

一般口演 | 第40回医療情報学連合大会（第21回日本医療情報学会学術大会） | 一般口演

## 一般口演14 医療データ解析

2020年11月20日(金) 14:00 ~ 15:40 G会場 (イベントホール・特設会場2)

### [3-G-3-05] 時系列マイニングによる多量の医療データを用いた医療プロセスの可視化

\*山崎 友義<sup>1</sup>、松尾 亮輔<sup>1</sup>、小川 泰右<sup>1</sup>、荒木 賢二<sup>1</sup>、Lei HieuHanh<sup>2</sup>、横田 治夫<sup>2</sup>（1. 宮崎大学医学部附属病院 IR部, 2. 東京工業大学情報理学院）

\*Tomoyoshi Yamazaki<sup>1</sup>, Ryouyusuke Matsuo<sup>1</sup>, taisuke Ogawa<sup>1</sup>, Kenji Araki<sup>1</sup>, HieuHanh Lei<sup>2</sup>, Haruo Yokota<sup>2</sup>（1. 宮崎大学医学部附属病院 IR部, 2. 東京工業大学情報理学院）

キーワード：Sequential pattern mining, Health care process, Frequent health care pattern

目的) クリニカルパス（以下パス）は医療プロセスを可視化し、そのプロセスを標準化した医療工程表である。パス作成と改善に利用できる医療データは自施設だけであり、多量の医療データを用いていない。我々は2013年より、電子カルテや医事記録を含む膨大な医療データを用い、時系列的にデータを解析し、疾患ごとで最も基本的な医療プロセス（以下頻出医療パターン）を可視化するシステムを開発している。今回はその成果とともに、次世代医療基盤法で得られる膨大な医療データを用いて疾患ごとでの頻出医療パターンの可視化について検討した。

対象と方法) 2013年から2016年まで宮崎大学医学部附属病院（以下当院）の経尿道的膀胱腫瘍摘出術（TUR-Bt）242例を対象とした。症例ごとで入院から退院までの医療支指示を、時間情報や薬効等でシーケンシャルパターンマイニング（SPM）を実施し、医療パターンを抽出した。抽出した医療パターンを時系列で整理し、頻出医療パターンを作成した。

結果) 患者ごとで異なるパターンを時系列で頻出医療パターンを作成した。頻出医療パターンに追加されたi医療パターンがあり、このパターンをバリエーションとした。バリエーションの設定により、複数ある頻出医療パターンより、必須の医療指示を時系列的に抽出できた。当院のTUR-Bt施行の95%はパスで実施し、パス分析が行われている。パスに記載された主要な医療パターンと頻出医療パターン差異がなかった。パス分析によるバリエーションと、今回のSPM解析のバリエーションとを比較した。在院日数延長に影響するバリエーションが共通する6例であり、術後予定外の抗生剤投与であった。SPM解析では迅速と緊急の検査種別や術日の浣腸施行の有無をバリエーションとした。

結語) SPM解析による頻出医療パターン作成は、膨大なデータを用いて新たなパス作成・改善を行える可能性を示唆できた。

# 時系列マイニングによる多量の医療データを用いた医療プロセスの可視化

山崎 友義<sup>\*1</sup>、松尾 亮輔<sup>\*1</sup>、小川 泰右<sup>\*1</sup>、荒木 賢二<sup>\*1</sup>、Le Hieu Hanh<sup>\*2</sup>、横田 治夫<sup>\*2</sup>

<sup>\*1</sup> 宮崎大学医学部附属病院 IR 部、<sup>\*2</sup> 東京工業大学情報理学院

## Sequential pattern mining-based visualization of health care process using big clinical data

Tomoyoshi Yamazaki<sup>\*1</sup>, Ryosuke Matsuo<sup>\*1</sup>, Taisuke Ogawa<sup>\*1</sup>, Kenji Araki<sup>\*1</sup>, Hieu Hanh Le<sup>\*2</sup>, Haruo Yokota<sup>\*2</sup>

<sup>\*1</sup> The Institutional Research Department for Hospital Management, Faculty of Medicine, University of Miyazaki Hospital, <sup>\*2</sup> Department of Computer Science, School of Computing, Tokyo Institute of Technology

We visualized the pattern of health care process by using sequential pattern mining for each disease from big clinical data. We could extract the health care pattern from frequent health care pattern by the visualization. This suggests that it has potential to create clinical pathways from big clinical data by extracting essential health care pattern.

Keywords: Sequential pattern mining, Health care process, Frequent health care pattern

### 1. 結論

医療データが電子化され、デジタルデータの作成、収集、管理する新しい医療情報技術システムを導入する施設が増加している。その結果、施設の臨床医、管理者、研究者が医療情報技術システムで利用できるデータの量は増加し続けている。しかし、電子カルテやレセプトデータ等の異なるシステム間を包括的に利用することが困難とされてきた<sup>1)</sup>。

膨大な医療データを用いて、新たな知見や医療安全に効果的な手法を得られる研究が行われている。ただ、有効なデータ解析手法を用いないと、誤った結果を招く恐れがあると指摘されている<sup>2)</sup>。我々も次世代医療基盤法で構築される大規模医療データを利用し、網羅的な解析を行い、有効で効果的な知見を得るための手法を開発している。

今回は、開発当初より行っている入院から退院までの一定の制約で医療プロセスを時系列的に記載した臨床パス(以下パス)を利用した症例の医療データの解析による医療プロセスパターンの可視化を行っている<sup>3,4)</sup>。医療プロセスパターンを解析することで、DPC14 桁毎で必須医療プロセスと新たに追加した医療プロセス(バリエーションとした)の関係性も可視化した。

### 2. 目的

パスは医療者の頭の中にある医療プロセスを可視化し、そのプロセスを標準化した医療工程表である。パス作成と改善に利用できる医療データは自施設だけであり、多量の医療データを用いていない。我々は 2013 年より、電子カルテや医事記録を含む膨大な医療データを用い、時系列的にデータを解析し、疾患ごとに最も基本的な医療プロセス(以下頻出医療パターン)を可視化するシステムを開発している。今回はその成果とともに、次世代医療基盤法で得られる膨大な医療データを用いて疾患ごとの頻出医療パターンの可視化について検討した。

### 3. 方法

2013 年から 2016 年まで宮崎大学医学部附属病院(以下当院)の経尿道的膀胱腫瘍摘出術(TUR-Bt) 242 例を対象とした。症例毎のデータを表1の方法で、人工知能 DWH 構築と同様なデータ処理を行った。処理したデータを用いて入院から退院までの医療指示をシーケンシャルとし、時間情報や薬効等でシーケンシャルパターンマイニング(SPM)を実施し、医療パターンを時間間隔ごとで抽出した。各医療指示間の時間は固定でなく、医療指示間の時間間隔を統計的に処

理した可変時間間隔でおこなった。抽出した医療パターンを時系列で整理し、頻出医療パターンを作成した。

表 1 データ処理方法

欠損値保管	例：入院時身長は、前後 6 ヶ月以内の身長を代用する
ペイロード型置換	縦持ちデータの項目をフィールドに変換
日次まるめ・基準日設定	入院日、手術日等を基準日とする
相対日付	基準日に沿って、日付を相対日付に変換
粒度調整	術式、DPC、薬剤、材料、検査項目を適切な粒度別に階層化
離散化	例：体温を発熱あり/なしに変換

### 4. 結果

患者ごとで異なるパターンを時系列で頻出医療パターンを作成し図 1 に示す。頻出医療パターンに追加された医療パターンがあり、このパターンをバリエーションとした(図 2)。バリエーションの設定により、複数ある頻出医療パターンより、必須の医療指示を時系列的に抽出できた。



図 1 頻出医療パターン抽出結果

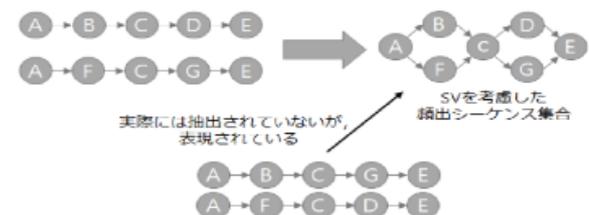


図 2 医療指示の時系列表示からシーケンシャルバリエーション(SV)の表示

## 5. 考察

従来の SPM は医療指示等を順列シーケンスとし、パターンを抽出していた<sup>5)</sup>。医療では医療指示の時間間隔は重要であり、時間間隔を加味した SPM が必要であった。時間間隔を考慮した SPM も開発されている<sup>6)</sup>。しかし、時間間隔は固定であり、医療用 SPM には不向きである。我々は時間間隔を固定しない SPM を開発し<sup>7)</sup>、今回の研究に用いることで時系列的に頻出医療パターンを抽出することが可能になった。

当院の TUR-Bt 施行の 95% はパスで実施し、パス分析が行われている。パスに記載された主要な医療パターンと頻出医療パターン差異がなかった。パス分析によるバリエーションと、今回の SPM 解析のバリエーションとを比較した。在院日数延長に影響するバリエーションが共通する 6 例であり、術後予定外の抗生剤投与であった。SPM 解析では迅速と緊急の検査種別や術日の浣腸施行の有無をバリエーションとした。検査種別は医師の術前、術日の検査を緊急検査でオーダーすることが慣行であった。浣腸は術当日の朝に排便の有無で浣腸することがパス作成記録で確認されていた。このことから、これらは医療的にバリエーションとしなかった。

## 6. 結語

SPM 解析による頻出医療パターン作成は、膨大なデータを用いて新たなパス作成・改善を行える可能性を示唆できた。

## 参考文献

- 1) Jesus J Caban, David Gotz : Visual analytics in healthcare - opportunities and research Challenges, Journal of the American Medical Informatics Association, 2015 : Vol. 22, Issue 2 ; 260-262
- 2) J Archenaa, E.A Mary Anita : A Survey of Big Data Analytics in Healthcare and Government, Procedia Computer Science , 2015 ; Vol. 50 : 408-413
- 3) 保坂智之, 浦垣啓志郎, 荒堀喜貴, 串間宗夫, 山崎友義, 荒木賢二, 横田治夫. 医療履歴の時系列解析におけるシーケンス間類似度評価による時間間隔調整の導入 DEIM Forum 2016. 2016 ; G7-5
- 4) 本田 祐一, 串間宗夫, 山崎友義, 荒木賢二, 横田治夫. 典型的医療指示シーケンスの共通部分検出とその可視化. 第9回データ工学と情報マネジメントに関するフォーラム論文集, 11-1, 2017.3.
- 5) R. Agrawal, R. Srikant. Fast algorithms for mining association rules in large databases. Proceeding of the 20th International Conference on Very Large Data Bases, 1994; 487-499
- 6) J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, M. Hsu. PrexSpan: Mining sequential patterns efficiently by prex-projected pattern growth. Proceeding of 2001 international conference on data engineering, 2001; 215-224
- 7) Keishiro Uragaki, Tomoyuki Hosaka, Yoshitaka Arahori, Muneo Kushima, Tomoyoshi Yamazaki, Kenji Araki, Haruo Yokota : Sequential Pattern Mining on Electronic Medical Records with Handling Time Intervals and the Efficacy of Medicines , First IEEE Workshop on ICT Solutions for Health, Proc. of the 21st IEEE International Symposium on Computers and Communications (ISCC 2016), 2016 ; 20-25