

一般口演 | 医療アセスメント

一般口演2 医療アセスメント

2019年11月22日(金) 09:00 ～ 11:00 E会場 (国際会議場 3階中会議室301)

[2-E-1-01] 地域の医療費分析を支援する視覚化システムの試作

○豊田 修一¹、橋本 いずみ¹（1. 上武大学看護学部）

キーワード：Information Visualization, Data Analysis, Health Insurance Claims Data

近年、医療費分析は、国レベルだけでなく、地域での取り組みも重要となってきた。対象となる医療費データは、その構造が複雑であり、量は膨大である。医療費の分析においては、①人口構造上の上位集団との比較、②同一階層内（地域間）での比較、③対象疾病の絞りこみ、④年齢範囲の絞りこみなどが行われる。一方、大量のデータを分析する必要がある医療費分析において、情報視覚化は有効な技術である。そこで、本研究では、地域で医療費分析を担う情報処理を専門にしない職種を支援する視覚化システムを提案する。

本研究で用いるデータは、レセプトの詳細データではなく、医療機関毎、患者毎、入院・外来別、月単位でまとめられたレセプトの集約データである。また、視覚化する統計量は、受診率、一人当たりの医療費、1件あたりの医療費、人口指標等である。

本研究では、視覚化の視点として、疾病指向と空間（地理的）指向の2視点を選択した。疾病指向の視点では、平行座標(PCP)や散布図行列を利用した視覚化を実現した。平行座標では、地域を軸に疾病を折れ線に対応させ、データを表示する。年齢範囲を変化させることもできる。空間指向の視点では、コロプレスマップ（地理空間表現）や累積分布表示(CDF)を利用した視覚化を実現した。コロプレスマップの表示では、単一の疾病を対象にして、データ集計の階層（都道府県＞医療圏＞市町村）に沿って、視覚化を実現した。

本システムでは、医療費全体の把握と共に、年齢範囲間、疾病間、地域間、及び、異なる空間的階層との間での比較を視覚的に行うことができる。これにより、若年者と高齢者の指標の相違や疾病毎の医療費の分布傾向の把握が容易になる。さらに、地域における生活習慣病の予防活動などにも有効であると考ええる。

地域の医療費分析を支援する視覚化システムの試作

豊田修一、橋本いずみ

上武大学看護学部

Prototype of Visualization System for Regional Medical Expenditure Analysis

Shuichi Toyoda, Izumi Hashimoto

Faculty of Nursing, Jobu University

Abstract

Recently the regional medical expenditure analysis becomes important in Japan. This study proposes the interactive visualization system that contributes to the grasp of the outline of the regional medical expenditure. This system visualizes the medical expenditure from two aspects (a disease aspect and a regional aspect) by using visual primitives. The parallel coordinates plot (PCP) and the scatterplot matrix (SPM) visualize the data from disease-oriented aspect. The geographic information system (GIS) and the cumulative distribution function (CDF) plot visualize the data from region-oriented aspect. This system is effective to understand the medical expenditure differences between a young generation and aged generation. In addition, it is effective for the regional preventive activity of the lifestyle disease by public health nurse (PHN).

Keywords: Information Visualization, Data Analysis, Health Insurance Claims Data.

1. 緒論

医療費データは、その構造が複雑であり、量は膨大である。医療費の分析においては、①人口構造上の上位集団との比較、②同一階層内(地域間)での比較、③対象疾病の絞りこみ、④年齢範囲の絞りこみ、などが行われる。このようなデータ特性から、医療費の分析は専門性の高い業務であった。しかしながら、近年、生活習慣病の増加、高齢化、医療技術の進展などに伴い、医療費が増大し、保険者である地方自治体では医療費分析のみならず、医療費の分析に直接かかわりをもたない職種(保健師等)の医療費構造の理解の促進が求められている。また、大量のデータを分析する必要がある医療費分析において、情報視覚化は有効な技術である¹⁾。

2. 目的

医療費分析における情報視覚化の適用は、大きく2つに分けることができる。具体的には、①医療費分析の専門職に対する分析支援を目的とした視覚化、②医療費分析を専門としない職種(非専門職)に対する情報把握を目的とした視覚化、である。専門職を対象とするシステムでは、医療費データのデータキューブを構成し、データ属性の階層に沿ったドリルダウンやロールアップなどの操作を視覚的に行うビジュアルアナリシス機能を実現する必要がある。一方、非専門職を対象とするシステムでは、利用者は医療費データに馴染みが薄く、多様な視点から医療費データの把握を支援するための視覚化を実現する必要がある。

そこで、本研究では、地域の保健医療活動を担うなど情報処理(医療費分析など)を専門としない職種を対象とした医療費分析支援システムを提案する。

3. 情報視覚化

情報視覚化方式は、いくつもの方式があるが、ここでは、本システムで使用した、主な技術について述べる²⁾。

平行座標(PCP:Parallel Coordinates Plot)は、多次元のデータをコンピュータ画面に表現する方式として有効である。2次元のコンピュータ画面では、大規模な表や5, 6次元以上の多次元データの表示には、限界がある。平行座標は、2次

元空間に多次元の情報を表現する方式である。平行座標では、データの次元の増加と表現の複雑性の増加が同じ速度で進行する。具体的には、データが1次元追加されると、表現軸が1本増加するだけである。横方向に軸を並べることが多いので、実際は、横方向に1本分大きくなる。PCPは、データ空間や時間的空間などに着目する。PCPでは、多次元データ表現において、どの次元にも均一した重みで表現できる。

散布図行列(SPM: Scatter Plot Matrix)では、散布図をマトリックス状に配置して、複数の変数の関係性を視覚化する。複数の変数間のデータの相関関係を俯瞰できる。

地理空間システム(GIS: Geographic Information System)を利用した特性分析では、疾病毎にコロプレスマップを使用して表現する。コロプレスマップでは、統計量を濃淡で表現する。コロプレスマップの表示では、複数変数の同時表示ができないため、表形式で他の指標を同時に表示する必要がある。

累積分布プロット(CDF:Cumulative Distribution Function Plot)は、値の大きい順に累積した値をプロットしたものである。一部のデータに大きな値が集中している場合は、CDFプロットは、急峻な立ち上がりを示す。データ間の大小差が小さな場合は、CDFプロットは、直線的なグラフとなる。

4. システム概要

4.1 視覚化の概要

本システムでは、視覚化の視点として、疾病指向と空間(地理的)指向、年齢指向の3視点を選択した(図1)。疾病指向のデータ階層には、ICD-10に基づく階層や、厚生労働省が使用しているレセプトの分類階層がある。疾病指向の視点では、データベースの全体を対象にして、特徴抽出を絞り込み機能などで実現した。空間指向のデータ階層には、(都道府県>医療圏>市町村)がある。空間指向の視点では、単一の疾病を対象にして、データ階層に沿って、視覚化を実現する。年齢情報は、対象データの絞り込みにも使用する。

4.2 データセット

本研究では、レセプト(診療報酬明細書)データを用いた医療費の分析支援を試みる³⁾。本研究で用いるデータは、レセプトの詳細データではなく、医療機関毎、患者毎、入院・外

来別、月単位でまとめられたレセプトの集約データである。この集約データからは、患者の個人情報除去されており、データ利用時に個人情報への配慮が必要なくなる。具体的には、群馬県国保連が公開しているデータを利用した。

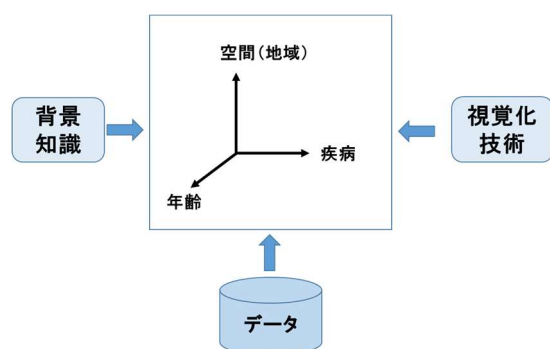


図1 視覚化の概要

4.3 視覚化基本要素

医療費データセットのサイズは非常に大きい。このため、対話的視覚化システムを開発することは、困難である。この課題に対処するために、本研究では視覚化基本要素 Visual Primitives を導入した。視覚化プリミティブは、レセプトデータから作成する。視覚化プリミティブの主要素は統計値データである。視覚化プリミティブを使用により、医療費の視覚化に必要とされるデータサイズは、大幅に縮小される。また、視覚化情報を生成するコンピュータ負担も大きく軽減できる。これにより、医療費データの対話的視覚化を実現する。

4.4 アーキテクチャ

本システムは、Constructor と Visualizer モジュールから成る。Constructor モジュールの役割はデータ変換と蓄積である。Constructor は、視覚化プリミティブの処理や未加工データと視覚化プリミティブが共存するデータベースを作成する。視覚化プリミティブは、1 年単位で統合的に管理される。Visualizer モジュールは、情報の視覚化を担当する。

4.5 インタラクション

本システムで実現したインタラクション(Interaction) 機能の一つに、ツールチップ(tooltip)がある。ツールチップ機能は、マウスポインタと連動した機能である。ユーザがカーソルを何かの項目に合わせたとき、その項目に覆いかぶさるような形で小さな枠が出現し、その枠内には選択された項目に関する補足情報を表示する。

4.6 ページリンク

本システムは、メニューページを基点として、リンクレベルの浅い構造とした。さらに、メニューページには、多くの項目を準備した。浅いリンクにすることで、データ検索が容易に再現できる。これにより、データ構造になじみの薄い初心者にも、利用しやすい構造となっている。

5. システム評価

5.1 疾病指向の視覚化

疾病指向の分析支援では、SPMとPCPの利用について述べる。疾病間の関係把握にはSPMを用意した。

多次元データを対象としたPCPを利用した視覚化では、疾病分類(130次元程度)と地域区分(40次元程度)で構成される多次元データを、2次元空間(コンピュータディスプレイ)に

一括表示して、分析支援を実現した。地域を縦軸に疾病を折れ線に対応させ、データを表示する。PCPは、データのパターン検出に有効である。表示する疾患は選択可能にした。

特定の地域と地理的空間での上位階層(県)との比較や、2地域間での比較においては、縦軸を移動することで、比較対象の相違をパターンで把握することができる。また、年齢区分を利用して対象データをフィルタリングする機能も実現した。これにより、年齢範囲を変化させることで、若年者と高齢者の指標の相違を的確に把握できる。図2は、疾患別の受診率をPCPで表現したものである。図は54歳以下のデータを表現したものであり、疾患の受診率は全て4%以下である。図3は75歳以上のデータを表現したものであり、30%を超える受診率を示す疾患の存在が確認できる。このように、表示対象の年齢を絞りこむことで、対象年代の特徴を表すことができる。

SPMでは、糖尿病や慢性腎疾患などの生活習慣病における保険者間の相関関係を表現する。各相関図は、受診率(水平な軸に対応)と、一人当たりの医療費(縦軸に対応)の関係を表す。

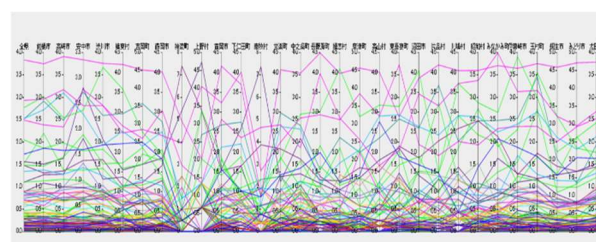


図2 受診率(54歳以下:最大4%)

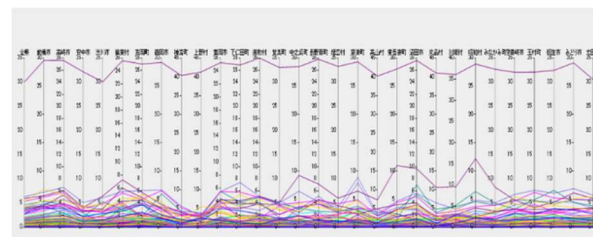


図3 受診率(75歳以上:最大35%)

5.2 地域指向の視覚化

地域指向の分析支援では、GIS、CDF、集合棒グラフ(Grouped Bar Chart)の利用について述べる。地域毎の空間的な特性の把握は、地理空間(コロプレスマップ)上に、医療費関連データの濃淡を表示することで実現している。地理空間上の空間的粒度は、市町村と広域市町村(医療圏)の2段階を提供した。また、複数のデータを同時に表示できないというコロプレスマップ特有の制約は、該当空間上に複数のデータを tool tip で表示することで、解消した。ツールチップの第1の使用法は、within data である。これは、表示データの詳細の表現である。第2の使用法は、between data である。これは、表示データと他のデータの関係表現である。図4は、医療圏単位に糖尿病の受診率を表示したものである。関連データの表と共に表示している。また、本システムは、地域の保健師への医療費情報の提供を目的にしたため、市町村単位での情報表示を行った。しかしながら、人口200万人の県においても、人口の少ない町村では、該当疾患の患者数が極めて少ないため、個人情報保護の点から、表示を保留したケースが複数あった。

生活習慣病を対象にして、疾病毎の医療費の分布の把握を累積分布CDFプロットで支援する。CDFプロットでは、グラフの形状が、急峻な立ち上がりを示す疾患、直線的なグラフ

を示す疾患などに分類できる。図5の右側は、虚血性心疾患の医療費の分布であり、急峻な立ち上がりをするグラフの例である。右側は、高血圧症疾患の医療費の分布であり、直線的な形状を示すグラフの例である。

集合棒グラフは、全体の傾向把握に有効である。図6は、市町村単位での4年間の受診率の変化を表したものである。市町村単位では、人口の少ない市町村があるため、受診率の変化が不規則に見える。しかし、医療圏単位では、受診率の増加傾向が明らかになる。

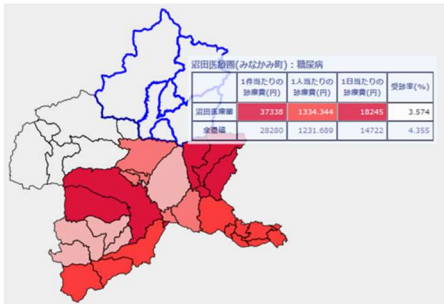


図4 医療圏毎の受診率と関連するデータの表示例

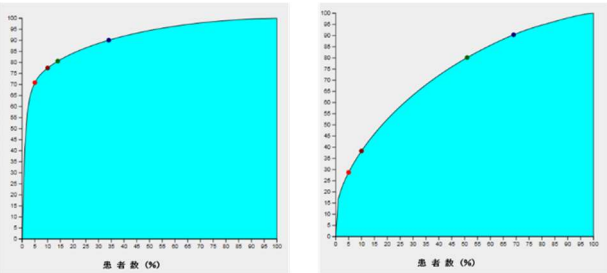


図5 医療費のCDFプロットでの表示例

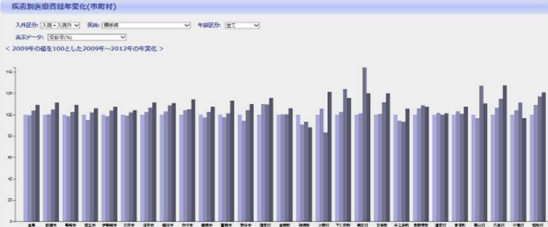


図6 市町村単位の受診率の変化

表1 視覚化方式の比較

タイプ	上段：特徴（下段：操作）
PCP	150 次元以上のデータを同時表示
	データのパターン検出 表示データのフィルタリング
SPM	疾病間の相関関係を表示
	回帰直線の表示
GIS	地理的関係の表示
	異なる空間粒度での一貫表示 Tool chip の利用
集合棒 グラフ	全体の傾向把握 時間経過の把握

6. 考察

Community Health Map⁴⁾ は、健康データ分析の専門家を

対象に、地理空間に健康管理データを視覚化したシステムである。表示対象データの動的フィルタリング機能や健康データと人口データの結合機能などを提供している。本システムでは、地理空間上での関連データの表示をマウスの動き追従して動的に表示する機能を実現することで、利用者が容易に関連データを得られるインターフェイスを実現している。

PopSciGrid⁵⁾ は、喫煙予防データを対象に、健康情報ポータルの開発を試みたシステムである。本システムは、生活習慣病の情報提供を重視したシステムである。これにより、保健師の役割の一つである地域の生活習慣病の患者への支援に貢献すると考えられる。

本システムを操作した保健師からは、複数の種類のデータを容易に把握できるとの評価があった。さらに、死亡率等のデータも、表示して欲しいとの要望もあった。

7. 結論

本システムは、医療費全体の把握と共に、年齢範囲間、疾病間、地域間、及び、異なる空間的階層との間での比較を視覚的に行うことができる。具体的には、若年者と高齢者の指標の相違や疾病毎の医療費の分布傾向の把握などが容易になる。保健師の地域における生活習慣病の予防活動などにも貢献できると考える。しかしながら、市町村単位の情報視覚化においては、対象データの該当人数を確認して、個人情報保護の配慮が必要になることがわかった。

参考文献

1) David Gots, David Borland. Data-Driven Healthcare: Challenges and Opportunities for Interactive Visualization. IEEE Computer Graphics and Applications 2016 : 90-96.

2) Robert Spence. Information Visualization: Design for Interaction (Second Edition). Pearson Education 2007.

3) 満武巨裕. 日本のレセプト情報・特定健診等データベース (NDB)の有効利用. 情報処理 2015 ; 56(2) : 140-144.

4) C. Tominski, P.Schulze-Wollgast. Community Health Map: A geospatial and multivariate data visualization tool for public health datasets. Government Information Quarterly 2012.

5) Deborah L. McGuinness, Tim Lebo. Towards Semantically-Enabled Next Generation Community Health Information Portals: The PopSciGrid. IEEE 46th Hawaii International Conference on System Sciences 2013.