

期間情報を用いた投薬パターンマイニング

杉山 祐一¹, 吉川 正俊¹, 岡本 和也², 黒田 知宏², and 濱崎 暁洋³

¹ 京都大学大学院情報学研究科 社会情報学専攻

² 京都大学医学部附属病院 医療情報企画部

³ 公益財団法人 田附興風会 医学研究所 北野病院 糖尿病内分泌センター

1. 序論

医療現場への電子カルテの導入によって、莫大な医療データの蓄積が行われている。それに伴い、電子カルテの二次利用による医学研究の支援が期待されている。

カルテの二次利用の形態として、日常臨床における症例経過や治療の特徴を抽出し、臨床的疑問を後ろ向きに検証する臨床研究が以前より行われてきている。電子カルテの導入によって、膨大なカルテデータからのデータ抽出および整理にかかる負担の大幅な軽減を実現しうる処理手段が医療者によって求められるようになっており、こうしたデータ処理手段が臨床研究の大きな推進力となることへの期待も大きい。

後ろ向きのデータ解析にあたっては、個々の症例の身体計測値や検査値、投薬・処置といった治療行為の前後関係や期間などの時系列が整理された形での抽出を行うことが必要であるが、実際の電子カルテ上では、一般的に処方箋の情報しか記録されおらず、服用状況をそのまま把握するのは難しい。そのため、本研究では処方箋データから服用状況を再現し、データを解析に適した形で格納することにより、時系列を意識して投薬データを解析できるようにした。そのうえで、投薬歴の変化の種類を四つ定義し、さらに、どのような医薬品の追加パターンが特異であるかを抽出した。

2. 投薬パターンの検索

投薬に関する検索要求は、一般的なデータベースで要求されるような、特定の医薬品が投与されたか、一定期間投薬されたかといった、ある時点での投薬状況の検索だけではなく、それに加えて次の四つのような時間軸上での変化の情報も検索されると考えられる。

- 医薬品の切り替え
- 医薬品の追加
- 医薬品の中止

Medication Pattern Mining with Time Interval
 Yuichi SUGIYAMA, Dept. of Social Informatics, Graduate School of Informatics, Kyoto University
 Masatoshi YOSHIKAWA, Dept. of Social Informatics, Graduate School of Informatics, Kyoto University
 Kazuya OKAMOTO, Div. Medical IT & Admin. Plan., Kyoto University Hospital
 Tomohiro KURODA, Div. Medical IT & Admin. Plan., Kyoto University Hospital
 Akihiro HAMASAKI, Center for Diabetes & Endocrinology, Kitano Hospital, The Tazuke Kofukai Medical Research Institute

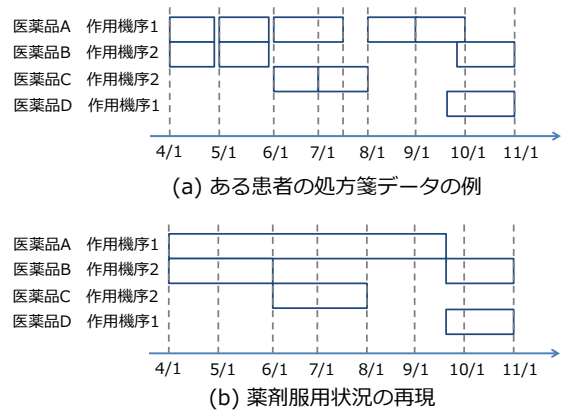


図 1: 処方箋からの薬剤服用状況の再現

● 新規の医薬品の投薬

本研究ではこの四種類の検索を区別できるように処方箋データから実際の服用状況を再現し、データマートを設計する。

3. データマート

データマートは、あらかじめデータを集計・統合することによりデータ分析効率の向上を図るものである。電子カルテ上で各患者が ID によって識別され、ある患者への処方箋データが図 1(a) のように記録されているとする。本研究では、まず処方箋データを図 1(b) のように薬剤服用状況に再現する。なお同じ作用機序(効果)をもつ医薬品は同時に服用されることがないと仮定をおく。

そして、2 節で述べたような検索要求を実現するために、投薬パターンが変化する日(投薬の開始日および終了日)に着目した。2 節の検索要求は、投薬パターンが変化した日にその時に投与が開始された医薬品の組み合わせを示す「追加」、その時継続して服用されている医薬品の組み合わせを示す「継続」、その時に投薬が中止された医薬品組み合わせを示す「中止」、その日時を示す「開始日」をそれぞれ付与してデータマートに記録することで実現できると考えた。表 1 は、図 1 を本研究データマートに変換した例である。

3.1. データマート作成手法

上記のデータマートの作成方法について述べる。

3.1.1. 閾値 ϵ を用いた不等号

長期にわたっての医薬品の服用状況は、単一の医薬品の処方箋が複数連続している状態であると考えられる。しかし、実際には診察のタイミングの問題で処方

表 1: 提案データマート

pID	追加	継続	中止	開始日
0001	{A,B}	∅	∅	4/1
0001	{C}	{A}	{B}	6/1
0001	∅	{A}	{C}	8/1
0001	{B,D}	∅	{A}	9/15
0001	∅	∅	{B,D}	11/1

箋が連続しないケースも存在する。そのため、処方箋から服用状況を再現するためには、処方箋の開始日・終了日を比較する際に [1] のようにマージン ϵ を用いる必要がある。マージンを用いた不等号の定義は、二つの時点 t_1, t_2 が存在するとき、

$$t_1 =^\epsilon t_2 \iff |t_1 - t_2| \leq \epsilon, t_1 <^\epsilon t_2 \iff t_2 - t_1 > \epsilon$$

とし、 $>^\epsilon, \leq^\epsilon, \geq^\epsilon, \neq^\epsilon$ も同様に定義する。

3.1.2. 薬剤服用状況の再現

薬剤服用状況の再現は、処方箋を以下の条件により統合・修正し行う。なお処方箋 m に対して、処方開始日を m^+ 、終了日を m^- と表す。

- 2つの同医薬品の処方箋 m, m' が、 $m^+ \leq^\epsilon m'^+ \leq^\epsilon m^-$ であるとき、処方開始を $\min(m^+, m'^+)$ 、処方終了を $\max(m^-, m'^-)$ として2つの処方箋を統合する。
- 2つの同じ作用機序で、異なる医薬品の処方箋 m, m' が、 $m^+ \leq^\epsilon m'^+ \leq^\epsilon m^-$ であれば、 $m^- = m'^+$ とする。
- 2つの異なる作用機序の処方箋 m, m' が、 $m'^+ \leq^\epsilon m^-$ であり、処方箋 m が条件1により統合されることがない場合、 $m^- = m'^+$ とする。
- 2つの異なる作用機序の処方箋 m, m' が、 $m^+ =^\epsilon m'^+$ であり、処方箋 m が条件1により統合されることがない場合、 $m^+ = m'^+ = \min(m^+, m'^+)$ とする。
- 2つの異なる作用機序の処方箋 m, m' が、 $m^- =^\epsilon m'^-$ であり、処方箋 m が条件1により統合されることがない場合、 $m^- = m'^-$ とする。

3.2. 投薬に関する検索要求の実現方法

薬剤服用状況を求めた後、本研究で提案するデータマートへ変換する。2節で述べた投薬の変化の情報の抽出のためのクエリの実現方法をそれぞれ次に示す。

- (中止) $\neq \emptyset \wedge$ (追加) $\neq \emptyset$
 \Rightarrow 医薬品 (中止) から医薬品 (追加) への切り替え
- (継続) $\neq \emptyset \wedge$ (中止) $= \emptyset \wedge$ (追加) $\neq \emptyset$
 \Rightarrow (継続) \wedge 医薬品 (追加) の追加
- (追加) $= \emptyset$
 \Rightarrow 医薬品 (中止) の中止
- (継続) $= \emptyset \wedge$ (中止) $= \emptyset$
 \Rightarrow 医薬品 (追加) の新規投薬

4. 実装と投薬パターンマイニング

上記のデータマートを京都大学病院で2000年9月から2013年9月の糖尿病患者に発行された、一般的に糖尿病治療に用いられている27種類の医薬品の処方箋データから実装した。マージン ϵ の値は14日とした。この中から特異な投薬パターンを抽出する。

特異な投薬パターンは、投薬後の病状の変化が他の投薬パターンと異なると考えた。そのため糖尿病での検査値として代表的なHbA1cの値を用いて、特異な投薬パターンを求めることにした。投薬パターン開始時からその3か月後に検査値が0.5%以上下がっていた場合そのパターンを「低下」のパターン、変動量が-0.5%~0.5%であった場合「一定」のパターン、0.5%以上上がった場合「増加」のパターンとする。なお投薬が3か月以上連続していないものは今回は考慮しない。

実装したデータの中から「医薬品の追加」である投薬パターンを抽出したところ、709パターンあり、低下・一定・増加パターンはそれぞれ、174件、464件、71件であった。

各「医薬品の追加」である投薬パターンごとに「継続」を作用機序ごとにまとめ、低下・一定・増加パターンの件数を抽出した。その中で、 χ^2 検定で全体件数中の変化内訳から有意な差のあった投薬パターンを特異な投薬パターンとし、件数の合計が10件以上あり、 $p < 0.05$ で有意であったものを表2に示す。

表 2: 特異な「医薬品の追加」である投薬パターン

No.	追加	継続	低下	一定	増加	p 値
1	アマリール	DPP-4 阻害薬	7	2	1	0.003
2	メトグルコ	α -GI	0	14	0	0.024
3	アクトス	SU 薬	17	24	1	0.026
4	スターシス	SU 薬	10	22	9	0.035
5	グリミクロン	グリニド薬	2	7	4	0.042
追加パターン全体			174	464	71	

表2のうちNo.1とNo.3は作用機序の異なる経口血糖降下薬を追加した際のHbA1c低下作用の知見がよく再現される抽出が得られ、No.4とNo.5は薬剤機序的に通常併用を行わない組み合わせの処方方も実際には行われていることが明らかとなるとともに、そうした組み合わせでの追加効果は乏しいことが示唆された。

5. 今後の課題

投薬パターンの変化として四種類を定義し、処方箋から服用状況を再現し、特異な投薬パターンのマイニングを行った。今後、作成したデータマートの上でのマイニング手法の精査が課題として挙げられる。

参考文献

- [1] Moskovitch, Robert, and Yuval Shahar. "Medical temporal-knowledge discovery via temporal abstraction." AMIA Annual Symposium Proceedings. Vol. 2009. American Medical Informatics Association, 2009.