病院情報システムに基づく診療サービスの創出に向けて
データマイニングによるアプローチ

Towards Service Innovation in Hospital Information System—Data Mining Approach—

津本 周作
Shusaku Tsumoto
島根大学医学部医学科医療情報学講座
Department of Medical Informatics, Shimane University, School of Medicine.

平野 章二
Shoji Hirano
(同 上)
hirano@ieee.org, http://www.med.shimane-u.ac.jp/med_info/hirano/

阿部 秀尚
Hidenao Abe
(同 上)
abe@med.shimane-u.ac.jp, http://www.med.shimane-u.ac.jp/med_info/abe/

Keywords: data mining, hospital information system, hospital service innovation.

1. はじめに

診療情報の電子化に伴い、すべての診療行為がヒストリとして記録され、その結果がコンテンツとして病院情報システムに蓄積されるようになってきた。データが長期間蓄積されることによって、今まで見えてこなかった病院のさまざまな側面が計量化できるようになってきている。例えば、島根大学医学部附属病院（来院者数1000人/日、病床数41)は2006年10月からすべての診療情報が電子化され、さまざまなサービスを提供する部門がデータを蓄積し、全体として分散型データベースシステムが構成されている。現在稼働しているシステムでデータを取得したところでは、診療報酬に伴う診療行為がオーダーとして平日1日当たりに約7000件登録され、約半分が外来、約半分が入院診療に当たれている。これらの診療行為の記録は、1日当たり約100MByteのテキストデータと約10GByteの画像データとして蓄積されている。これらは1年で、それぞれ10GByte、3TByteのデータ相当となり、これらのデータの有効活用は今後の病院情報システムでの大きな課題となっているが、有効活用の方法として、インターネットにおけるデータマイニング研究のアプローチが次に示される。

§1 コンテンツの解析：診療支援

これまで、我々は慢性疾患の病像をスナップショット的にしか観察することができなかった。長期間に蓄積された慢性疾患のデータをマイニングすることで、その慢性疾患の全体像を把握することが可能となり、より質の高い診療を支援できる。§2 ヒストリの解析：病院管理

一方、院内感染のような診療リスクについても、診療行為の電子化によって蓄積されたデータをマイニングすることことで、これらの事象のパターンを抽出し、感染・事故を防止することが可能となってきている。さらに、津本08において、用いられるデータは診療情報の配置に関わる病院管理のデータであり、マイニング技術の適用はデータに基づいた病院管理という新たな展開を生む。

著者らは、この二つの視点に基づき、前者については、実際に病院情報システムに蓄積された検査データを元にしたデータマイニングによる診療支援、および後者については、オーダー履歴による病院の動態把握として、プロトタイプ的な研究を重ねてきた。前者については、[Tsumoto 05a,津本 05b]を中心に、慢性疾患に関するマイニング技術の適用成果をマイニング結果のリスク管理へ応用する可能性について[Tsumoto 10]に報告した。

本稿では、後者のヒストリ解析の着目点、これらの研究成果およびデータマイニングに基づく病院情報システムでの新たなサービスの創出の可能性について概説する。

2. データマイニングからサービス創出へ：基本的な考え方

病院におけるサービス科学の目標は、「電子計算機を用いることによって、医療というサービスをいかに拡張し、
医療の質を高めるためのサービスを創出できるか」である。これまで、医療は、医師・看護師・検査技師といった医療従事者の個々の技量によって高い医療水準が維持されてきた。彼らの提供する医療はその技量に高度に集中化し、その技量を効率的に高めることができるよう、さまざまなサービスに集約的で大量の医療情報を配置することによって、サービスの質を高めることがいうことに対して実現であった。いう、医療を診断的に高めるマネジメント的視点がないで、医療の質を高めることに限界が生じていた。

我々は、上記の推進の目標をより具体的な問題設定として、以下の4点に着目し、人材資源の集約的な専門のサービスを提供する病院を対象として、情報基盤を構築・開発することを目標としている。（1）病院情報システムに蓄積されたデータを利用して、大規模システムとしての病院の特性を明らかにする。2）病院管理者に対して、病院の動態を情報的に観察するためのツールの開発を行い、医療管理のための新たなサービスを提供する。（3）医療スタッフに対して、一歩的な診療行為では見えてこない院内感染検査・医療事故のリスクの観点、情報を医療の質を低下させることの予防を支援するサービス、また横断的なデータ解析による各病棟のより詳細な把握を可能とするサービスを提供する。（4）診療サービスを受ける患者に対する、診療のペネフィットの観点に関して、データに基づき、患者個々の嗜好に合わせた情報提供などのサービスを創出する。データマイニングの技術を基盤として、これらのサービスと病院情報システムとの関係を階層的に図示したもののが、図1である。

図1 病院情報システムを基盤とした病院サービス創出の概念図

これらは、医療の質を高めるための新たなサービス創出の基盤を与えるだけでなく、これから計量的な観点で、それまでを詳細に観察されない病院規模システムに、1,000人以上のさまざまな職種が相互に連携して、サービスの実現を行う大規模システムにおいて、データに基づき新たなサービス創出に方法論的な基盤を提供できると考える。

3. 病院情報システム

病院情報システムの基盤はオーガニーシストシステムであり、すべての診療行為をオーガおよびそのオーガの実施結果とみなしている。これは、診療行為に対する診療報酬を対象とする医事会計の仕組みより由来するものであり、もともと医事会計を容易にするための仕組みとして登場した。電子カルテリシステムは、このようなオーガニーシストシステムによるオーガの実行結果を見出し、診療情報を対象とするという仕組みであり、オーガントロイ、電子カルテリ、医事会計システムが病院情報システムを構成する基本的な要素となる。診療行為とオーガの関係を患者と主な医療スタッフという関係は図2のように図示できる。

図2 オーガを対象とした医療スタッフと患者との関係

3.1 オーガントリシステム

オーガントリシステムは通常POSシステムを含めた販売管理関係するシステムの総称であるが、病院ではこれとは若干異なる形で特殊化したうえ、電子的な診療報酬と組み合わせることで、病院情報システムへと進化していた。

電子化以前は、異なる医療スタッフ間の指示は、伝票（オーガ）の形で伝達され、そのオーガを実行する形で、診療行為が実行されていた。例えば、処方箋は医師から薬剤師への治療薬投与に関する指示伝票とみなされる。これらの伝票は、外来では患者自身が、病棟では伝票配達のためのスタッフが、各部署に届けるという形がとられてきたが、このような届ける作業がネットワーク経由で行うというのが、オーガントリシステムの直接の動機であり、医療におけるプロセス管理とみなすことができる。例えば、薬の処方の流れは図3のようなプロセスに分解できる。

現在、オーガントリシステムは、実際のオーガの発信からオーガの実行結果を含めた、オーガに関する発注・履歴に関するデータを蓄積するシステムとして実装されている。このようなシステムは、病院の電子商取引とは独立した形で開発されてきたが、その類似性から、電子商取引で発展してきた仕組みを取り入れるような形を進化しつつある。処方のプロセスを診療全体のフローとして図示すれば、図4のようになる。
図3 処方の流れ

図4 処方オーダのプロセス、長方形で囲まれた部分が病院情報システムの関わる領域

図5 処方オーダシステムのチャート

図6 注射オーダのプロセス、長方形で囲まれた部分が病院情報システムの関わる領域

図7 注射オーダシステムのチャート

図5は実際の処方オーダの実装をチャートで示したものである。各ノードが必要なプロセスを示しており、各プログラム実行時に实行結果を伴ったヒストリを吐き出す。図からわかるように、これらのプロセス間は新たなプログラムを埋め込むことで、情報システムの中で新たなデータ収集あるいは新たな支援サービスを展開できると考えられる。

診療行為によって、このようなフローは異なるため、それぞれの行為によって、オーダエントリーの仕組み、そしてプログラムは大きく異なることになる。例えば、処方オーダは比較的単純なプロセスであるが、注射は図6に示すように介在する医療スタッフが増えるうえ、そのオーダ実施にいくつかのプロセスが付随することになる。

特に、オーダ実施が実際の手技を伴い、副作用などのリスクを常に考慮しなければならないオーダ実施のプロセスは複雑化する傾向にある。必然的に、そのオーダをサポートするシステムでのフローも図7のごとく複雑になる。

このような形で診療のワークフローを詳しく分析することで、オーダエントリーシステムを開発され、それぞれのプロセスがどのようなコンポーネントで構成されているかについては基本的な考え方で面着してきた。いわば

＊1 この図は、図4で記述されたプロセスに加えて、看護システムとの連携が記されている。外来処方ではこの部分は必要ないが、入院処方では、基本的には看護師の管理下にあり、そのための管理に関わるプロセスが存在する。このように、オーダによっては、入院・外来によって、そのプロセスが異なることがある。
オーダエントリシステムの開発は、医療行為のプロセス管理機能の過程をみなしむことができ、また、ソフトウェアをサービスの実現過程であるとみなせば、オーダエントリシステムは診療行為を支援するサービスともみなせ、このシステムを改良していくことは、診療行為の支援に繋がる。

したがって、オーダについてその全体的な傾向、各オーダの時系列的な性質や、病院での診療行為を見る基本的なデータとなり得る。これまで、オーダの結果、つまり処方などの治療経過、検査結果などがデータとして注目されてきたが、これは病院情報システムを構成するデータの一部であり、これらの実行過程が履歴として保管され、それが利用されなくなっていっている。

3.2 業務基幹系データベースから情報系データベースへの
病院情報システムをはじめとした業務で用いられ、データを蓄積するシステムはその業務内容に最適化された形で蓄積されており、可読性は低い、そのデータを閲覧する際に、データを変換し、人間が読める形に変換する必要がある。このように、業務のために最適なデータ収集を行うデータベースシステムを業務基幹系データベースという。業務基幹系データベースでは、そのデータ発生源に対して、これに関係するデータをできるだけ多く収集し、蓄える。例えば、検査機器に関係する業務基幹系データベースでは、その検査の実行時間、実行のステータス、実行結果を含めて、その機器に関するデータを実行時（ログ）の形で保存する。業務基幹系データベースは、このまでのデータ解析に向かない、可読性を高めて、データ解析を容易に変換する形に変換する必要がある、ある解析の目的を定めて、業務基幹系データベースからデータを収集して、一つの可読性の高いデータベースにまとめることがデータウェアハウスがあるが、この部分をいかにして効率よく行うかについては統一的な議論はない。

4. 基礎統計からのマニング
現在稼働しているシステムで調査したところでは、診療報告に伴う診療行為が平日に1日当たり7000件、休日は1500件登録されている。
表1に平日1日当たりのオーダ数の平均を、表2に、患者1人当たりの1日オーダ数の平均を示した。平日1日当たりのオーダ数は、外来が約3600、入院が約3400であり、処方、次回の予約、処置、記載、診療書面（診断書などの作成）が10％以上を占めており、これらが全体の70％以上になる。したがって、外来の業務の中心はこれらに集中しているといえる。

一方、入院は、記載、看護、注射が10％以上を占めて、全体の60％以上となっており、外来とは異なる診療行為が行われていることがわかる。

5. 長期経過観察患者に関するマニング
ある条件を設定することで、このようなデータから、来院患者の特性を見ることができる。表3は5年以上、大学病院の外来受診している患者のデータを用いて、診断を除いた総オーダ数、1回当たりの平均オーダ数および上位3つのオーダ七についての総オーダ数と平均オーダ数を示した。表2では、外来1人当たりの平均は約27であり、長期患者では2回と減っている。さらに、

*2 記載をオーダに含めないシステムがあるが、記載に診療行為が記載している場合を考慮して、オーダの一部として含めている。

*3 平均4オーダから診療記載分を引いたもの。
表3 長期観察患者の外来1回当たりの平均オーダ数

<table>
<thead>
<tr>
<th>オーダ種</th>
<th>平均オーダ数</th>
<th>外来1回当たりの平均オーダ数</th>
</tr>
</thead>
<tbody>
<tr>
<td>総オーダ数</td>
<td>219.75</td>
<td>2.042</td>
</tr>
<tr>
<td>処方</td>
<td>90.80</td>
<td>0.845</td>
</tr>
<tr>
<td>血液検査</td>
<td>51.89</td>
<td>0.483</td>
</tr>
<tr>
<td>診療予約</td>
<td>39.81</td>
<td>0.371</td>
</tr>
</tbody>
</table>

図8 長期観察患者データに関する決定本

全体平均よりも、処方・血液検査・診療予約で合計1.7程度あり、80%以上はこの三つのオーダで構成されており、このように安定した外来での経過観察がどのような診療内容に集約していくかを概観できる。

この抽出されたオーダに関するデータについて、診療科をはかの変量から推測ができるだろうか？このデータに決定本を適用すると、図8のような結果が得られる。これによって、外来オーダ数によるデータを使っただけでも、病院の特性を表すパターンを抽出することが可能となっている。

6. 履歴データの時系列マイニング

6.1 オーダの時間的推移

オーダの数の時間的推移を表、病院での診療の活動の統計をおおまかに把握することができる。図9に

図10 オーダの時間的推移（1週間：2008.6.1～6.7）

2006年8月から2008年2月までの1年間の1日当たりオーダ数の推移を示した。これは、平日・休日での周知的な傾向を示している。図10は、2008年6月1日から6月7日まで1週間のオーダ数の推移をオーダ種ごとに分けたものである。この中で、火曜日の1日オーダ数推移を取り出したものが図11である。図でおぼろげながらわかるのは、午前と午後の2相に病院のオーダ人りのピークがあるということ、また、処方、検査はおおむね午前中にピークがあるが、注射、看護関係は午後後にピークがありそうである。これらとはこの病院では、午前・午後において、その診療行為の性質が異なることを推測させる。

6.2 オーダの時系列変化的共起性

前節の個々のオーダの時系列変化のみならず、これらへのオーダの二つをにとって、二次元的な時間変化をプロットした軌跡を描き、その軌跡の類似性を考えることで、時系列マイニングを行うことができる。

例えば、上記図11において、診療記録を縦軸に、縦軸に処方、血液検査のオーダ数をとると、図12のような軌跡を描くことができ、これらの軌跡が類似していることがわかる。

これらを二次元的な軌跡の時系列マイニング [Tsumoto
される。このクラスタに含まれるものは、診療記録と診療予約の組合せのほかに、診療録と処方、検査、放射線、予約などの組合せが類似の形の軌跡をなす。

§2 病棟型

図15に、診療記録と看護とのプロットの軌跡を示したが、これが第2のクラスタ（グループ）に含まれる例である。まず、午前中は、診療記録は増加するが、看護オーダー数は増えない。午後になって、診療記録は最初減少し、看護オーダー数は変わらないが、ある時点で看護オーダー数が急増する。したがって、この軌跡は四つの相からなっており、前二相が外来、後二相が病棟、特に後半の一相目が病棟での日常・深夜の活動の交差時期の活動を反映していると考えられる。このクラスタに含まれるものは、診療記録と看護の組合せのほかに、診療録と射線の組合せが類似の形の軌跡をなす。

以上のところ、病院に蓄積されたヒストリを時系列計数データとみなしてマイニングすることにより、全体としての病院の動態を把握することができる。特に、オーダーの指定方法は大きく外来型と病棟型に分かれており、それは午前と午後の病院での診療行為の動態を反映したものである。これらのマイニング結果をさらに精査し、データを洗練化することで、病院の動的なモデルをつくり、そのモデルによって、診療・病院管理を支援するサービスを生み出すことが可能となる。これにより院内のサービスのためのワークフローといえ、上にあげたマイニングはその第1ステップである。

7. 端末ログからのマイニング

前章ではオーダ数の時系列変化のマイニング結果を示した。これは病院についてのグローバルな分析とみなせる。では、ローカルにおけるはどのような分析が可能であろうか？我々は2006年から病院情報システム各端末（約750台）からのデータベースのアクセス、ジョブ起動の履歴をとるシステムを導入し、ログを保管している。データは1週間で約1Gbyteとなり、過去5年で約120Gbyte程度まで蓄積している。

*4 樹状図上には、より多くのグループのクラスタに分けるべきであるが、ここは形状から最終的に二つのクラスタと解釈した。
7.1 端末の使用状況
図16は外来、病棟での端末の使用状況をプロットしたものであり、縦軸はそれぞれの設置台数に対する割合（％）を示している。これから明らかのように、外来は午前後、病棟は午後前に集中した一時的な分布であるが、病棟は午前・午後に二つのピークをもった二峰性の分布を示している。端末が利用されているときにオーダがされていると仮定すれば、図10が実はこの二つの動態を重ね合わせたものであることがわかる。

図16 端末使用状況（2007.8.15 ～ 2008.2.10）

7.2 ログ履歴からの系列マイニング
それぞれ、病棟でのログ履歴からどのようにパートーンが得られるであろうか？ここでは、各患者のオーダ履歴に系列マイニング [Abe 11]の方法を適用し、処方図17は外来端末における患者の診療スタート時間、終了時間、患者が外来受付をした時間に対応付けたデータから、待ち時間=診療スタート時間-受付時間、診察時間=診療終了時間-診療スタート時間として計算したものを時間別・日別でプロットしたものである。図からわかるように、午後の期間に見られるうえ、待ち時間のピークは11～13時付近に見られるが、診察時間はほぼ一定であることがわかる。

オーダの変更の有無にクラスターとして、どのような系列パターンが生じるかを分析した結果を示す。

使用データとしては、オーダエントリ調査データで実際の診療情報を含めており、このデータから診療の動態を把握することが可能である。

図17 ある端末における診療待ち時間（2006.2.2 ～ 6.8）

7.3 リハビリテーションの分類
図18は外来、病棟でのリハビリテーションの分類をプロットしたものであり、縦軸はそれぞれの設置台数に対する割合（％）を示している。これから明らかのように、外来は午前後、病棟は午後前に集中した一時的な分布であるが、病棟は午前・午後に二つのピークをもった二峰性の分布を示している。端末が利用されているときにオーダがされていると仮定すれば、図10が実はこの二つの動態を重ね合わせたものであることがわかる。

図18 リハビリテーションの分類（2007.8.15 ～ 2008.2.10）

7.4 マイニングの応用
マイニングの応用として、外来、病棟での端末の使用状況をプロットしたものであり、縦軸はそれそれぞれの設置台数に対する割合（％）を示している。これから明らかのように、外来は午前後、病棟は午後前に集中した一時的な分布であるが、病棟は午前・午後に二つのピークをもった二峰性の分布を示している。端末が利用されているときにオーダがされていると仮定すれば、図10が実はこの二つの動態を重ね合わせたものであることがわかる。

図19 マイニングの応用（2007.8.15 ～ 2008.2.10）

8. おわりに
本稿では、全診療情報の電子化およびその電子化情報を利用することにより、どのような成果が得られるかについて、基礎統計からのマイニング、時系列マイニング、系列マイニングの結果について紹介した。これらは医師・看護師の活動だけでは見つけなかった、病院の業務の全体像を把握するための手段を提供し、個々の診療行為がどのような位置にあるのかを推定に役立つ。

しかしながら、これらのマイニングは、ヒストリとして蓄積されたデータのごく一部を使用したものであり、今後の可能性を追求する必要がある。
解析し、新たなモデルをつくるという試みはまだ始まったばかりであり、今後さらにマイニングを重ね、新たな電子的サービスの創出を試みたい。

謝辞

本研究の一部は科学研究費補助金基盤研究「B」研究課題番号：21300052「診断プロセス管理のための時系列アクティブマイニングに関する研究」による補助を受けて、この研究は島根大学医学部医学科医療情報学講座安田晃、岩田春子、島根大学医学部附属病院小林善泰病院長の協力を得た。さらに、病院情報システムからのデータ抽出については、東芝住電医療情報システムズの青木浩二、吉本則夫、重藤章男氏を含めた病院情報システム開発チームの協力によるものである。ここに謝意を表したい。

参考文献


2011年2月10日 受理

著者紹介

津本 周作（正会員）