要求に対応するために、次期版であるHL7バージョン3(以下HL7V3)の開発が現在進められている。HL7V3では、全てのメッセージが参照情報モデル(RIM)という一貫した情報モデルに基づき設計され、そのメッセージが表現するセマンティクスはRIMによって厳格に定義される。しかしながら、個々の具体的な文脈で交換されるデータをHL7V3メッセージとして表現するためには、RIMを始めとするHL7V3の情報モデルに精通している必要があり、個々の実装者が独自に行なってしまうと、モデルの誤った解釈が起こりうる。その場合、たとえ標準化規格を利用していても互換性を損なうことになりかねない。従って、HL7V3標準化規格を有効に用いるためには、我が国のユースケースを反映した実装レベルで参照することのできる実装ガイドラインが提供されることが望ましい。そこで本研究では、HL7バージョン2規格において、よく利用されている領域の一つである薬剤領域を対象とし、我が国におけるHL7V3標準化規格の実装に関する参照的なガイドラインを議論するための基礎とならうHL7V3処方オーダーメッセージを開発することを目的とする。

2. 方法

投票のために公開されている最新のHL7V3仕様では、臨床検査や患者事務といった20の領域で150以上のメッセージが定義されている¹⁾。仕様には薬剤領域も含まれており、薬剤領域の全てのメッ

ッセージの導出元となる情報モデルである領域情報モデル(D-MIM)、代表的な2つのユースケースに関してメッセージ内容を定義した詳細化情報モデル(R-MIM)及びメッセージ型、そのメッセージ型を使用する文脈を定義したメッセージインタラクションといった成果物が定義されており、現在試用のための草稿版として公開されている。

HL7V3はRIMに基づきメッセージを開発するためのメッセージ開発方法論を定義しており、その方法論に従って要求を満たす処方オーダーメッセージを一から定義することも可能ではある。しかしながら、類似するメッセージが国や施設、プロジェクト毎にそれぞれ作成されることは、標準化という観点では適切ではない。一方でHL7は、地理的、政治的な要求に従って標準規格を拡張するレムムと呼ばれる仕組みを提供している。そこで、本研究では、まず、これらの既に定義されているメッセージの適用可能性を検討する。次に、既存のメッセージ情報モデルでは対応できない要求事項に関してのみ、HL7V3で規定されたメッセージのローカル化手法を適用し、我が国での要求に対応することのできるメッセージインタラクション、及びR-MIM、メッセージ型を定義する。基礎とするHL7V3仕様として、2004年3月に公開された投票用公表パッケージ7を用いる。

開発する処方オーダーメッセージは、病院情報システムから薬剤部門システムへ送信される処方指示メッセージと、薬剤調剤システムへ送信される調剤指示メッセージを対象とし、投薬や調剤の実施情報を表現する処方結果メッセージは対象外とする。

3. 結果

3.1 メッセージインタラクションの同定

神戸大学医学部附属病院及び麻生飯塚病院を例として、処方オーダーに関する業務分析を行ない、結果として「入院定期処方新規オーダー」や「外来処方変更オーダー」といった、処方オーダー業務に関連する18個のユースケースを洗い出した。これらの各ユースケースに対して、HL7V3メッセージ開発方法論に従い、HL7メッセージが必要とされる現実世界の事象の経過を簡潔かつ写實的に描写するストーリーボードを記述し、そのストーリーボードで使用されるメッセージインタラクションを同定した。このとき、現行のHL7バージョン3仕様で定義されている処方インタラクションを基礎として、さらに、臨床検査領域といったHL7V3の他の領域のメッセージインタラクションの定義や命名規則に可能な限り従い、汎用的な定義となるように努めた。

例えば、現行のHL7V3仕様で定義される処方メッセージは大きく地域診療(Community Practice)と施設内診療(Institutional Practice)の2つに分類することができ、それぞれ外来処方オーダーのユースケースと入院処方オーダーのユースケースに対応付けた。結果として表1上に示す17のメッセージインタラクションを定義した。

3.2 メッセージ型の定義

処方オーダーに必要な用法や用量、薬剤名といった各項目を、HL7V3で提供されるRIMへ対応付け、処方オーダーメッセージR-MIMを定義した。2節で述べたように、現行のHL7V3薬剤領域のR-MIMで定義されているモデルを基礎とし、対応できない要求事項に関してのみ、HL7で規定されているローカル化を適用した。一例として用法情報のRIMへの対応付けを表2に示す。このようにして定義した4つのR-MIM及びそのR-MIMによって定義される表1下に示す4つのメッセージ型を定義した。

4. 考察

現行のHL7V3仕様はまだ開発段階であり、薬剤領域で定義されている現時点のメッセージインタラクションやメッセージ型では、必要となるインタラクションやデータ項目を全て包含することができなかった。しかしながら、RIMは非常に汎用的かつ包括的なモデルであるため、HL7で規定されたメッセージ開発方法論に従い、RIMに基づいて必要なメッセージ情報モデルを定義することができた。現行の仕様に新たに追加する必要が生じた要素には、例えば、処方オーダーに関連する健康保健情報や、処方箋全体に対する投薬コメントと薬剤払い出

しコメント、それらのコメントのコード化された指示があった。これらの結果は、我が国でのユースケースをさらに検討するとともに、必要であれば標準仕様へ反映させていくべきである。

5. おわりに

開発したメッセージは、神戸大学医学部附属病院の電子カルテシステムで使用するために、HL7バージョン3メッセージングライブラリ²⁾を用いて実装中である。今後、国内の他の電子カルテシステムで使用しメッセージの実用性と汎用性について評価を行なうとともに、我が国における薬剤領域の標準的なHL7V3メッセージを議論する際の基礎としたい。

参考文献

- [1] HL7 Version 3 Standard Ballot Site, HL7, Inc HL7 Version 3 Standard, Health Level Seven, Inc. Available at <http://www.hl7.org/v3ballot/html/index.htm>, 2004.
- [2] 増田剛, 他: HL7バージョン3メッセージングライブラリの開発, 医療情報学, 23(Suppl.), 512-515, 2003.

表1 開発した処方メッセージインタラクションとメッセージ型

HL7V3 メッセージインタラクション	結果物ID
入院処方オーダー発行依頼(確認)	PORXJN231010.F01
入院処方オーダー発行依頼(印刷)	PORXJN231020.F01
入院処方オーダー削除依頼	PORXJN231220.F01
外来処方オーダー発行依頼	PORXJN131020.F01
外来処方オーダー削除依頼	PORXJN131220.F01
入院調剤オーダー発行通知	PORXJN231050.F01
入院処方オーダー削除通知	PORXJN231250.F01
外来処方オーダー発行通知	PORXJN131050.F01
外来処方オーダー削除通知	PORXJN131250.F01
入院調剤オーダー発行依頼	PORXJN221020.F01
入院調剤オーダー削除依頼	PORXJN221220.F01
外来調剤オーダー発行依頼	PORXJN121020.F01
外来調剤オーダー削除依頼	PORXJN121220.F01
入院調剤オーダー発行通知	PORXJN221050.F01
入院調剤オーダー削除通知	PORXJN221250.F01
外来調剤オーダー発行通知	PORXJN121050.F01
入院調剤オーダー削除通知	PORXJN121250.F01
HL7V3メッセージ型	結果物ID
外来処方オーダー	PORXJMT232000.F01
入院処方オーダー	PORXJMT132000.F01
外来調剤オーダー	PORXJMT122000.F01
入院調剤オーダー	PORXJMT122000.F01

表2 用法情報のRIMへの対応付け

項目名	RIMへの対応付け
用法区分	1用おに対応するSubstanceAdministrationRequest.code
用法コード	1用おに対応するSubstanceAdministrationRequest.instructionCode MedicationAdministrationInstruction.code
1日分量	1用おに対応するSubstanceAdministrationRequest.quantity SubstanceAdministrationRequest.quantity
用法名称	1用おに対応するSubstanceAdministrationRequest.text
処方日数/回数	日数 1用おに対応するSubstanceAdministrationRequest.effectiveTime 回数 1用おに対応するSubstanceAdministrationRequest.doseQuantity
日数回区分	日数(MedicationSupplyQuantity)で表現
不均衡存在時	日数(MedicationSupplyQuantity)に表現するための不備
処方用量	1用おに対応するSubstanceAdministrationRequest.quantity
処方変更相対開始日	相対開始として1用おに対応するSubstanceAdministrationRequest.effectiveTime
処方変更相対中止日	相対終了として1用おに対応するSubstanceAdministrationRequest.effectiveTime
Med指示	薬剤供給指示として1用おに対応するMedicationSupplyInstruction.code
投与指示	1用おに対応するSubstanceAdministrationRequest.instructionCode

HL7Ver3による臨床検査システム接続

森山 和好¹⁾ 東 重幸¹⁾ 星本 弘之²⁾ 坂本 憲広²⁾

株式会社麻生情報システム 医療システム事業部 先端医療情報システム研究所¹⁾
神戸大学医学部附属病院 医療情報部²⁾

Clinical Inspection System Connection by Using HL7Ver3 Message

Kazuyoshi Moriyama¹⁾ Shigeyuki Higashi¹⁾ Hiroyuki Hoshimoto²⁾ Norihiro Sakamoto²⁾

Laboratory for Advanced Medical Information Systems, Aso Information System Inc.¹⁾
Division of Medical Informatics, Kobe University Hospital²⁾

Abstract: "Ground Design for Computerizing Healthcare Information System Field" by Ministry of Health, Labor and Welfare's Healthcare Information System Investigative Commission has defined HL7Ver3 as a standard to be implemented as well as DICOM.

Based on this announcement, Aso Information System has developed the HL7-based interface for Hospital Information System along with the replacement of clinical examination system in Kobe University Hospital.

We have developed 11 interactions for six months, including examination request, examination result, confirmation on arrival of samples, patient basic information, prescription, injection, admission, discharge, transfer of clinical departments, transfer of wards, and supply ordering.

We would like to introduce our development process and outcomes for this six-month development.

Keywords: HL7Ver3, standardization, Clinical Inspection System

1. はじめに

HL7Ver3は厚生労働省保険医療情報システム検討会が発表した「保健医療分野の情報化に向けてのグランドデザイン」においてDICOMとともに使用するべき標準とされている。

これを踏まえ麻生情報システムは、今後の診療情報の標準化に対応できる病院情報システム開発を目指し、第一段として神戸大学附属病院の検査部門システムリプレースに伴い、病院情報システムと臨床検査システムの接続をHL7Ver3により行った。

2. 開発工程

今回開発したインタラクションは、検査依頼、検査結果、検体到着確認、患者基本情報、処方実施、注射実施、入院、退院、転科、転棟、物品発注の計11インタラクションで、プログラムは各インタラクションとも、今後の他システムへの流用を考慮して、HISデータベースのRead/Writeプログラムと、HL7メッセージ解析及び送受信プログラムで構成した。(図1)

稼動までの作業フェーズは以下のようにっており、約6ヵ月で開発作業を行った。

2.1 送受信項目の選定

リプレース前の送信項目をベースにして、項目個々の意味を明確にするとともに、送信不要な項目の洗い出し作業を行った。

2.2 インタラクションの決定

業務ストーリーを考慮し、トリガイベント、アプリケーションロールより、インタラクションを決定した。

2.3 HL7メッセージマッピング

送受信項目の選定で明確になった項目の意味をもとに、神戸大学医学部附属病院医療情報部の協力を頂き、HL7メッセージへのマッピング作業を行った。

2.4 プログラム製造

開発言語は.net C#、通信プロトコルは、XMLとの親和性と双方のサーバー環境が限定されないことを考慮し、HTTPを使用したSOAP(Simple Object Access Protocol)を採用した。

また、メッセージ実装に関しては、プログラムがメッセージ情報モデルやXML実装仕様を意識することなく作業ができるように、HL7ライブラリを使用した。

2.5 テスト

テスト環境を使用し、メッセージ送受信テストと、検査部の協力を頂き、運用の流れにそったリハーサルテストを実施した。

3. 稼動状況

今回、通信プロトコルとしてHTTPを使用したSOAPを採用するにあたり、既存システムで使用していたSOCKETを使った独自プロトコルと比較しての通信レスポンスに不安があったが、開発当初のレスポンス調査、テストフェーズでのレスポンス測定、及び、本稼動を通して、独自プロトコルと遜色ない計測値を出している。

(本稼動での参考値)※計測値は、メッセージ生成からAckを受信するまで

検体検査依頼情報: 平均メッセージサイズ10Kb
計測値: 0.3981Sec

最小メッセージサイズ4Kb

計測値: 0.0916Sec

最大メッセージサイズ78Kb

計測値: 0.8114Sec

一日で最もトラフィックの集中する15時からの入院患者一括処理においても、平均766件のメッセージを約6分で処理しており、通信レスポンスによる運用面での問題は発生していない。

システム面においても、HL7Ver3実装段階でHL7ライブラリを利用したことによる、標準規格に沿ったメッセージ交換の実現と、送受信項目の意味を明確にしたことで、ベンダー間での意味の取り違いが無くなったことによって、品質の高いシステムの構築につながり、リリース後にトラブルが発生することもなく順調に稼働している。

4. 開発を通して

当初、HL7Ver3を採用するにあたり、メッセージへのマッピング作業に不安を感じていたが、臨床検査システム接続の作業を通して最も労力を必要としたのは、送受信項目の洗い出しと、項目の意味を明確にする作業であり、HL7とは直接的に関係ない作業であった。

検査依頼を例に挙げると、送信項目の意味を明確にすることにより、患者の情報、医師の情報、検査項目の情報、検体の情報に区分され、必然的に、それぞれPatient、Author、Observation、Specimenにマッピングを行えば良いことが明確となる。

結果的にメッセージへのマッピング作業は、それほど困難な作業ではないと感じた。

しかし、今回、麻生情報システムがHL7Ver3に取り組めたのは、神戸大学医学部付属病院医療情

報部を始めとして、HL7Ver3に精通した方々のご支援が頂ける環境であったことが大きく、抵抗無くベンダーがHL7Ver3に取り組む為には、実装手順書等の資料を充実させる必要性を強く感じた。

5. 今後の取り組み

今回の開発により習得した知識と開発ノウハウを基礎として、診療情報の標準化に対応できる病院情報システムの開発を目指し、HL7Ver3メッセージを全面的に採用していく予定である。

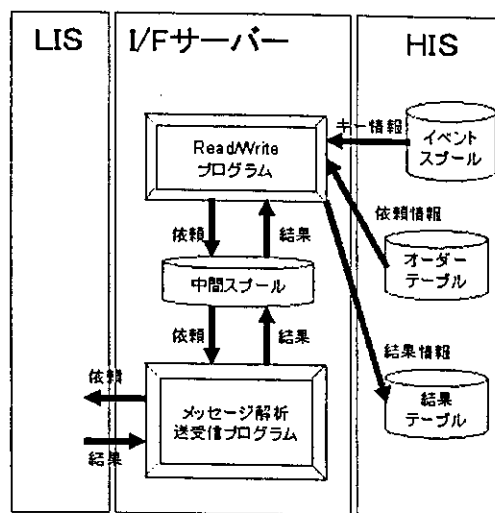


図1 システムイメージ

薬品マスタのHL7バージョン3による記述

広井嘉栄¹⁾ 増田剛²⁾ 星本弘之¹⁾ 坂本憲広¹⁾

神戸大学 医学部附属病院 医療情報部¹⁾

(財)先端医療振興財団 臨床研究情報センター 遺伝子データベース研究部²⁾

A Description of a Medication Master File in HL7 Version 3

Kaei Hiroi¹⁾ Go Masuda²⁾ Hiroyuki Hoshimoto¹⁾ Norihiro Sakamoto¹⁾

Department of Medical Informatics, Kobe University Hospital¹⁾

Laboratory for Clinical Genome Informatics, Foundation for Biomedical Research and Innovation²⁾

Abstract: Recent increase in the amount and complexity of drug information has the possibility of interfering with proper use and safety management of medicine. Therefore, drug information needs to be standardized. Standardization of Medication Master Files is particularly important, because it is used as a basis for drug selection in a doctor's prescription design. However, it is very difficult because of its extremely complex structure. In this study, we have defined a model for the standard Medication Master File by using HL7v3 which is one of the standard protocols for exchanging clinical information. We have applied the model to the description of Medication Master Files in Kobe University Hospital and Aso Iizuka Hospital. Using this model, we were able to describe the Medication Master files which are currently operated in both hospitals. We think that this model becomes the standard Medication Master File by examining Medication Master Files in other clinical facilities in our future work.

Keywords: HL7 Version3, Medication Master File, Drug information

1. はじめに

近年、医薬品情報は増大化、複雑化しており、その取り扱いが非常に困難なものとなっている。その為、医薬品の適正使用や安全性の確保に支障をきたす恐れがあり、医薬品情報の標準化が望まれている。特に、薬品マスタについては、各医療施設の医薬品の情報源であり、医師の処方設計において医薬品選択の直接の根拠となるため、早急に対応する必要がある。しかし、薬品マスタを始めとする医薬品情報は非常に複雑な構造を持つため標準化が難しい。HL7[®]は保健医療情報の国際的な標準化規格であり、特に現在バージョン3(v3)はバージョン2に比べ記述力が優れており、医薬品情報にも対応可能であると考えられる。また、今後、電子カルテシステム、薬歴管理システム等への応用が期待されている。

そこで、本研究では、HL7v3薬品マスタモデルを定義し、実際に神戸大学病院、麻生飯塚病院の薬品マスタを記述することで、標準的な薬品マスタの開発の可能性について検討した。

2. 方法

2.1 メッセージ定義

メッセージの開発にあたり、現行のHL7v3投票用パッケージ6では薬品マスタのためのメッセージは定義されていない。しかし、HL7では、地理的、政治的な要求に従って標準規格を拡張するレルムと呼ばれる仕組みを提供している。本研究では、そのようなHL7v3で規定されているメッセージ開発方法論に従い、薬剤領域の領域メッセージ情報モ

デル(D-MIM)から新規に詳細化メッセージ情報モデル(R-MIM)を定義した。

2.2 マッピング

定義したモデルに対し、神戸大学病院と麻生飯塚病院の薬品マスタの記述を試みた。概念の重複する項目、不必要な項目等は整理し、新規に追加すべき項目についても対応した。

2.3 値集合

HL7では、あるコード化属性の取りうる概念の集合としてボキャブラリドメインを定義している。そして、その属性がメッセージ内で使用される場合、実際に属性が取りうるボキャブラリドメインの概念と特定のコード化体系との関係付けを値集合として定義しなければならない。なお、ある特定のレルムで使用される特定のコード化体系を導入することも可能であり、例えば1つの病院内のような閉じられた通信範囲において使用されるローカルなコード化体系の使用も可能となっている。そこで、本研究で定義したモデルにおけるコード化属性の取りうる概念の整理を行い、値集合を定義した。その際、既存のボキャブラリドメインで対応できないものに関しては、新規にコード化体系を作成し対応した。

3. 結果

神戸大学病院の10の薬品関連マスタは薬品マスタ、医療材料マスタ、相互作用マスタの3つのR-MIMに集約され、各マスタに含まれていた161の項目は約90にまとめられた。更に、麻生飯塚病院の薬品マスタの記述も可能であった。なお、相互作用マスタに関しては薬品マスタに含めることも

できるが、更新頻度等、運用上の都合により別マスタとした。これら3つのR-MIMは、レルム「日本(JP)」を付けバージョン番号01として定義した。また、既存のボキャブラリドメインを考慮しつつ、29の値集合を定義した。そのうち、既存のボキャブラリのみで対応可能なものは4つであった。図1に、本研究で定義した薬品マスタR-MIM図を示す。

4. 考察

本研究で開発したモデルは、神戸大学病院、麻生飯塚病院で運用されている薬品マスタに適用可能であり、薬品コードとしてローカルなコードと共にHOTコード²⁾を採用しているため、今後、他施設の薬品マスタについても検討し、必要な情報を整理、追加していくことにより標準的な薬品マスタになりえると考えられる。また、一薬品一ファイルであるため、医薬品に付随する情報の一元的な管理が可能であり、医薬品の適切な使用や安全性の確保、流通・在庫管理業務の効率化に寄与すると考える。なお、開発した薬品マスタを、神戸大学病院で現

在開発中の電子カルテシステムで実際に使用し、実用性の評価を行っていく予定である。

5. 課題

今回は既存のボキャブラリドメインに概念の存在しないコード化値についてはローカルなコード化体系を参照する値集合を定義し対応した。HL7v3のボキャブラリドメインでは既に保健医療分野における各領域の概念が体系的に整理され、コード化がなされている。しかし、そこで定義されているコード化値を日本におけるメッセージ定義の際にも使用できるとは限らない。日本特有の概念が存在するのも事実であり、今後、HL7との対応をとりつつ、日本でのコード化体系を整理していく必要があろう。

参考文献

- [1] HL7 Version 3 Standard, Health Level Seven Inc., Available at <http://www.hl7.org/>, 2004.
- [2] 標準医薬品名マスター, Available at http://www.medis.or.jp/4_hyojyun/download/index.html, 2004.

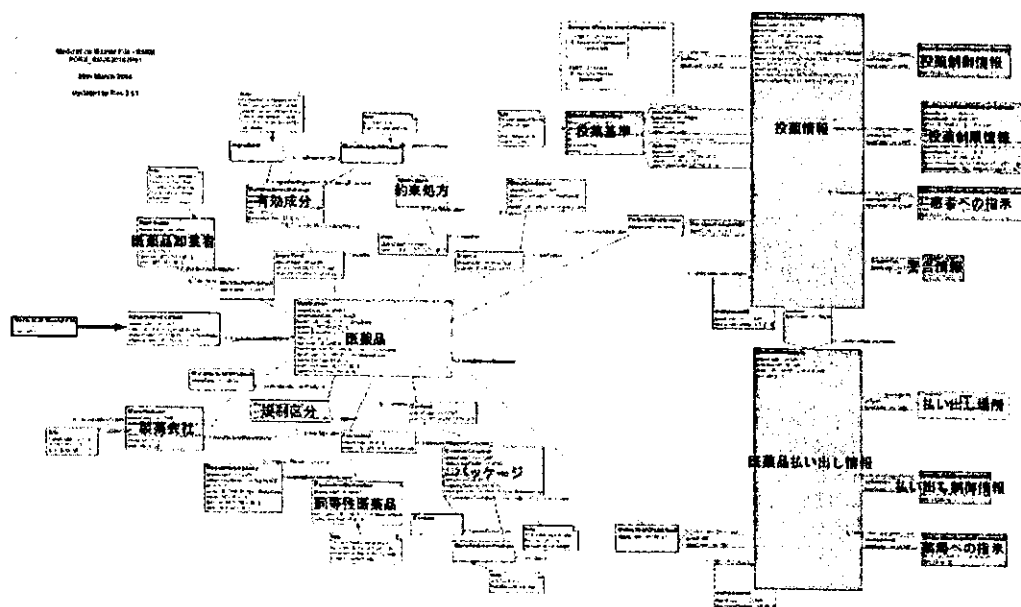


図1 Medication Master File R-MIM

電子化に向けた糖尿病カルテの分析

山本 さつき 前田 英一 星本 弘之 坂本 憲広

神戸大学 医学部附属病院 医療情報部

Analysis of The Progress Records of The Diabetic for Electronic Medical Records

Satsuki Yamamoto Eiichi Maeda Hiroyuki Hoshimoto Norihiro Sakamoto

Division of Medical Informatics, Kobe University Hospital

Abstract: To investigate an efficient way to the structural progress recording, we analyzed terms in the current paper-based medical records. Frequently used terms were picked up from each record of 100 diabetic subjects, and calculated the ratio of utilization of them in all patients. Most highly utilized term among all subjects was 'Plasma Glucose' (60%), however, majority of highly utilized terms were about general findings mainly in the objective findings record. In suffered subjects with the specific diabetic complications, terms about their complication were appeared in every day's records, however, their utilization ratio in all subjects were low. It is considered that suitable combinations of terms that are categorized by different characters in utility are useful for electronic progress recording.

Keywords: Electronic Medical Records, Standardization

1. はじめに

診療録の電子化において、経過記録用紙の記載は、その多様な記載様式のため標準化が進んでおらず、しばしば、自由文を中心とした記録となっている。経過記録用紙には、患者の主訴や、他覚的所見など診療時の症状のほか、医師の診断過程や診療計画案、治療評価などが記載されている。それらは、医療情報として重要な部分であり、経過記録を自由文として記載にすることは、電子診療録開発目的の一つでもある。記述の標準化や構造化、データマイニングへの利用性を損なうことになり、解決すべきテーマの一つである。我々は、糖尿病という罹患率が高く臨床像がよく知られた疾患をモデルとして、現在の紙診療録の記載を基に、経過記録用紙が実際にどのような形で記録されているかを検討し、項目の標準化について考察した。

2. 方法

平成15年度に、本院糖尿病内科に入院した患者の入院診療録を対象として、個々の診療録内で使用頻度の高い項目を抽出し、それらの項目が、全調査対象診療録間において普遍的に使用頻度の高い項目として用いられているかを検討した。まず、診療録ごと、項目ごとの使用頻度の分布を概観するために、最初に20症例分の診療録において、医師による経過記録用紙に記載されている全用語を項目立てし、項目ごとに、毎日の経過記録の記載の中での出現回数を計数して入院期間中の全記載日数に対する出現回数の割合を求めた。その結果を基に、単一診療録内で出現回数が20%以上記載のある項目を、「使用頻度の高い項目」と設定した。ついで、全調査対象診療録の経過記録から「使用頻度の高い項目」を抽出し、その項

目が全調査対象診療録において、どの程度普遍的に高頻度として用いられているかを、「再現率」として表した。なお、一冊の同一診療録であっても、担当医の交代などで表記方法の変化があった場合は、別の診療録として取り扱ったため、最終的に、対象診療録数100冊、同記載医師数17名となった。それらを基に、経過記録に記載されている項目を、共通項目として電子診療録に展開することが、可能かどうかを検討した。

3. 結果

各診療録から使用頻度の高い項目として抽出された項目は、全対象診療録100冊において30項目であった。その中で再現率が3%以上であった項目および再現率を表1に示した。最も再現率の高い項目は「血糖値」であり、60%の診療録において、日々の経過記録中に20%以上出現する「使用頻度の高い項目」であった。ついで再現率の高い項目は、「BP」(血圧)、「BT」(体温)、「HR」(心拍数)といったバイタルサインに関する項目、「heart」「S1S2clear」「no murmur」(心音)、「N. V. S.」「no rale」(胸部聴診所見)、「abdomen」「soft & flat」「グル音」(腹部所見)といった一般的他覚的所見に関するものであった。主訴に関する項目の再現率は、「no new complaint」(「著変なし」「no remarkable」「stable」)の17%を除いては、他覚所見に関するものと比べて全体に低く、3%以上の再現率を示した項目は「浮腫」「口渇」のみであった。「ふらつき」「浮遊感」「倦怠感」「冷や汗」などの低血糖症状に関する所見、「手足のしびれ」などの末梢神経障害に関する所見は、再現率は1%程度と高くはなかった。また、患者の訴えや感想等に関する記載は、同様の記載がほぼ毎日記述されており、単一の診療録に

においては高頻度で出現したが、それらは、患者自身の言葉に基づく多様な表現により記述されていたため、異なる患者間での再現性は高くならなかった。

評価・立案の項目の記述に関する項目は、「BSコントロール」におけるインシュリン投与についての記載が再現率29%であった。なおインシュリン投与の記載については、表記が様々であり、パターンがいくつか見られたが、今回の調査ではまとめて扱った。

4. 考察

今回の検討は、複雑な記載がなされる経過記録を構造化・標準化するための最初のステップとなる用語の項目立てが可能か否か、その効果はどの程度のものかを考えるために、糖尿病という比較的頻度が高い疾患を用いて行った。その結果、単一診療録において使用頻度の高い項目は、その診療録の中ではほぼ毎日記載されるなど繰り返し使用される傾向にあったが、全調査対象診療録を通して見た場合、再現率の高い項目は、主に他覚所見、なかでも疾患に特徴的ではない一般的な他覚的所見に集中する傾向が示された。これは糖尿病といえども一般的な内科疾患でもあり、通常の内科診療で施行される他覚所見が高率であったことは不自然なものではなく、これら再現率の高い項目を電子診療録の標準的項目とすることは意義が高いものと考えられた。一方、糖尿病性合併症など、糖尿病に特徴的な所見に関する項目は、単一診療録においては極めて繰り返し使用されていたが、全対象診療録の中では再現率が高くなかった。これは、糖尿病合併症という糖尿病に特徴的なものであっても実際にはそれほど頻度の高いものではないためと思われる。従って、これらを全患者に適用する標準的な項目として扱う意義は必ずしも高いものではないが、必要な症例においては利用可能とすることは、意義が高いものと考えられた。さらに、「腹部エコー」「グルカゴン負荷」等の内科領域、糖尿病領域で頻繁に行われる検査項目は、ほぼ全ての対象診療録において一度は用いられる項目であったが、今回の調査では日々の経過記録を対象としたため表1には挙げられていない。しかし、これらは出現頻度こそ低いものの、診療過程においては、必ず使用される項目であり、日々の経過記録とは別のグループとしての項目化は必要と考えられた。このように、項目の用いられ方の性格に併せて、項目化の方法を工夫することにより、多くの記載が構造化・標準化できるものと思われる。今回の調査で

は、毎日診察が行われていたにも関わらず、経過記録には記載があまりない診療録も少なからず見られた。これらは、糖尿病の症状には日々大きな変化がなく、変化がなければ所見があっても経過記録として特に記述しないケースがあるためと思われるが、このようなケースであっても、項目を適切に設定することにより、日々記録される所見情報の量的質的向上が期待されるものと思われる。

一方、主訴や所見についての記述には微妙な表現も多い。例えば、「変わらない」という患者の言葉は、症状がないという意味と、前日との変化がなく、依然として神経障害などの持続的症狀が続いている場合も含まれている。また患者によって表現の仕方も多様であり、その意味合いも微妙に異なっている。主訴のニュアンスをそのまま表現することもまた必要であると思われ、自由文入力も不可欠と考えられた。今後、本来の意義を損なうことなく診療録の電子化を進めて行くにあたり、これらをいかに効果的に組み合わせていくかが検討される。

表1 高再現率項目リスト

項目	再現率
血糖値	60 %
BP	35 %
BT	35 %
BSコントロール	29 %
HR (PR)	30 %
heart (no rale)	17 %
heart (no murmur)	14 %
abd (soft&flat)	13 %
PR	13 %
heart (S1S2 clear)	13 %
abd グル音	12 %
lung (N. V. S.)	12 %
no new complaint *	17 %
BW	7 %
浮腫	5 %
口渇	3 %

* 「著変なし」「no remarkable」「stable」などを含む

PKIを用いた広域対応の臨床試験情報収集システムの構築

西脇 清行¹⁾ 増田 剛²⁾ 坂本 憲広¹⁾

神戸大学医学部附属病院 医療情報部¹⁾

財団法人先端医療振興財団 臨床研究情報センター 遺伝子データベース研究部²⁾

Construction of a wide area clinical trial information reporting system with PKI

Kiyoyuki Nishiwaki¹⁾ Go Masuda²⁾ Norihiro Sakamoto¹⁾

Department of Medical Informatics, Kobe University Hospital¹⁾

Translational Research Informatics Center, Foundation for Biomedical Research and Innovation²⁾

Abstract: Paper based management of clinical trial data is inefficient, because the data is collected from nationwide. Therefore, construction of efficient data management system with online data collection would be cost effective. To assure the system's security, Public Key Infrastructure is suitable for the purpose. However, a large sum of cost is needed to construct CA in the medical institutions in the whole country. Then, we construct a RootCA, and the certificate is issued from HPKI compliant SubCA in each medical institution. The digital signature can be verified, because RootCA is the same. And reductions in cost are possible. In our practical experiment, HPKI compliant SubCA in Kobe University Hospital and TRI were used. SSL client authentication was introduced in the clinical trial data collection server, and digital signature was added to the clinical trial data. So security countermeasures were done. It was possible to register and inspect the clinical trial data, it was also possible to verify the electric signature in the demonstration. However, there was a problem with operativeness deteriorates by SSL client attestation, addition of the digital signature, etc. Therefore, in the future, it is important to construct system that users can use without the consideration of PKI.

Keywords: PKI

1. はじめに

全国多施設共同で実施される臨床試験のデータは、紙ベースで授受されてきたためにデータ管理の効率が悪く、インターネットを使用したオンラインシステムによる効率のよいデータ管理を行うシステムの構築が不可欠である。また、臨床試験データには個人情報が含まれるためユーザ認証・通信の暗号化と、データが誰によって作成されたものなのか確認する必要もあり、普及しつつあるPKI(公開鍵基盤)を利用する方法がある。

しかし全国に広がる医療機関から独自に電子証明書を発行・管理するに、RootCA設置、CP(証明書ポリシー)やCPS(運用規定)の独自策定などで多大な費用が必要となり、データを収集・利用する際には複数のRootCAを信頼する手間がある。

この問題を解決するにあたり、共通RootCAを設け、各施設にSubCAを設置し、RootCAの信頼、一箇所のCAから多施設のユーザへの証明書の発行の手間を省く方法がある。

そこで本研究ではMEDIS-DCのHPKI(ヘルスケアPKI)準拠RootCAと、それをRootとするTRI((財)先端医療振興財団臨床研究情報センター)、神戸大学医学部附属病院のHPKI準拠SubCAを使用し、その有用性を検証した。

2. 方法

TRI、京都大学医学部附属病院、大阪大学医学部附属病院、株式会社エスアールエルの関係者にはTRIより電子証明書を発行し、神戸大学医学部附属病院の関係者には院内のSubCAより電子証明書を発行した。また臨床試験情報サーバはTRIに設置し、各施設のPCから模擬的な臨床試験データを入力してもらうことにより実験を行った。

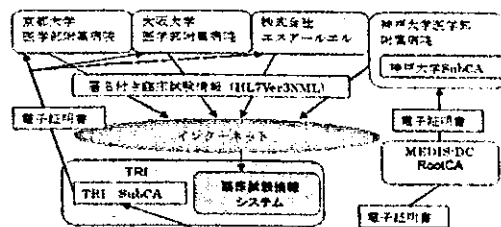


図1 システム構成図

2.1 臨床試験情報収集サーバシステム

収集サーバはMicrosoft Windows 2003 Serverを使用し、各施設の臨床試験データ登録・閲覧サイトはSharePoint Serviceにて構築した。セキュリティ対策としてWebサーバへの認証には、SSLクライアント認証を使用し、各施設のWebサイトへの認証はユーザ名・パスワードを用いた。SSLクライアント認証にて各施設のサイトへの認証もIIS(Internet Information Services)の証明書マッピング機能を使

用することで可能であるが、ユーザ数が多くなるとユーザと証明書のマッピングに時間を要するために見送った。

2.2 臨床試験情報収集クライアントシステム

臨床試験データを入力するクライアントアプリケーションには Microsoft InfoPath 2003 を使用することにより、入力しやすい GUI と、HL7 Version3 に準拠した XML を作成しサーバへ保存する機能を提供した。

臨床試験データ登録時は、サーバに登録されている InfoPath フォームテンプレートを IE (Internet Explorer) 上から開くことにより InfoPath が起動され、データを入力することが可能である。閲覧時は InfoPath により作成された XML ファイルを IE 上から開き、InfoPath にて閲覧を行う。

作成された XML の真正性の確保と作成者・認証者の確認を行うため、サーバに保存する際には作成された XML に電子署名の付加が必須となるように構成した。さらに、サーバに保存されている XML を閲覧する際は、電子署名の検証を行い、かつ電子署名に使用された証明書が HPKI 準拠であるか検証を行うことにより、XML の真正性と作成者・認証者の確認を行えるようにした。下記は証明書の検証時に行った確認事項である。

- 公開鍵証明書の発行 CA 署名検証
- 公開鍵証明書信頼チェーンの確認
- 公開鍵証明書の証明書ポリシーを確認
- 公開鍵証明書有効期限の確認
- CRL の確認
- KeyUsage、BasicConstraints の確認
- EE 証明書の hcRole

3. 結果

同じ施設のユーザ間では、同じ SubCA から証明書が発行されているため、他のユーザが登録した臨床データを閲覧する際の電子署名の検証が正常に行うことができるのはもっともであるが、他施設のユーザ間で異なる SubCA から発行された証明書を使用した電子署名でも、正常に検証可能であった。また、Webサーバによる SSL クライアント認証も RootCA、各 SubCA を信頼しているため、各施設のユーザが SSL クライアント認証を行うことができ、権限の無い者による登録・閲覧は不可能となり、かつ通信の暗号化・改ざん防止が可能であった。

臨床データの登録・閲覧時に使用する証明書が CA により失効された場合は、SSL クライアント認証は不可能となり、登録・閲覧時が不可能となった。また、すでに登録されている臨床試験データの電子署名の検証も失敗した。

4. 考察

臨床試験データに付加した電子署名の検証では、証明書が一度失効されると以前に登録していたデータも検証できなくなるため、中・長期的な臨床試験のデータを収集するには問題が生じる。よって、運用方法または電子署名の付加方法を見直す必要がある。

また、実験終了後に行った電子署名の必要性(図1)と今後のネットワーク型電子カルテの利用(図2)についてのアンケートでは、実証実験に参加したすべての人が電子署名は必要であると回答した。しかし、今後のネットワーク型電子カルテの利用については、積極的な回答は少なく、その理由として自由回答ではセキュリティ上の問題と手間の増加があげられた。これは、エンドユーザの多くが最近多発する個人情報漏洩事件を問題視しているが、実務の効率性が落ちることも同様に問題視していると思われる。よって、RootCA を一つにすることにより、各施設に CA を構築する際のコストの問題は解決できるが、ユーザが満足するセキュリティの確保と、PKI を使用した場合の利便性の低下の相反する二つの大きな問題が浮き彫りとなった。

このことから、今後はユーザが意識しなくともセキュリティの確保できるシステム構築も重要である。

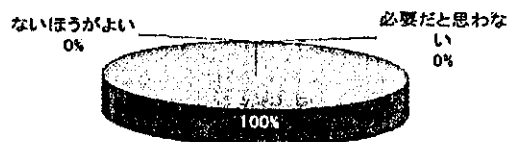


図2 電子署名は必要性があると思うか

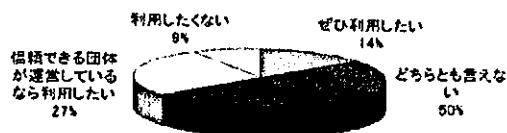


図3 今後、インターネット通信の「電子カルテ」システムを利用したいと思うか

保健医療情報標準化規格に基づく疾患関連遺伝子解析研究のための データ収集基盤

増田剛¹⁾ 広井嘉栄²⁾ 坂本憲広³⁾

財団法人先端医療振興財団臨床研究情報センター遺伝子データベース研究部¹⁾
神戸大学医学部附属病院医療情報部²⁾

A Data Gathering Infrastructure Based on a Healthcare Data Exchange Standard for Candidate Genes Analysis

Gou Masuda¹⁾ Kaei Hiroi²⁾ Norihiro Sakamoto³⁾

Laboratory for Clinical Genome Informatics, Foundation for Biomedical Research and Innovation¹⁾
Department of Medical Informatics, Kobe University Hospital²⁾

Abstract: Identification of disease related genes is one of the important research topics in post-genome era. Genome-wide association analyses with Single Nucleotide Polymorphisms (SNPs) are widely performed. In the analysis, it is pointed out that making use of detailed clinical information is essential in order to characterize and classify phenotypes. On the one hand, a lot of clinical data are collected electronically in healthcare institutions. However, it is difficult to analyze, compare and evaluate the data in a common measure because its format, code and normal value are different each other. A standardized method for clinical and genomic data is required. In this study, we develop a data gathering format for candidate gene analysis with clinical and genomic information based on HL7 Version 3 which is a standard protocol for healthcare information exchange. We then applied this format to the data set which is collected in the Japanese Millennium Genome Project. In addition that, we develop a piece of software library to develop any tools or applications based on the format. By using HL7 Version 3, the data gathering format becomes more complex than simple spreadsheet formats which are conventionally used. However, the software library makes it easy to handle such a complex format and to develop tools or applications based on the format.

Keywords: Clinical Genomics, Disease Related Gene Analysis, HL7 Version 3

1. はじめに

疾患の発症に関連した遺伝子を同定する疾患関連遺伝子解析研究はポストゲノム研究の重要な課題の一つである。現在、一塩基多型 (SNP) を代表とする遺伝子多型を利用した全ゲノム探索による相関解析が広く行われている。しかし、糖尿病のような多因子疾患では、異なる病因の患者が同一の集団として扱われてしまうため偽陽性が起こりやすいという問題が指摘されており、この問題を回避するためには詳細な臨床情報を利用し表現型をさらに詳細に分類する必要がある¹⁾。一方で、臨床の現場では診療情報の電子化が進み、詳細な臨床情報が収集されつつある。しかしながら、これらを疾患関連遺伝子解析研究に利用する場合、データが各施設に固有の方法で記述されており、データ項目の定義や基準値、コードが統一されていないため、収集されたデータを共通に解析・比較・評価することが困難である。臨床情報と統合化された疾患関連遺伝子解析を行なうためには、標準化規格に基づいたデータの収集形式が必要である。さらに、このような収集形式がデータ収集のための情報基盤として利用されるためには、データ形式を簡便に扱うためのツールやライブラリが必要不可欠である。そこで本研究では、疾患関連遺伝子解析研究

のためのデータ収集形式とその形式を簡便に扱うためのツールを構築することにより、疾患関連遺伝子研究のためのデータ収集基盤の構築を試みた。

2. 方法

保健医療分野における電子的な情報交換のための国際的な標準化規格の一つに Health Level Seven (HL7) がある。特に、HL7の最新版であるHL7バージョン3(以下HL7V3)は、より表現力の高い情報モデルRIMを有しており、検査や処方データといった臨床情報だけでなく、アレルやSNPといった遺伝子情報をも統一的なモデルで表現することを可能とする。そこで本研究では、標準化規格としてHL7V3を用いた疾患関連遺伝子解析研究のためのデータ収集形式を開発する。また、著者らが開発しているHL7V3メッセージを扱うための基盤ライブラリ²⁾を用いて、標準化規格の詳細を意識する必要なく、開発したデータ収集形式を扱うことを可能にするためのメッセージインターフェースを開発する。開発したデータ形式は、ミレニアムプロジェクト糖尿病疾患グループの収集データ³⁾をモデルケースとして、その有効性を評価する。

3. 結果

3.1 データ収集形式

疾患関連遺伝子解析研究のためのデータ収集形式は、HL7バージョン3のメッセージ開発方法論に基づき、HL7RIM から導出されるHL7V3メッセージ型として定義した。臨床情報に関しては、臨床検査領域で定義されている検査結果メッセージを基礎として用いた。その結果、ミレニアムプロジェクトで収集される年齢、BMI、HbA1c、FPGといった、計44の項目をHL7V3の詳細化メッセージ情報モデル (R-MIM) として表現することができた。一方、遺伝子情報に関して、現在HL7では、主に診療という観点からゲノム情報を扱うための標準化規格が検討されている。本研究では、その中で現在開発中である情報モデルを基礎として用いた。このモデルでは、SNPなどのゲノム情報は、観察という行為やその観察結果を表現するRIM Observationクラスを用いて、臨床検査結果と同様に統一的に表現される。このモデルを基礎として、疾患関連遺伝子解析研究で必要となる遺伝子やSNPの情報を、図1に示す通りHL7V3のR-MIMとして表現した。

3.2 データ収集基盤

標準化規格を用いた結果、従来までの単純なスプレッドシート形式と比較して、より複雑な形式となった。そのため、筆者らが研究を行なっているHL7V3メッセージングライブラリ³⁾を利用して、規格の詳細を意識することなく簡便に扱うためのメッセージインターフェースを開発した。例えば、1つのSNPデータに対応するインターフェースISnpObservationを定義することで、SNPの識別子やタイピング結果を、HL7V3メッセージの詳細を知ることなくメッセージ中に設定・取得することができる。

4. 考察

遺伝子情報に関して、HL7で現在定義されている情報モデルは、主に診療の観点で定義されたモデルであり、個人に対して非常に多くのSNPを扱うユースケースが考慮されていない。しかしながら、疾患関連遺伝子解析研究の場合には、例えば、個人の10万のSNPデータを解析対象とするミレニアムプロジェクトの例のように、大量のSNPデータをメッセージとして表現する必要がある。今回モデルケースとして使用したデータは、1サンプルあたり58個のSNPデータを含み、これをHL7V3メッセージで表現すると約70KBのXML文書となった。単純に計算すると10万SNPでは約100MBのHL7V3メッセージとなり、1メッセージとして扱うことは実用的ではない。今後、複数のSNPデータを1つのObservationクラスの値として表現するようなモデルを考える必要がある。

5. おわりに

標準化規格を用いた結果、従来までの単純なスプレッドシート形式と比較して、より複雑な形式となった。しかしながら、標準化規格をサポートするライブラリとメッセージインターフェースを提供することで、規格の詳細を意識することなく扱うことが可能となった。今回明らかとなった大量のSNPデータの扱いに関する課題は、今後検討していく必要がある。

参考文献

- [1] 坂本憲広: 臨床ゲノム情報学への誘い、医療情報学、23 (Suppl.), 138-139, 2003.
- [2] 増田剛, 他: HL7バージョン3メッセージングライブラリの開発、医療情報学、23 (Suppl.), 512-515, 2003.
- [3] 三宅一彰, 他: Perspective: ミレニアムプロジェクトと今後の展望、Molecular Medicine、40 (9), 2003.

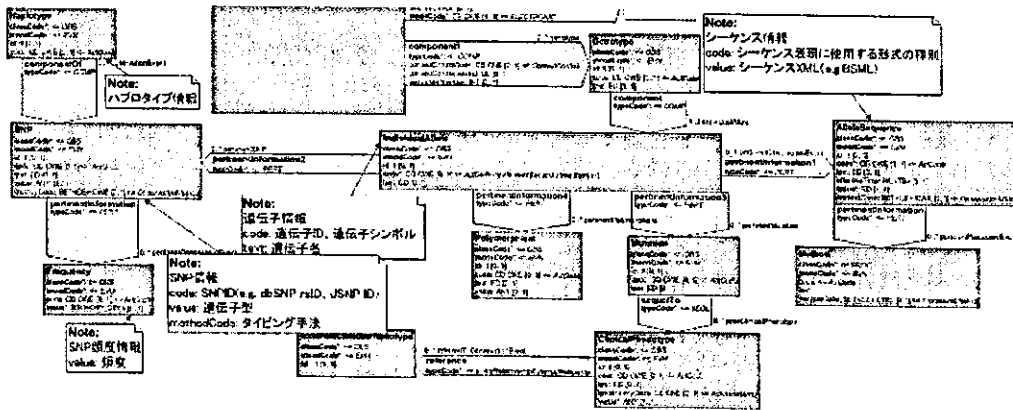


図1 開発した遺伝子情報のためのR-MIM (抜粋)

Towards the Construction of the Information Infrastructure for Genome Medicine

Norihiro SAKAMOTO

Department of Clinical Genome Informatics, Graduate School of Medicine, Kobe University

7-5-2, Kusunoki-cho, Chuo-ku, Kobe, JAPAN

nor@med.kobe-u.ac.jp

Abstract. One of the most expected applications of the human genome information is the development of genome medicine. Towards this goal, basic and applied biomedical laboratory researches and developments have been steadily in progress. However, clinical genome informatics and its applications in practice have been making slower progress compared to biomedical laboratory works. In order to facilitate the progress of clinical genome informatics and its powerful application to the development of genome medicine, the construction of the information infrastructure for genome medicine calls for urgent attention.

In this paper, we propose the architecture of the information infrastructure for genome medicine. It is required to provide three essential features: security mechanisms, comprehensive information models, and intelligent analyzers. To implement these features, we employed PKI as the security mechanisms and the HL7 Version 3 as the information models. To analyze information, knowledge discovery tools are expected to be implemented in addition to a lot of clinical and genetic statistical functions.

1 Introduction

The Human Genome Project was successfully completed in May, 2003. The project has produced the detailed map and the comprehensive dictionary concerning the human genome on a basis of about 3 billion nucleotide sequences. These results have changed the way researchers approach the life sciences and have pushed them to the post genomic era.

1.1 Clinical Genome Informatics

The major research topics in the post genomic era contain proteomics, system biology, genome-based drug discovery and other advanced life science areas. Among them, the research and development of genome medicine is one of the most expected applications of the human genome information. The research and development of genome medicine need both high throughput biological laboratory works and high performance computational power. Towards realizing genome medicine, basic and applied

biomedical laboratory researches and developments have been steadily in progress so far. However, informatics for genome medicine and its applications in practice have been making slower progress compared to biomedical laboratory works. Bioinformatics is the critical research field and the key information technology that supports genome medicine. Although bioinformatics seems to have been active in basic life science and genome based drug discovery among wide range of its scope, its activity in the application field of genome medicine seems to be lower. We have taken up and focused this research field of bioinformatics that supports the development of genome medicine and named it clinical genome informatics. In order to facilitate the progress of clinical genome informatics and its powerful application to the development of genome medicine, we consider that the construction of the information infrastructure for genome medicine calls for urgent attention. Towards the construction, we describe the architecture of the information infrastructure for genome medicine in this paper.

1.2 Research and Development Processes of Genome Medicine

As the first step of the research and development of genome medicine, genomic analyses for detecting disease susceptibility genes have been popular. The genomic analyses are mainly categorized into the following two approaches: the candidate gene approach and the whole genome scan approach. The whole genome scan approach for detecting disease susceptibility genes makes more use of the human genome information, particularly the single nucleotide polymorphisms (SNPs) and microsatellites (MSs), than the candidate gene approach. The whole genome scan approach is further categorized into the following subgroups based on the methodology of genetic statistics: the linkage analysis and the linkage disequilibrium analysis or the association study. In order to determine a disease susceptibility gene, the linkage analysis usually requires 10-200 samples or patients that are members of a few large families while the linkage disequilibrium analysis requires 200-2,000.

The main targets of genome medicine are so called common diseases that are very popular in today's society, for example, diabetes mellitus, hypertension, heart disease, cancer and so forth. The association analysis is more suitable for the genome analyses of such common disease because it is easier to collect a large number of samples of the common disease than to get a large family. The millennium genome project employed the approach based on the association study. The millennium genome project started in 2000 in order to detect susceptibility genes of the common disease. The project consists of the following 5 subgroups: Alzheimer's disease, asthma, cancer, diabetes mellitus (DM), and hypertension. The DM subgroups have been led by Prof. Masato Kasuga from Kobe University and we have been involved in the subgroup. Each of subgroups has collected and analyzed more than 200 patients (cases) and 200 normal controls on average. The DM subgroup collected 178 cases for the first screening and 752 cases and controls of the same number for the second and third screening. More than 15 medical schools have been cooperating to collect such a large number of cases and controls. This cooperation has been strongly

facilitated by using the Internet. We have set up a home page and shared analysis results. The accesses to the home page are secured through SSL.

Another example of such a large study was reported in the research paper on an affected sib pair study on DM by Prof. Hajime Nawata from Kyushu University [1]. They collected more than 100 cases and their siblings from more than 30 medical facilities. It is clear that the secure communication infrastructure through the Internet is indispensable to connect the participant medical facilities and researchers and share the information.

We need to handle various types of information in the research and development of genome medicine. These pieces of information include genome sequences, genomic structures such as promoters, exons and introns, amino acid sequences, protein structures and functions, genomic polymorphisms such as SNPs and MSs, metabolic pathways, signal transductions, and so forth. In the public biological databases, a part of them are usually managed in relational database systems. Such management is appropriate for the purposes of retrieving and browsing pieces of information. However, the efficient research and development of genome medicine requires the integrated management and analyses of the whole of those pieces of information. For this purpose, a relational model is not necessarily suitable but more powerful data model such as an object data model is preferable. Health Level Seven (HL7) is the international standards for the messaging and communications in the fields of healthcare and provides the reference information model (RIM). HL7 is now extending its scope to the field of clinical genomics.

The common diseases are usually considered as a multifactorial disease. More than 10 genes in addition to environmental factors are involved in the development process of the disease. The extent of the involvement of each gene varies from case to case. According to such difference, a case shows different phenotypes or clinical symptoms. Therefore, efficient analyses of disease susceptibility genes require the combined analyses of the genome information and the clinical information. The genome information and the clinical information have been usually analyzed separately. In the genome analyses of disease susceptibility genes, the clinical information is used only when cases are diagnosed. However, in the genome analyses regarding the common diseases, patients of which present different phenotypes, any pieces of clinical information are required to be used in order to categorize the cases into subgroups. The combined information of the genome information and the clinical information also requires the combined method of clinical statistics, genetic statistics, and data mining. Such integrated intelligent approaches are critical for the efficient research and development of genome medicine.

1.3 Intelligent Database

To summarize the above requirements, the architecture of the information infrastructure for genome medicine shall possess the following features: security mechanisms, comprehensive information models,

and intelligent analyzers. We recognize the information infrastructure for genome medicine with these features as a natural extension of the intelligent database.

The intelligent database represents the state of the art in the evolution of database technology [2]. The intelligent database handles a variety of functions ranging from data modeling and management to information discovery, summarization, and presentation. The basic architecture of the intelligent database consists of three layers: the intelligent database engine, the object model, and a set of high-level tools that tend to vary somewhat according to the application [3]. The intelligent database handles the storage, retrieval and management of information. It also handles extraction of data from heterogeneous databases and the various communication and transaction protocols required by this activity. In terms of the intelligent database, the information framework for genome medicine is the clinical genome specific one featured by secure communication on the Internet. Therefore, the architecture of the information framework for genome medicine can be primarily a natural extension of the intelligent database.

2 Method

We designed the architecture of the information framework for genome medicine based on the concept of the extended intelligent database. Here, we describe the approaches of design and implementation for the three outstanding features: the security mechanisms, the comprehensive information models and the intelligent analyzers.

2.1 Security Mechanisms

For making use of the genome information and clinical information across healthcare institutes, standardization of data formats and vocabularies of representing the information are essential. The detail regarding this issue is described in the next subsection. In addition to them, digital signature is a critical key technology for making the information to be exchanged secure and trustworthy [4]. Especially for the genome information, digital signature is much more important because of its reusability. Efficient exchanges of authorized information with a digital signature in healthcare information networks require a construction of a public key infrastructure (PKI).

The necessity of PKI in the healthcare domain has been internationally discussed and recognized [5]. However, there have been few reports on the implementation and practical use of PKI for healthcare. A plan to use PKI for out-patients' prescription is under way in Korea [6]. Japan is ahead in a point of view of implementation of PKI because some of the EPR projects have already implemented a PKI and been operating a CA for different purposes [7]. One of the purposes of using a PKI is user authentication. A second is an access control and privilege management. A third is a digital signature on clinical information. A composite usage of them can be a fourth purpose. We employed a PKI for the security

mechanism for the construction of the information infrastructure for genome medicine because it provides these various security solutions.

2.1.1 Community of the PKI

The community of a PKI is composed of the following entities: the root CA (Certification Authority), the sub CAs, and end entities. The end entities are in turn categorized into subscribers and verifiers.

- **Root CA:** The root CA authorizes the sub CAs by means of issuing PKCs (Public Key Certificates) for digital signature to them. It also authorizes itself as the root CA by a self-issued PKC.
- **Sub CAs:** The sub CAs authorize end entities by means of issuing end entity PKCs for digital signature. In this study, there are no intermediate sub CAs that issue PKCs to subjects that are CAs.
- **Subscribers:** The subscribers are end entities that are the subjects of PKCs, hold the PKCs and sign pieces of genome information and clinical information by using them. In terms of healthcare, they are referring doctors that send clinical information to a referred doctor.
- **Verifiers:** The verifiers are end entities that make use of PKCs and verify digital signatures. In terms of healthcare, they are research leaders that receive clinical information and genome information signed by a participant researcher of different medical institutes.

2.1.2 Policies and Statements

In order to manage the hierarchical PKI, coordinated security policies are essential. We defined the security policies to manage both the root CA and the sub CAs as two kinds of documents. These documents are referred to as Certificate Policy (CP) and Certification Practices Statement (CPS) [8].

A policy is a set of rules established to govern a certain aspect of organizational behavior. The CP addresses the components of a PKI in total. It describes the goals, responsibilities, and overall requirements for the protection of the CAs, PKCs and their supporting components. It is a high-level document that describes a security policy for issuing PKCs and maintaining certificate status information. This security policy describes the operation of the CA, as well as the users' responsibilities for the requesting, using, and handling of PKCs and keys.

Compared to the CP, the CPS is a highly detailed document that describes how a CA implements a specific CP. The CPS identifies the CP and specifies the mechanisms and procedures that are used to achieve the security policy. Each CPS applies to a single CA. The CPS may be considered the overall operations manual for the CA.

The policy information is indicated in the three policy extensions of the PKC: certificate policies, policy mapping, and policy constraints. Now, we have prepared the secure mechanism based on the PKI. Next we describe the comprehensive model that defines the forms of information to be communicated.

2.2 Comprehensive Information Models

The architecture of the intelligent database proposes the adoption of the object model to handle complex data. In the area of healthcare, HL7 is the international standards and have entirely adapted the object oriented technologies for the latest version, HL7 Version 3. Therefore, the architecture of the information structure for genome medicine is expected to follow HL7 Version 3. However, HL7 Version 3 has been strongly focusing on the clinical information. The HL7 information model for the genome information is not completed. Hence, in this paper we devise and propose the comprehensive information models that integrate the genome information and the clinical information by extending HL7 Version 3.

2.3 Intelligent Analyzers

The genome information and the clinical information require their specific analysis methods respectively. Namely, they are genetic statistics and clinical statistics. For efficient analyses of both the genome information and the clinical information, the analyzing system should provide the various kinds of statistical analysis functions.

2.3.1 Genetic Statistical Functions

The following functions are primary genetic statistical functions that the intelligent analyzers are expected to provide.

- Hardy-Weinberg Equilibrium
- Case-Control study
- Linkage Disequilibrium Analysis
- Haplotype Inference
- others

2.3.2 Clinical Statistical Functions

The following functions are primary clinical statistical functions that the intelligent analyzers are expected to provide.

- Paired t-test
- Student t-test
- Welch t-test]
- Mann-Whitney U-test
- others

In addition to these traditional statistical analysis methods, a sort of knowledge discovery approaches or data mining approaches are known to be powerful for efficient analyses of the integrated genome and