

病院情報システムのデータから見た病院

津本 周作, 平野 章二, 岩田 春子, 津本 優子

大規模病院において病院の診療情報の電子化が進んでいる。電子化の進展によって、慢性疾患をはじめとしたさまざまな病態の理解のみならず、病院サービスの品質を定量的に理解・評価されることが期待されている。本稿では、病院のサービスを向上させる仕組みとして、病院情報システムに蓄積されたデータを用いた病院の動態の可視化・計量化を試みたアプローチについて紹介する。指標として各診療行為のオーダー発行数を用い、その時間的な変化に基づいて、病院の動態を把握する解析を試みた。さらに、サービスの実行履歴を用いた待ち時間の解析とオーダー発行数との関連について調べた。これらのアプローチは病院の動態を可視化するための新たな手法論を提供していくことが期待される。

キーワード：病院情報システム, オーダーエントリシステム, 病院管理, データマイニング

1. はじめに

大規模病院は職員が1,000人以上が働く大規模なシステムであり、その動態を可視化、計量化することは困難であった。システムの動態をさまざまな計量的指標を使って測定することは、システム管理の第一歩であり、これまで製造業を含めて、品質管理 [1]・品質工学 [2] の分野において成果を収めている。しかしながら、サービス産業、特に多数のスタッフが集約的に働く大規模なシステムのサービスの質を把握することはこれまで困難であった。

病院管理においても、入院日数、稼働率、診療報酬稼働額等の指標を用いて、マクロな視点から病院全体のパフォーマンスを測ることが経験的に行われてきた。しかしながら、これらの指標は組織全体の粒度の粗い指標であり、例えば入院日数を減らすには、どのような改善が必要かについては、個々の部門のサービスに熟知しているものが経験的に知り得る情報を基に判断することに頼らざるを得なかった。これは、例えば、管理すべきサービスの品質が1つあるいは少数であれば機能するが、ある部門のフローを変えることで、他部門が影響を受ける複雑な組織においては、その影響を測ることすら難しい。全体としてある数値目標を達成するためには、長時間の議論と試行錯誤が余儀なくされ、効率が良いとは言えない。

本稿では、このような病院を管理し、サービスを向

上させる仕組みとして、病院情報システムに蓄積されたデータを用いた病院の動態の可視化・計量化を試みたアプローチについて紹介する。はじめにお断りしておきたいのは、本アプローチはある指標を用いて、病院の動態を把握しようというものであり、病院管理のツールとして成熟したものではないことである。著者らが言いたいのは、ある指標・評価の枠組みを定めれば、さまざまな手法を用いて、さまざまな角度からデータをとらえることができるということである。特に、医療のサービスは診療の過程を含めて、時間的な要素が本質的であり、指標による時間的な解析は必須である。

2. 病院情報システム

前世紀において病院では、医師が指示箋を書き、それを患者あるいはメッセージャーが各部門に送ることで、検査・処方などが行われてきた。指示伝票を送付する仕組みを電子計算機で肩代わりさせ、時間的負担を軽減することが目的で、1980年代後半から大規模病院でオーダーエントリシステムが導入された [3, 4]。その後、計算資源の著しい進展を背景に、オーダーエントリではなく、検査や処方などの実施歴、医師・看護師ほか医療スタッフの記載をも電子計算機で蓄積・管理することが提唱された。1990年代後半から、これらの仕組みを統合的に実現した病院情報システムいわゆる電子カルテが登場した [5]。この導入は徐々に進み、2012年現在では、ほぼ100%の大学病院、70%以上の大規模病院が電子カルテシステムを導入するに至っている。

つもと しゅうさく, ひらの しょうじ, いわた はるこ,
つもと ゆうこ

島根大学医学部医学科医療情報学講座
島根大学医学部看護学科基礎看護学講座
〒693-8501 島根県出雲市塩治町 89-1

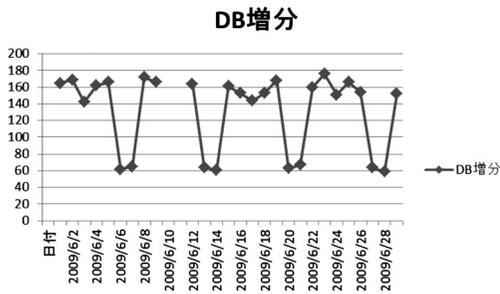


図1 データの1日あたり蓄積量
データベース (DB) の純増分を MB で計測している。収集期間は 2009 年 6 月のデータであり、2009 年 6 月 9～11 日はシステムのバージョンアップのため稼働を停止していたため、データは蓄積されていない。

3. 病院情報システムに蓄積されたデータの解析

病院情報システムでは、すべての診療行為が履歴として記録され、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。例えば、島根大学医学部附属病院 (外来: のべ 800～1,000 人/日, 病床数 600) 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調査したところでは、診療報酬の伴う診療行為がオーダーとして平日 1 日当たり約 7,000 件登録され、約半分が外来、約半分が入院診療に充てられている。これらの診療行為の記録は、1 日当たりで約 160 MB のデータベースと約 10 GB の画像データとして蓄積されている。例えば、図 1 に 2009 年 6 月のデータベースの増分を示した。平日においては、ほぼ毎日 160 MB の蓄積がされている。これは 1 年で、それぞれ 10 GB、3 TB のデータ相当となり、これらのデータの有効活用として、大きく次の 2 つが考えられ、著者らはデータマイニングの枠組みで、これらのデータの解析方法を研究してきた。

3.1 コンテンツの解析：診療支援

これまで、われわれは慢性疾患の病像をスナップショット的にしか観測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。例えば、[6] では慢性肝炎に関する蓄積された検査データの時系列を類型化し、病態をグループに分類することを試みた。図 2 に血清アルブミン (ALB) と血小板数 (PLT) の時間的変移を 2 次元上に軌跡としてプロット、軌跡マイニングによ

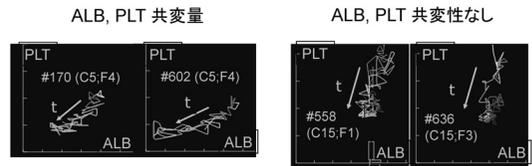


図2 軌跡マイニングの結果
血清アルブミン (ALB) と血小板数 (PLT) の時間変化を 2 次元平面上にプロットし、非類似度を計算し、クラスタリングした。時間的進展は図中の矢印で示したとおりである。

て、患者ごと軌跡を類型化した結果の 1 つを示した。左側のグループでは ALB と PLT が時間の経過に伴って両者が減少していくのに対し、右側では PLT は減少するが、ALB は減少していない。ここで、左側と右側とで予後に相違があり、左側の予後が悪いとすれば、この 2 種類の変量の時間的変化は患者の経過を観察するうえで重要な道具となりうる。

3.2 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た [7, 8]。ここで用いられるデータは診療資源の配置にかかわる病院管理のデータであり、ここでの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可能性がある。

著者らは、この 2 つの視点に着目し、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基にしたデータマイニングによる診療支援として、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 [9, 10]。以下では、後者のヒストリ解析としてオーダー履歴に着目し、病院の診療行為を俯瞰的に可視化する手法の研究・開発について紹介する。

4. オーダー履歴の解析

4.1 データ収集の仕組み

病院情報システムはもともとオーダーエン트리システムを元に開発されている。オーダーエン트리システムはそれまで伝票で運用されていた仕組みを計算機によって肩代わりするものとして導入された。例えば、図 3 は処方箋による外来での処方の流れを示したものである。各診療行為について、このようなオーダーの流れを分析し、電子化できるものを計算機上に実装したものが、オーダーエン트리システムであり、これに

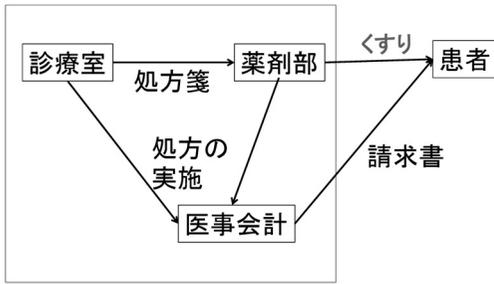


図3 処方オーダーのフロー
正方形内の各矢印の部分にはオーダーエントリーシステムが介在し、その実行結果はサーバーにログとして蓄積される。

診療記載を加えたものが、通称電子カルテと呼ばれる病院情報システムの仕組みであり、その実態は業務系基幹システムである。

図4に業務系システムにおけるデータ収集の仕組みを示した。まず、各端末においてどのような時間にどのようなジョブが起動し、オーダーが発行されているかについてのログをとる。オーダー処理サーバにおいて、オーダー処理がいつ集まり、どの部門において処理され、結果がどうであったかについてのログをとる。最後に、会計における処理時間についてのログをとる。これらは業務系システムに随時蓄積されるが、それぞれのデータが下記に示すデータウェアハウス (DWH) 上に蓄積されるような仕組みを実装している。これらによって、端末のレベルから、各診療科、病院全体についてのさまざまな粒度におけるジョブ実行、オーダー発行についての時間的記録を保持できる。

4.2 データウェアハウス (DWH) [12]

柔軟なデータ解析のためには、業務系と情報系の相互通信は interactive にならざるをえないが、このようなデータ抽出は、業務系システムに負担をかける可能性がある。われわれは2次DWHからのデータ解析によって、globalな傾向を取り出し、その精査を1次DWHで行えるよう図5で示した仕組みを構築し、できるだけ業務系システムのデータベースアクセスを減じるように工夫した。第1段階として、各患者に関するオーダー履歴を抽出し、1次DWHを作成する。この1次DWHにさまざまな抽出条件をつけ、その条件にあてはまるレコードを検索、条件に関するデータを抽出し、2次DWHの形にまとめる。この2次DWHのなかから必要なデータ集合を作成している。

4.3 データ解析の枠組み：オーダー数

以上のようなデータを用いて、病院の動態を描出しようというわけだが、以下では、1時間当たりのオー

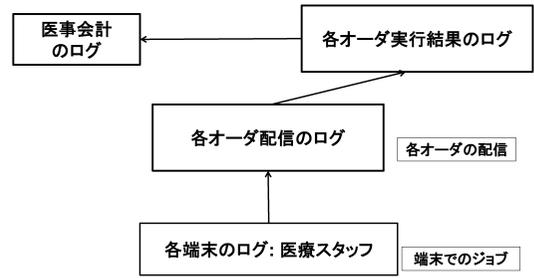


図4 データ入力とログの蓄積のフロー

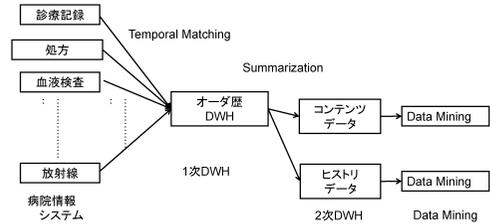


図5 データウェアハウス (DWH) の概要図

ダー発行数を基本的指標とする。1時間当たりのオーダー発行数は、1時間内で指示されたオーダーの数であり、医師あるいは看護師の診療行為の活動を示すとともに、診療行為の初動と見なすことができる。このオーダーの実施には、診療記録・看護記録のように即実施されるものもあれば、1~2週間後とタイムラグがあって実行されるものもある。即実施の場合は、発行数と実行数は等しくなるが、予約されたオーダーは、偶発的な要因に影響されやすいため、一定数実行されないものが発生する。したがって、診療の活動の度合いを見るためには、オーダーの発行数を見たほうが攪乱因子によって影響されない指標が得られる。もちろん、これが病院で行われている診療行為のすべてを把握できているわけではないが、大まかにその活動のある視点から計量的に眺めることができる [10, 11]。それは、以下の解析結果からも示されている。

5. オーダー履歴の解析結果

5.1 基礎統計

この小節では、基礎統計を示す。表1に、2006年10月~2011年2月までの各オーダーの1日平均総オーダー発行数、患者1人当たりのオーダーを外来・病棟別に示した。外来では1日4オーダー、病棟では8オーダーが出されている。診療記載を除けば、外来では、診療予約、処方、検査が上位に含まれ、入院では、看護、注射、処置が上位に含まれている。これらは発行数で

表 1 オーダー 1 日平均数 (一人当たり 2006.10~2011.2)

	外来オーダー		入院オーダー	
	1 日平均 (診療日)	患者平均	1 日平均 (診療日)	患者平均
処方	403.0095012	0.45733701	259.8551069	0.6261987
検査	287.9532858	0.326770695	221.3729216	0.53346435
生理	78.59540776	0.089190425	45.61995249	0.10993494
放射線	218.7814727	0.248274208	56.35154394	0.13579592
手術			12.85273159	0.03097251
輸血	1.167854315	0.001325286	11.08709422	0.02671767
食事			186.2929533	0.44892866
病理	21.90894695	0.024862372	16.73713381	0.04033314
新注射	84.41409343	0.095793496	469.3942993	1.13114613
予約	663.192399	0.752593743	77.46476643	0.18667455
診療文書	454.4845606	0.515751141	156.9659541	0.37825647
看護	11.08709422	0.012581685	846.3341251	2.03949553
処置	422.7181314	0.479702453	99.04196358	0.23867127
記載	951.9548694	1.080282704	963.0174188	2.32067887
リハビリ	3.047505938	0.003458324	3.064924782	0.00738585
入退院			55.40300871	0.13351014
総計	3602.315123	4.087923541	3480.855899	8.3881647
患者数	881.2090261		414.9722882	

表 2 各オーダー予約と実施のタイムラグ

	平均	中央値
処方	0	0
検査	28.55	14
生理	29.97	7
放射線	29.22	9
輸血	8.897	6
病理	-0.003	0
注射	9.799	2
予約	41.85	28
看護	1.397	0
処置	-0.0003	0
リハビリ	0	0

あり、実際には、実施のタイミングとの間にタイムラグがある。表 2 にそのタイムラグがどれぐらいかの統計 (平均値, 中央値) を示した。例えば、検査オーダーであれば、実施は平均で 28 日、中央値は 14 日であり、発行されたオーダーの約半数は 14 日以内に実施されている。図 6 は、次回の診療予約について、このタイムラグをヒストグラムとして表したものである。予約のピークは 14, 21, 28 日であり、最頻値は 28 日となっている。その後は、56 日、84 日といったところにも小さなピークがある。これは、同一の診療医が同一曜日に診察を行っていることによるものと考えられる。

これらは実際の総発行オーダーについての統計であるが、これらを 1 時間ごとに区切ってみれば、オーダー発行数の時間的変化が観察できる。図 7 では、全オー

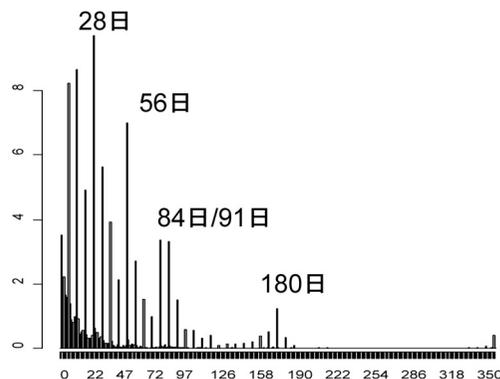


図 6 診療オーダーの予約と実施のタイムラグのヒストグラム
収集期間は 2006 年 10 月から 2011 年 2 月。

ダー発行数の平均数を曜日毎 1 時間ごとにプロットした。本院では、火・木にオーダーが集中し、月・金は比較的少ない。図 8, 9 では、このなかで、A 内科と B 内科の火曜日についての各オーダー平均発行数の時間的推移を抽出したものを示す。これらの図から、このオーダー平均発行数の時間的推移は、各診療科外来の特徴を反映していると考えられる。例えば、A 内科では、すべてのオーダーが診療記載よりも少なく、また、そのピークは一峰性である。一方、B 内科では、診療記載とは異なるピークが存在するとともに、診療記載の発行数がかなり長い期間続いている。これは、A 内科外来は比較的短い時間で終了しているのに対し、B 内科外来では、診療に時間がかかっていることが推察

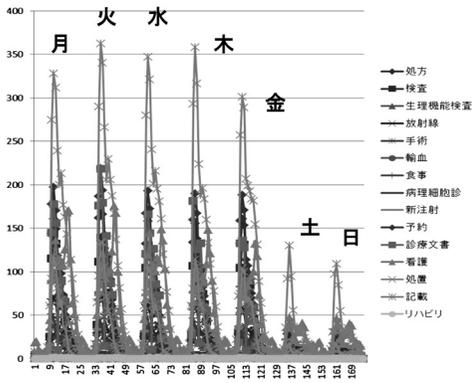


図7 平均オーダー数の時間的変化
収集期間は2010年4月から2011年3月。

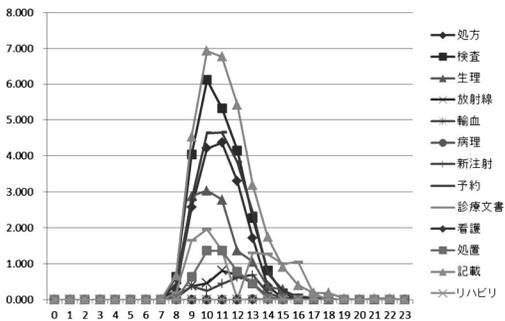


図8 A内科の火曜日での外来平均オーダー数の時間的変化
収集期間は2010年4月から2011年3月。

できる。

図10, 11では、各内科外来の外来受付から診療開始までの時間を患者の診療開始時間に合わせてプロットしたものである。横軸は各患者の診察が開始した時間を示し、縦軸は、受付時間と診察開始時間との差を示している。A内科では、11時以降待ち時間は劇的に増加しているが、一方、B内科では、10時から11時にかけては待ち時間は長い、その傾向は1例を除き、14時以降減少傾向にある。この1例は、複数回の検査の後、再度、外来を受診した例となっており、例外値と見なせる。表3にこれらを含めて、受付から検査、検査から診察、受付から診療科医師時間、診察時間、受付から予約時間、予約時間から診察開始時間という差分を統計値（平均値と中央値）として示した。この表から、以下のことがわかる。受付から検査までの時間、受付から予約時間までにほぼ時間的差がないことから、患者はほぼ予約の1時間前に病院を訪れ、まず血液検査を受けている。しかし、予約時間から診察開始時間については、両者には差があり、平均ではA内科42分、B内科は66分中央値ではA内科16分、B内科

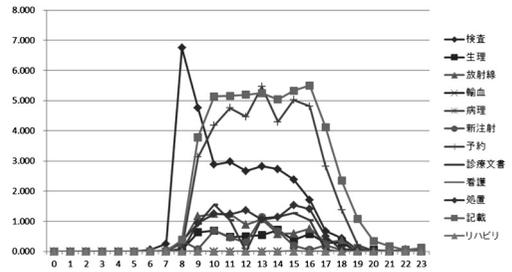


図9 B内科の火曜日での外来平均オーダー数の時間的変化
収集期間は2010年4月から2011年3月。



図10 A内科の火曜日での受付から診療開始までに必要な時間の推移（平均、収集期間は2010年4月）

は48分となっている。待ち時間の少ない50%の患者においては、ほぼ3倍の待ち時間の違いがあり、B内科では「待たされている。」という感覚が生じる可能性がある。ただし、図9, 10を見てもわかるように、各患者での待ち時間の差はB内科のほうが少なく、B内科では一様に待たされていると感じられる。

5.2 類型化

5.2.1 オーダー種の類型化

これらのオーダー発行数の時間的経緯によって、オーダーあるいは診療科はどのように類型化できるだろうか？このためには、時間的経緯を時系列と見なし、系列間の非類似度を算出し、クラスタリング [13] することが最も容易である。図12には、火曜日の時間別平均総オーダー発行数を時系列と見なして、クラスタリング (Ward法) を適用した結果を示した。この結果から、高さ5,000で見ると以下の4グループに類型化できることがわかる。グループ1: 処方, 処置, グループ2: 予約, グループ3: 看護, 輸血, リハビリ, 生理, 病理, グループ4: 検査, 注射, 放射線。ここからグループ2は外来, グループ3は病棟のオーダー発行を反映していると考えられ, グループ1, 4は外来と病棟の混合型で, どちらかという、グループ4は病棟優位であると解釈できる。

さらに、われわれは [3] において、オーダーの時間

表 3 待ち時間 (A, B 内科外来: 2010 年 4 月)

	A 内科		B 内科	
	平均値	中央値	平均値	中央値
受付～検査ラベル	30.64285714	8	31.37243402	7
検査ラベル～診察	87.88690476	80	103.1671554	92
受付～診療開始	116.7197279	90	143.5544431	106
診察時間 (分)	8.462549277	6	10.32727273	9
受付～予約時間	72.0221169	57	72.37642045	54
予約時間～診察	42.57661927	16	66.79403409	48.5

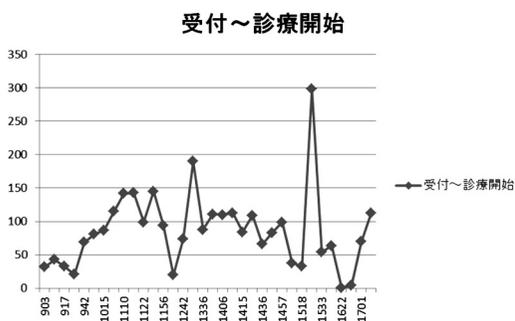


図 11 B 内科の火曜日での受付から診療開始までに必要な時間の推移 (平均, 収集期間は 2010 年 4 月)

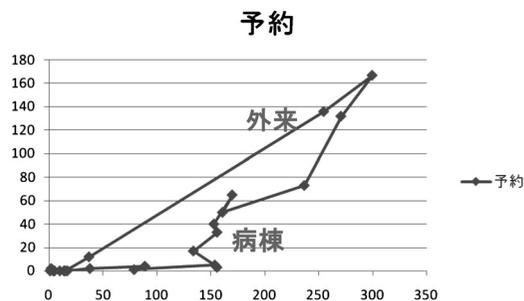


図 13 火曜日平均オーダー数の 2 次元軌跡の類型化の結果 横軸は診療記載, 縦軸は処方, 検査, 放射線, 予約で類似の形状を示した.

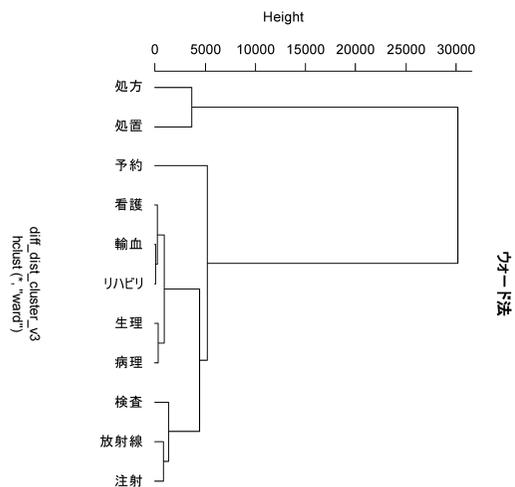


図 12 火曜日平均総オーダー数の時間的変化によるオーダー種に関するクラスタリング (Ward 法) の結果 算出には, 統計ソフトウェア R (ver. 2.15) を用いた.

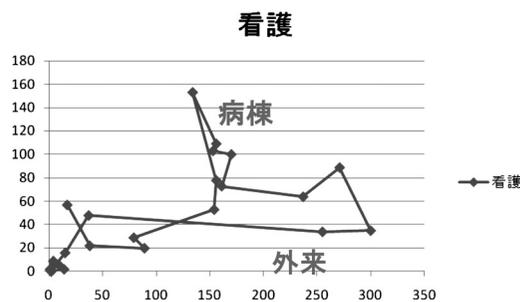


図 14 火曜日平均オーダー数の 2 次元軌跡の類型化の結果 横軸は診療記載, 縦軸は看護, 注射で類似の形状を示した.

的な傾向は外来型・病棟型に分かれ, 外来型については, 各オーダーが診療記載とほぼ同じような時間的変化を取っており, 外来オーダーに関しては, それらをすべて合わせた総オーダー数の時間的な変化が指標となりうることを示した. 図 13, 14 はその結果の概略を示している. 図 13 は外来・病棟の混合型で, 午前中の動きは外来, 午後の動きは病棟を反映しているも

のと考えられる. 一方, 図 14 は病棟型であり, 看護オーダー, 注射オーダーが増える夕方にそのピークを有している. これらは, 診療のプランがどの時点で行われているかを示している.

5.2.2 診療科の類型化

さらに, 外来総オーダー平均数の時間別集計をそれぞれ属性, 各診療科を 1 レコードとして, クラスタリング (Ward 法) を適用した. 図 15 に示すように, 高さ 2 で診療科は大きく 3 グループに分かれることがわかった. 1 つ目のグループは火曜日に外来患者数が極端に多い診療科, 2 つ目のグループは主として外科系, 3 つ目のグループは主として内科系からなる. こ

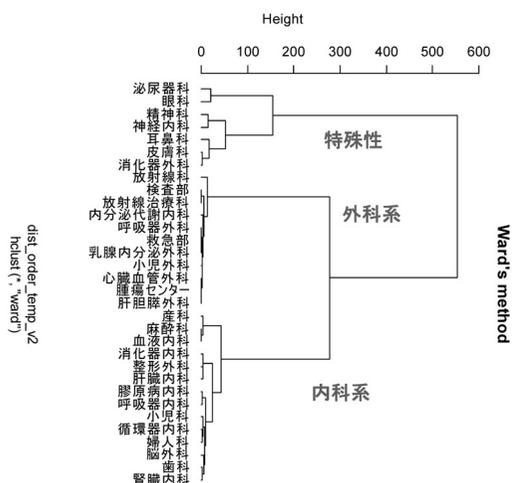


図 15 火曜日平均総オーダー数の時間的変化による診療科のクラスタリング (Ward 法) の結果
算出には、統計ソフトウェア R (ver. 2.15) を用いた。

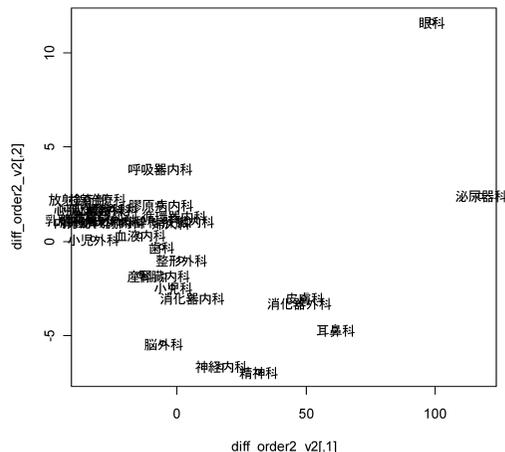


図 16 火曜日平均オーダー数の時間的変化による多次元尺度構成法 (MDS, Torgerson 法) の結果
算出には、統計ソフトウェア R (ver. 2.15) を用いた。

のクラスタリングの結果を補完する形で多次元尺度構成法 (MDS, Torgerson 法) [13] を適用した結果が図 16 である。この結果からは、2次元プロットとしては、4つのグループに分かれていることが予想される。左上の領域には、火曜日にオーダー数の多い診療科 (眼科・泌尿器科)、中央下の領域には神経内科、整形外科、精神科、密集した領域の外には内科系、密集した領域には外科系を含む診療科が含まれている。

火曜日のどの時間帯がこのグループ化に寄与しているであろうか。クラスタリングで得られたラベルを目的変数として、決定木 [14] の手法を全属性について適用した。図 17 に示すように、変数としては、火曜日 9 時 (T9) と 12 時 (T12) が選ばれ、これらの変数がグループ分けを最もよく説明するものであることがわかった。

6. おわりに

本稿では、病院情報システムに蓄積されたデータを用いた病院の動態把握を、オーダー数を指標とした枠組みで解析する方法とその結果を提示した。オーダー数という数値データを用いることで、従来の多変量解析、データマイニングの手法を無理なく使える。解析には、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るが、これらは医師・看護師の活動だけでは見えなかった、病院動態の全体像を把握する機会を与えてくれる。ただ、本稿で示した方法は一つの指標による視点を見せただけにすぎない。複雑な大規模な病院のシステムの動態をさらに深く把握してい

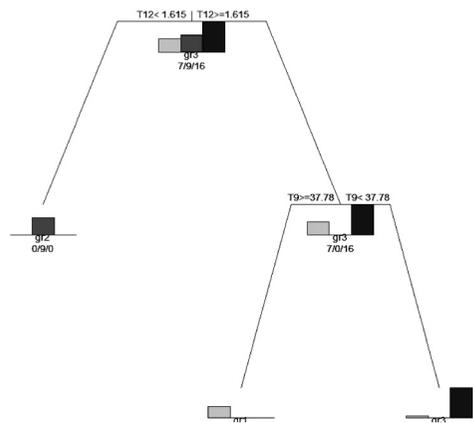


図 17 クラスタリングの結果によるラベル (3 グループ) を目的変数とした決定木の結果
算出には、統計ソフトウェア R (ver. 2.15) を用いた。

くには、単なるオーダー数ではなく、別の指標を用いることが必要であろう。どのような指標によって、病院を測定・管理していくべきかは、まだまだ試行錯誤していく必要があり、それがさらに病院の質の評価を深化させていくはずである。病院の診療情報の電子化はこの深い解析を支える情報基盤を与えるものであり、指標の創出は今後さらに容易になっていくに違いない。

参考文献

- [1] 特集『デミング賞受賞企業における方針管理の発展と課題』, 品質, **42**(1), 5-57, 2012.
- [2] 田口玄一監修, 品質工学便覧, 品質工学会, 2012.
- [3] 松本武浩, 宮下光世, 本多正幸, 病院情報システム導入効果の検討, 病院管理, **43**, 129, 2006.
- [4] 篠田英範, 病院情報システムの現況と今後の動向, 医療機器システム白書 2008~2009, 261-265, 2008.

- [5] 月刊新医療編集部, 電子カルテ & PACS 白書 2011~2012, エム・イー振興協会, 2011.
- [6] S. Tsumoto and S. Hirano, Detection of Risk Factors using trajectory mining, *Journal of Intelligent Information Systems*, **36**(3), 403–425, 2011.
- [7] 和泉多映子, 新見喜洋, 比良野正孝, 松岡喜美子, 横山茂樹, 津本周作, 原尾政, 医療 データマイニングを用いたリスクマネジメント, 日本臨床検査自動化学会誌, **32**, 26–33, 2007.
- [8] S. Tsumoto and S. Hirano, Risk Mining in Medicine: Application of Data Mining to Medical Risk Management. *Fundamenta Informaticae*, **98**(1), 107–121, 2010.
- [9] 津本周作, 平野章二, 津本優子, 医療リスクの計量化, 計量生物学, **29**, 155–176, 2008.
- [10] 津本周作, 平野章二, 阿部秀尚, 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, 人工知能学会誌, **26**(2), 139–146, 2011.
- [11] S. Tsumoto, S. Hirano, H. Iwata and Y. Tsumoto, Characterizing Hospital Services Using Temporal Data Mining. *SRII Global Conference 2012*, 219–230.
- [12] 千葉幸司, 宮内隆, 石井孝子, 中川裕道, 杉崎信彦, データウェアハウスの応用研究, 情報処理学会全国大会講演論文集 第 55 回平成 9 年後期 (3), 495–496, 1997.
- [13] 齋藤堯幸, 宿久洋, 関連性データの解析法—多次元尺度構成法とクラスター分析法, 共立出版, 2006.
- [14] 藤井良宜, カテゴリカルデータ解析 (R で学ぶデータサイエンス 1), 共立出版, 2010.