

治療クラスタ遷移と医学的判断の紐づけによる敗血症治療ナビゲーションシステム実現可能性の検証

Treatment strategy of Septicemia based on DPC data analysis by using pLSA method

山下 和也 *¹ 阪本 雄一郎 *² 櫻井 瑛一 *¹*² 本村 陽一 *¹*²
 Kazuya Yamashita Yuichiro Sakamoto Eiichi Sakurai Yoichi Motomura

*¹ 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology

*² 佐賀大学

Saga University

This paper shows our study about treatment strategy of septicemia based on DPC data by using pLSA, probabilistic latent semantic analysis, which find hidden segments in data. Our aim of using this method is modeling patients' change along with time. In this study, we estimate segments from the data that consists of times when each patient had some medical treatment and name of treatments. We compared each segments from the points of death rate, medical cost and patients' hospitalization. We find some patients' movement pattern between segments. We also find that these movements show the change of treatment strategy.

1. はじめに

医療の質を向上させながら増え続ける医療費の圧縮が求められているなか、DPC (Diagnosis Procedure Combination=診断群包括分類) データの蓄積が進み、これら医療ビッグデータの重要性が益々認識されその利活用に大きな期待が寄せられている。DPC データを用いた集中治療関連の報告は散見されるが、治療法の効果などを検証した結果が大半であり、治療戦略への応用などの報告は認められないのが現状である。本研究では患者にとって重要な「1. 予後（特に死亡率）」、医療制度を支える国にとって重要な「2. 医療費」、安定的に患者を受け入れる為に病院にとって大事になる「3. 入院日数」の3つの変数に注目し、他の変数2つが同じならもう1つの変数はより「生存率が高く」、「安価で」、「短い入院日数」となるような治療戦略を持つことを目的とした上で、多彩な治療薬が存在する敗血症に対してDPC データで得られた治療薬を pLSA (Probabilistic Latent Semantic Analysis) 手法 [Hofmann, 1999] によりクラスタ分類するとともに同一患者におけるクラスタの時間遷移を解析する事で、DPC 入力データが敗血症治療ナビゲーションとして治療戦略に有益である可能性を検証した。

pLSA を用いた解析の実社会への応用としてはスーパーの購買履歴である ID-POS データを用いた顧客理

解・来店人数予測 [山下, 2017a], [石垣, 2011]、需要予測 [本村, 2011] などマーケティング分野への例がある。更に顧客 ID と月を組み合わせた変数により購買行動について時間軸上での変化のモデリングを行った研究 [原田, 2016] がある。同様の手法を医療分野に適応した初の例として [山下, 2017b] があり、「入院初日から3日まで」、「4~30日まで」、「31日目以上」と3期間における患者の時間変化も考慮した解析である。本研究では上記のように固定した期間に離散化する代わりに1日毎の時間変化を解析し、患者の変化の変わり目をデータから検出出来るようにした（動的離散化）。

2. 時間変化も考慮した pLSA によるモデリング手法

pLSA は、2つの変数 x, y の背後に共通の特性、潜在変数 $z \in Z = \{z_1, \dots, z_k\}$ があると仮定し、 x, y が各 z に関する確率をデータから期待尤度最大化法によって推定し、その確率値が最大となる潜在変数 z に x と y を分類する。pLSA では、観測される変数の分布に対する事前の仮定が不要であり、異なる分布や形式の変数を同時に厳しい制約なく扱えるため、本研究のような医療に関する変数を扱う際に適している。

pLSA の実行には産業技術総合研究所の知財ソフトウェア「PLASMA」を用いた。

本研究では、DPC データを用い上記の x, y のうち x を患者 ID と診療行為実施日（入院初日から 1 日毎）を組み合わせた変数“患者 ID_実施日”とした。[山下, 2017b]では、 y を注射薬剤(分類名が注射の診療行為名称)としていたが、本研究では、注射以外にも、投薬、処置、手術等も含めた診療行為名称を用いた。上記の変数 x, y を入力として pLSA を行い、同じような診療行為を受けている患者及び実施日と、同時に行われやすい診療行為を同時に分類した。使用した患者 ID 数は 458 (患者 ID_実施日の総数は 19658)、診療行為数は 4406 であった。得られた潜在変数 z を診療クラスタと呼ぶこととし、このクラスタ毎に、死亡率（死亡の有無の平均）、出来高金額合計（医療費）、在院日数を集計した。更に同一患者の時間経過によるクラスタ遷移を調べた。

3. 敗血症治療戦略への時間変化も考慮した pLSA の適応結果

3. 1. 診療クラスタ毎の死亡率・医療費・在院日数の集計

2章で示した方法で pLSA を行ったところ診療クラスタ数は 9 で AIC 最小となった。その各 9 クラスタに所属する患者 ID について、死亡率・医療費・在院日数の平均を求めた結果を表 1 に示す。死亡率の低い順に aC1, aC2, .. とクラスタに名前を付けその順に示している。

表 1. 診療クラスタ毎の死亡率・医療費・在院日数

クラスタ	死亡率	医療費(万円)	在院日数(日)	実施日	データ数
aC1	0.05	7074508	126		63
aC2	0.11	2920674	46		14
aC3	0.15	5580585	95		57
aC4	0.19	5138048	78		35
aC5	0.20	5394294	83		48
aC6	0.22	4550132	57		15
aC7	0.24	7836502	106		63
aC8	0.51	9178218	79		28
aC9	0.57	10470061	130		64

上記のクラスタ aC1~aC9 を、死亡率-医療費の平面にマッピングした結果を図 1 に示す。図 1 をみると概ね死亡率が高いクラスタ程、医療費も高くなっている傾向が見られ、aC2 のように死亡率が低く医療費も安価なクラスタ、aC8 や aC9 のように死亡率が高く医療費も高いクラスタがある。一方、aC1 のように死亡率は最も低いが比的医療費が高いクラスタもある。

図 2 は、各クラスタを死亡率-在院日数の平面にマッピングした結果である。aC2 は在院日数が短く aC9 は長い。aC8 は aC9 よりは在院日数が短く、aC1 は死亡率は低い但在院日数は長い。

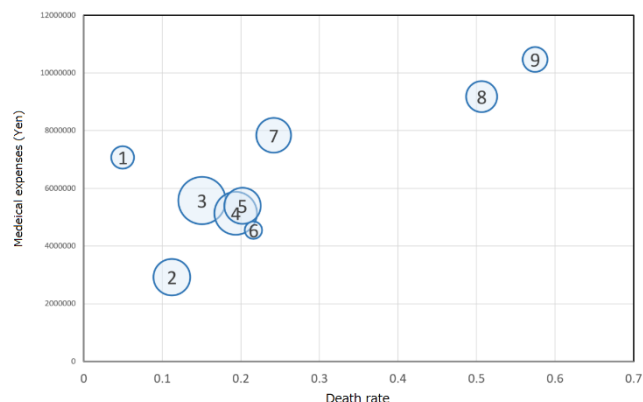


図 1. [死亡率-医療費]平面への診療クラスタマッピング
クラスタの円の大きさはデータ数を示す

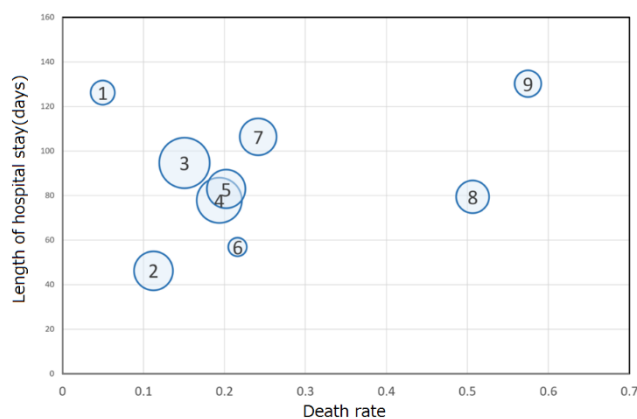


図 2. [死亡率-在院日数]平面への診療クラスタマッピング
クラスタの円の大きさはデータ数を示す

3. 2. 各診療クラスタの診療行為

3. 1. 節では変数 x の患者 ID 部の集計を行った。本節では変数 y の診療行為の分類結果を示す。前節で見出した死亡率が最も低い aC1 クラスタと、死亡率が低く医療費が安価で在院日数が短い aC2 の結果を表 2 に示す。

3. 3. 同一患者の時間経過に伴うクラスタ遷移

本節では、同一患者の時間変化に伴うクラスタ遷移についての結果を示す。2章に示したように本研究では同一患者も日毎に別の変数 x として（日が変われば“別患者”のように）取り扱った上で pLSA を行っているが、改めて同一患者がどのようにクラスタ遷移をしているかを見ると概ね以下のような遷移パターンが見られた。

表 2. 各診療クラスタの診療行為名称

(a) aC1 クラスタ

入院_特定機能病院15対1入院基本料精神病棟
入院_看護補助加算1
入院_罹患の危険性が高い>
入院_超重症児(者)入院診療加算(6歳未満)
入院_<その他重篤な状態>
入院_<ア吐血で全身状態不良>
入院_<イ易感染性により、感染症
入院_<イ易感染性により、感染症腫
処置_▽幼児加算(喀痰吸引)
処置_イソゾール注射用0.5g500mg(溶解液付)
処置_カウンターショック(非医療従事者向け自動除細動器を用いた場合)
処置_ノボ・ヘパリン注5千単位/5mL5,000単位
処置_ペネトリン吸入液0.5%
手術_心不全(NYHAⅢ度以上のものに限る。)の患者
手術_抗悪性腫瘍剤静脈内持続注入用埋込型カテーテル設置(頭頸部その他)
手術_血管露出術
手術_造血幹細胞移植(臍帯血移植)
手術_静脈麻酔(短時間のもの)
手術_▽乳児加算(輸血)
手術_▽幼児加算(輸血)
投薬_セフゾン細粒小児用10%100mg
投薬_ダントリウムカプセル25mg
投薬_チラーヂンS散0.01%
投薬_テオドール錠100mg
注射_プリディオ静注200mg2mL
注射_プロイメンド点滴静注用150mg
注射_ペントシリン注射用1g
注射_メソトレキセート点滴静注液1000mg1,000mg40mL
画像_プロハンス静注5mL
画像_プールのシンチ注
画像_ポジトロン断層・コンピューター断層複合撮影(18FDGを用いた検査)
診察_<エクザール>
診察_<エンドキサン>
診察_<セルセプト>
その他_入院精神療法(I)
その他_急性脳症

(b) aC2 クラスタ

入院_救急医療管理加算
入院_看護職員夜間配置加算
入院_退院調整加算(一般病棟入院基本料等)(14日以内)
処置_センノサイド錠12mg
処置_ニフレック配合内用剤
処置_パルミコート吸入液0.25mg2mL
処置_ラキソベロン内用液0.75%
手術_硬膜外麻酔後における局所麻酔剤の持続的注入
投薬_プレマリン錠0.625mg
投薬_ミオナール錠50mg
投薬_ランドセン錠0.5mg
検査_口蓋扁桃
検査_唾液腺
検査_基本的検体検査実施料(4週間以内)
注射_セビドナリン静注用1g
注射_ナベルピン注1010mg1mL
注射_ヒューマログ注ミリオペン300単位
画像_尿管ステントセット(一般型・異物付着防止型)¥
画像_撮影用薬剤
画像_腎シンチグラム
診察_診療情報提供料(II)
その他_パーキンソン病急性増悪
その他_下肢蜂窩織炎
その他_右膝化膿性結核性関節炎掻爬術
その他_左膝化膿性結核性関節炎掻爬術

- ・死亡率の低いクラスタに留まり早期退院
- ・死亡率の高いクラスタから低いクラスタに遷移し生存
- ・初期から高いクラスタで早期に死亡
- ・段階的に高いクラスタへ遷移し死亡
- ・高いクラスタに長期に留まり死亡

上記のうち、死亡率の高いクラスタから低いクラスタに遷移し生存・退院したケースと、段階的に高いクラスタへ遷移し死亡したケースの例を図3にそれぞれ示す。

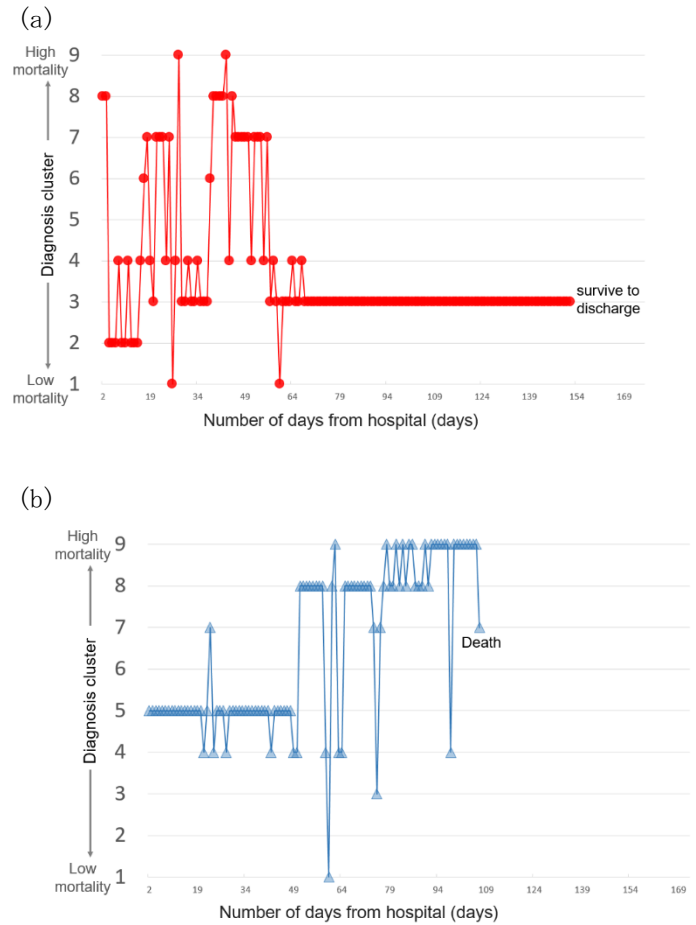


図 3. 同一患者の時間経過に伴うクラスタ遷移

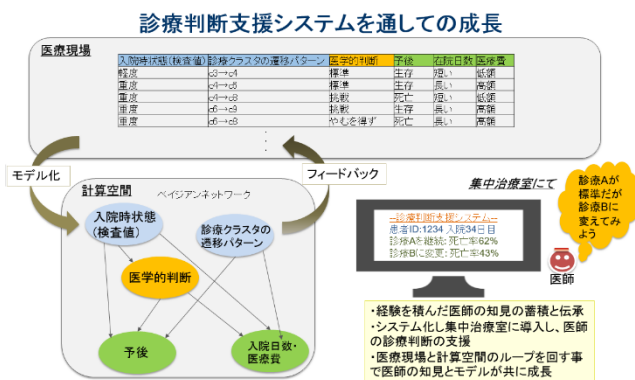
縦軸の数字は診療クラスタ aC1～aC9 の数字部

- (a)死亡率の高いクラスタから低いクラスタに遷移し生存したケース, (b)段階的に高いクラスタへ遷移し死亡したケース

4. 医学的判断の紐づけによる診療判断支援システムを通しての成長

図3に示したようにある日にクラスタが遷移する際には、医師が何らかの医学的判断により診療行為を変えた事を意味する。また同じクラスタが継続されているのも医学的判断の結果である。たとえば、「その際の患者の状態における標準的な診療を行っ

ている」や「集学的治療を行った」、「ある種の挑戦的な診療を行った」、または「with draw」や「with hold」などの医学的判断である。図4に示した概念図のように、これら医学的判断とクラスタ遷移、また検査値を紐づけ、さらにそれらの結果として、目的変数としての予後、在院日数、医療費の症例データを蓄積する事で、医療現場の医師の判断も含めた現象を、ベイジアンネットワーク等により計算空間にモデル化する事が出来る。そのモデルは診療判断システムとして実際の集中治療室で活用する事が出来、さらにその活用結果は新たな事例の蓄積として医療現場にフィードバックが可能である。このように医療現場と計算空間でのループを回していく事で、経験を積んだ医師の知見の蓄積と伝承が可能となり、全体の医師の知見とモデルの両方が共に成長していく事が可能となる。



参考文献

[Hofmann,1999] Hofmann, Thomas. "Probabilistic latent semantic analysis." Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., (1999)

[石垣,2010]石垣司, 竹中毅, 本村陽一: 百貨店ID付きPOSデータからのカテゴリ別状況依存的変数間関係の自動抽出法, オペレーションズ・リサーチ, 56, 2, (2010)

[石垣,2011] 石垣司, 竹中毅, 本村陽一: 潜在クラスモデルによる流通量販店舗の来店人数予測の精度改善の評価, 人工知能学会全国大会, (2011)

[原田, 2016] 原田奈弥, 山下和也, 本村陽一: ID付POSデータによる購買行動の季節変化の分析と視覚化, 人工知能学会社会におけるAI研究会 27回研究会,(2016)

[本村,2011] 本村陽一, 竹中毅, 石垣司: 条件付層別差分モデルによる需要予測の高精度化, 人工知能学会全国大会, 1B3-1, (2011)

[山下, 2017a] 山下和也,原田奈弥,黄冬陽,吉開朋弘,本村陽一: ID-POS データによる来店行動・購買行動の潜在的時空間意味構造分析, 人工知能学会全国大会,(2017)

[山下, 2017b] 山下和也,阪本雄一郎,櫻井瑛一,本村陽一: 時間変化も考慮した pLSA 手法による敗血症治療戦略への応用, 人工知能学会 社会におけるAI研究会 27回研究会,(2017)

5. まとめ

- 時間変化もモデル化の対象とした pLSA 手法を用いる事によって DPC データの敗血症治療戦略への活用可能性を検証した。
- 日々の患者と診療行為で pLSA を行い患者と診療行為を同時にクラスタリングしクラスタ毎に死亡率・医療費・在院日数を集計した。
- 同一患者の時間変化に注目すると死亡率の高いクラスタへ推移していく患者や日数経過後も同一クラスタに留まる患者、死亡率の低いクラスタへ遷移していくケース等のパターンが見られた。
- 本手法を用いる事で症例の集積と医学的判断の関与によって DPC データが治療戦略に関与し得る可能性が認められ、今後診療判断支援システムとして医療現場に導入して活用を始めることで、医師の知見・経験と症例が蓄積・学習され、医師とシステムが共に成長する仕組み実現への道を開いた。

謝辞

本研究は NEDO 委託事業「人間と相互理解できる次世代人工知能技術の研究開発」の支援を受けて行いました。