

図8 検査の実施状況（ポータブル vs. 胸部単純）

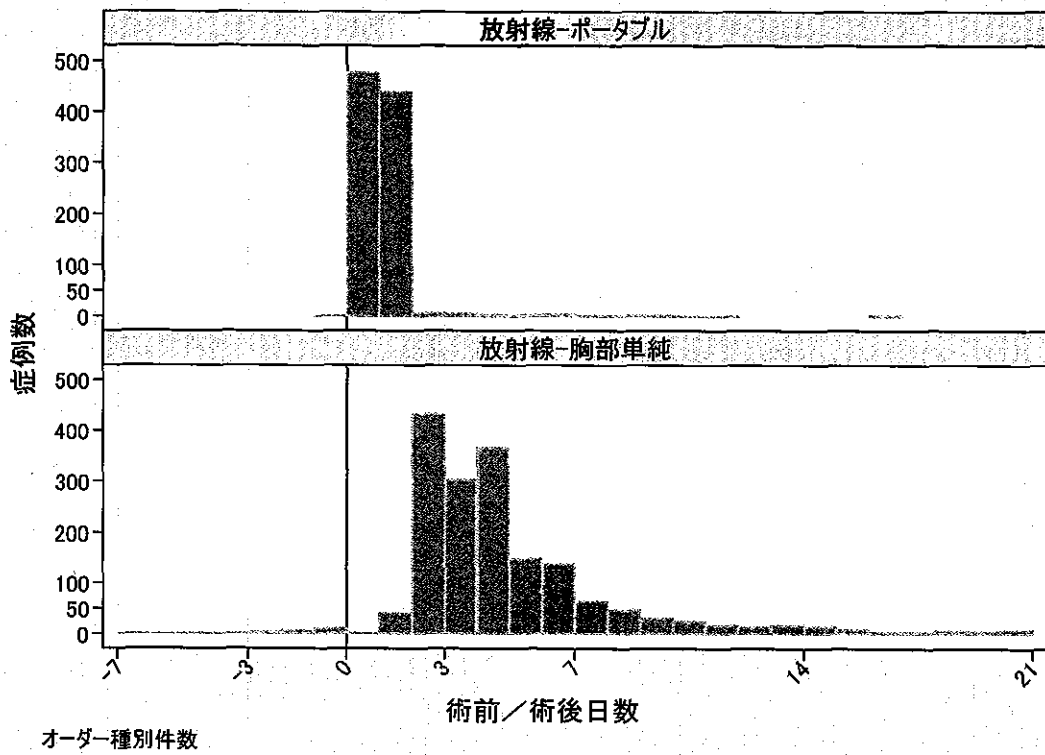


図9 検査の実施状況（術前・肺機能検査）

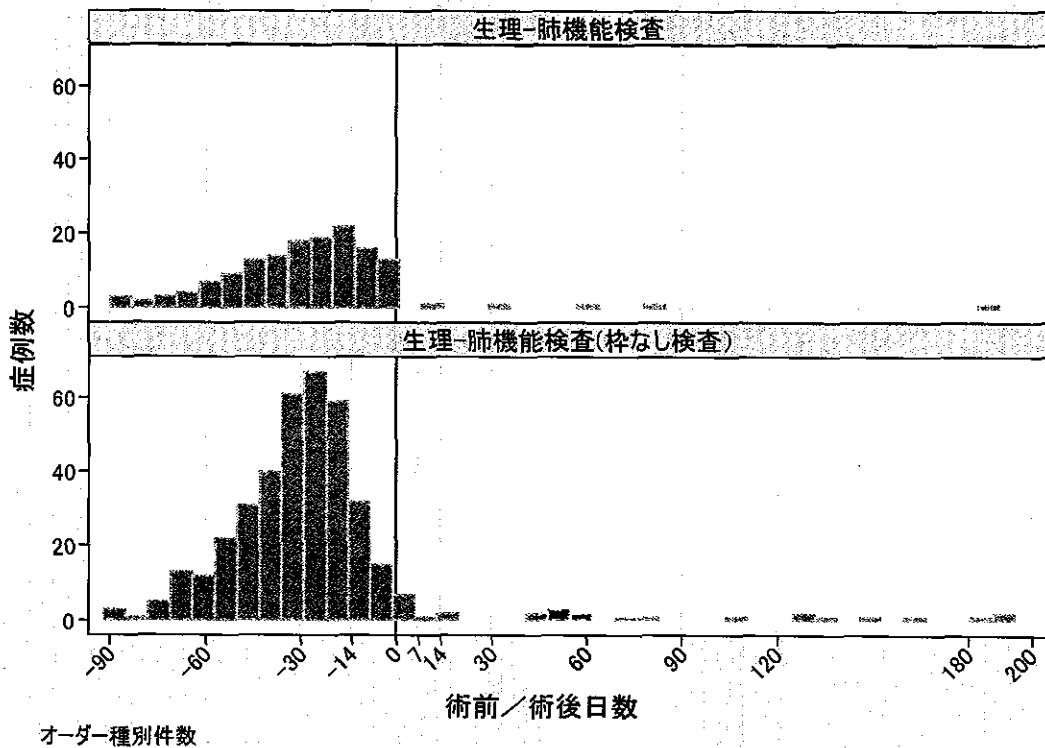


図 10 検査の実施状況（気管支鏡・針生検）

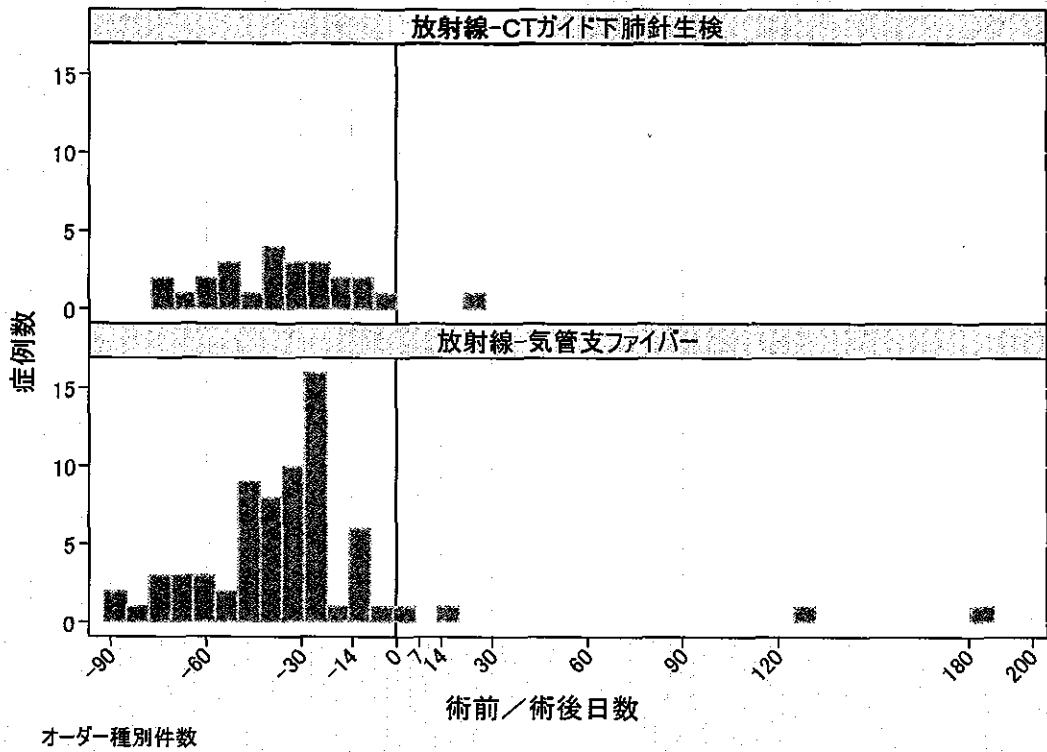


表 1 各種検査の実施率（外来～術前）

検査名称	術前			再掲：入院前			再掲：入院後		
	症例	%	回数	症例	%	回数	症例	%	回数
生理-心電図(枠なし検査)	372	99.7%	1.2	368	98.7%	1.1	19	5.1%	1.1
放射線-胸部単純	372	99.7%	2.1	360	96.5%	1.4	277	74.3%	1.1
放射線-CTスキャン検査	370	99.2%	1.5	356	95.4%	1.3	102	27.3%	1.1
細菌検査	364	97.6%	3.7	206	55.2%	1.4	354	94.9%	3
生理-肺機能検査(枠なし検査)	287	76.9%	1	284	76.1%	1	7	1.9%	1.1
生理-肺機能検査	114	30.6%	1.1	109	29.2%	1.1	6	1.6%	1
放射線-気管支ファイバー	54	14.5%	1	53	14.2%	1	2	0.5%	1
内視鏡-X線透視下気管支鏡	54	14.5%	1	53	14.2%	1	2	0.5%	1
放射線-MRI検査	52	13.9%	1.3	40	10.7%	1.2	15	4.0%	1.3
生理-腹部超音波	29	7.8%	1	23	6.2%	1	6	1.6%	1
放射線-CTガイド下肺針生検	15	4.0%	1.1	12	3.2%	1.1	3	0.8%	1
生理-心臓超音波	15	4.0%	1	10	2.7%	1	5	1.3%	1
内視鏡-非透視下気管支鏡	13	3.5%	1	10	2.7%	1	3	0.8%	1
放射線-一般撮影	12	3.2%	1	6	1.6%	1	6	1.6%	1

表2 各種検査の実施率（手術当日）

検査名称	症例	%	回数
放射線-ポータブル	372	99.7%	1
細菌検査	21	5.6%	2
放射線-胸部単純	2	0.5%	1
放射線-CTスキャン検査	2	0.5%	1
生理-心臓超音波	2	0.5%	1
放射線-一般撮影	1	0.3%	1
生理-肺機能検査(枠なし検査)	1	0.3%	1
生理-心電図(枠なし検査)	1	0.3%	1

表3 各種検査の実施率（術後～）

検査名称	術後			再掲：入院中			再掲：退院後		
	症例	%	回数	症例	%	回数	症例	%	回数
放射線-胸部単純	373	100.0%	10.8	370	99.2%	4	367	98.4%	7
放射線-ポータブル	353	94.6%	1.3	346	92.8%	1.1	19	5.1%	3.1
放射線-CTスキャン検査	247	66.2%	2.2	39	10.5%	1.6	240	64.3%	2
細菌検査	125	33.5%	3.4	98	26.3%	2.6	44	11.8%	3.7
放射線-MRI検査	57	15.3%	1.8	16	4.3%	1.2	45	12.1%	1.8
放射線-一般撮影	50	13.4%	2.1	18	4.8%	2	36	9.7%	1.9
生理-心電図(枠なし検査)	32	8.6%	1.2	6	1.6%	1.2	28	7.5%	1.1
生理-肺機能検査(枠なし検査)	20	5.4%	1.1	2	0.5%	1	18	4.8%	1.1
内視鏡-非透視下気管支鏡	15	4.0%	2.3	4	1.1%	1	13	3.5%	2.4
生理-腹部超音波	13	3.5%	1.1	3	0.8%	1	10	2.7%	1.1
内視鏡-上部・通常	12	3.2%	1.6	4	1.1%	1.3	9	2.4%	1.6
生理-肺機能検査	6	1.6%	1.2	1	0.3%	2	5	1.3%	1
放射線-その他造影	5	1.3%	1.6	4	1.1%	1.8	1	0.3%	1

表4 術前のCT部位

●				199	53.8%	胸部のみ
●	●			87	23.5%	胸部+腹部
●	●	●		71	19.2%	胸部+腹部+脳
●		●		10	2.7%	胸部+脳
●	●		●	2	0.5%	胸部+腹部+頸部
●			●	1	0.3%	胸部+頸部
胸部	腹部	脳	頸部	症例数	%	部位組み合わせ
370	160	81	3	370		
100.0%	43.2%	21.9%	0.8%	100.0%		

## がん診療支援データベース構築に関する研究

分担研究者 松谷 司郎 東京大学大学院医学系研究科 CBI

### 研究要旨

H18年度までに、病院情報システムに蓄積されている臨床情報の中からがん診療に関する診断、検査、治療に関するデータを抽出し、実施されている診療内容について集計・解析を行うシステムのコーポレートモデルを作成し、診療実態データを基にしたがん治療指針を支援するプロトタイプシステム（データベースシステム）を構築する。データベース化された診療内容は生物統計的手法、及びデータマイニング手法を用いて解析可能とする。

H16年度は、がん診療データベース構築に向けて、1) クリニカルパスやガイドライン作成、およびそれらと診療実態の差異把握などに有効と考える、時制データベース（医療行為などの時間を統一的に管理するための基本情報）の現状調査、2) 検査データを用いた臨床情報データマイニングのフィージビリティスタディ、3) データベース及び解析システム開発のための基盤整備、を実施した。本報告では、1)及び2)についてまとめた。

### A. 研究目的

#### 1. 時制データベース

時制データベースはその性質上、患者情報のデータベースである電子カルテ、病院内のワークフロー業務、意思決定支援システム、ガイドラインと診療実態との検証比較システムなど本質的に現在・過去・未来の時間を扱う業務の基本情報となると考えられる。このため、時制データベースの医療情報処理への適用課題をまとめ、がん診療データベースへの適用可能性を調査することを目的とした（学会発表 1）、2)）。

#### 2. 臨床情報データマイニング

がん診療データベース構築後に利用を想定しているデータ分析の手法のひとつであるデータマイニング（今回は関連性分析(アソシエーションとも呼ばれる)と呼ばれる手法を使用)が、現実的な課題を解決するために有効であるか検証することを目的とした。

### B. 研究方法

#### 1. 時制データベース

時制データベースの操作言語のひとつである、ATSQL2 の臨床情報への適用課題をまとめるために、研究用の時制データベースのひとつ

である時制仲介型(時制を扱う部分のみを実装し、DBMS としての基本機能は既存の DBMS を使用するタイプ) のフリーソフトウェア TimeDB を使用した(図 1)。医療情報処理(特に臨床情報処理)においては、データについての頻度、持続、順序、時区間の操作が重要であるが、今回は ATSQL2 でサポートされている時区間の操作を検証するため、ある検査結果のダミーデータを作成し検証した。

#### 2. 臨床情報データマイニング

三徴候は医師が患者を診断する際に用いる疾患を代表する三つの徴候のことである。あるコホート研究(患者数 477 人)の検査データを使用し、その研究対象の疾患の三徴候をデータマイニングにより抽出可能かどうかのフィージビリティスタディを実施した(学会発表 3)）。関連性分析により、患者グループと検査結果のすべての組合せを評価し、「検査値の特異な組合せパターン」の発見を試みるとともに、関連性分析を実施する前までのデータ洗浄の手順についてもまとめた。

DBMS 及びデータマイニングツールは日本 NCR(株)の Teradata 及び Teradata Warehouse Miner を使用した。

## C. 研究結果

### 1. 時制データベース

国際規格であるリレーショナルデータベース操作言語 SQL-92 の上位互換である ATSQL2 は、時区間をもつデータ（例えば、ある症状の持続期間）の操作については、SQL-92 と比較して直感的で、SQL 文が簡潔になる（データ操作のためのアプリケーションプログラム開発の効率が高くなる）ことがわかった。但し、データベースの設計や運用上、留意すべき点があった。

### 2. 臨床情報データマイニング

関連性分析を用いて、既知の徴候組合せをデータから導出し確認した。また、これまで知られていなかった検査結果（徴候）の組合せをいくつか、確認した。

なお、検査データに対するデータ洗浄（欠損値の扱いやデータ項目の分類、集約、カテゴリー化など）に、データマイニングの過程（工数）の大部分を占めることがあらためて判明した。

## D. 考察

### 1. 時制データベース

1) ATSQL2 は SQL-92 の上位互換を謳っているが、混在して使用する場合、構文上と意味論上に齟齬が生じる場合があった。このため、ATSQL2 の導入には、注意深いデータベース設計と運用が必要となる。

2) DBMS として ATSQL2 をフルサポートしている製品（商用、フリーウェアを含め）がないため、ORACLE などの既存の DBMS のひとつのアプリケーションとして実装することが、時制をサポートするひとつの解であると考え（TimeDB のアプローチ）。

また、時制データベースを医療情報処理に適用する場合、ATSQL2 言語の仕様をすべてサポートする必要はないと考えるので、最低限必要とする機能を抽出し、既存の DBMS 上で「軽い仕様」の操作言語を開発してもよいと考える（Stanford 大の ChronusII システムがその例）。

3) 但し、TimeDB のような ATSQL2 の実装では、

トランザクション処理の性能や、障害管理の面で問題があるため、データ更新が少ないデータウェアハウスのような使用が望ましいと考える。

4) 時間の管理は、リレーショナルモデルのような一階述語論理を基本とするモデルでは表現しきれない部分があるので、オブジェクト指向モデルのような、柔軟なモデルも今後検討していく必要がある。

### 2. 臨床情報データマイニング

1) 今回の関連性分析手法が、がんの予防（検査データの新しい解釈と患者に対する指導）や予後の予測などのがん診療の支援にも有効であるが今回の実験では定量的に把握できなかったため、H17 年度以降、がんセンターの病院情報システムのデータを整備し、開発予定のプロトタイプシステムに適用して検証する予定である。

2) 関連性分析は、対象データ間の組合せを網羅的に探索できる優れた手法であるが、データ数の増加による探索空間の爆発（組合せ数の爆発）の問題を抱えており、解決すべき課題によっては異なった手法（統計的手法や他のデータマイニング手法）を用いたほうがよい場合があるので、これも実際のがんセンターのデータを用いて手法の適用と限界を見極めていく必要があると考える。

## E. 結論

### 1. 時制データベース

時制データベース及び（ATSQL2 のような）その操作言語は、現状の TimeDB2 のような実装では、既存の標準 SQL ベースのデータベースとの共存時には注意深いデータベース設計と運用が重要であるが、時間を扱うデータの操作には一定の効果（直感性と簡潔性）がある。

### 2. 臨床情報データマイニング

ある疾患の三徴候（既存のものと、今まで知られていないもの）を、データマイニング手法（今回は関連性分析を使用）から抽出することができた。

## F. 研究発表

### 1. 学会発表

1) 松谷司郎, 小山博史, 篠原信夫:

医療情報処理における SQL-92 と ATSQL2 の「時制」の扱いやすさの比較, 第 24 回医療情報学連合大会 (第 5 回日本医療情報学会学術大会), 2004. 11. 26-28

2) 松谷司郎, 小山博史, 篠原信夫:

医療情報処理への時制データベースの適用課題 -ATSQL2 とその実装 TimeDB を例として- (第 9 回日本医療情報学会春季大会), 2005 (予定)

3) 松本伸哉, 今村知明, 神奈川芳行, 田島文一, 松谷司郎, 小山博史:

医療における関連性分析を用いた三徴候の識別方法, 第 5 回計測自動制御学会システムインテグレーション部門講演会, 2004. 12. 17-19

G. 知的財産権の出願・登録情報

4. なし.

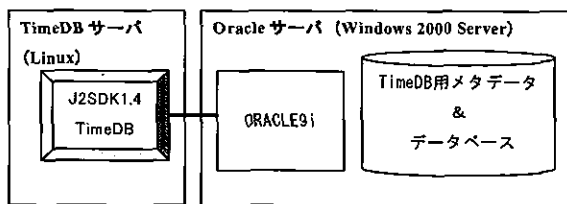


図 1 時制データベースのシステム構成

研究成果の刊行に関する一覧表

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Yamashiro K., Wakao F, et al	Telecytology in Hokkaido Island, Japan: result of primary telecytodiagnosis of routine cases.	Cytopathology	15	221-227	2004
若尾文彦	画像検査部門システム の概要と導入の考え方	IT vision	6	12-16	2004
飯沼 元, 若尾文 彦 他	CT, MRI 検査におけ る大腸癌の術後再発診 断と新たな診断技術	早期大腸癌	8	139-144	2004
飯沼 元, 若尾文 彦 他	大腸癌術前における multi-detector row CT colonography の有用性	臨床放射線	49	409-418	2004
飯沼 元, 若尾文 彦 他	消化管造影検査におけ る FPD-DR	カレントセラ ピー	23	17-21	2005
Goto, K., Shinkai, T. et al	Multi-institutional phase II trial of irinotecan, cisplatin, and etoposide for sensitive relapsed small-cell lung cancer	Br. J. Cancer	91	659-665	2004



Takigawa, N., <u>Shinkai, T.</u> et al	Clinical and pharmacokinetic study of docetaxel in elderly non-small-cell lung cancer patients.	Cancer Chemother Pharmacol	54	230-236	2004
佐伯俊昭、 <u>新海哲</u> 他	新薬展望 2004-第一部 治療情報を臨床の場でどう生かすか、がん専門医の立場から	医薬ジャーナル	40	249-252	2004
別所昭宏、 <u>新海哲</u>	肺癌臨床試験の最近動向-進展型小細胞肺癌の治療	呼吸器科	5	408-413	2004
畝川芳彦、 <u>新海哲</u>	医師主導型治験の問題点	医薬ジャーナル	40	1659-1662	2004
高田三郎、 <u>新海哲</u>	肺癌患者に対する緩和ケア	呼吸器科	6	276-280	2004
野上尚之、 <u>新海哲</u>	肺癌-I I I 期肺癌に対する手術 con	Medical Prectic	21	1365-1369	2004
松谷司郎 他	医療情報処理における SQL-92 と ATSQL2 の「時制」の扱いやすさの比較	第 24 回医療情報学連合大会 (第 5 回日本医療情報学会学術大会)	11	26-28	2004
松谷司郎 他	医療情報処理への時制データベースの適用課題-ATSQL2 とその実装 TimeDB を例として	第 9 回日本医療情報学会春季大会			2005 (予定)

Igaki H, <u>Kato H</u> , et al.	Improved survival for patients with upper and/or middle mediastinal lymph node metastasis of squamous cell carcinoma of the lower thoracic esophagus treated with 3-field dissection.	Annals of Surgery	239	483-490	2004
Yachida S, <u>Kato H</u> , et al.	Adenosquamous carcinoma of the esophagus. Clinicopathologic study of 18 cases.	Oncology	66	218-225	2004
Narikiyo M, <u>Kato H</u> , et al.	Frequent and preferential infection of Treponema denticola, Streptococcus mitis, and Streptococcus ngiosus in esophageal cancers.	Cancer	95	569-574	2004
<u>石川ベンジャミン</u> <u>光一</u>	DPC とがん診療	癌と化学療法	31 (8)	1169-1173	2004
<u>石川ベンジャミン</u> <u>光一</u>	がん化学療法の医療経済学的背景と外来治療へのシフト	MEDICO	35 (7)	19-22	2004
<u>石川ベンジャミン</u> <u>光一</u>	包括医療制度とがん化学療法	化療ニュース	13 (1)	1-3	2004
<u>石川ベンジャミン</u> <u>光一</u>	包括医療の導入とがん化学療法	コンセンサス 癌治療	3 (3)	160-162	2004

石川ベンジャミン 光一	診断群分類の特徴を可視化するDPCポートフォリオ	第24回医療情報学連合大会 論文集			2004
川村 直樹、山城 勝重 他	インターネットを利用した遠隔細胞診の診断成績と課題	日本臨床細胞学会雑誌	43(22)	205-213	2004
加幡晴美、福田治 彦 他	National Cancer Institute-Common Toxicity Criteria (NCI-CTC version2.0) 日本語訳 JCOG 版の信頼性の検討	癌と化学療法	31	1187-1192	2004