

甲状腺疾患の薬物療法における 臨床判断支援システム用ユーザインタフェース

加田 昇[†] 吉川 正俊^{††} 濱崎 暁洋^{†††}

杉山 治^{††††} 岡本 和也^{††††} 黒田 知宏^{††††}

[†] 京都大学工学部情報学科 〒606-8501 京都府京都市左京区吉田本町

^{††} 京都大学大学院情報学研究科 〒606-8501 京都府京都市左京区吉田本町

^{†††} 北野病院 糖尿病内分泌センター 〒530-8480 大阪市北区扇町 2-4-20

^{††††} 京都大学医学部附属病院 〒606-8507 京都市左京区聖護院川原町 54

E-mail: [†]n-kada@db.soc.i.kyoto-u.ac.jp, ^{††}yoshikawa@i.kyoto-u.ac.jp,

^{†††}{hamasaki,sugiyama,kazuya,tomo}@kuhp.kyoto-u.ac.jp

あらまし 薬物療法において次に行う処方の推薦を行ってくれるアルゴリズムは、医師にとっては大きな一助となると考えられる。しかしながら、計算機によって計算された評価値が高いという理由だけで医師が処方選択等の医療行為の判断を行うには難点があると考えられる。本稿では、頻度だけでなく医師の多様性も考慮して処方の評価を行う既存のアルゴリズムを元に、その評価結果がどういった基準で計算されているのかとその処方が実際に症状を快方に向かわせやすいのかを、可視化を用いて分かりやすく示すことによって評価結果の信頼性を強化することを目的としたユーザインタフェースを提案する。本研究では、提案インタフェースを実際に医師に使ってもらうことによって評価を行った。

キーワード 医療情報システム, ユーザインタフェース

1. はじめに

近年、カルテの電子化が進んでおり [1], 膨大な量の電子カルテを蓄積したデータベースを利用して医師の支援を行う臨床判断支援システム (CDSS: Clinical Decision Support System) への関心が高まっている [2] [3]. 本研究では、薬物療法における CDSS に着目した。薬物療法における CDSS では、電子カルテのデータベースから入力患者と似た症例の患者のデータを抽出し、それを元に独自のアルゴリズムを用いて各処方の評価を行い、その評価値が高いものから順に推薦が行われるのが一般的である。しかしながら、そのシステムのユーザとなるのは患者に対して実際に医療行為を行う医師である以上、単に計算機によって計算された評価値が高いという理由だけでユーザがその処方を選んでくれることは通常無いと考えられる。アルゴリズムによって推薦される処方をユーザに納得して選んでもらうためには、その処方が信頼できるものであることを示す根拠となるデータも同時に示す必要があると考えられる。例えば、その投薬パターンを用いた他の患者においてどのくらいの割合で症状が良くなっているのかを見ることができれば、ユーザである医師は計算機が推薦する結果をただ鵜呑みにするのではなく、実際のデータを自分の目で見て、自分の判断で処方を選択することができる。そのためには、データを適切に集約し、効果的に可視化を行うことが重要だと考えられる。

本研究では、投薬パターンの推薦の根拠となる情報の提供を目的とし、森田らによる医師の多様性も考慮した投薬パターン評価アルゴリズム [4] を利用したユーザインタフェースの設計・

開発を行う。このアルゴリズムでは、頻度に基づく評価値として処方患者割合を表す Support, 医師の多様性に基づく評価値として処方医師割合を表す D-rate と医師ごとの処方患者割合の滑らかさを表す D-disp が用いられている。このアルゴリズム自体は様々な疾患に適応できる汎用的なものであるが、現時点では甲状腺疾患のみを対象としているため、本研究においても甲状腺疾患に焦点を定める。

評価結果の根拠情報としては様々なものが考えられるが、本研究では評価結果の理解を促すデータと他の患者において同じ処方を選んだ場合の検査値のデータの二つを用いることとした。評価値の理解を促すデータには評価値の計算に用いた元データとなる、医師ごとの処方患者数グラフや処方患者割合グラフを用いた。ユーザは処方患者数の総和を見ることで Support の大きさを、医師の人数を見ることで D-rate の大きさを、処方患者割合のグラフの滑らかを見ることで D-disp の大きさを理解することができるので、結果として数値を見ずともグラフの形を見ることで直感的に評価値の大きさを理解することができる。他の患者において同じ処方を選んだ場合の検査値のデータには処方日前後における検査値の変遷を折れ線グラフで表したものをを用いた。検査値が良くなったか悪くなったか、もしくはそのどちらでもないかによって変遷の色分けを行なったため、その処方がどれだけの患者の症状を改善させたのかをユーザに一目で理解してもらえると期待される。また、さらなる判断材料としてもらうために、医師ごとの処方患者数グラフや処方患者割合グラフをクリックすることによって、その医師が処方した場合の検査値の変遷がハイライトされる機能も

加えることとした。以上の設計案を固めた後、Web アプリの形式でユーザインタフェースの実装を行なった。

実装後には、実際に専門医の方々にユーザインタフェースを使ってもらって意見を伺う評価実験を行なった。実験を通して専門医独自の考え方について知ることができ、設計側からは見えなかった意見を収集することができた。

2. 関連研究

2.1 可視化の技術

可視化の手法には様々なものがあるが、数値データやカテゴリデータ等データの形によって相性の良い可視化の手法は異なるので、ユーザに分かりやすく伝えるためには適した手法を選択する必要がある [5]。まず、データの分類には大きく分けて三つあり、個数や長さ等といった数値の形で表される定量的データ、日付やランキング等といった大小関係や順序関係で表される順序データ、個体名や種族名といった名称や分類で表されるカテゴリデータに分けることができる。定量的データは大きさが存在し、二つのデータの差を計算することがあるので、図形の長さや数、座標平面上での位置などで可視化を行うのが相性が良いとされている。順序データは定量的データと同じく大きさが存在するが、二つのデータの差を計算することはないため、図形の色の強度や数、座標平面上での位置などで可視化を行うのが相性が良いとされている。カテゴリデータには大きさは存在せず、データ同士の区別がつけばいいため、図形の色相や追加マーク、形状などで可視化を行うのが相性が良いとされている。

2.2 医療用システムの課題

電子カルテの導入には「医療の質の向上」や「医療サービスの向上」といった利点がある一方で、「セキュリティ」や「医師の負担」といった課題が指摘されている [1]。これらの課題は電子カルテを利用したシステムを設計する際にも関わってくる課題であり、患者情報のプライバシー保護やシステムによる医師への負担の軽減は医療用システムを設計する際に重視すべき課題と言える。また、操作ミスや誤動作をしてしまうと医療ミスに繋がるためシステムの使いやすさも重視すべき課題とされている [2]。使いやすいシステムの設計のためには、はっきりとしたボタンを使うことやダブルクリックしてしまった場合に予期せぬ結果に繋がらないようにボタンの位置を調節することなどが挙げられる [6]。

2.3 がん診断支援システム Watson

既にアメリカやカナダで活用されている CDSS として、“Watson for Oncology” [7] が挙げられる。これはがん専門医の支援を目的とした AI で、患者の医療記録の分析を行うことで各治療法の評価を行い、その評価値が高い順にランキング形式で治療法の推薦を行う。このシステムの特徴としては、その治療法を推薦する根拠情報として、その治療法に関連する記述のある文献のリストを提供していることが挙げられる。信頼できる文献はかなり強力な根拠だが、文献のマイニングが必要なので結果が出るまでにかなりの時間がかかってしまうことや、文献の内容を確認するのに時間も必要としてしまうことといった欠

点が存在する。また、文献の内容を確認する際、専門的な知識が必要となることも考えられる。これらの特徴は、命にかかわる疾患であるため慎重な判断と専門医による診察を必要とするがんを対象とすることを考えた場合、さしたる問題とはならない。しかしながら、本研究で対象とする甲状腺疾患は一般的で患者数も多い疾患のため、症状が重くない限り診察にそこまで時間をかける必要性が薄いことと専門医以外が診る場合が多いことの二点から、信頼できる文献を根拠情報とするのは相性が良くないと考えられる。これらのことから、Watson とは違って専門医以外の医師でも簡単に推薦の根拠が理解できるようなシステムを考案する必要があるといえる。

3. ユーザインタフェースの設計と実装

3.1 評価アルゴリズム

評価アルゴリズムには森田らによって提案された医師の多様性も考慮した投薬パターン評価アルゴリズムを活用することとした。このアルゴリズムには既存研究では頻度のみを考慮して評価を行っていたところを医師の多様性も考慮しているという特徴があり、その評価結果は頻度に基づく Support と医師の多様性に基づく D-rate, D-disp の三つの評価値で表される。投薬パターン α に関する Support, D-rate の値は次のようにして計算される。

$$Support(\alpha) = \frac{\text{入力患者と似た症例の時に}\alpha\text{を選んだ患者数}}{\text{全患者数}}$$

$$D_{rate}(\alpha) = \frac{\text{入力患者と似た症例の患者に}\alpha\text{を選んだ医師数}}{\text{全医師数}}$$

すなわち、Support はその投薬パターンを処方された患者の割合を表し、D-rate はその投薬パターンを処方した医師割合を表している。D-disp の値は、次の三つの値 d_i , $\bar{d}$, V_{UB} を用いることで次のようにして計算される。なお、 n は全医師数を表し、医師には 1 から n の番号が振られているものとする。

$$d_i = \frac{\text{入力患者と似た症例の時に医師 } i \text{ が}\alpha\text{を選んだ患者数}}{\text{医師 } i \text{ が処方した全患者数}}$$

$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n}$$

$$V_{UB}(\alpha) = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2$$

$$D_{disp}(\alpha) = \frac{1}{V_{UB}(\alpha)}$$

すなわち、D-disp は、その投薬パターンを処方した患者割合を医師ごとに計算した d_i の分散である V_{UB} の逆数ということになる。最終的な結果となるランキングでは Support, D-rate, D-disp の三つの値の単純な積の大きさによって順位付けが行われるため、その投薬パターンを用いた患者割合と医師割合が高く、医師による偏りのない投薬パターンほど高いランクに位置することとなる。

このようにアルゴリズム自体はどのような疾患にも適応できる汎用的なものになっているが、現時点では甲状腺疾患に絞っ

て研究が進められている。甲状腺疾患が選ばれた理由としては、患者数が多いが専門医が少ない疾患であるため症状がさほど深刻でない限り専門医以外の医師が診察を行うことが多く、CDSS に対する需要が高いことが挙げられる。また、甲状腺疾患の診察の際には症状や体重などといった身体的特徴はあまり考慮されず、主に検査値だけを参考に診察が行われるため、システムの扱いやすいという理由もある。アルゴリズムでは、入力患者と類似した状況のカルテの抽出の際にこの特徴が用いられており、前回の処方薬が同じであること以外に、TSH(甲状腺刺激ホルモン)、FT3(トリヨードサイクロニン)、FT4(サイクロキシン)の三つの検査値の正常異常の組合せが同じであるかどうかによって類似性の判定を行なっている。

3.2 根拠情報として利用するデータ

1. で述べたように、本研究の目的は、処方ごとの評価結果とともにその処方を選ぶ根拠となる情報の提供を行うユーザインタフェースを設計することである。したがって、根拠情報として何を見せるのかが争点となってくる。その際に気をつけるべきこととしては、ユーザに自分で判断して処方を決めてもらわなくてはならない以上、誤った印象を与えないようになるべく実データを見せるようにした方が良いことが挙げられる。根拠情報として利用するデータとして次の二つのものを考案することとした。一つ目は、評価値の元となるデータである。3.1 節で、アルゴリズムによって出力される三つの評価値の説明を行なった。これらの評価値は確かにその処方がどれほど広く用いられているかを示すのに重要な値といえるが、評価値をただ見ただけではあまりに抽象的で、実際のデータが具体的にどういった点で高い評価を受けているのかを理解するのは難しい。特に、患者割合を医師ごとに計算した時の滑らかさを表す D-disp は 800 や 1400 といった値を取りうるのだが、この差がどれくらい滑らかさに違いを与えているのかをユーザにイメージしてもらうのは困難であろう。したがって、実際に評価値を計算するとき用いた患者数や患者割合といったデータをグラフなどで表現することで、評価結果を直感的に理解できるようになるのではないかと期待できる。二つ目は、その処方を選んだ他の患者における検査値の変遷である。1. で述べたように、その処方を選んだ他の患者において症状が良くなった患者がどのくらいの割合なのかを示すだけでも十分な根拠となると考えられるが、より詳細な情報を得ることができた方が良いだろうと考え、検査値の変遷自体を表示させることとした。また、入力患者との比較を容易にするため、入力患者における前回から今回にかけての検査値の変遷もグラフの中に書き加えることとした。

3.3 評価値の元となるデータの表し方

評価値の元となるデータの表し方には、各評価値の意味を考慮して、医師ごとの処方患者数グラフ(図1)と医師ごとの処方患者割合グラフ(図2)の二つのグラフを用いることとした。この二つのグラフはどちらも棒の一本一本が医師一人一人に対応しており、一見同じグラフに見える。しかしながら、縦軸が異なっており、図1ではその医師が入力患者と似た症例の患者を診察した時にその処方を選んだ患者数が縦軸に、図2はその

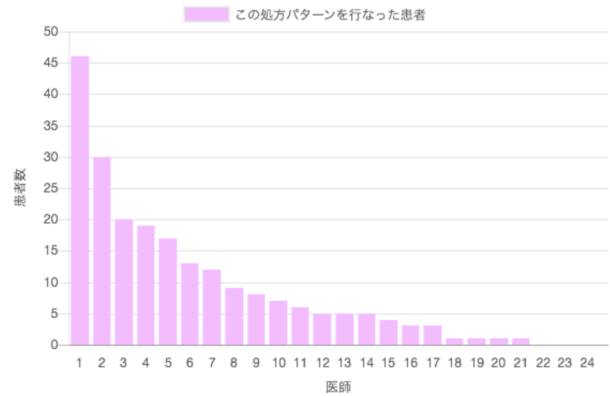


図1 医師ごとの処方患者数グラフ

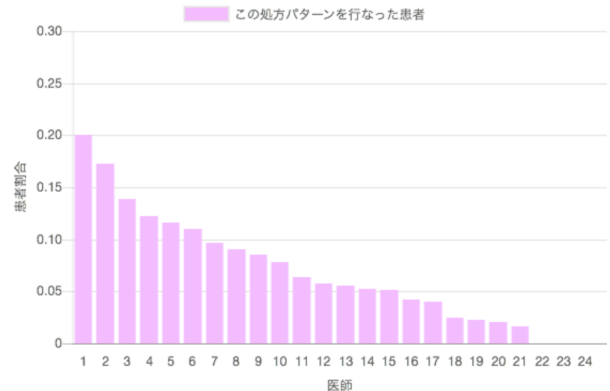


図2 医師ごとの処方患者割合グラフ

医師が入力患者と似た症例の患者を診察した時にその処方を選んだ患者の割合が縦軸になっている。このグラフを見ることでどのようにして評価値を直感的に理解できるのかについて説明する。

まず、一つ目の値である Support はその処方を受けた患者数の全患者数に対する割合を表していた。つまり、その処方を受けた患者数を全患者数で割ることで計算されるのだが、全患者数はどの処方を選んだ場合でも一定であるため、Support の値はその処方を受けた患者数のみに影響されることになる。そして、その処方を受けた患者数は図1のバーの総和で計算することができる。すなわち、Support の大きさは図1のバーの総和を見ることで理解することができる。

続いて、二つ目の値である D-rate はその処方を行った医師数の全医師数に対する割合を表していた。すると、Support の時と同じように D-rate の値はその処方を行った医師数のみに影響されることが分かる。その処方を行った医師数はバーの本数と同じなので、D-rate の大きさは図1と図2のどちらでもバーの本数を見ることで理解することができる。

最後に、三つ目の値である D-disp はその処方を行った患者の割合を医師ごとに計算した時の偏りの逆数を表していた。医師ごとの患者の割合をグラフに示したものが図2だったので、D-disp の大きさは図2の山のなだらかさを見ることで理解することができる。

以上の理由から、図1を見ることで Support と D-rate の大きさが、図2を見ることで D-rate と D-disp の大きさが分かる

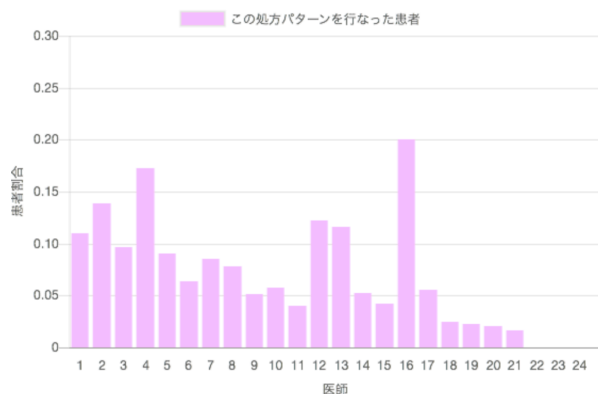


図 3 医師ごとの処方患者割合グラフ (処方患者数順ソート)

ことになる。

図 1 と図 2 では評価値との関係性の説明の分かりやすくするために大きい順にソートしたグラフを紹介したが、このままだと多くの患者にこの処方パターンを用いている医師は患者割合も大きいという間違っただけの印象を与えてしまう可能性がある。その対策として、図 1 での医師の並び順、すなわちバーの並び順を変えずに図 2 のように縦軸で患者割合を表すようにした場合のグラフを作ってみると図 3 のようになる。図 1 と図 3 は医師の並び順が同じなので、同じ番号のバーの長さを見比べることにより、同じ医師における患者数と患者割合の比較ができるようになっていく。例えば、この処方パターンを使った患者数が最も多かった医師は一番目のバーで表されていることが図 1 から読み取ることができるが、医師の並び順は変えていないのでその医師における患者割合も同様に図 3 での一番目のバーで表されていることが分かる。図 3 の一番目のバーを見てみると患者割合がそれほど高くないことが分かり、このことからこの処方を選んだ患者数が最も多かった医師はそもそも担当している患者数が多かったためにこの処方を選んだ患者数も多くなったのだらうと推測することができる。このように、医師の並び順を同じにした二種類のグラフを見比べることによってより正確な情報を得ることができる場面も考えられるため、医師の並び順に関しては操作によって患者数か患者割合のどちらかによってソートするかを選べるように設計することとした。また、医師の並び順が同じであることを理解しやすいように、二種類のグラフは並べて表示するのではなく操作によって切り替えられるように設計した。以降、医師ごとの処方患者数グラフと医師ごとの処方患者割合グラフの二つをまとめて医師ごとの処方患者グラフと呼称する。

3.4 検査値の変遷の表し方

検査値の変遷の表し方には検査値変遷グラフ (図 4) を用いることとした。これは入力患者と似た症例の患者がその処方を選んだ場合の処方日前後の検査値の変遷のグラフとなっている。3.1 節で述べたようにアルゴリズム上では検査値には三つの値を用いていたが、一度にそのすべてを表示すると情報量が多くなりすぎるため、セレクトボックスによる表示する検査値の切り替えで対応している。「上限」、「下限」と記されている二本の水平線はその検査値における正常値の範囲を表している。変遷

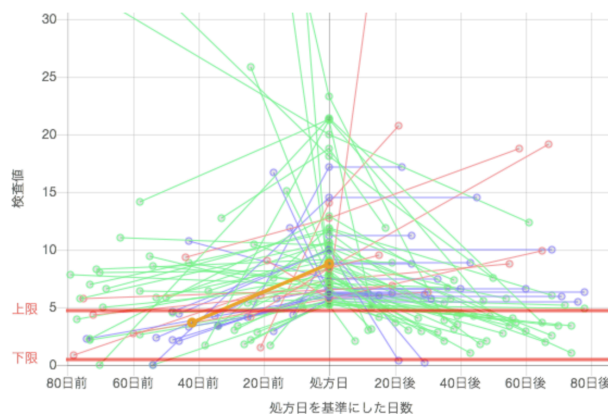


図 4 検査値変遷グラフ

を表す折れ線は処方日から次の検査日にかけての検査値の経過が良好なものは緑色、悪化しているものは赤色、そのどちらでもないものは青色と色分けがなされている。それぞれの定義を明確にすると、次のようになる。

- 経過が良好とは「処方日も次の検査日も正常値」または「処方日は異常値だったが次の検査日には正常値に近づいた」のどちらかを表す。
- 悪化とは「処方日は正常値だったが次の検査日には異常値になった」または「処方日は異常値だったが次の検査日にはさらに正常値から離れた」のどちらかを表す。
- どちらでもないとは「処方日も次の検査日も異常値で値までも同じ」または「処方日も次の検査日も異常値だったが正常値範囲を跨って変化してしまっている」のどちらかを表す。また、太い橙色の線で表されている変遷は入力患者における前回の検査日から処方日にかけての変遷を表しており、他の患者における検査値の変遷との見比べを可能としている。このグラフを用いることにより、その処方を選んだ場合に快方に向かいやすいかどうかを色の割合を見ることで理解し、同じような検査値の変遷をたどっている患者においてどのような処方が好まれているのかを知ることができると期待できる。

3.5 医師ごとの処方患者グラフと検査値変遷グラフの関係性について

評価値の元となるデータを表すグラフとして 3.3 節で紹介した医師ごとの処方患者数グラフと医師ごとの処方患者割合グラフは操作によって切り替えるようにしているため、実装するユーザインタフェースで表現されるものは医師ごとの処方患者グラフと検査値変遷グラフの合計二つとなる。単純にこれら二つのグラフを独立して表示させるのでもいいのだが、より正確な判断を医師に行ってもらうためにはシステムが複雑になりすぎない範囲でより多くの情報を提供できた方が良くと考え、医師ごとの処方患者グラフ上のバーをクリックすると図 5 のようにクリックしたバーに対応する医師が処方した場合の検査値の変遷がハイライトされるように設計を行なった。

3.6 全体的なユーザインタフェースのデザイン

上述した二つのグラフを中心に、機能切り替え用のセレクトボックスを配置することにより、全体的なデザインは図 6 のようになった。3.1 節で述べたように、元となる評価アルゴリズム

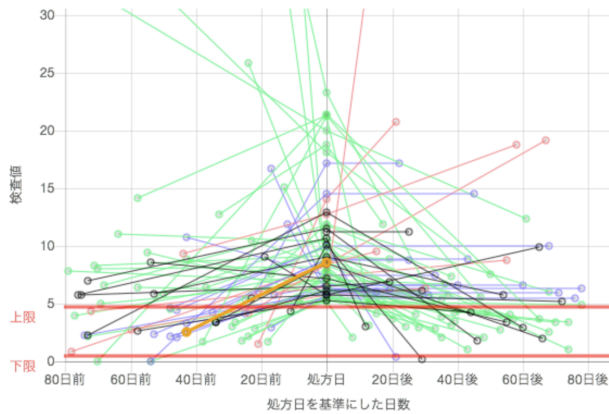


図 5 ハイライト付き検査値変遷グラフ

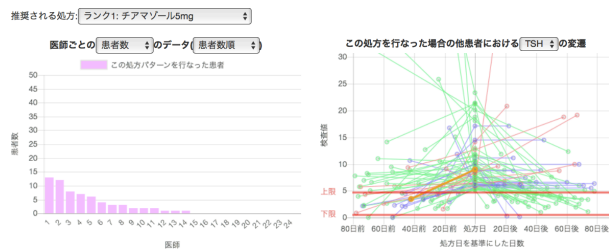


図 6 全体的なユーザインタフェースのデザイン

ムは患者の電子カルテを入力してランキング結果を出力するため、このページは電子カルテを入力した後に表示されるページで電子カルテを入力するページ自体は別個に作成することとした。そのため、「推奨される処方」の項目には受け取った電子カルテを評価アルゴリズムに入力して得られたランキングがすでに出力されており、医師がその中から自分が気になる処方を選ぶと、それに応じて下のグラフの内容が更新されるようになっている。その下の、医師ごとの処方患者グラフのタイトルの部分に位置する二つのセレクトボックスは、それぞれグラフの縦軸や医師の並び順を患者数と患者割合のどちらに対応させるかの選択肢になっており、図 6 では縦軸は患者数、並び順は患者数順となっているが、縦軸で患者割合を代わりに表したり、医師の並び順を患者割合順に並び替えることもできるようになっている。検査値変遷グラフのタイトルの部分に位置するセレクトボックスはどの検査値の変遷を表示するかを選択肢になっており、TSH の他に FT3、FT4 の中から自由に選ぶことができるようになっている。

3.7 実装に用いた言語について

3.6 節のデザインを元に、javascript を用いて Web アプリの形でシステムの実装を行なった。その際に用いたライブラリを表 1 にまとめた。

4. 評価実験

評価アルゴリズム及び提案するユーザインタフェースの有用性を評価するために、専門医四名による評価実験を行なった。本章ではその実験の方法とその結果を記す。

表 1 使用したライブラリ

ライブラリ名	バージョン	URL
chart.min.js	2.1.6	http://www.chartjs.org
bootstrap.min.js	3.3.6	http://getbootstrap.com
jquery.js	2.2.4	http://jquery.com

前回の処方: 前回の検査日: 日前
 前回のTSH: 今回のTSH:
 前回のFT3: 今回のFT3:
 前回のFT4: 今回のFT4:

図 7 患者情報の入力フォーム

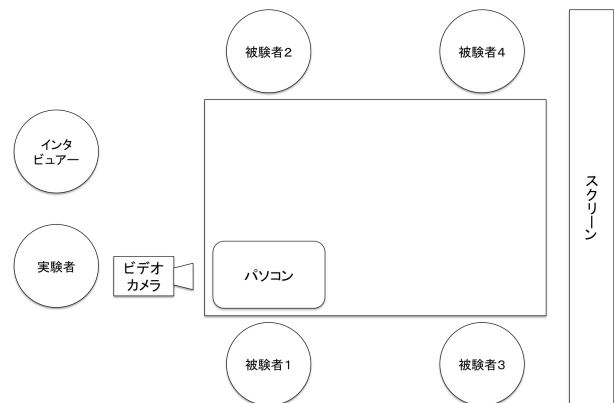


図 8 実験環境

4.1 実験内容

4.1.1 実験対象

実験は、糖尿病・内分泌領域の専門医診療に従事している医師四名に参加してもらい、提案するユーザインタフェースに対する評価を対話形式で行なった。参加した医師のうち一名は特に甲状腺疾患の専門診療に携わる甲状腺専門医であった。

4.1.2 実験に用いたデータ

京大病院の甲状腺疾患の患者カルテのデータベースから抽出してきた 28098 回分の投薬データと検査データを使用した。本来インタフェースに用いるデータは、対応する患者の ID に基づき、電子カルテシステムから抽出し、オンタイムで医師に示すものであるが、個人情報の観点からデータを匿名化する必要があったため、図 7 に示すように必要な情報だけを入力する入力フォームを用いた患者情報の検索を行なった。

4.1.3 実験環境

図 8 のようにパソコン一台と大型スクリーン一台、ビデオカメラ一台を設置した実験室を用意した。パソコン上にはシステムを実装した Web ページがすでに開いてあり、ソフトを用いてその画面の録画を行うようにした。大型スクリーンには、パソコンを操作している人以外にもパソコン画面が見えやすいようにパソコンと同じ画面を常に表示させるようにした。ビデオカメラは、図 9 のように操作中の手元の様子撮影と周囲の声などの録音のために用いた。

4.1.4 実験方法

システムと実験内容の説明を行なった後、四人分の患者情報

表 2 患者情報の組合せ

患者番号	前回の検査日	前回の処方	TSH
1	42 日前	チアマゾール 5mg	正常値
2	57 日前	チアマゾール 5mg	異常値 (High)
3	72 日前	チアマゾール 5mg	異常値 (Low)
4	65 日前	チアマゾール 10mg	正常値



図 9 ビデオカメラで手元の様子を撮影したもの

を記載した紙を被験者全員に配布した。これらの患者情報にはシステムを実際に操作してもらう際に入力する必要のある情報が含まれており、被験者それぞれに一人ずつ割り振りを行なった。被験者には自分に割り振られた患者に対してどういった処方を行うのが適切かをシステムを通して判断してもらった。操作中の被験者には操作中に考えていることや操作の意図などを常に口頭で話すようにしてもらい、操作を行っていない被験者にはその様子を観察して、気になったことや意見などがあれば自由に発言してもらった。これらの議論は全てビデオカメラにより録画した。また、意見を引き出すためにインタビューを用意し、合間合間に質問を挟んでもらうようにした。

なお、使用した患者情報は実際の患者データを参考にして作った架空のデータであり、次のような形式で表現した。

患者 1 (仮名: 田中 幸江)

前回の処方: チアマゾール 5mg

前回の検査日: 42 日前

前回の TSH: 3.09(正常) 今回の TSH: 3.82(正常)

前回の FT3: 2.54(正常) 今回の FT3: 2.78(正常)

前回の FT4: 0.978(正常) 今回の FT4: 1.33(正常)

様々なパターンの患者について実験を行うため、前回の処方や前回の検査日、今回の TSH の値が表 2 のような組合せとなるように患者情報の設定を行なった。なお、前回の検査値や今回の TSH 以外の検査値の値は全て正常値とした。

また、実験終了後に個別に意見を伺うために次の四つの項目について自由記述の形式でアンケートを行った。

(1) システムを使っていて、使いにくく感じた部分をお答えください。

(2) システムを使っていて、こういったデータもしくは機能がある方が便利だろうと感じたものはありますか。

(3) 二つのグラフを通してランキングの結果に納得することはできましたか。

(4) その他、ご意見ご感想など自由にお書きください。

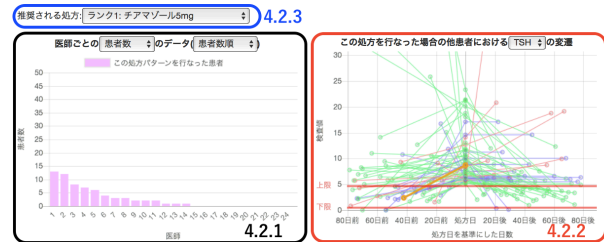


図 10 システムの部分ごとの節の分け方

4.2 実験結果

実験を通して、システムに対する専門医目線での様々な意見を収集することができた。本節では、インタフェース中の部分ごとに意見をまとめ、それに関する考察とどのように改善すべきかを述べる。それぞれの節がシステム中のどの部分に対応しているかについては、図 10 にまとめた通りとなっている。

4.2.1 医師ごとの処方患者グラフ

医師ごとの処方患者数グラフについて出た意見をまとめると次のようになる。

(1) 患者割合の分母である「その医師が処方を行なった全患者数」は、「入力患者と似た状態の時に診察した全患者数」に絞った方が良いのではないかな。

(2) 単なる棒グラフではなく、帯グラフや積み上げグラフにした方がランキング間の比較がしやすいのではないかな。

(3) ランキングや検査値変遷グラフに注視してしまうので、このグラフを活用する機会がないのではないかな。

(1) は医師にとって非常に重要な問題だったようで、(3) のように医師がこのグラフを活用しようとしなない理由の一つになってしまっていたとも考えられる。確実に改善が必要な部分といえるだろう。

(2) については、いくつかの問題点があり実装を見送った。一つ目は、400 人近くを診察した医師がいる一方で 15 人程度しか診察していない医師がいるように、医師間で処方を行なったことのある患者数に開きがあるため、グラフが潰れかねないというものである。しかしながら、この問題点はズーム機能などを取り付ければ解決できるため、さほど大きな問題点ではない。二つ目は、患者・前回の処方・検査値の変遷の全てが同じであっても異なる処方を行う場合があるため、ランキングに跨って同じ患者を数えてしまう場合があるというものである。まったく同じ状況であっても異なる処方を行う理由としては、同じ処方を続けても変化がなかったで別の処方を試してみるといったもの、実際の症状や前回よりさらに前からの検査値の変遷といったアルゴリズム上では考慮されていないファクタを考慮した結果といったものが考えられる。患者数を数えていく際にランクが異なる処方と同じ患者を何度も数えていってしまうと、例えば患者割合を積み上げグラフで表現すると積み上げた結果が 100% を超えてしまうといった事態が発生し、ユーザに混乱を招く危険性がある。そのため、今回は積み上げグラフではなく単純な棒グラフを利用し、異なるランキング間の比較はグラフの切り替えによって行うようにした。帯グラフや積み上げグラフの採用のためには、同じ対象を何度も数えてしまわ

ないように患者数ではなく処方数を表示するといった工夫が必要となるだろう。

(3) は、実験に用いたデータが悪かったために生じた意見ともいえる。元にした投薬パターン評価アルゴリズムは、単純に処方患者数だけを考慮する先行研究のアルゴリズムとは違って医師の多様性も考慮することによって特定の医師だけがたくさん行なっているような処方はあまり強く推薦しないようにするといった特徴があった。そのため、このグラフには医師ごとの処方患者割合に偏りが見られた場合に注意喚起するという目的もあるのだが、今回の実験に用いたデータにはそういった偏りのあるデータは存在しなかった。

4.2.2 検査値変遷グラフ

検査値変遷グラフに関して出てきた意見をまとめると次のようになる。

(1) もっと前の検査値も見みたい。

(2) 良くなったか、悪くなったか、そのどちらでもないかで変遷を色付けするだけでなく、それぞれの割合を具体的な数値で示して欲しい。

(3) 良くなった変遷には目立った色をつけてもらいたい。

(4) ズーム機能や要約機能を取り入れて欲しい。

(1) については、元となる評価アルゴリズムでは処方日とその前の検査日での検査値のみを考慮していたため、それに合わせる形で前の処方のみを表示することになったという経緯が存在する。しかしながら、評価アルゴリズムの側には前々回の検査日まで考慮し始めるとデータ数が極端に少なくなってしまうというやむを得ない事情があるのに対し、ユーザインタフェース側には前々回の検査日まで遡って表示することによって生じる不都合はほとんど存在しない。グラフが見えにくくなるといった問題は生じるだろうが、(4) で触れるようにズーム機能を取り入れれば十分解決可能である。そのため、この意見に関しては素直に取り入れるべきと思われる。発症時期が知りたいという意見も見られたので、その解決にも繋がるだろう。しかしながら、この意見を取り入れようとするのとどこまで遡るべきかという問題点が生じる。これに関しては、診察を行う際にカルテをどのくらい遡るのかについて専門医に調査を行って確認を行うべきであろう。

(2) については、本数だけで十分どういった割合になっているか理解できるだろうと考えていたが、(3) でも触れることになるのだが、色による目立ち度合いの差もあり具体的にどのくらいのケースで良くなっているのかを理解しにくく感じられたようであった。具体的な実装としてグラフの近くに帯グラフの形で割合を示して欲しいといった意見も見られた。その際にはやはり数値で示した方がイメージしやすいということで、帯の長さだけでなく実際の数値も書き添えて欲しいとのことだった。

(3) は、ユーザの要望と設計側の設計意図との差異から生じた意見ともいえる。デザインを考案する際に色の分け方を考えるにあたり、検査値の変遷が良好なものには安全なことを表す緑で、悪化したものには危険なことを表す赤で、どちらともいえないものには二色の中間的な色として青で表すこととした。これには、検査値が良くならなかった変遷に印象の強い色を配

置ることによって注意喚起を行うといった意図も存在していた。しかしながら、ユーザである医師からすると悪化したものは何かしらの例外が起こったのだと認識し、どれくらいの患者が良くなったのかの方が重要に感じるということだった。そのため、良くなった変遷に対しては黒や青といったもっと目立った色にして欲しいというご指摘を受けた。とはいえ、悪くなったものを完全に無視してしまうわけではないということだったので、悪化した変遷は継続して赤で表すこととし、良くも悪くもならなかった変遷の青色やハイライトに用いる黒色も含めて、悪化した変遷以外の色の割り振り方を考慮し直すこととする。

(4) については、TSH が非常に不安定な値であることから正常値範囲の上限が5であるのに対して140といった非常に高い異常値を取ることにすらあったため、非常に重要な問題といえる。今回はそういった例外的な値を除いた時の最大の値までをグラフ上に表示するようにグラフの上限を設定したが、処方数が多いに多い投薬パターンだと一本一本の変遷がどうなっているのか認識することができなくなることすらあったため、ズーム機能や要約機能は優先的に実装すべきだろう。

4.2.3 ランキング

ランキングに関しては、ユーザインタフェースに関する意見と評価アルゴリズムに関する意見の二種類が見られた。そのため、ここではそれぞれを別々に分けてまとめることとする。

a) ユーザインタフェースについて

ランキングのユーザインタフェース面に関して出てきた意見をまとめると次のようになる。

(1) 異なるランクのデータを並べて表示してほしい。

(2) ランキング外の処方でのデータも見みたい。

(3) ランキングを普段折りたたむのではなく、常に全てが見えるようにして欲しい。

(1) について実装を行わなかった理由としては、検査値の変遷では表示するデータ数が非常に多いため、並べて表示させようとするとデータが細かくなってしまったといったものが挙げられる。解決案として被験者である医師の一人が提示したものに、各データを別窓で表示するというものがあったため、その方向性で改善を行うこととする。

(2) については、ランキングに載っている処方を使用した履歴が残っている以上ある程度の信頼性があると考え、特に実装を行わなかった。しかしながら、実際には医師にとっては想定外な処方が多く含まれてしまっていたようで、こういった意見が寄せられることとなってしまった。自由度を高めるためにも必要な機能といえるため、何らかの形で実装を行うこととする。

(3) については、ランキング間の比較の不便さに繋がっているように実験を通して感じられたため、(2) と合わせてランキングのより良い見せ方を練っていくこととする。

b) 評価アルゴリズムについて

投薬パターン評価アルゴリズムに関する研究は本研究の対象外ではあるが、今回の実験は評価アルゴリズム自体の有用性の評価という目的も含んでおり、また、評価アルゴリズムに変更を加えると必然的にユーザインタフェースにも変更を加えなく

てはならないため、評価アルゴリズム面での意見に関しても触れることとする。評価アルゴリズム面に関しては、専門医からすると想定外な処方があることが主に問題視されていた。想定外な処方が高く評価されている要因として考えられるものについて医師に尋ねたところ、様々なものが挙げられた。それらをまとめると次のようになる。

(1) 医師の多様性の影響が強すぎるのではないかな。

(2) 特殊な事例かどうか疾患名などで症例に絞り込みをかけるべきではないかな。

(3) 他の検査値も考慮すべきではないかな。

(4) 発症からの期間も考慮すべきではないかな。

(5) もっと前の処方や検査値も考慮すべきではないかな。

(1) については、今回用いた評価アルゴリズムの特徴である医師の多様性の要素が強く影響してしまっているのではないかと考えられていた。評価結果における医師の多様性の要素とは、その処方を選んだ医師の割合である D-rate と、医師ごとに処方患者割合を計算した時のなだらかさである D-disp であった。専門医にとって想定外とされた処方では、その処方を行なった医師数はさほど多くなかったため D-disp の要素が強く影響を及ぼしていたのだと考えられる。D-disp の値に重みをつけたり、アルゴリズムの計算手法を見直してみるなどして、専門医にとって評価結果が出力されるように調節する必要性があると考えられる。

(2) については、甲状腺疾患は様々な疾患の総称であるため、今回用いた評価アルゴリズムではそれらをひとまとめにして扱ってしまっているのが影響しているのではないかと考えられていた。今回用いた評価アルゴリズムでは、TSH と FT3、FT4 の三つの検査値と前回の処方薬のみを参考にするため、疾患名の考慮まではなされていなかった。他の検査値などを考慮するなどして、特別な症例は弾くようにすることが必要だと考えられる。

(3) については、具体的な検査値名として TSH(甲状腺刺激ホルモン) レセプター抗体が挙げられていた。TSH レセプター抗体はバセドウ病の判定に使用される検査値であり、(2) の解決のために利用することができる。あまりに多くの検査値を考慮してしまうとデータ数が極端に少なくなり、正確な結果が得られないといった事態が発生することも考えられるため慎重な判断が必要だろう。

(4) については、投薬を長く続けてきて傾向が把握できてきた患者と、ごく最近に発症して様々な処方を試している段階の患者を区別していないのが影響しているのではないかと考えられていた。これに関しては、(2) 同様、抽出の段階で絞り込みを行うのも良いと考えられるが、データ数が足りなくなると正確な結果が得られない可能性もある。アルゴリズムで評価を行う際には絞り込みは行わず、検査値変遷グラフにおいて発症からの期間を表示するといった解決策が考えられる。

(5) については、処方日前後の検査値や処方のみを考慮しているのが影響しているのではないかと考えられていた。4.2.1 節でも触れたように、同じ処方を続けても変化が見られない場合など、前回より前の処方が影響を与えている状況というのは

様々なものが考えられる。診断の際にはカルテをなるべく遡るという発言を被験者からいただいてもいるため、二回分の情報だけで処方を決めてしまうのは危険と考えられる。しかしながら、何度も触れているように抽出の条件を厳しくすることはデータ数の縮小化にも繋がってしまう。なるべく患者データの抽出条件は変えないままに専門医の想定に沿うよう評価結果を調節する必要があるだろう。

5. おわりに

5.1 ま と め

本研究では、既存の投薬パターン評価アルゴリズムを元にその評価結果の根拠情報を提供することを目的としたユーザインタフェースの設計と Web アプリの形式での実装を行なった。評価結果の根拠情報には、評価結果を直感的に理解してもらうことを目的とした医師ごとの処方患者グラフとその処方を選んだ他の患者における検査値の変遷を折れ線グラフで表した検査値変遷グラフを用いた。また、実装したシステムを専門医に操作していただき、システムに関する意見を伺う評価実験も行なった。検査値変遷グラフに関してはいくつか拡張案が挙げられていただけて概ね良評価が得られていた。しかしながら、医師ごとの処方患者グラフについては、これを見ただけではランキング結果に納得できないといった意見があり、また、そもそもの評価アルゴリズムに大幅な変更が予想されるため、大幅な見直しが必要だと考えられる。

5.2 今後の課題

まず、実験中で得られた改善点に関して実装を行う。また、実験を通して専門医独自の考え方について知ることができた部分もあるため、それを参考にさらなる改善の余地がないかどうかシステムの見直しを行う。

文 献

- [1] 奥山文雄, 橋本則男, 河村徹郎. “電子カルテの動向と課題”. Medical Imaging Technology, Vol.25, No.3, p.143, 2007.
- [2] 井口竜太, 佐藤元, 中村謙介, 松原全宏, 軍神正隆, 石井健, 中島勸, 矢作直樹. “米国の救急外来における電子カルテシステムと臨床診断意思決定支援システム”. 保健医療科学, Vol.62, No.1, p.88-97, 2013.
- [3] Amit X Garg, Neill KJ Adhikari, Heather McDonald, M Patricia Rosas-Arellano, PJ Devereaux, Joseph Beyene, Justina Sam, and R Brian Haynes. “Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review”. Jama, Vol.293, No.10, p.1223-1238, 2005.
- [4] 森田祐司. “投薬歴の構築と石の多様性を考慮した投薬パターンマイニング”. 修士論文, 京都大学. 2017.
- [5] Riccard Mazza. “Introduction to Information Visualization”. Springer, 2009. (中本 浩訳. “情報を見える形にする技術”. 株式会社ボーンデジタル, 2011.)
- [6] 山野辺祐二, 相澤志優, 本多正幸. “電子カルテシステム GUI の問題点”. IT ヘルスケア, Vol.2, No.1, p.28-31, 2007.
- [7] IBM. “IBM Watson for Oncology”. <http://www.ibm.com/watson/health/oncology/>. 2017 年 1 月 16 日アクセス.