

入院早期における在院日数予測手法の検討

仲濱 正大*¹ 納富 一宏*¹ 斎藤 恵一*² 黒田 史博*³ 外山 比南子*⁴

神奈川工科大学情報工学科*¹ 国際医療福祉大学情報教育室*²

国際医療福祉大学医療福祉・マネジメント学科*³ 国際医療福祉大学大学院*⁴

1 はじめに

2003年より始まった急性期入院医療を対象とした診療報酬の包括評価制度(DPC/PDPS)は、従来の出来高払い制度に代わるものとして特定機能病院等で導入され、2012年には1500余病院・約48万床が対象となっている。DPCにより患者の疾病別医療行為等の診療情報が標準化されたことで他施設との比較が可能となり、在院日数の適正化や診療プロセスの見直しなど医療の質の向上に貢献している。

我々は過去のDPCデータをもとに自己組織化マップを使って新しい患者の在院日数の予測を試みている。本研究では在院日数の予測可能性を検討し、精度評価を行うことを目的とする。今回は、自己組織化マップの学習に利用する属性ベクトルの構成と、マップ生成時の視覚化について検討し、在院日数の予測可能性を考察したので報告する。

2 在院日数予測の対象データ

マップ作成に用いた症例は、2010年までに複数の病院から収集したDPCデータ490,295ケースである。各ケースが持つ属性は、主傷病病名、性別、年齢、併存症、合併症など119項目である。

マップ作成前の分析として各主病名の週ごとの在院日数分布を調査した。多くの主病名で入院1週目までに退院する患者が約5割を占めることがわかった。図1に胃体部癌、図2に胃腸炎における週当たりの退院する患者の割合を示す。特定期間において多くの患者が退院する場合は、その期間を在院日数の予測値として与えれば良いが、他の項目が在院日数に与える影響を無視することになる。

本稿では週ごとの在院日数分布が偏っていない主症病名が胃体部癌である診断・診療データを対象にして在院日数の予測可能性を検討した。

3 自己組織化マップによる分析

3.1 自己組織化マップ

主症病名が胃体部癌である診断・診療データを自己組織化マップ(SOM: Self-Organizing Maps, 以下SOMという)に投入し、学習を行った。

SOMはn次元の入力ベクトルを2次元マップ上に配置する能力を持ち、類似度の高いベクトルが近くに配置される特徴を持つ^[1]。診断・診療データを構成する119項目の内、SOMの学習に使用した項目を表1に示す。

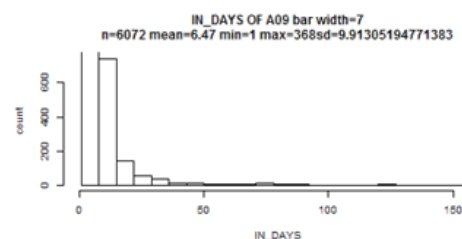


図1 胃体部癌における週当たりの在院日数割合

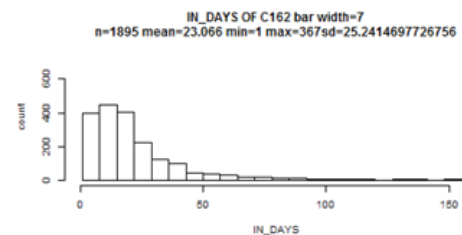


図2 胃腸炎における週当たりの在院日数割合

表1 SOMに投入する属性ベクトルの構成要素

番号	説明	型
1	年齢	値
2	在院日数	値
3	併存症コード1	文字列(ICDコード)
4	併存症コード2	文字列(ICDコード)
5	併存症コード3	文字列(ICDコード)

表1の併存症コード1~3は疾病及び関連保健問題の国際統計分類(ICD: International Statistical Classification of Diseases and Related Health Problems, 以下ICDと略)第10版を用いてコーディングされた値となっている。ICDコードは、先頭のアルファベット1文字と数字2文字~4文字によって表される。

3.2 ICDコードのエンコード方法

併存症コードは文字列で構成されているため、SOMに投入するには数値にする必要がある。本稿ではICD10コードを構成する頭文字A~Zを1~26の数値に対応させ、符号化した。

3.3 SOMの学習条件

主症病名が胃体部癌である診断・診療データの在院日数が1~6週目のデータを抽出した。そこから各週60個のデータを無作為に抽出し、50個のデータを学習用データ、10個のデータを検証用データとした。さ

Study on Prediction Method of Length of the Hospital Stay in Early Days

*¹Masahiro Nakahama, *¹Kazuhiro Notomi,

*²Keiichi Saito, *³Fumihiko Kuroda, *⁴Hinako Toyama

*¹Information and Computer Sciences, Kanagawa Institute of Technology

*²Education Center of Medical Informatics, International University of Health and Welfare

*³Department of Health Service Management, International University of Health and Welfare

*⁴International University of Health and Welfare Graduate School

らに、検証用データから在院日数の項目を省略した。

学習用データを 0~1 の値に正規化し、5次元の入力ベクトルとして SOM に投入した。総ユニット数 4,900 (マップサイズ 70×70)、学習回数 100,000 回で学習を行った。SOM の学習が行われる出力層はトーラス型 (マップの上端と下端、右端と左端が繋がっている) となっている。

3.4 出力された 2次元 SOM マップ

3.2 節で提示した条件で作成した SOM マップを図 3 に示す。薄い色が在院日数 1~3 週間、濃い色が在院日数 4~6 週間の領域である。マップ上の数字は、投入された入力ベクトルの在院日数が何週目かを表している。

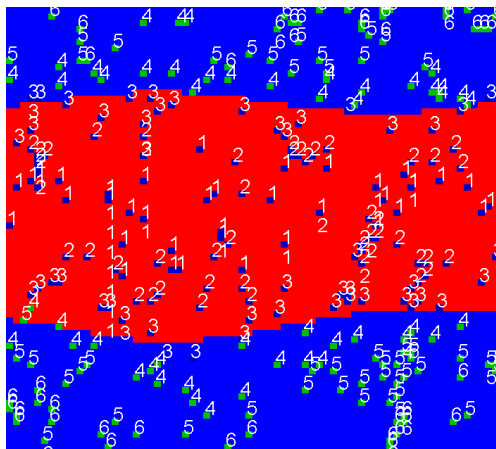


図 3 前節で提示した条件で作成した SOM マップ

3.5 在院日数予測手法の検討

得られた SOM マップに検証ベクトルを投入し、検証ベクトルの本来の在院日数が 1~3 週間である場合は薄い領域、4~6 週間である場合は濃い領域に配置された場合、予測が成功したと判断する。

検証ベクトルは在院日数の値を省略しているため、仮の値を入れた。今回の実験は、仮の値を 0.0, 0.5, 1.0 の 3 段階に分けてそれぞれ 10 個ずつ、合計 30 個を SOM マップに投入し、在院日数予測精度を比較した。その結果、仮の値が 0.0 の場合は 50%、0.5 の場合は 55.0%、1 の場合は 50% となった。

4 考察

精度評価の結果は在院日数の予測期間を広く設けたにも関わらず、在院日数予測精度は約 50% となった。これは入力ベクトルの属性が在院日数との相関が低いために起きたと考えられる。

在院日数予測精度向上のため、在院日数と相関のある可能性が高い 2 つの項目を以下に示す。

1 つ目は合併症数である。図 4 に x 軸が合併症数、y 軸が在院日数の箱ひげ図を示す。箱ひげ図の長方形の下辺は第 1 四分位数、上側の辺は第 3 四分位数、中央の線は中央値となっている。図 4 より合併症数が上がるにつれて中央値が大きくなっていることがわかった。

2 つ目は DPC コードである。DPC コードは 14 桁で表される。各桁は意味づけされており、第 1~6 桁は

疾患コード、第 7 桁は入院目的 (現在は未使用)、第 8 桁は年齢・体重・JCS 条件、第 9~10 桁は手術の分類、第 11 桁は手術・処置等 1、第 12 桁は手術・処置等 2、第 13 桁は副傷病、第 14 桁は重症度等を表す^[2]。入力ベクトルの属性に入院後の処置の情報である DPC コードの 9~10 桁、12 桁、13 桁は在院日数と相関があると考えられる。

合併症数と DPC コードの項目を追加した属性ベクトルの構成要素を表 2 に示す。

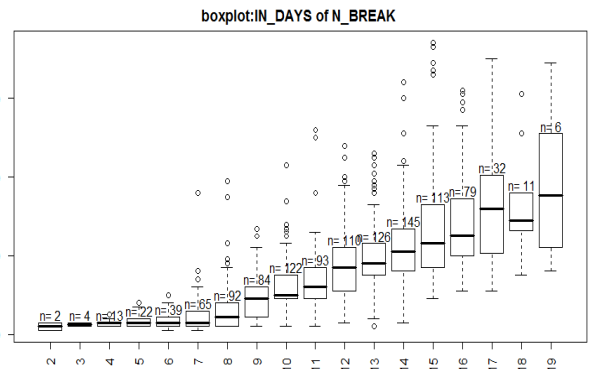


図 4 合併症数毎の在院日数

表 2 拡張された属性ベクトルの構成要素

番号	説明	型
1	年齢	値
2	在院日数	値
3	併存症コード 1	文字列
4	併存症コード 2	文字列
5	併存症コード 3	文字列
6	合併症数	値
7	DPC コード 9~10 桁	値 (名義尺度)
8	DPC コード 12 桁	値 (名義尺度)
9	DPC コード 13 桁	値 (名義尺度)

5 おわりに

本稿では SOM を用いて診断・診療データの分析をし、在院日数予測可能性の検討を行った。その結果、SOM を用いた診断・診療データの分析による在院日数の予測には、在院日数と相関のある項目を学習ベクトルの属性として含める必要があることを確認した。また、在院日数と相関のある項目の候補として合併症と DPC コードを挙げた。

今後は表 2 に示した属性ベクトルの構成を用いて SOM マップを作成し、在院日数予測手法の検討を進めていき、より高い精度での在院日数予測の提案を行っていく。

参考文献

- [1] T.Kohonen: 自己組織化マップ, シュプリンガー・フェアラーク東京(1996), 徳高平蔵 他.
- [2] 厚生労働省大臣官房統計情報部: 疾病, 傷害及び死因分類の正しい理解と普及に向けて <http://www.mhlw.go.jp/toukei/sippe/dl/fukyuubon.pdf> (2007)