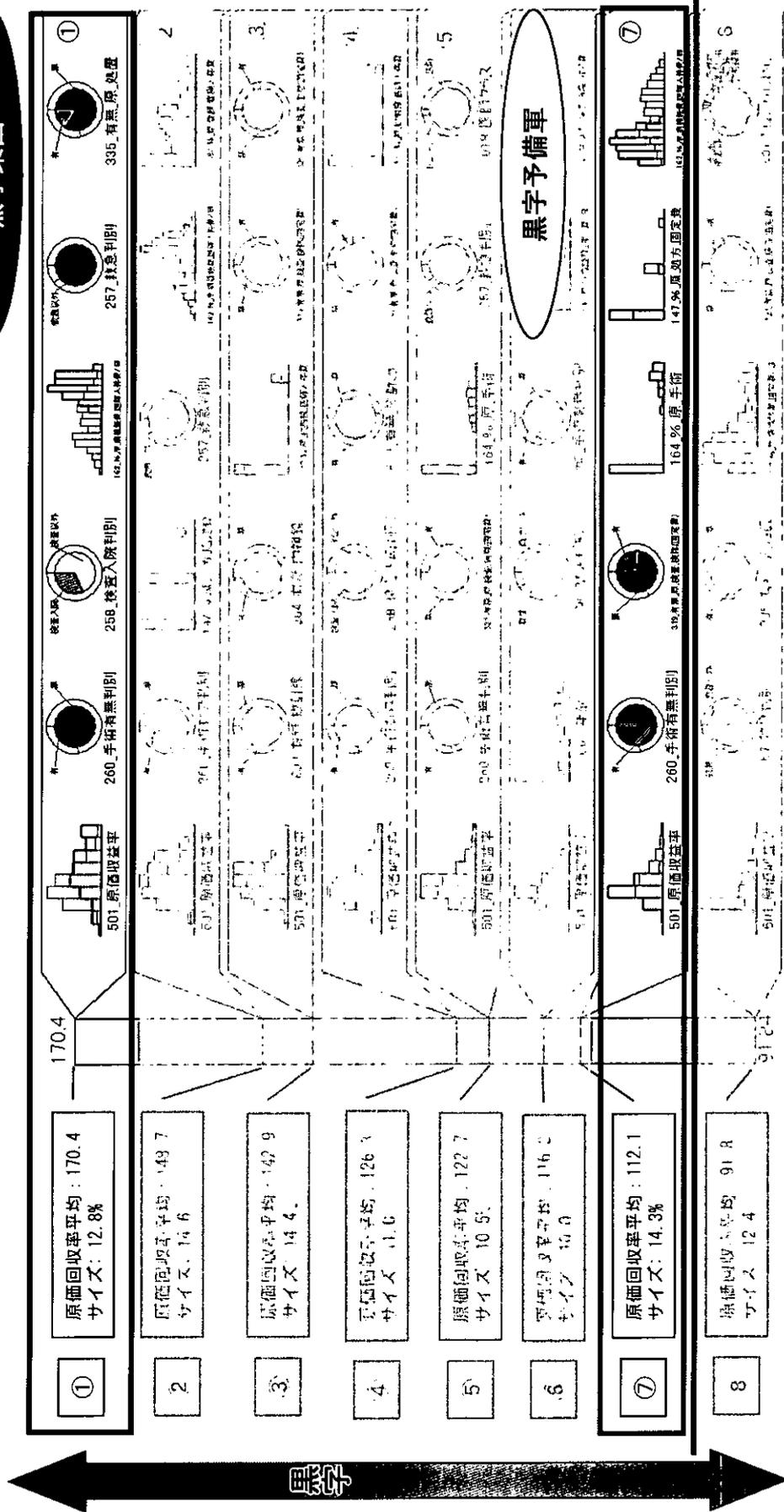


RBFによる分析試行結果 (5)

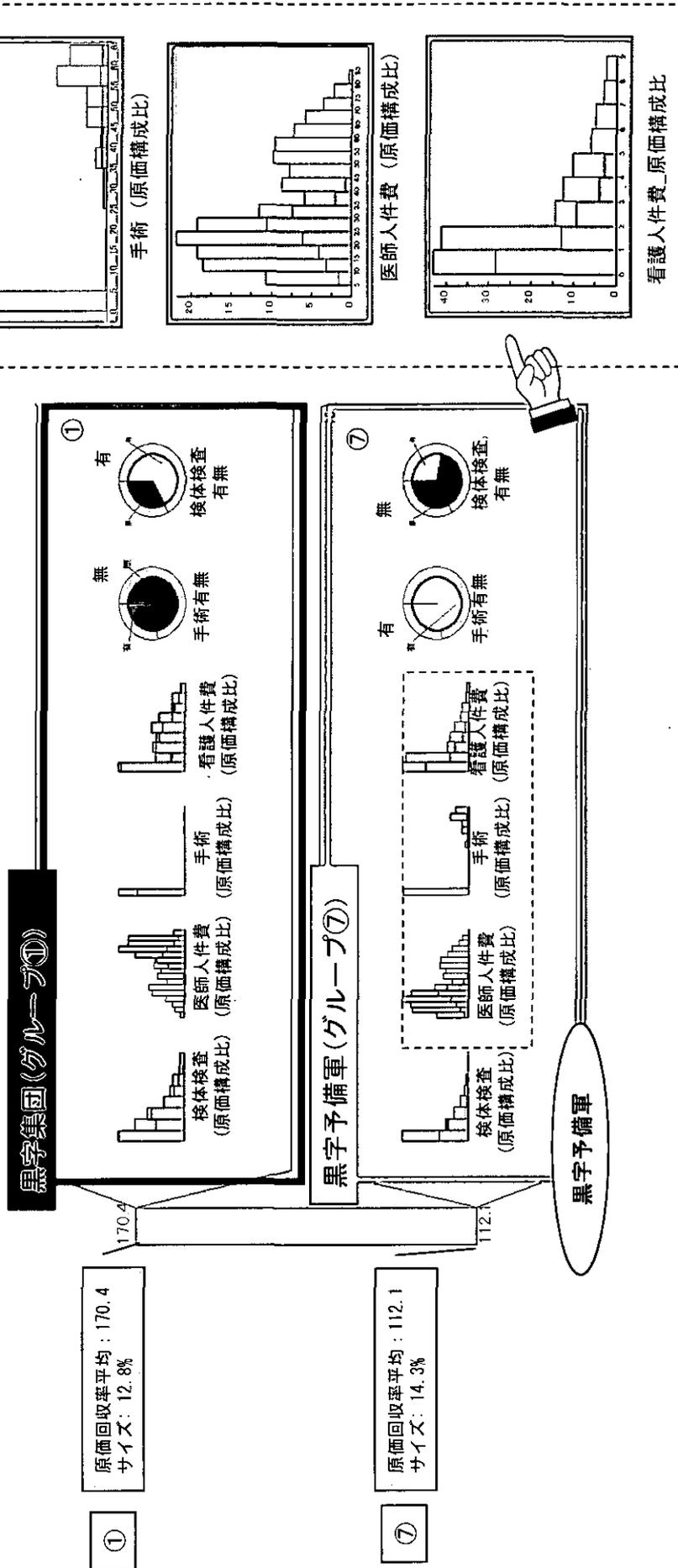
赤字グループ⑧は除き、最も赤字に近いグループを「黒字予備軍」として、このグループを、最も原価回収率の高いグループ（黒字集団）と比較します。

黒字集団



RBFによる分析試行結果 (6)

黒字集団と黒字予備軍のグループを、原価回収率を元に比較します。



この黒字予備軍の特徴としては、手術が有り、検体検査をしていないグループで、手術の原価構成比が高く、医師人件費、看護人件費の構成比が低いことが分かります。

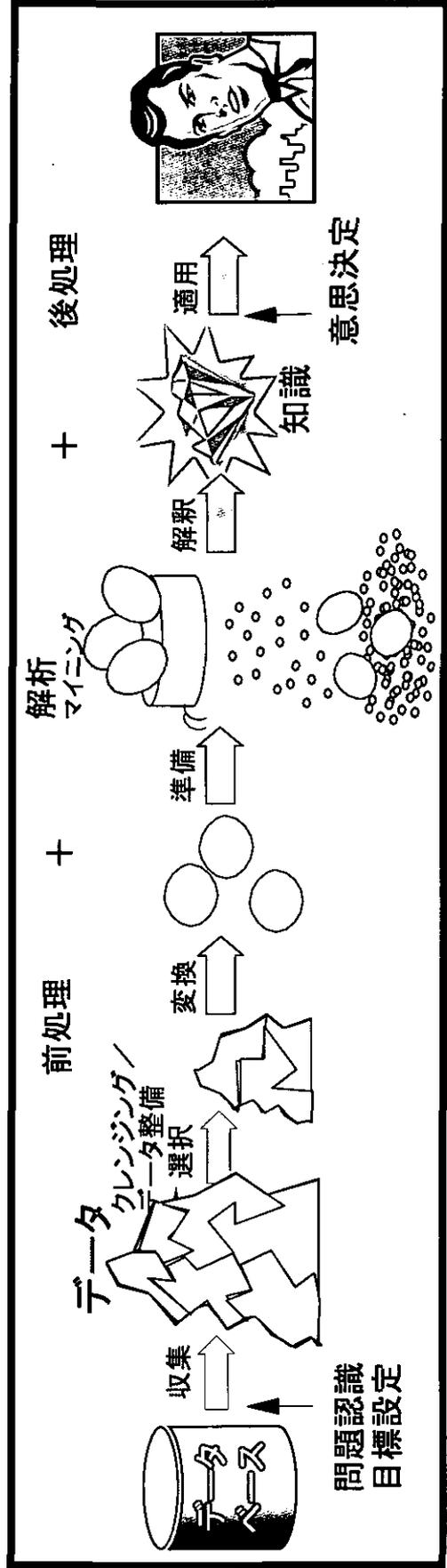
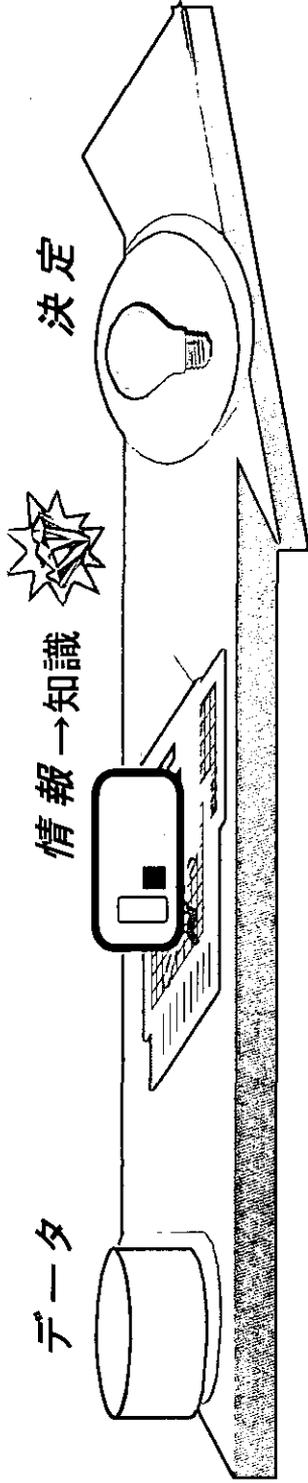
データマイニングとは

データマイニングについて

データマイニングとは、

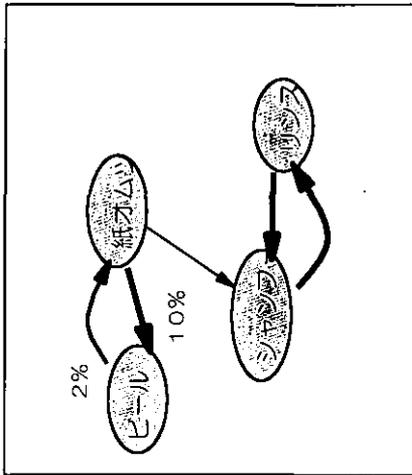
「以前には知られていなかった、有効で、活用できる知識を大規模データベースから抽出し、その知識を用いて重要なビジネスの決定を行うプロセス」である。

(P. キヤベナ他著、河村佳洋他監訳、データマイニング活用ガイド、トッパン、1999)



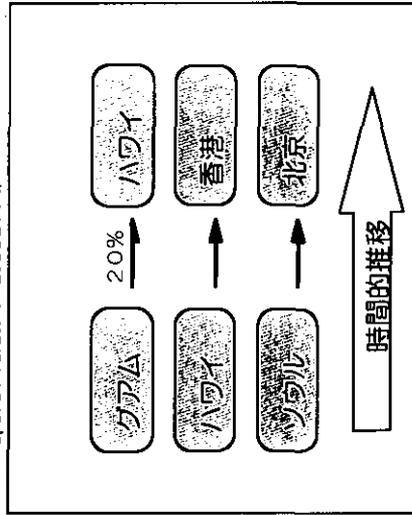
データマイニング6つの手法 (IBM Intelligent Miner)

①相関関係 (Association)



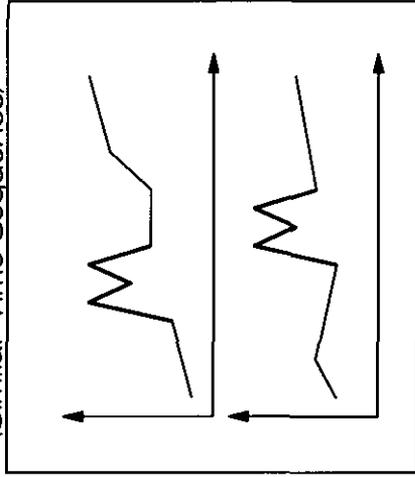
紙オムツを買う顧客の10%は同時にビールも購入する

②時系列順パターン (Sequential Pattern)



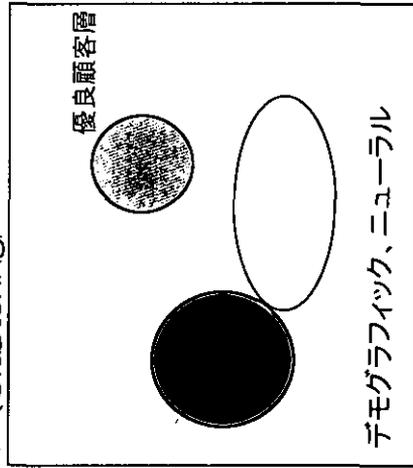
グアムに旅行した若い女性は、その次は20%の確率でハワイに行く

③類似時系列 (Similar Time Sequence)



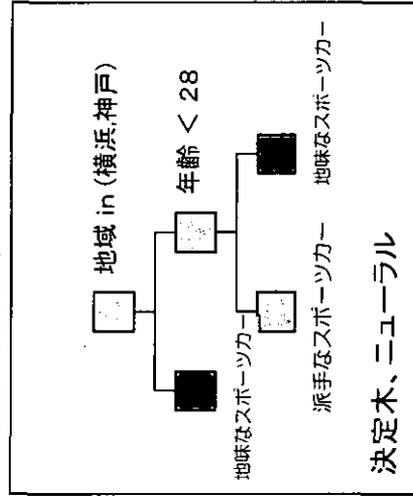
過去の商品別売上時系列データから新商品の売上パターンを予測する

④クラスタリング (Clustering)



100万人の顧客を50のグループにセグメンテーション

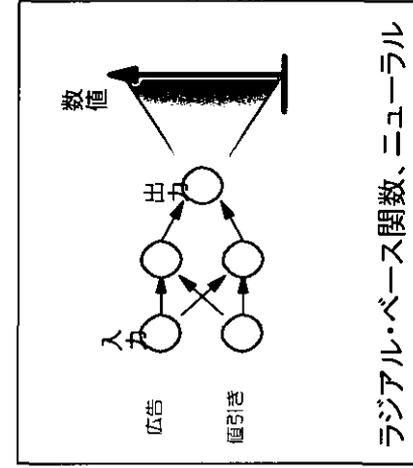
⑤クラス判別 (Classification)



決定木、ニューラル

派手なスポーツカーを買うのは郊外に住む若いプロフェッショナルである

⑥予測 (Value Prediction)



ラジアル・ベース関数、ニューラル

広告、値引き、エンド陳列をパラメータとしてインスタントコーヒーの販売量を予測する

肝疾患ネットワークにおける知識発見型大規模診療情報自動解析（データマイニング）システム
の臨床応用に関する研究

分担研究者 八橋 弘 国立病院長崎医療センター 臨床研究センター治療研究部長

研究要旨 L-net（肝疾患政策医療ネットワーク支援システム）を用いて肝疾患患者の診療情報を集積する上での問題点を明らかにした上で、その解決策を考案した。L-net 運用の問題点としては、1) L-net 運用の目的が不明確。2) L-net 運用の利点が不明確。3) 診療データ入力作業が煩雑。4) 患者登録の際の同意書の取得の問題点などが考えられた。それぞれの問題点の解決策の中でも、L-net そのものを研究目的とした研究を新たに始めるとともに、臨床評価指標の評価システムとして L-net を用いることを決定した。以上の解決策によって、今後本システムが有効に利用されることが期待され、データマイニング（Dm）解析を開始する環境が整備されうると考えられた。

研究協力者

松本 武浩 国立長崎医療センター
石橋 大海 国立長崎医療センター

A. 研究の背景と目的

データマイニング（Data mining: Dm）とは、過去のデータを分析し、その中から有効な規則性を発見して、重要な意思決定支援を促進する先進的情報解析システムで、金融ビジネス流通分野において広く導入されている手法である。従来の統計解析手法は限られたサンプル数を用いて仮説を検証する方法であることから、網羅性、迅速性において難点があるも、Dm では膨大なデータから網羅的に高速探索し、精緻な解析が可能である。ランダムイズを行わなくとも迅速に知識を発見することが可能なことから、米国 NIH では既に Dm による医療医学情報の解析が導入されているが、わが国の医学分野では SNP などの遺伝子情報解析や臨床研究分野での試みが始まったばかりである。

昨年度、本研究班の中において、肝疾患患者死亡者を対象に生存時の患者情報と随時の血

液検査所見を Dm を用いて解析をおこなうことで肝疾患患者の生命予後を明らかにし、さらに Dm を用いることでの予後予測が可能であることを報告した。この肝疾患患者生命予後予測モデルの構築は、従来の統計解析の手法では不可能であることから関連学会での発表では高く評価された。

昨年の Dm 解析では、当院という一施設での患者診療情報をデータベースとして用いたが、今後は、複数施設、多施設の診療情報を用いての Dm 解析を予定している。国立病院内では全国 23 箇所の施設が肝疾患の専門医療施設として実績があることから、これらの施設をネットワーク施設として従来から共同研究をおこなってきた。一方、2001 年から L-net（肝疾患政策医療ネットワーク支援システム）というシステムを用いて、これらの 23 箇所の肝疾患専門医療施設を中心に肝疾患診療情報を自動集積、解析する方向でシステム構築をおこなってきたが、実際の運用面での問題点が多いことが次第に明らかとなった。

本年度は、データマイニング（Dm）解析を

開始する環境を整備するという観点で、L-net (肝疾患政策医療ネットワーク支援システム) 運用の問題点を明らかにし、その解決策を見出すことを研究目的として検討をおこなった。

B. 研究方法

各施設での L-net への登録数を調査するとともに、L-net 構成員にアンケート調査をおこなうことで、問題を明らかにした。なお、L-net の概要は以下のとおりである。

参考資料 1: L-net: 肝疾患政策医療ネットワーク支援システム概要

はじめに

肝疾患政策医療ネットワーク支援システム (L-net) は各肝疾患政策医療ネットワーク参加施設にて診療を受けている肝疾患患者を登録し多施設臨床研究を支援し得られた解析結果をもとに EBM に基づいた医療を実践するものである。

目的

肝疾患政策医療ネットワーク関連施設にて診療を受けている肝疾患患者のうち同意書の得られた者をデータベース登録し得られた患者情報を解析し臨床研究、診療支援を行なう。

対象

肝疾患政策医療ネットワーク高度専門医療施設、専門医療施設、協力施設にて診療を受けている同意書の得られた肝疾患患者

管理方法

肝疾患政策医療ネットワーク支援システム管理要綱に従う。管理運営方針は肝疾患政策医療運営協議会にて決定し倫理委員会の承認を受けるものとする。

システム

国立国際医療センターに設置したデータベースサーバーにすべての診療情報を保存する。各施設からの情報のやり取りは国立病院・療養所情報ネットワーク (HOSPnet) 上で行なう。

利用できる端末は HOSPnet に接続された windows NT 4.0 workstation もしくは windows 2000 Professional をオペレーションシステムに持つ端末とする。使用するソフトウェアは無料で配布されている Netscape Communicator 4.75 および JAVA plug in 1.3.0 とする。

対象とする肝疾患

B 型慢性肝炎のラミブジン治療例から登録を開始し、登録状況を加味しながら急性肝炎、C 型慢性肝炎と対象範囲を広げる。

利用者資格取得手順

肝疾患政策医療ネットワーク利用者 ID 申請書を総括管理者宛送付し電子メールにて利用者 ID および仮パスワードを取得する。

運用報告

肝疾患政策医療ネットワーク支援システムシステム委員会および肝疾患政策医療ネットワーク運営協議会にて年 1 回報告する。

データ解析方法

疫学研究利用申請書を総括管理者宛提出し、個人を識別可能な情報を削除したデータを MO ディスクにて取得する。得られた結果は肝疾患政策医療ネットワーク運営協議会にて報告する。

C. 研究結果と考察

L-net 構成員からのアンケート調査結果からは、以下の 4 点が L-net 運用の問題点として指摘された。1) L-net 運用の目的が不明確。2) L-net 運用の利点が不明確。3) 診療データ入力作業が煩雑。4) 患者登録の際の同意書の取得の問題点。それぞれの問題点の解決策として以下のような対策をおこなった。

1) L-net 運用の目的が不明確という問題点に対しては、L-net そのものを研究テーマとして国立病院政策医療共同研究ネットワーク研究班を 2003 年から立ち上げた。その課題名は、L-net を用いた肝炎、肝癌治療の成績向上に関

する研究（班長；松本武浩）である。この研究班の中では、L-net の運用の問題点も論議しながら、肝炎、肝癌治療の治療成績を明らかにすることを目的としており、国立病院の独立行政法人化後も引き続き本研究をおこなう。

また一方で、国立病院の独立行政法人化後、国民に提供するサービスの質に関し、全病院が統一的な指標を用いて、自らのサービスを評価し改善していくことを目的に、新たに臨床評価指標を設置することが決定した。表1のような9つの評価指標項目を肝疾患分野では設定したが、そのうち、5つの指標はL-net を用いてデータを集計するシステムである。このようにL-net そのものを研究目的とした研究を新たに始めるとともに、臨床評価指標の評価システムとしてL-net を用いることで、各施設のL-net の運用が必須のものになり、L-net の活発な利用が今後期待される。

1. C型慢性肝炎に対するIFN治療患者数	L net
2. C型慢性肝炎に対するIFN治療による有効患者数	L net
3. C型慢性肝炎に対するIFN治療患者の新規肝癌発生率	L net
4. B型慢性肝炎に対するLAM治療患者数	L net
5. B型慢性肝炎に対するLAM治療による臨床的治療患者数	L net
6. 新規に発生した肝細胞癌の入院患者数	
7. 肝細胞癌に対する肝動脈塞栓術(TAE)件数	
8. 肝細胞癌に対する超音波局所療法(PEI, RFA)件数	
9. 肝細胞癌に対する肝切除件数	
(10. 肝細胞癌の生存率)	L net

2) L-net 運用の利点が不明確。この点に関しては、昨年度、Dm を用いて肝疾患患者の生命予後予測モデルが可能であることを既に実証したことから、Dm の価値観、その可能性に関しては L-net 参加施設構成員に十分理解されたと思われる。

3) 診療データ入力作業が煩雑。この問題点に関しては、検査結果値の手入力ではなく自動取り込みサーバーでの検査結果の自動入力が可能であれば、その問題点の多くは解決される見込みである。一部の施設では、既に自動取り込みサーバーが設置されているが、多くの施設

においては、設置の見込み、予算処置がおこなわれなかった状況であった。しかし、平成16年度、国立病院の独立行政法人化後、本システムのあり方が評価され、独立行政法人国立病院機構内での研究費予算内で自動取り込みサーバーが設置される見込みとなり、順次この問題点は解決されると思われる。

4) 患者登録の際の同意書の取得の問題点。この問題点に関しては、2000年当初、文書での同意書は必ずしも必要ないという議論もおこなわれたが、個人情報保護の問題がクローズアップされる最近の情勢からは、煩雑ではあるも書面でのインフォームドコンセント取得は厳守すべきであり、むしろ時代を先取りしていたというL-net 参加施設構成員からの意見も聞かれるようになった。この点に関しては、当初の予定どおり引き続きおこなう。

参考資料2：肝ネット登録同意書

当院は、厚生労働省が政策医療として推進している肝疾患の専門病院に指定されています。各種情報があふれている現在、正確な事実に基づいた高度な医療をおこなう為には、多くの患者様のデータを集積、解析して、その成果を日常診療に還元することが求められています。この度、コンピューターを利用して肝疾患専用のデータベースシステム（肝疾患政策医療ネットワーク支援システム）を厚生労働省、国立病院、国立療養所内で新たに構築することになりました。よりよい肝臓病の診断法と治療法を見出すことを目的として本ネットワークは運営されます。このような主旨をよく御理解いただいた上で、ひとりひとりの肝臓病の患者様の所見を本ネットワーク内に登録することにご協力いただきたいと思います。

1. システムの目的

このシステムは、患者さんのプライバシー保護を図りながら、患者さんの診療情報の一部を肝疾患政策医療ネットワーク支援システムの

ホストコンピュータに入力し、ホストコンピュータと参加施設をネットワークで結ぶことにより、診療情報を共有し、肝臓病診療を円滑にし、各地から得られた多くのデータを元に真に肝臓病診療に役に立つ治療法の検討や効果の判定の研究を行い質の高い診療を可能にすることを目的としています。

2. システムに参加することのメリット

このシステムに参加すると、患者さんの診療情報がホストコンピュータに入力され、前回の受診時の状況や図表化された治療歴の情報が活用でき治療の経過やその効果をわかりやすく説明を受けることができます。また、情報を蓄積し、研究に活用することにより、我が国の肝臓病医療を向上させ、ひいては、患者さん自身が質の高い肝臓病医療を受けられるようになります。

3. 個人情報の安全確保

このシステムでは、患者さんの診療情報を守るために次のような対策を講じています。

- (1) このシステムでホストコンピュータに入力された患者さんの診療情報を見ることができるのは、当該患者さんが同意した病院の医療関係者のうち、肝疾患診療支援ネットワークシステム総括管理者が許可した者に限られます。
- (2) このシステムは、外部からの不正な侵入に対して、銀行などが電子商取引を行うときと同じように厳格に情報を保護しています。
- (3) 研究目的で診療情報を活用する場合は、肝疾患政策医療ネットワーク運営協議会に諮り、肝臓病医療の向上に役立つと認められた研究に限り許可することとし、患者さんの氏名・住所・電話番号を除く診療情報を提供します。

4. 個人情報の閲覧請求権

このシステムによりホストコンピュータに入力された診療情報は、当該患者さんの希望により、診療時間内であればいつでも患者さん又

は患者さんの指定する代理人が無償で閲覧することができます。入力情報に誤りがあった場合には、主治医を通じて、訂正を求めることができます。

5. 参加をやめたいときは

このシステムに参加した後で参加を取りやめなくなった場合には、システム全体の参加を中止することも、特定の医療機関での参加をやめることもできます。その場合には、下記のシステム管理者宛、備え付けの書面で申し出て下さい。

ただし、このシステムによりコンピュータ入力された診療情報は、改ざん等を防ぐため、削除することはできません。患者さんから参加の同意撤回の申し入れがあった場合には、診療情報の入力は停止されます。

6. 最後に

このシステムへの参加は、患者さん一人一人の自由な意思によります。このシステムについて主治医から説明を受け、目的、意義、安全性確保等に納得された方に、ご参加いただきます。もし、参加されなかった場合や途中で参加を取りやめた場合でも、今後の診療に何ら不利益をこうむることはありません。

(問い合わせ先)

〒856-8562

長崎県大村市久原 2 丁目 1001 番地 1

国立病院長崎医療センター

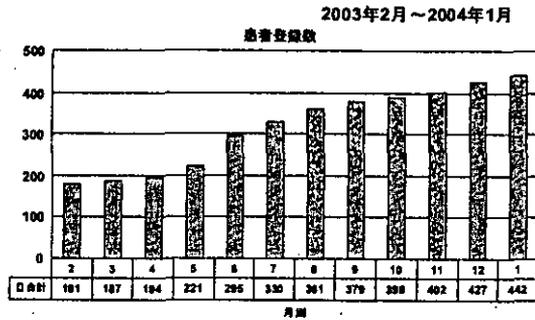
臨床研究センター

Tel:0957-52-3121(内線 3273)

Fax:0957-53-6675

以上の 4 つの問題点に対する解決策を考案した。図 1 に示すように L-net への登録数は、時間とともに増加しており、今後一層活用化されることが期待される。

図1. L-net入力状況



D. 結論

L-net（肝疾患政策医療ネットワーク支援システム）を用いて肝疾患患者の診療情報を集積する上での問題点を明らかにした上で、その解決策を考案した。L-netそのものを研究目的とした研究を新たに始めるとともに、臨床評価指標の評価システムとして L-net を用いることを決定したことにより、今後本システムが有効に利用されることが期待され、データマイニング（Dm）解析を開始する環境が整備されうると考えられた。

E. 健康危険情報

なし

F. 研究発表

1. 論文発表
2. 学会発表

八橋弘、秋山昌範、河村佳洋、知識発見型大規模診療情報自動解析（データマイニング）による肝疾患肝癌患者死亡予測、予後予測モデルの構築：パネルディスカッション：肝細胞癌の診断・治療の今後の展開、第39回日本肝臓学会総会、福岡、2003年5月22日。

G. 知的財産権の出願・登録状況

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

腫瘍外科学における知識発見型大規模診療情報自動解析システムの応用に関する研究

分担研究者 大内 憲明 東北大学大学院医学系研究科腫瘍外科学分野 教授
研究協力者 武田元博、多田寛 東北大学大学院医学系研究科腫瘍外科学分野

研究要旨

データマイニングシステムの運用による新たな知識発見のモデルとして乳癌検診における被験者のデータウェアハウスを非線形解析することによる、乳癌の新たな危険因子の発見を目指す。今年度は、我々と研究協力関係にある宮城県対がん協会において、乳癌発症リスクと関連するといわれる家族歴、既往歴等がデータベース化されていることを確認した。今後、データの抽出項目を始め、順次データ解析のための準備をすすめていく予定である。

A 研究目的

データマイニングシステムの運用による新たな知識発見のモデルとして乳癌検診における被験者のデータウェアハウスを非線形解析によって乳がんの新たな危険因子の発見し、本手法の有用性を検証する。

B 研究方法

1. 対象：平成10年以降の宮城県対がん協会が行った乳癌検診を受診した被験者（年間約2万人、5年間）のデータウェアハウス。
2. 方法：上記データに対してデータクレンジングを行い、抽出項目を定めた後、乳癌検診被験者のデータを個人情報保護に配慮し、個人名、ID、住所を除いて抽出項目のみまとめる。
抽出項目に対して非線形解析によるデータマイニングを行い、乳癌検診における乳癌のリスクファクター、経過観察を行うべきハイリスク項目についての知見を得る。

倫理面への配慮

本研究はデータウェアハウスを対がん協会から持ち出す際、氏名、ID、住所等の項目を完全に削除し、完全に非連結匿名とするため個人情報には完全に保護される。

C 研究結果

乳癌検診受診者の問診表の中で乳癌発症リスクと関連するといわれる家族歴、既往歴等、データベース化されていることを確認した。さらに他の項目で必要と思われる項目についてデータクレンジングを完了した。宮城県対がん協会において、データマイニングの手法について研修を行うとともに、今後のデータの相互利用を始めとした研究協力体制について討議した。その結果、1検診当たり200項目にわたるデータの中から50項目に絞り込むデータクレンジングを行った。今後解析を行いつつ、解析の結果をフィードバックしながら、さらにデータクレンジングを行い、解析の妥当性について検証を進める予定である。しかし一方

でこれまでデータベース化されているものは項目が限られ、年間 2 万人、5 年分のデータのうち半分は、これからコンピュータに入力しなければならないことが判明した。現在データ入力を行うべく人の配置等準備を行っている。

D 考察

宮城県対癌協会のデータは比較的欠損が少なく、対象症例数、データの項目数とも優れている。今回のデータ解析により、新たな乳癌リスクファクターなどの知見の発見が期待できる。

E 結論

乳癌検診データウェアハウスはデータマイニングの非線形解析による新たな知見を得るための対象として妥当である。今後、必要なデータ入力のために人材を確保し、データ入力から解析まで行う必要がある。

F 健康危惧情報

本研究は被験者そのものではなく、過去のデータが対象となるため、健康に対する害は生じない。

G 研究発表

1. 論文発表

Amari M, Moriya T, Harada Y, Ishida T, Ohnuki K, Takeda M, Sasano H, Horii A, Ohuchi N. LOH analyses of asynchronous lesions of ductal carcinoma in situ and invasive ductal carcinoma. *Jpn J Clin Oncol*, 33(11) 556-562, 2003

Ishida T, Ohuchi N, Moriya T, et al. Pathological assessment of intraductal spread

of carcinoma in relation to surgical margin state in breast conserving surgery. *Jpn J Clin Oncol*, 33 (4). 161-166, 2003

Sakayori M, Ohuchi N, Kawahara M, et al. (total 11, 8th) Evaluation of the diagnostic accuracy of the stop codon (SC) assay for identifying protein-truncating mutation in the BRCA1 and BRCA2 genes in familial breast cancer. *J Hum Genet*, 48: 130-137, 2003

Minami Y, Ohuchi N, Tsubono Y, et al. (total 5, 4th) The increase of female breast cancer incidence in Japan: emergence of birth cohort effect. *Int J Cancer*, 108:901-906,2004

Ohuchi N, Ishida T, Ohnuki, Takeda M. Advances in diagnosis of breast cancer: mammography for screening and mri for breast-conserving surgery. R. Ros, T Kakizoe (eds,) "Innovative achievements in cancer imaging". The 33rd International Symposium of the Princess Takamatsu Cancer Research Fund. pp. 30-34, 2003

Takeda M Tishida T, Ohnuki K, Suzuki A, Sakayori M Ishioka C, Nomizu T, Noguchi S, Matsubara Y Ohuchi N. Collaboration of Breast Cancer Clinic and Genetic Counseling Division for BRCA1 and BRCA2 Mutation Family in Japan. *Breast Cancer*, 11(1), 30-32, 2004

De La Cruz C, Moriya T, Endoh M, Watanabe W, Takeyama J, Ming Y, Oguma M, Sakamoto K, Suzuki T, Hirakawa H, Orita O, Ohuchi N, Sasano H. Invasive micropapillary carcinoma of the

breast. Clinicopathological and immuno-histochemical study. *Pathol Int*, 54: 90-96, 2004

Shiraishi K, Kato S, Han S-Y, Liu W, Otsuka K, Sakayori M, Ishida T, Takeda M, Kanamaru R, Ohuchi N, Ishioka C. Isolation of temperature-sensitive p53 mutations from a comprehensive missense mutation library. *J Biol Chem*, 279 (1): 348-355, 2004

大内憲明, 武田元博, 川添良幸, 佐竹正延*, 粕谷厚生. クラスタ利用ナノ医療。ナノ学会会報, 1 (2): 101-105, 2003

武田元博, 石田孝宣, 大貫幸二, 大内憲明. 乳癌手術におけるリスクマネジメント。外科病棟・手術室のリスクマネジメント。中外医学社。183-194, 2004

庄子由美, 梅澤昭子, 星 邦彦, 根本建二, 大内憲明. 医療の安全文化をいかに醸成するか。医療マネジメント学会雑誌, 4 (3), 445-449, 2003

大内憲明. マンモグラフィースクリーニングの現状と課題。日本産婦人科学会雑誌, 55 (8), 959-965, 2003

大内憲明, 久道茂. がん検診、臨床腫瘍学, 443-449 2003

大内憲明, 石田孝宣, 大貫幸二, 武田元博, 石橋忠司. 非浸潤性乳管癌: 放射線科医師に直面する課題と展望。臨床放射線, 48 (7), 803-809, 2003

鈴木昭彦, 石田孝宣, 森谷卓也, 大内憲明, 乳腺症と乳癌。産婦人科治療, 87 (6), 642-646, 2003

2. 学会発表

石田孝宣, 大貫幸二, 武田元博, 鈴木昭彦, 田澤篤, 中島護雄, 宇佐美伸, 多田寛, 大内憲明, 当科における非浸潤癌の乳癌温存療法の実績。第103回日本外科学会

石田孝宣, 大貫幸二, 武田元博, 大内憲明, 当科における両側乳癌症例の検討。第11回日本乳癌学会

武田元博, 石田孝宣, 大貫幸二, 酒寄真人, 石岡千加史, 野水整, 野口眞三郎, 福島雅典, 松原洋一, 大内憲明家族性乳がん保因者における遺伝子診療体系の確立。第11回日本乳癌学会

大貫幸二, 石田孝宣, 武田元博, 鈴木昭彦, 鈴木昭彦, 宇佐美伸, 大内憲明, 初回受診者と繰り返し受診者の発見率から見た乳癌検診の精度管理, 第13回日本乳癌検診学会

宇佐美伸, 大貫幸二, 石田孝宣, 武田元博, 鈴木昭彦, 鈴木昭彦, 大内憲明, 検診発見の微細石灰化病変に対する診断方法についての検討, 第13回日本乳癌検診学会

II. 知的財産権の出願登録状況

本研究に関してなし。

医療技術評価総合 研究事業

「電子カルテシステムにより集積したデータの診療情報解析(データマイニング)に関する研究」平成15年度研究報告書

「国立病院における病院業務分析に関する検討 —電子カルテシステムを利用したIgA腎症に関するデータマイニングの試み—」

分担研究者 吉村光弘 国立金沢病院第二内科医長

研究要旨 データマイニングによる有益な知識の発掘は新たな臨床応用や研究に利用できると評価されている。厚労省国立病院ネットワーク(HOSP ネット)内に置かれた腎ネットにより、全国6施設から集めた2000例を超えるIgA腎症をデータベース化した。その予後をIgA腎症の予後判定基準別に検討した結果、はじめて大きな規模の症例数で判定基準の検証が行え、問題点を明らかにした。急速に普及する電子カルテシステムから安全かつ容易にデータを書き出せる環境を整えることが新たな研究手法を提供する可能性が高い。本研究では、国立病院の臨床研究、政策医療を実施するための電子カルテシステムのカスタマイズについて検討した。

研究目的

電子カルテの開発が進み、データの収集に関する技術的な問題点が解決しつつある。世界のトップレベルにある日本の電子カルテシステムによって蓄積された診療情報をデータベース化し、これを利用した決定木形式の治療指針を作成したり、莫大なデータから検査変動値と疾患や特定の医薬品との関連を見いだしたりといった手法での臨床研究の進展が期待される。

しかし、現時点ではこれまで臨床医が手作業で収集した症例データベースを電子カルテから容易に作成することがもともと現実的であり、情報漏洩事件が頻発する現状を打開しつつ、臨床医が容易にデータを利用できるシステム作りが整えられることが先決であろう。

腎ネットを構成する国立病院6施設のIgA

腎症をデータベース化し、新たなIgA腎症予後判定基準の作成を試みた。この研究をとりあげ、新規症例の追加やさらなる追跡調査などデータ収集が容易に実施できるような電子カルテのシステム構築について検討を行いたい。

B.研究方法

IgA腎症データベースは、国立病院7施設(西札幌、佐倉、千葉東、東京医療センター、金沢、大阪、米子)および厚生連高岡病院、黒部市民病院の2施設において、IgA腎症患者の同意を得て、発症日と最終観察日(透析導入日)における年齢、一日尿蛋白量、血清クレアチニン、クレアチニンクリアランスを調査し、ファイルメーカープロに入力して解析した。

電子カルテは、平成16年10月より本格稼

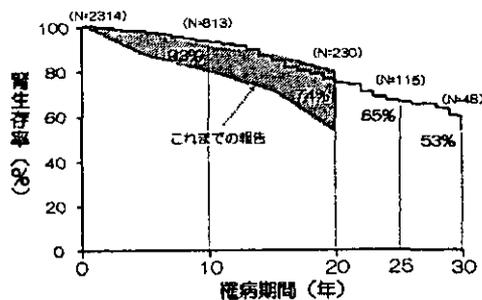
働の富士通社製 EG-MAIN で、臨床研究データが活用しやすいように病院独自のカスタマイズを行った。

C. 研究結果

1) IgA 腎症データベースに関する研究

2314 例(平均発症年齢 27 歳, 尿蛋白量 1.1g/日, 観察期間 9.6 年)の IgA 腎症データベースを作成し, 予後判定基準別の腎生存率を検討した。透析に導入された 218 例(9.4%)の導入年齢 43.2 歳は, 透析医学会の全国集計データ IgA 腎症 2371 例の透析導入年齢 43.6 歳とほぼ同じであった。

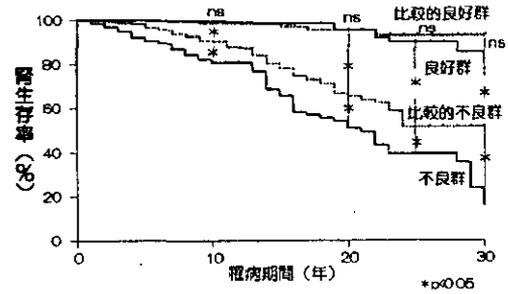
図1 IgA腎症2314例の腎生存率 (Kaplan-Meier法)



Kaplan-Meier 法による腎生存率は 10 年目で 93%, 20 年目で 74%, 25 年目で 65%と, 斜線で示す帯状の範囲で示すこれまでの本邦での主な報告とほぼ同等の成績であった(図 1)。

予後判定基準別にみると, 25 年目の生存率は予後良好群が 90%, 予後比較的良好群が 93%, 予後比較的不良群が 51%, 予後不良群が 39%であった(図 2)。

図2 予後判定基準別の腎生存率



予後良好群と予後比較的良好群の生存曲線がほとんど重なっていた。これは, 予後判定の補助基準の尿蛋白量, すなわち 0.2g, 0.5g, 2.0g で 4 群に分けて腎生存率をみても同様であった。すなわち, 予後良好群と比較的良好群に相当するそれぞれ尿蛋白 0.2 未満と 0.5g 未満の 2 群の生存曲線はほとんど重なっていた。

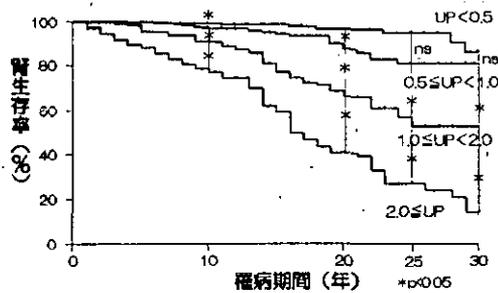
予後比較的不良群の尿蛋白量には 0.5g 以上 2.0g 未満というかなり大きな幅があった。これまでの多くの報告結果から尿蛋白量 1.0g 以上が予後不良因子とされているので, 1.0g で予後比較的不良群を 2 つに分け, かつ, 尿蛋白 0.5g 未満として予後良好群と予後比較的良好群をあわせて 1 群として扱ってみた(表 1)。このように病理所見と臨床データを組み合わせた新しい分類方法にすると, 尿蛋白量が増えるにつれて予後が不良となり, 4 群間で有意な差が認められた(図 3)。

表1 予後判定基準と尿蛋白量による分類

予後判定基準	×増*	分節**	尿蛋白量(補助基準)	尿蛋白量 (g/日)
良好群	無	0	<0.2g/日	<0.5
比較的良好群	軽	<10%	0.2g/日<0.5g/日	
比較的不良群	中	10-30%	0.5g/日<2.0g/日	0.5≤ <1.0 1.0≤ <2.0
不良群	高	30%<	2.0g/日	2.0≤

* ×増: メサンジウム細胞増殖・基質増加
** 分節: 糸球体硬化・半月体形成・ホルモン薬との併用

図3 尿蛋白量 (UP) 別の腎生存率



2) 電子カルテからのデータ書き出しシステム

IgA 腎症の追跡は患者カルテからデータを拾い集めて、数ヶ月から1年間の数値を平均したものを採用する方法とした。電子カルテでは、血清クレアチニン値などの調査項目をCSV形式でデータ出力が可能にすることによってたやすくデータ収集ができるものと推測される。

平成16年10月本格稼働の当院の電子カルテシステムの仕様には、CSVファイルで出力可能となる書き出し口を、医事会計システム、物流管理システム、輸血システム、検体検査システム、電子カルテ、オーダリング、看護支援システムに配置した(表2)。

表2 CSV形式のデータ出力

医事会計システム
物流管理システム
輸血システム
検体検査システム
電子カルテ
オーダリング
看護支援システム

(病歴センターでのみ書き出しが可能)

ただし、この書き出し操作は情報の大量漏洩につながる危険性があるために、病歴室からのみ可能となるようにプログラムを組み、データの書き出しの許可申請書の提出を必

須とし、審査・許可の後、原則として病名室の担当事務官によって操作を行うこととした。

書き出し項目名の選択や日付の指定は新規のプログラム開発となり、費用と開発時間がかかるためあらかじめざるを得なかった。基本的には患者のすべての検査データの羅列したCSVファイルで書き出し、これをデータベースに変換して処理することとした。なお、輸血システムについては専用ソフトが開発されているので、輸血・廃棄・日赤返却の使用別薬剤リスト、在庫薬剤リスト、在庫薬剤リスト、輸血歴の指定を可能とした。

腎生検所見の病理報告書の参照や、来院歴、他科受診状況などの確認は、採血検査が行われていないときの腎死(透析導入)を判断する材料ともなる。また、電話番号や住所などの情報は、電話による聞き取り調査を格段に容易とするので、これらに関しても合わせてデータ書き出しが行えるように設定した。

問題となったのは、IgA 腎症患者の病名が、「IgA 腎症」(標準病名マスターによるIgA 腎症の ICD10 コード:N028, 病名交換用コード:E1KB)の他にも、「慢性糸球体腎炎」(N039, J4F2), 「慢性腎炎症候群」(N030, DSUN), ときには「持続性蛋白尿」(N391, NPUH), 「顕微鏡的血尿」(R31, U1J2)といった広義の病名や症候名で代用されていた点であった。このため、ICD10 に準拠した病名への変更(一括変換), マスターからの類似病名の削除, 原則としてワープロ病名の禁止といった病名の標準化や整備が進行中である。

D. 考案

1995年に提唱(2003年改訂)されたIgA腎症の予後判定基準では、予後良好群と予後比較的良好群の病理学的所見や1日尿蛋白量の差がわずかで予後に有意な差を生じなかった。つまり、予後判定基準という主目的を果たせていない可能性がある。また、本邦での多くの報告では予後不良因子とされている尿蛋白量1.0g以上が区分されておらず、予後比較的不良群の尿蛋白量が0.5g~2.0gと大きな幅になっており、予後に大きなばらつきが生じている。病理組織のサンプリングによる誤差、retrospective検討であること、治療の影響や医師の判定技術の問題などが加わることも判定を曖昧にさせる要因と考えられた。予後判定基準を病理所見のみで分類せずに、頻回に検査できる尿蛋白量という臨床データをも加えて、尿蛋白1.0gで予後比較的不良群をさらに2つの亜群に分けることが有用と考えられた。

重要な点は、2000例を超えるIgA腎症患者で予後判定基準別の生存率を再調査できたということである。十分な症例数が集まらなると4群に分けての腎生存率の検討は難しい。項目を限定し、コンピューターシステムを利用した平易な入力環境を整備することで、膨大なデータから診断基準やガイドラインなどの正当性を検証するという新たな研究方法が生まれるのではないだろうか。

データマイニングは、本来膨大なデータから有益な知識を発掘し、臨床応用や研究に利用することが目的とされる情報処理の新たな手法である。ただし、臨床研究を行う一般病院にとってのデータマイニングは、それまで収集した症例データベースを電子カルテから容易に作成できるような環境作りが先決である

ところが、現在、急速に導入が進んでいる電子カルテシステムにはデータマイニングの手法を取り入れるための標準化作業やデータの書き出し口の設定が配置されていない。それがあっても、多くは経営分析のための統計資料作成を目的とした側面を有している場合が多く、臨床データを取り出す機能は十分ではない。

なかでも、病院や医師の間での病名や検査成績などの取り扱いに関して全国的な標準化が遅れていることが問題である。臨床評価項目などの国の医療政策(表3)や国立病院ネットワークを利用した全国規模の共同研究を行っていく上で、病院間での病名の標準化は今後の大きな課題として残された。

表3 政策医療における臨床評価指標(8分野)

例：がん	胃癌患者総数
	胃癌治療関連死数及び率
	胃癌切除例5年生存率 (stageⅢ)
	EMR施行例(実施件数)
	乳癌全患者数
	乳癌の乳房温存手術率
	乳癌治療関連死率
	乳癌の10年生存率(stageⅡ) など

データを取り出し易くすることは、反面、セキュリティが甘くなる危険性をはらんでいる。多量の情報が漏洩する事件が相次いでいる。最近では、2003年11月のファミリーマートの会員情報の流出(18万人)や2004年2月のヤフーBB(240万件)、2004年3月ジャパネットたかた(30万人)の顧客情報流出、金沢医科大学病院の医師が医学生にパスワード漏洩した問題など大きく報道されている。これらの事件が電子カルテでも起こるのではないかと危惧する声が高まるにつれて、ともすると患者データの書き出しに関して非積極的な意見が出されかねないのが現状で

ある。

たとえば、電子カルテシステムに組み込まれた病歴要約(サマリー)を院外へ持ち出すことについては、医療を専門としない外部監査委員の指摘で当院独自の規制案が検討されている。しかし、内科学会や外科学会の認定医・専門医の申請にあたっては、それまでに受け持った患者のイニシャル、年齢、性、生年月日、職業、居住地(■■県××市)、カルテ No、入院日、病名、現病歴、既往歴、家族歴などといった個人情報を入力したサマリーを提出しなければならない。これは、虚偽の申請が防止するためであろうが、2004年3月15日現在、学会に確認したところ、これまでどおり患者の個人情報の記入は不可欠との返答を受けた。

個人情報になりうるデータは、「生存する個人に関する情報」で「特定の個人を識別できるもの」と個人情報保護法(第2条)に定められており、具体的には氏名、住所、電話番号、画像や音声データ、メールアドレスなどが該当し、5000件以上の患者情報を保有している場合に処罰の対象となる個人情報取扱事業者となるとされており、大半の病院がこれに該当する。患者の同意を得ることはむろん当然であるが、患者データの持ち出しに関しては各病院の倫理委員会による規定作りがこれから必要と考えられる。

最後に、当院で10月に導入される富士通の最新版の電子カルテシステムでも、データの書き出しに際して、項目名を指定して抽出したり、ソートしたりといった汎用データベースでできる編集作業のほとんどができないことが判明した。今後は、プログラムの修正あるいは電子カルテ上で行える抽出機能を組み合わせてどの程度のデータ書き出しが可

能かどうかを検討していきたい。

E. 結論

電子カルテシステムを利用したデータマイニングを一般病院の臨床研究に使用するには、病院内ならびに病院間でのデータの標準化や臨床医が目的とするデータを容易に書き出せること、データ漏洩に対する十分なセキュリティ対策が兼ね備えられていることなど多くの問題点が残されている。

F. 研究発表

1. 学会発表

IgA 腎症に対するステロイドパルス療法の有用性について

吉村光弘, 原 章規, 伊勢拓之, 木田 寛
46 回日本腎臓学会学術総会, 東京 5, 24, 2003

2. 論文

The point of no return during the course of IgA nephropathy

Mitsuhiro Yoshimura, Akinori Hara,

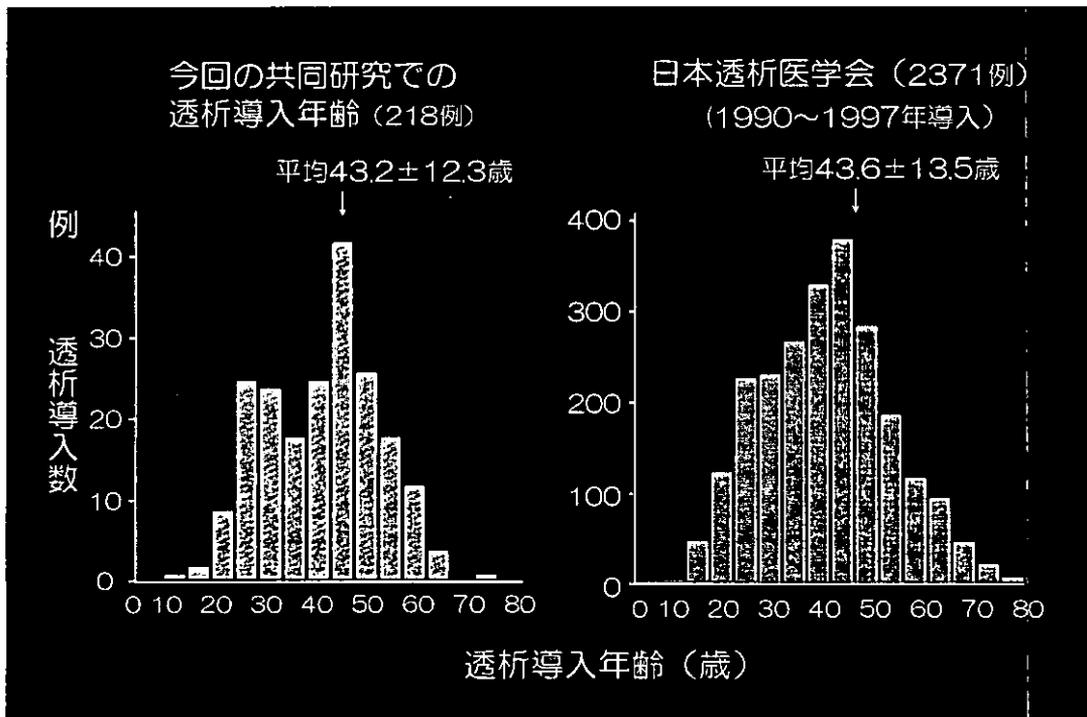
Takuyuki Ise, Hiroshi Kida, Hisao Sasaki,

Takashi Wada, Hitoshi Yokoyama, Ken-ichi

Yamada Nephrology A104-6, 8, 2003

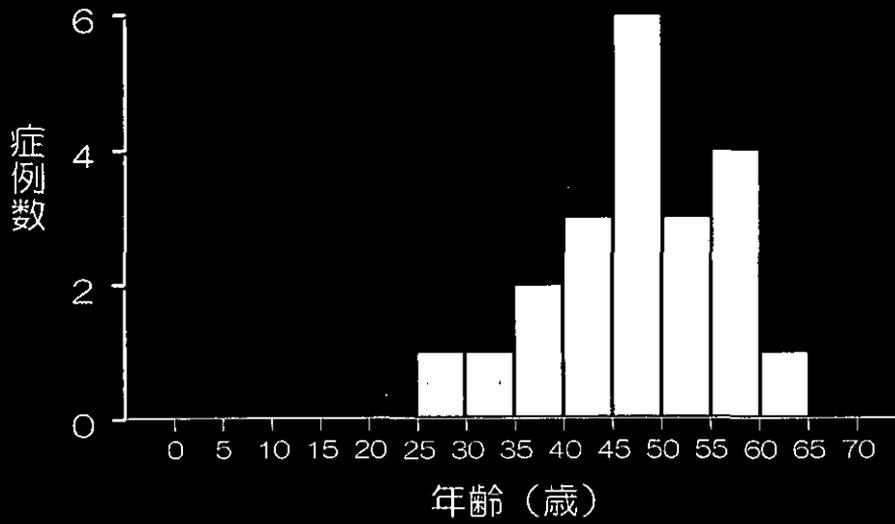
国立病院総合情報ネットワークシステムによる
IgA腎症データベース（腎専門8施設）

症例数 2314 例
 男/女 0.93/1
 発症年齢 27 ± 16 歳（1~86歳）
 尿蛋白量 1.1 ± 3.7 g/日
 観察期間 9.6 ± 7.9 年
 入力項目 初診および最終観察時の尿蛋白量，尿潜血，
 血清Cr，Ccr，転帰，光顕所見（予後判定基準）



燃え尽き 年齢

(Cr \geq 1.3mg/dl, 尿潜血(-), 瘢痕性病変)



発症から透析導入までの期間 (年)

