

医療情報 DATACUBE の活用

～医療情報システムにおける蓄積から利用へのパラダイムシフト～

山本康仁
都立広尾病院小児科

医療情報がデジタル化され蓄積、活用される段階で、医療連携分野にその活路を見出すのは、医療経済的にも合理的であり、従来の郵便やFAXを利用した連携では実現出来ない成果が期待出来ると同時に、個人情報保護の観点からも望ましいだろう。しかし、医療情報システム導入が当初夢見た事務処理の省力化の達成や、臨床情報の二次利用が順風満帆と言えるのだろうか。

昨今のICT技術の発達は目覚しく、遠隔地の医療者同士の連携を可能にしたテレカンファレンスや、遠隔ロボット治療など突出した成果が出始めている。また医療DWHの構築、そしてその利用も進んでいる。宇都らは(1)、診療報酬請求制度の変化に影響されず原価計算を継続評価することで病院経営戦略に寄与するとした。また中島らは(2) 包括評価制度に重症度を考慮した分類が必要であり、臨床的アウトカム、患者要因、併存疾患を抽出し医療の質を継続的にモニターできるDWHの構築を提唱している。

本邦での医療DWHの構築やOLAPなどによる利活用は確かに活発化しているだろう。(3) しかし、病院経営上の戦略決定より個別患者視点で、遠隔ロボット治療よりはマクロの医療情報の利活用は進んでいるだろうか。吉村らは(4) DWHとOLAPを利用し病院経営支援システムの開発を発表し、その中でDWHを用いて診療支援を実現したとしているが、時系列で集計表示することにとどまっている。ここで欧米に目を向けると Ledbetterらは(5) 積極的にエビデンスの創成とその活用にDWHとOLAPを利用しているという。

国内の病院が置かれている状況を省みると、既に大量のデジタル化情報が蓄積されていることに気づく。意思決定のための医師の思考経路の記録まで含めると、一般の業務システムと同様、或いはより詳細で多くの情報が記録されている。証拠能力を期待され、職種別の厳密な権限によって半ば強制的に記録されるそれらは、現業系の職務の中では圧倒的な質と量となって保存されている。それらを、業務遂行と病歴記録だけに利用を留めるのはもったいないはずだ。

しかし1990年に著書(6)のなかでDWHを提唱したInmonは2007年にData Warehousing in the Healthcare environmentで(7)医療DWHが他の職種と異なりその価値を享受していないと指摘している。通常の業態では顧客と数量が変化するだけで業務パターンが一定であり、情報活用期間が比較的短いに対し、医療では状況が多彩でパターン化しにくく、取り扱われる情報も冗長な文章でそのままの利用が難しく、何十年にもわたりその情報を収集しなければならないことがある。Inmonは医療情報DWHで成功をおさめるのなら、専門用語や略語の統一が必要で、専門領域と語彙の関連情報を利用したり、自然言語処理を備える必要があると述べている。

このような、医療DWHの状況や問題をふまえ、我々のDWHとDataCubeの構築に関して、その利用状況を含め報告する。

方法

我々は医療情報システムに流れるすべてのトラフィックを網羅的に収集し、リアルタイムDWHと、そのMOLAP環境を実現するためDataCubeを構築し、医療安全のための臨床判断支援システム (CDSS)を運用して

いる。東京中央部に近い467床の急性期病院で、2003年3月から長期診療情報DWHとして運用開始し、2005年から富士通製HOPE/EGMAIN-EXと相互接続した。2007年からはOLAP上に知識処理を加え、携帯電話端末へ情報送信を開始している。

病院情報システムが持つトランザクションデータは膨大である。Coddは(8)、業務系トランザクションデータと意思決定データベースが物理的、速度的に両立できないとして、DWHとOLAPをもって多次元データ解析を効率よく行えることを示した。我々は、現在進行する医療行為に対する判断支援を提供するために、トランザクションデータを発生と同時に得ることにしたが、情報追跡性を持たせるため発生環境やプロパティを含め170フィールドから3000フィールドに及ぶ情報を一日3.2~3.8万トランザクション処理している。情報処理のリアルタイム性を確保するため、処理サーバーを並列、パプライン化してサーバファームを構築し処理を行っている。

サーバー間通信と排他処理を減らすために、実施情報を病棟と外来で分割するなど、容量や関連性を考慮してテーブルの水平分割 (horizontal partitioning) を行った。分割と同時にオーダーに関連する処方、処方コード、処方量などの関連データはあえて垂直分割 (Vertical partitioning) を行わず、一つの可変長文字フィールドに、構造を保ったまま索引を設定せずに結合格納した。分散処理を前提に、一連の分割と結合を同時に行っている。

電子カルテへの入力作業より医療行為が優先するために、必ずしも電子情報が過不足無く現実を示しているとは限らない。現実と電子情報の一時的相同性 (Eventual Consistency) が成立したと同時に、知識処理を実行、その結果をすぐに利用できるようにシステムを構築している。並列処理の過程でデータの整合性、可用性、分散性は同時に達成することはできないとしたCAP定理(9)の示す通り、一時的相同性を考慮した構造となっている。また知識記述は既存の言語が中間状態を記述できないため、新たに知識処理パラメータ、処理進行状況、中間結果などを表記できるEDOL(Even Driven Objective Language)を策定し実装している。

情報入力の多くは基幹システム端末に依存し、正規化されているものも含まれるが、医師記載のそれは充分とはいえない。そこで記載から必要な情報を抽出するために自然言語処理を行っている。医師記載の特殊性を考慮し、実際の記載135万件から該当するコーパス20万件を抽出、シソーラスを含む言語処理装置を新たに制作している。こうして得られた情報をもとに、症候群サーベイランスを実施、新型インフルエンザ流行検出を行った。

ヒューマンインターフェースとAugmented Cognition

人間の注意、能力、意思決定の限界などを計算し、インターフェースにフィードバックし、最終的に意思決定能力を改良するAugmented Cognition(拡張知覚)がある。(10)我々は、情報を提示する相手の状況、知り得る情報の種類や量、職種、専門性などのパラメータを考慮し、携帯電話端末を用いて音声およびプッシュ型メッセージとして情報を示している。従来の電子カルテ画面を利用する場合と異なり、Eventual Consistencyが成立した瞬間に伝達が可能なのが利点である。情報伝達は一人ではなく、複数の医療者に伝達することが可能であると同時に、専門家とのコンサルテーションを促し、直接電話交換をおこなって両者を接続するなどの能動的制御も合わせている。これは、もとの拡張知覚よりも踏み込んだあらたなヒューマンインターフェースといえるだろう。

医療者の状況を計算するにあたっては、職種に合わせて、外来、手術、検査予約やそれに関連する患者移動情報、あるいは手術室入室情報、過去の業務解析による行動パターンの抽出から算出する。これにはPHS電話機の移動ベクトル情報と業務解析によるリアルタイム情報もあわせることで、精度を高めている。

結果

自然言語処理の精度を測定するために、あらたに抽出した65万件の記載から、対象コーパスを含む83654記載を抽出した。誤差2%、信頼度5%、母比率10%としてサンプル数を865件に設定し、乱数を用いて選別した例文を用いて正確度を測定した。文節から主語や述語、定型的な言い回しや仮定文などを考慮することができるため、自然文から症候を95.3%の正確度で識別できた。こうして得られた「発熱、下痢、嘔吐、発疹、痙攣」という代表的症候の24時間以内受診総数を患者移動情報とあわせて集計した。2008-2009シーズンインフルエンザの迅速診断陽性例数と症候症例数をJMP8.0、パーテーションを用いて解析を行い予測式を算出した。この予測式をMOLAP内部に組み込み、カルテ記載後5秒程度で過去24時間の集計、解析、可視化を行い、インフルエンザ流行状況か否かを判別、判別結果をTwitterする装置を設置した。2009年4月28日から前方視的に装置を稼働させ、通常とことなる季節に流行を迎えた2009-2010新型インフルエンザを観察している。この装置は病名やインフルエンザ迅速診断の施行の有無や結果には影響されない。4月28日から3月8日までの316日間、延べ244500記載を解析した。幸いにも新型インフルエンザは迅速診断にて診断が可能であり、4582回の施行に対して1177回の陽性例をスタンダードとした。インフルエンザ発生が24時間以内に2名を超える状況を検知するとしたとき、特異度97.5%、陽性反応的中度93.5%であった。

特定抗菌剤の開始日時や使用量に関して感染制御科チーム（ICT）に情報を提供して、院内感染制御の効率向上にも寄与している。薬剤投与実施前にICTが患者の把握を簡便にするため、横断的患者状況を収集してあわせて送付している。6ヶ月間で459件（1日平均2.5件）の情報提供がICTに対して行われ48件の介入が行われた。これは全依頼経路の53%を占めていて、導入前と比較するとTDM(Therapeutic Drug Monitoring)の月間平均件数が0.42から8.43へ向上（ $p<0.05$ ）し、感染制御薬剤師が処方を確認するまでの日数が1.90日から1.19日に短縮した（ $p<0.05$ ）

手術室の入室にあたって、抗凝固剤の薬歴や左右の手術部位の嫌疑に関して注意を促す装置は、管理看護長に注意喚起を行うだけでなく、看護長が多忙のときは即座に手術室の看護師への連絡を行えるよう、電話交換制御を行っている。6ヶ月間に手術申し込み部位の注意喚起を15回行い、そのうちの1件の申し込み記載ミスを事前に発見し注意が有効であった。

考察

DWHとOLAP

Ledbetterら(5)は、臨床DWHを積極的に利用している事例としてトロントのUniversity Health Network (UHN)を例に挙げている。これはトランザクションを処理しルールベースの警報や介入を行うために、トランザクションOLAPが用いられ、その評価も同時に行っている。それとは別にETLを介して非リアルタイム系の臨床判断支援がROLAPで構築されている。ROLAPであり、自由度が高くその情報は診療に使用されるだけでなく、システムアドミニストレーターに提供され、ルールベースの開発に役立つ。(fig.1)

われわれのシステムは、診療判断支援で示す長期診療系の情報をレトロスペクティブではなく、リアルタイムDWHから再構築している。リアルタイム処理が必要であることから、処理負荷の高いROLAPではなく、多次元DataCubeを用いたMOLAP構成とし、トランザクションOLAPもMOLAPに一元化した。(fig.2)

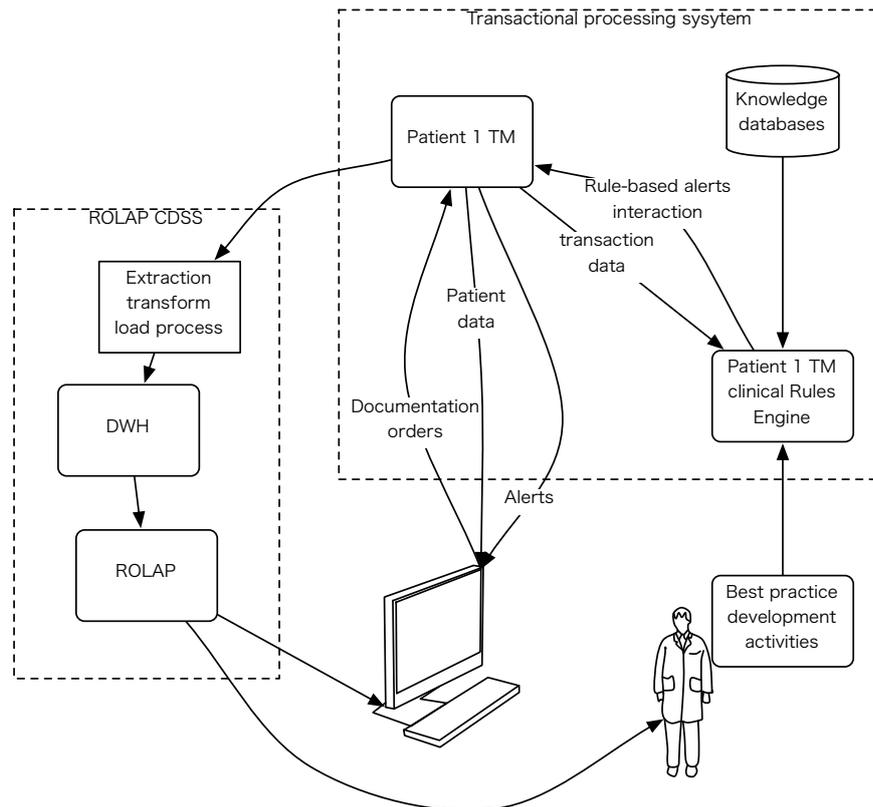


Fig.1 UNHシステム

OLTPの挙動を決定するのは、別に用意されたROLAPからの情報をもとにしている。ROLAPのETLは手動あるいは一日一回のバッチ処理で、過去の情報を利用する。

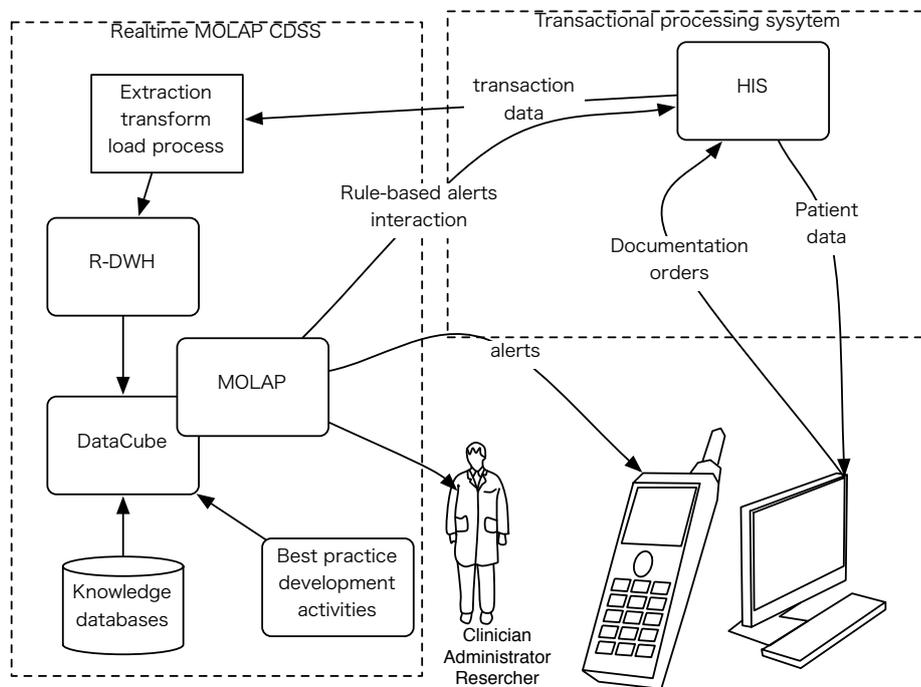


fig.2 広尾病院におけるMOLAP活用

業務システムとの情報連携は情報発生と同時とすることでETLフェーズを含めた、MOLAPと警報の送出、配信までをリアルタイムとした。OLAPとOLTPは区別されず一元化している。

UHNの方法と比較すると、長期診療系の蓄積情報が常に最新であり、患者集合から得られる情報をすぐに患者個人に還元できるという利点がある反面、DataCubeに厳密な設計が必要になり、ROLAPほどの柔軟性を持たない。

そもそも、OLAPは実務系のトランザクションデータベースと意思決定データベースが物理的、速度的に両立しないため、データを冗長格納することから始まっている。(8) OLAPの利点が明らかになれば、おのずとOLAPを利用する人と、その種類が増えることから、柔軟な解析処理に対応するために考案された。われわれが構築したMOLAPとDataCubeをCoddらが示したOLAPの評価に当てはめるのなら、多次元であるとか透明性、パフォーマンスでは充分だが、柔軟性や各次元へのアクセスビリティがフラットであるとする項目ではあきらかに劣っているといわざるを得ない。これはMOLAPの構築にあたって、異なる考えがあるからである。大量のデータを処理するには、並列化だけではなくパイプライン化も必要である。分散構築されたMOLAPの情報は、全体では時間的に整合性が得られない。トランザクションデータの到達と同時に、その局所での一時的な相同性が成立するといえる。トランザクションデータの到達=イベント発生にあわせて、知識処理を実行するために、中間状態を記載、保持するのに適した言語EDOLを利用した。MOLAPレベルでの全体的相同性を求めず、知識処理レベルでの局所的かつ理論的相同性を確保した。このことは、情報伝達まで徹底してイベントドリブンとすることで初めて実現する。

CDSSにおける認知空間への医療知識マッピング

CDSSが行う診療判断支援は、何を、どこまで支援すべきなのか。CDSSの使用目的には、医療プロセスの改善があるが、それが患者個人への最適な医療の提供なのか、患者集団への能率的な医療なのかは、状況に応じて臨床医が判断する必要がある。判断を行うのも責任をとるのも医師であるから、CDSSは臨床医の判断を補助するのみで、臨床医に替わって判断を行うことは、望まれない。また、非専門家がCDSSの示す誤った示唆を、誤っていると判断できなくなることも避けなければならない。従って、厳密にはCDSSは判断をしてはならない。

人間の行動は、人と環境で決定される。環境を人間の心理的空間に投影し、その心理的空間で人は合理的であり、理解の範疇では正義を行うのである。医療におけるリスクは、行動目的に由来するのではなく、環境の投影を失敗することに起因するといえる。医療の置かれている状況は、中断作業が多く、多重タスクで、状態が個々に異なり、異常状態に置かれている物事に対応するものである。常に緊張が強いられ、特に緊急時において人間の信頼性は下がる傾向がある。医療リスクの低減をCDSSの一つの目的とするなら、心理空間への投影を補助するものが望まれる。

心理空間への投影の材料となる情報が、単純に列記、あるいは一定の操作後に表示されるような場合には、個人の資質をもって情報の再配置が必要になる。ここに心理空間への投影が失敗するリスクがある。そこで、機械が人間の認知能力を補助するために、空間的、物理的、心理的に対象となる人がどのような状況に置かれているのかを考慮し、情報を線形に再配置し、文章として理解できるように加工した。伝達される側の状況によっては、示される情報以外に先入観であるとか、リスクを誘発する心理空間へのマッピングを既に済ませている可能性がある。そこで、先入観の少ない関係者に対して、判断する情報を提供し討議を誘導することで、リスクを軽減することも行っている。

医療プロセス中に定義される問題となる行為や結果に対して、どの程度の頻度で干渉するかは、干渉者の繁忙度、職務内容で変化する。医療者が物理的、思想的にどのような状態におかれているのかを求める

必要がある。被干渉者の選定には、プロセスの計画者、あるいは実施者を追跡することになる。干渉者の判断に必要な情報を、線形にマッピングするための設計図は事前に用意するが、情報を受ける側の経験や状況によって、これらの設計は動的に変化しなければならない。必要であればグラフなどの可視化も必要であり、判断を短時間で出来るよう工夫する。

医療プロセスを検知し、行われている医療行為を出来るだけ多く、予定情報も含めて収集し、患者の状況だけではなく、医療者の行動も解析が出来る装置が必要である。医療行為の立案や実施と同時刻に収集し、医療者に密着し、どのような状況でも伝達可能な装置が必要である。従って、医療行為を一旦デジタル化したあと、音声会話を可能とする電話交換機の制御等の一連のシステム情報を伝達するインターフェースシステムとしてCDSを再構築した。CDSは情報を提示するにとどまり、判断は行わない。判断するのは医療者であり、それらの判断は医療者間のディスカッションによってはじめて生じるのである。

さて、CDSSが医療プロセス中の問題となる行為や結果を見つけることは、医療判断にならないのか疑問が生じる。どの程度の情報を示すかの選別の閾値は、情報を受ける側の経験値や状況によって動的に変化し、たとえば、専属者には多くの情報が示される。ここでは、CDSSは討議を持ちかけるのであって、実際に討議するのか、しないのかを判断するのは医療者側にゆだねられている。しかし、情報が開示されなければ人にとって、その情報は存在しないことになり、開示性と到達性をCDSがコントロールしているのなら、CDSが「開示を判断」したことにはならないのか。従来、CDSに組み込まれた判断閾値の多くは、学会や病院などがガイドラインや経験から、絶対的なルールを決定して組み込んできた。しかし患者の置かれる状況が多彩だけではなく、医療者の状況も刻々と変化するに対し、絶対的なルールを決めることは困難である。一方、患者、医療者双方に個別な注目点や閾値を医療者の記載や行動から抽出することは可能である。すなわち、医療者別の置かれている繁忙度、注目点が異なることを念頭に、到達性=発露を定義するのである。発露の定義は、医療者と患者の双方の状況と合わせて、抽象化して個別に格納する。発露の定義は、個別の閾値の集合であるが、同様の状況が発生したときに再利用が可能になるだろう。従来のインタフェースの改良や発展が、機械の中に格納されている情報を、人間の心理空間によりよく投影するのことに傾倒しているに対し、ここでいう新たなCDSSは、人間の心理空間の代わりにCDS内部に仮想空間を作り、発露の定義を再利用する。多くの医療者が生み出した発露の定義が交換され、「気付き」が強化されるインタフェースは、伝播し再利用される仮想空間で、発露のための仕組みを動かすことは、機械と人のあり方、インタフェースの発展の方向、CDSの立場を変化させ、よりよい共存が可能となるだろう。

参考文献

- (1) 宇都由美子、村永文学、宇宿功市郎、熊本一朗. 品質管理・コスト管理のツールとして有効な患者別原価計算システムの開発—病院DWHを利用した DPCごとの患者別原価計算—. 医療情報学 2003 ; 23 (1) : 23-31
- (2) 中島和江、松村泰志、桑田成規、長浜宗敏、武田 裕. 医療の質評価システム構築を目的とした入院日数の分布とその影響要因に関する分析. 医療情報学 2003 ; 23 (1) : 15-22
- (3) 稲岡則子、紀ノ定保臣、宇都由美子、石原 謙、伏見清秀. データウェアハウスとデータ利活用. 医療情報学 2007 ; 27 (3) : 261-268
- (4) 吉村明伸、紀ノ定保臣、梅本敬夫、白鳥義宗、白木由香、中島義憲. OLAP 技術を活用した診療支援と病院経営支援システムの開発. 医療情報学 2003 ; 23 (2) : 159-164
- (5) Ledbetter CS、Morgan MW. Toward best practice : Leveraging the electronic patient record as a clinical data warehouse. J of Healthcare Info Manage 2001 ; 15 (2) : 119-131
- (6) Inmon WH. Building the Data Warehouse(First Edition).USA:John Wiley & Sons Inc,USA,1990
- (7) Inmon WH. DATA WAREHOUSING IN THE HEALTHCARE ENVIRONMENT. [http://inmoncif.com/registration/whitepapers/DATA WAREHOUSING IN THE HEALTHCARE ENVIRONMENTR1.pdf](http://inmoncif.com/registration/whitepapers/DATA_WAREHOUSING_IN_THE_HEALTHCARE_ENVIRONMENTR1.pdf) 2007
- (8) Codd EF、Codd SB、Salley CT. Providing OLAP to User-Analysts : An IT Mandate http://www.minet.uni-jena.de/dbis/lehre/ss2005/sem_dwh/lit/Cod93.pdf 1993
- (9) Seth Gilbert, Nancy Lynch : Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. ACM SIGACT、New Volume 33;Issue 2 :51-59
- (10) Augmented Cognition <http://www.augmentedcognition.org/>