

平成 27 年 6 月 10 日現在

機関番号：16101

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24790502

研究課題名(和文)医療情報の二次利用のためのOLAPシステム構築

研究課題名(英文)Construction of OLAP system for secondary use of medical information

研究代表者

玉木 悠 (TAMAKI, Yuh)

徳島大学・大学病院・助教

研究者番号：40563987

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では、電子カルテシステム、地域医療連携システム等に蓄積されている医療情報について、任意の項目を容易に組み合わせる集計、検索できるインタフェースを提供し、データベース等に関する技術的知識が無くとも容易に分析を可能にする、OLAP(On-Line Analysis Processing)システムを構築した。構築されたOLAPシステムを用いて、糖尿病症例について治療パターン毎に抽出し、検査結果値から治療経過を評価するダッシュボードを実装、評価した。これにより、プログラミング・データベース等の技術的知識を有さずとも、容易に医療情報を分析可能にするシステム基盤を実現した。

研究成果の概要(英文)：In this study, I built the On-Line Analysis Processing (OLAP) system which analyzed accumulated medical information into a hospital information system, an electronic health record system. The user interface of the OLAP system can understand operation visually and can operate it easily. Even the person who does not wear a programming skill enables data analysis easily. Using a built OLAP system, I implemented the dashboard which classified cases in each treatment pattern and evaluated an episode of care and evaluated it. In this way, I realized the system infrastructure which made medical information analyzable easily.

研究分野：医療情報学

キーワード：二次利用 OLAP BI

1. 研究開始当初の背景

我が国においては、中・大規模病院を中心として電子カルテシステムやオーダエントリシステムといった医療情報システムの導入、普及が進んでいる。これらのシステムは診療や業務といった医療情報の一次利用を目的としたシステムである。これに対し、二次利用を目的とした DWH (Data Warehouse) の構築も広まってきている。これは、臨床研究や病院経営を目的に、電子カルテシステムやオーダエントリシステムに蓄積されたデータを分析するためのデータベースシステムである。病院情報システムの普及により、医療情報が電子データの形で保管、集積されるようになった事に伴い、DWH も病院情報システムの一部として導入されることが多くなった。

医療情報の二次利用としては大きく分けて「病院経営」と「医療の質評価・向上」の2点が考えられる。多くの医療機関では、DWH を活用して病院経営分析を行っており、DWH から必要なデータを吸い上げ、集計しグラフ化する病院の経営分析のための可視化ツールが一般的にも整備されている。これに対し、臨床研究のための症例抽出や、クオリティ・インディケータ作成といった、医療の質評価・向上への活用のための可視化ツールやシステムの整備は進んでいない。医療の質評価のための分析には、病名、処置・処方や検査結果といった診療行為と結果の内容によって症例を検索、集計する必要がある。しかし、従来の病院情報システムでは、症例を患者 ID 等の番号で検索する仕組みしか備えておらず、診療行為と結果の内容によって症例を検索するには、データベースへ直接アクセスするしかなく、分析する知見を持つ医療者が、直接取り扱うには技術的ハードルが高かった。

2. 研究の目的

従来の病院情報システムでは、症例を患者 ID 等の番号で検索する仕組みしか備えておらず、処置・処方や検査結果といった診療行為と結果の内容によって症例を抽出するには、データベースへ直接アクセスするしかなく、医療者にとっては技術的ハードルが高かった。本研究では容易に任意の項目を組み合わせ集計、検索できるインタフェースを提供し、技術的知識が無くともシステムに蓄積された医療情報の集計、分析を可能にする OLAP (On-Line Analytical Processing) システムの構築を目的とする。

3. 研究の方法

(1) データマートを備えた BI (Business Intelligence) システムを活用した OLAP システムの設計・構築

一般的に DWH は、各基幹業務システムで運用されているデータを収集し、蓄積する機能を持つ。病院情報システムにおいては、医療

機関における業務を行う機能を持つだけで無く、患者に対する治療を記録として全て保持、保管する機能を持っている。そのため、病院情報システムのデータベースをそのまま DWH として活用することも可能である。ただし、レスポンスの低下など業務への影響を考慮し、データベースの複製を DWH とすることが一般的である。

そのため、病院情報システムの DWH には全患者の全治療履歴が含まれており、これから直接、個別の分析毎に集計処理を行うことは、処理に要する時間や、扱うデータ項目数の観点から現実的では無い。そこで、DWH から分析目的ごとに必要とするデータ項目を抽出し、処理し易くするために加工した中間データを保持するデータベースであるデータマートを用いる。

また、分析においてはデータベースに関する技術的知識を持たない医療者が、容易に直接操作、参照可能とするため、直感的に理解可能なユーザインタフェースが求められる。そのため、抽出・集計条件を容易に設定、操作可能なユーザインタフェースを備え、抽出・集計結果をグラフ及び表により可視化してユーザに提示するシステムとして BI システムを用いる。

これらデータマート、及び BI システムの機能を有する OLAP システムを設計、構築する。

(2) OLAP システムによる症例抽出ダッシュボード構築

構築した OLAP システムを、徳島大学病院の DWH と接続し、症例分析のためのダッシュボードを作成し、OLAP システムの活用を実証する。

分析対象としては、糖尿病患者とし、処方、注射内容から治療内容をパターン分けし、治療状況を HbA1c により評価するダッシュボードの構築を行う。症例抽出アルゴリズム作成にあたっては、徳島大学 糖尿病臨床・研究開発センターの糖尿病専門医の協力を得、抽出条件の作成、作成したダッシュボードから得られた分析結果の評価を行う。

(3) 地域連携コホートデータベースにおける OLAP システムの実証

徳島大学では、糖尿病を対象とした地域医療連携ネットワークである徳島糖尿病克服ネットワークを運営している。当該ネットワークでは、レセプト電算データと、医療情報の標準データ形式である SS-MIX (Standardized Structured Medical record Information eXchange) により検査結果を収集し、蓄積する仕組みを有しており、地域医療連携ネットワークの DWH として、コホートデータベースを運用している。コホートデータベースは一般的なデータベースの仕組みである RDBMS (Relational DataBase Management System) で構築されている DWH と異なり、より検索も高速で大規模データの

処理に適している KVS (Key-Value Store) 型データベースである Cassandra により構築されている。

構築された OLAP システムについて、このコホートデータベースへ接続することで、より広域かつ過去にわたる大規模データを扱うことを実証する。

4. 研究成果

(1) データマートを備えた BI システムを活用した OLAP システムの設計・構築

データマート、及び BI システムを備えた OLAP システムとして、下図に示すシステムアーキテクチャを設計した。

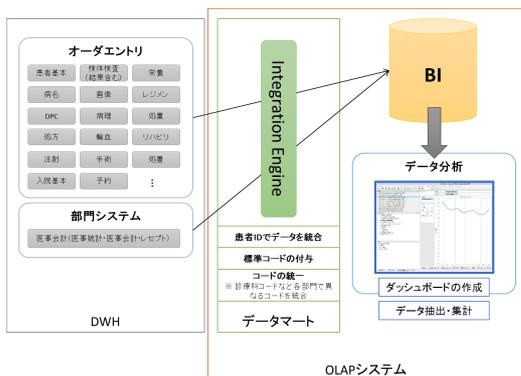


図1 OLAP システム概念図

データマートでは、単純に分析対象となるデータを一時保管するだけでなく、分析に有用な分類コード等のDWHに含まれていないデータを付与する機構を加えた。医療情報で用いられる、病名における ICD-10 (International Classification of Diseases) や、薬剤における HOT 番号は標準化されたコード、分類であり、これらを付与することで抽出条件の設定、集計結果の比較が容易になる。

これらの機能を実現可能な BI システムとして、Qlik 社の QlikView を用いて OLAP システムを構築した。QlikView は、オンメモリ型データベースを備えた BI システムである。本研究においては、QlikView が有する、オンメモリ型データベース上にデータマートを構築、ODEC ドライバにより DWH と接続し、必要なデータをデータマートに抽出する仕組みとした。データマートに蓄積されたデータは、QlikView によって集計、可視化される。データ分析に際しては、QlikView の BI 機能を用いてダッシュボードを作成し、抽出条件の設定に必要なインタフェースを備えた画面を構築する。ダッシュボード上のインタフェースを操作することで、任意の条件での症例抽出、集計結果の可視化が可能になる。

(2) OLAP システムによる症例抽出ダッシュボード構築

構築された OLAP システムを用いて、糖尿病患者症例集計ダッシュボードを構築した。構築したダッシュボードを以下に示す。

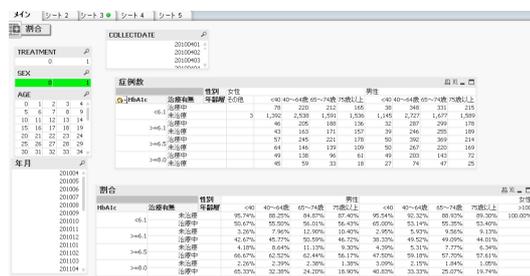


図2 糖尿病症例集計ダッシュボード画面



図3 糖尿病症例集計ダッシュボードグラフ

糖尿病症例集計ダッシュボードでは、画面左側の抽出条件パネルを操作することにより、性別・年齢・対象年月日・薬物療法の有無、により条件を絞り込み糖尿病症例を抽出、集計が可能になっている。また、表・グラフ上に表示されている集団を直接クリックすることで、対象集団の詳細をドリルダウンして表示することを可能にしている。

本ダッシュボードは、徳島大学 糖尿病臨床・研究開発センターの協力を得て開発された症例抽出アルゴリズムを用いて実装され、疫学研究でのデータ集計に用いられた。

症例抽出アルゴリズムの開発においては、その妥当性を検証するため、抽出結果と電子カルテの記録内容を比較し、アルゴリズムにおける条件設定の更新した。更新に伴い、ダッシュボードの実装も更新する必要があったが、データマート内における抽出条件の変更のみで対応できるため、容易なアルゴリズム開発が可能になった。

しかしながら、データマートとDWHの接続、及びデータマートに取り込むデータ項目の設定については、QlikView 独自のプログラミング言語で記述する必要があり、症例抽出、及び分析に用いるデータ項目が追加となる場合は、技術的知識が要求される課題があった。

(3) 地域連携コホートデータベースにおける OLAP システムの実証

地域連携コホートデータベースは KVS 型データベースの Cassandra で構成されており、RDB (Relational DataBase) との接続を前提とした BI システムは、直接接続することはできない。また、前項で明らかになった、データマート設定において、技術的知識が要求

される課題を踏まえ、システムアーキテクチャの見直しを行った。下図に示す。

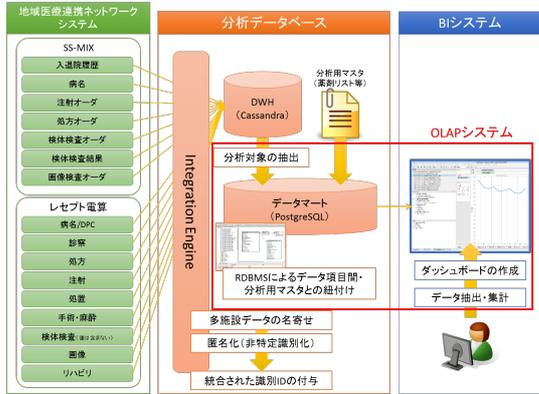


図 4 OLAP システム概念図

地域連携コホートデータベースはそれ単独でデータ検索・抽出の機能を持つ。しかしながら、KVS型データベースであるため、同一患者のデータ項目間を紐付けて検索する機能はない。そのため、データマートをRDBで構成し、地域連携コホートデータベースから分析対象期間、分析に必要なデータ項目等の粗い条件で抽出したデータをデータマートに蓄積し、データマート上でデータ項目間や、各種標準分類・コードとの紐付けを行う構成とした。また、BIシステムが接続するデータベースはRDBである必要があるが、データマートをRDBで構成することで、解決される。

RDBにはPostgreSQLを用い、BIシステムには、RDB上におけるデータ項目の紐付け、抽出条件を操作するユーザインタフェースを有するTableau Desktopを用いた。データマートであるRDBを操作するユーザインタフェースを備えることで、プログラミング言語やデータベース操作言語等の技術的知識が無くとも、データマートを操作することを可能とした。

構築されたOLAPシステムを用い、糖尿病症例集計ダッシュボードを実装した。実装したダッシュボードの例を以下に示す。

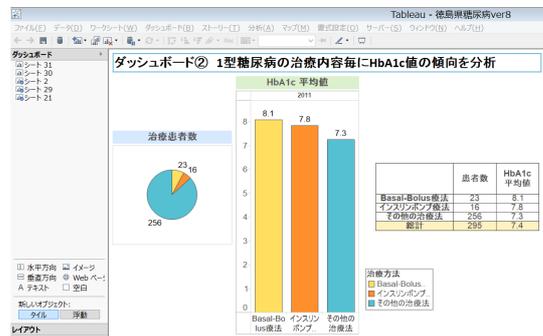


図 5 1型糖尿病の治療内容毎にHbA1c値傾向を分析

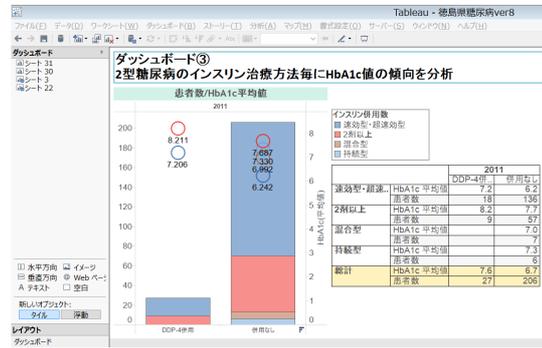


図 6 2型糖尿病のインスリン治療方法毎にHbA1c値傾向を分析

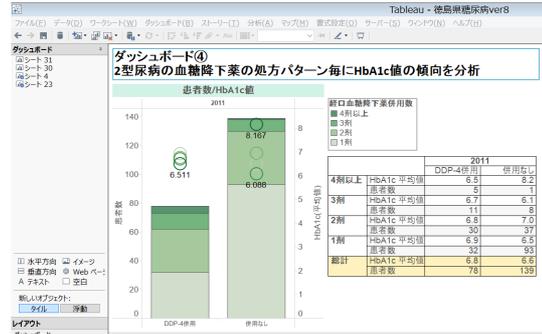


図 7 2型糖尿病の血糖降下薬の処方パターン毎にHbA1c値傾向を分析

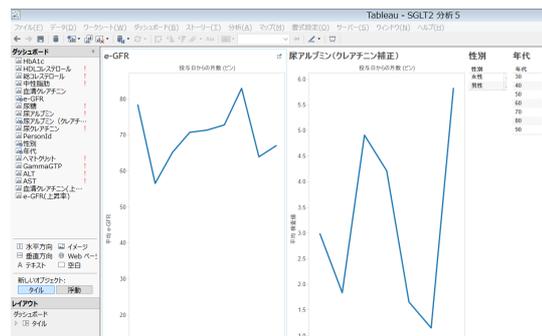


図 8 DPP-4 阻害薬投与開始前後の腎機能検査結果を評価

これらのダッシュボードでは、糖尿病症例を、処方・注射等で投与されている薬剤の組み合わせにより抽出し、パターン毎にHbA1c値傾向や、特定薬剤の投与開始前後での腎機能検査値傾向をグラフで可視化している。このような複雑な抽出条件であっても、OLAPシステムにより、プログラミング言語・データベース操作言語等の技術的知識が無くとも、ダッシュボードを作成し、扱うことが可能であることを示した。

5. 主な発表論文等なし

6. 研究組織
(1)研究代表者

玉木 悠 (TAMAKI, Yuh)
徳島大学・病院・助教
研究者番号: 40563987