

一般口演

一般口演18

看護情報・データ解析・評価

2017年11月22日(水) 08:45 ～ 10:15 J会場 (3F イベントホールA)

[3-J-1-OP18-6] 同一 DPCコード内症例の類型化を考慮した看護クリニカルパス生成支援システムの開発津本 周作¹, 岩田 春子², 木村 知広³, 平野 章二¹ (1.島根大学医学部医学科医療情報学, 2.島根大学医学部附属病院入退院管理センター, 3.島根大学医学部総務課)

著者らは、これまで看護オーダーの実施歴から、クラスタリングを用いて看護クリニカルパスを生成する仕組みについて提案してきた。この中で、同一 DPCコード内の症例を用いてパスの評価を行ったが、眼科領域等外科領域ではこれまで看護師が作成したパスに類似のものが生成できたのに対し、脳梗塞・肺がんといった症例では複雑なパスになった。これは DPCコード内に複数の病態が含まれ、それぞれの病態によって、入院日数が異なってくる場合があり、コーディングと臨床の病態の間との乖離がそのままパスに投影された形となってしまったことによる。今回、この問題を克服するべく、同一 DPCコードの症例に対して、入院日数を指標として、混合分布モデル(いわゆる EMクラスタリング)を適用し、DPCコード内の症例の分類を行った上で、退院時要約を用いて特徴付けし、分類モデルを生成、生成したモデルを用いて、改めて当該 DPCコード内の症例をサブクラスに分類、分類後に、看護クリニカルパスの生成法を適用する形での拡張を試みた。具体的には、まず症例の選別を自動化するために、同一 DPCコード内での疾患の類別を入院日数を指標として、混合分布モデルを用いて行い、得られたクラスについて、対応分析を用いて、得られたクラスにおける性質に対応するキーワードを退院時要約から抽出した。さらに、これらのキーワードを用いて、判別モデルを深層学習, Random Forestで構築した。次に、構築したモデルを用いて、DPCコード内の症例をサブクラスに分類し、それぞれのサブクラスの看護実行歴のデータに分割、それぞれに対して、これまで提案してきたクラスタリングに基づく看護クリニカルパスの生成法を適用した。本研究では、本プロセスの性能について2015年度の退院時要約のうち、症例数の多い10コードについて評価を試み、それぞれのサブクラスで生成されたパスは、従来法で得られたパスよりも簡易なものが得られた。

同一 DPC コード内症例の類型化を考慮した 看護クリニカルパス生成支援システムの開発

岩田春子^{*1}, 木村知広^{*2}, 平野章二^{*3}, 津本周作^{*3}

*1 島根大学医学部附属病院入退院管理センター, *2 島根大学医学部医療サービス課

*3 島根大学医学部医学科医療情報学

Clinical Pathway Generation with Classification of Cases with the same DPC code assigned based on Cluster Analysis

Haruko Iwata^{*1}, Tomohiro Kimura^{*2}, Shoji Hirano^{*3}, Shusaku Tsumoto^{*3}

*1 Center for Bed-control, Shimane University Hospital,

*2 Medical Services Division, Faculty of Medicine, Shimane University,

*3 Department of Medical Informatics, Faculty of Medicine, Shimane University

Abstract

This paper proposes the following two-fold clinical care generation method. First, the system proposes how clinical cases with the same DPC code are characterized by mixture model clustering, and construct classification model by the analysis of discharge summaries. Then, cases are classified by using the classification model and a clinical pathway is generated for each new class. The proposed method was evaluated on the datasets extracted hospital information system, whose results show that plausible clinical pathways were obtained, compared with previously introduced methods.

Keywords: Clinical Pathway, Agglomerative Hierarchical Clustering, Mixture Model

1. はじめに

著者らは、これまで看護オーダーの実施歴から、クラスタリングを用いて看護クリニカルパスを生成する仕組みについて提案してきた[1-6]。この中で、同一 DPC コード内の症例を用いてパスの評価を行ったが、眼科領域等外科領域ではこれまで看護師が作成したパスに類似のものが生成できたのに対し、脳梗塞・肺がんといった症例では複雑なパスになった。これは DPC コード内に複数の病態が含まれ、それぞれの病態によって、入院日数が異なってくる場合があり、コーディングと臨床の病態の間との乖離がそのままパスに投影された形となってしまうことによる。今回、この問題を克服するべく、同一 DPC コードの症例に対して、入院日数を指標として、混合分布モデル(いわゆる EM クラスタリング)を適用し、DPC コード内の症例の分類を行った上で、退院時要約を用いて特徴付けし、分類モデルを生成、生成したモデルを用いて、改めて当該 DPC コード内の症例をサブクラスに分類、分類後に、看護クリニカルパスの生成法を適用する形での拡張を試みた。具体的には、まず症例の選別を自動化するために、同一 DPC コード内での疾患の類別を入院日数を指標として、混合分布モデル[7]を用いて行い、それぞれのクラスへの帰属確率を推定する。次に、帰属確率を用いて、DPC コード内の症例をサブクラスに分類し、それぞれのサブクラスの看護実行歴のデータに分割、それぞれに対して、これまで提案してきたクラスタリングに基づく看護クリニカルパスの生成法を適用した。本研究では、本プロセスの性能について 2015 年度の退院時要約のうち、症例数の多い 10 コードについて評価を試み、それぞれのサブクラスで生成されたパスは、従来法で

得られたパスよりも簡易なものが得られた。

2. これまでの方法

著者らは電子的に蓄積された看護オーダー実行歴を頻度情報として用いて、頻度情報の距離の差に基づくクラスタ分析を適用し、データ解析でパスの作成支援を提唱してきた[1-5] (図 1)。

この方法では、まず病院情報システムから DPC を指定し、与えられた DPC をもった症例について、入院日を基準として、各症例の看護オーダー数を算出し、行を看護オーダー、列を入院日数とした表形式データを入力データとして出力する。この入力データを元に、各オーダー間の距離を算出し、距離行列を計算する。この距離行列を用いて、多次元尺度構成法、階層クラスタリング[8]を適用し、看護オーダーの類型化を試みる。この段階では、病日数を属性として用いるが、これらの属性のクラスターで与えられたクラス分類の寄与度を考えるために、決定木で使用する Information Gain を用いて、各属性の寄与度を求める。これによって、日数が Information Gain の大きさによってグループ分けが可能になる。この寄与度によるグループ分けを用いて、元のデータを分割する。この分割した各データについて、距離行列の計算、クラスタリングを行う。クラスタ分析が収束するまで、この手続きを繰り返し、最終的に、分割した入院日数での類型化した看護オーダーを統して、パスを構築する。

しかし、この手法では、指標の値によって、属性のグループ分けは複雑になる可能性があるため、これを避けるため、属性の選択をクラスタリングによる属性の類型化による方法、Dual-Clusteringを開発した。(図 2)

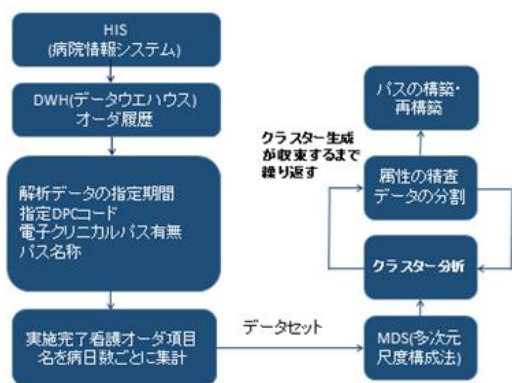


図1 属性選択によるパス自動作成手順

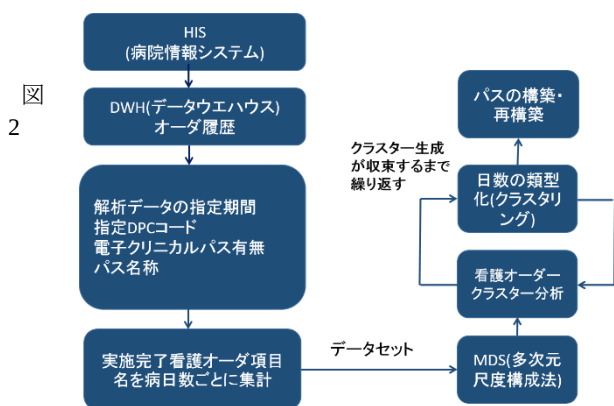


図2 Dual-Clustering によるパス自動作成手順

3. 混合正規分布モデルによる類型化

3.1 混合正規分布モデル

DPC は診療に医療資源を一番投入した疾患によって、コーディングされる。したがって、例えば、治療薬は同じであれば、違う病態を持ちうる疾患群でも、同じコードに属する可能性がある。したがって、DPCをキーとして検索し、それによって得られたデータについての解析を行う場合、複数の病態をできるだけ分離することが必要であると考えられる。

同一コード内に複数の病態が含まれる時、各病態が独自の時間発展を遂げると考えると、コードでの入院日数の分布は複数の病態を反映した混合分布となると考えられる。もし、入院日数の分布が正規分布で表現できるとすれば、これらを分離する方法として、混合分布モデル[7]が適用できる。もともと本手法は判別分析における方法として提案されたが、現在、EM アルゴリズムを用いたモデルベースのクラスタリングの一つと位置づけられている。

混合正規分布では、パラメータ θ で x が出現確率は:

$$p(x, \theta) = \sum_{m=1}^M \pi_m \phi(x; \mu_m, \Sigma_m)$$

で与えられる。ここで、 M はクラスタ数、 μ_m, Σ_m はそれぞれクラスタ m の平均と共分散行列を示し、正規分布の線型結合で表現される。

3.2 類型化

ベイズの定理から、クラスタ m が現れる確率(事後確率)は、

$$p(m|x, \theta) = \frac{p(x, m; \theta)}{p(x; \theta)} = \frac{\pi_m \phi(x; \mu_m, \Sigma_m)}{\sum_{m=1}^M \pi_m \phi(x; \mu_m, \Sigma_m)}$$

によって与えられる。

この事後確率の尤度から最尤推定法によって、 μ_m, Σ_m を推定するのが、E-M アルゴリズムであり、次のようなステップで構成される。

E-step:

ある $\theta^{(t)}$ を用いて、モデルの事後確率に基づく対数尤度(Q関数を計算する。)

tステップでの対数尤度は

$$Q(\theta, \theta^{(t)}) = \sum_{i=1}^n \sum_{m=1}^M p(m|x_i; \theta^{(t)}) \log p(x_i, m; \theta)$$

で与えられる。

M-step:

Q関数を最大化する $\theta^{(t)}$ を求める。

$$\theta^{(t)} = \arg \text{Max}_{\theta} Q(\theta, \theta^{(t)})$$

このEステップとMステップが収束基準を満たすまで反復する。正規分布の場合、 θ は平均と共分散行列(分散)で与えられるので、各クラスについて、これらを推定することになり、クラスタ数を指定して、事後確率を最大化するようなパラメータを推定することで、各クラスの情報が得られる。

3.3 各症例の分類

EM-clustering では、各例については、あるクラスタに属する確率が算出される。しかし、この場合、正確に症例を排反的に分離できていない。このため、もともとのデータをクラスタ数に分割するため、以下のように帰属群を判定する。

- クラスタ数 k を与え、与えられた DPC に属する症例の入院期間のデータに対して、混合正規分布モデルによる EM クラスタリングを適用する。
- 入院日数(単位: 日)に対する各症例のクラスタへの帰属確率を算出する。
- 2つのクラスタが算出した確率が等しくなる日数、なければ、前後の日数の平均を2つのクラスタの分界点と定義。
- 分界点の前後によって、症例の分布クラスタを決定する。

4. パス生成プロセス

以上の方法と、Dual-clustering によるパス生成を組み合わせ、パス生成方法を図3のように拡張する。図中のクラスタ数の上限は10と設定しているが、上限の設定は任意である。

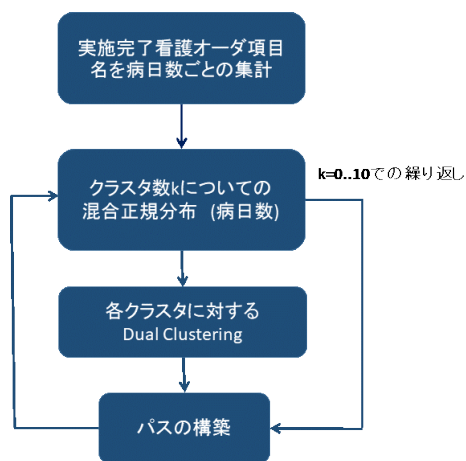


図3 パス自動作成手順

5. 実験

島根大学医学部附属病院の2015年度に入院した患者について、DPC使用頻度上位20位までの疾患(表1)を抽出し、その看護オーダー数を集計し、基礎データとした。プロセスはR3.4.1のスク립トとして記述し、混合正規分布モデルにはmixtools[9]を、階層クラスタリングにはR上のhclustを用い、距離行列はユークリッド距離を用いて計算した。

表1 DPC上位20位までの件数(2015年度)及び入院日数の分布の正規性・二峰性

No	DPC名称	件数	正規性	二峰性
1	白内障、水晶体の疾患 手術手術あり 重症度等片眼	445		
2	白内障、水晶体の疾患 手術手術あり 重症度等両眼	152		
3	2型糖尿病(糖尿病性ケトアシドーシスを除く)	145	○	
4	肺の悪性腫瘍 手術手術あり 処置等2なし	131	○	
5	子宮頸・体部の悪性腫瘍 手術手術なし 処置等24あり 副傷病名なし	121		
6	肺の悪性腫瘍 手術手術なし 処置等1あり 処置等2なし 副傷病名なし	120		
7	子宮の良性腫瘍 手術腹腔鏡下腔式子宮全摘術等	111		
8	肺の悪性腫瘍 手術手術なし 処置等1なし 処置等24あり	110		
9	妊娠期間短縮、低出生体重に関連する障害(出生時体重2500g以上) 手術手術なし 処置等2なし 副傷病名なし	110		
10	肘、膝の外傷(スポーツ障害等を含む。) 手術縫合術等	99	○	○
11	全身性臓器障害を伴う自己免疫性疾患 手術手術なし 処置等2なし	96		
12	非ホジキンリンパ腫 手術手術なし 処置等24あり 副傷病名なし	94		
13	肺炎、急性気管支炎、急性細気管支炎(15歳以上) 手術手術なし 処置等2なし	86		
14	肺の悪性腫瘍 手術手術なし 処置等1なし 処置等2なし	85		
15	慢性腎炎症候群・慢性間質性腎炎・慢性腎不全 手術手術なし 処置等1なし 処置等2なし 副傷病名なし	83	○	○
16	肝・肝内胆管の悪性腫瘍(続発性を含む。) 手術その他の手術あり 処置等2なし	82		
17	胆管(肝内外)結石、胆管炎 手術限局性腹腔腫瘍手術等 処置等2なし 副傷病名なし	82		
18	脳梗塞(JCS10未満) 手術手術なし 処置等1なし 処置等23あり 副傷病名なし	80	○	○
19	網膜剥離 手術手術あり 重症度等片眼	75		○
20	胎児及び胎児付属物の異常 手術子宮全摘術等	73		

5.1 混合正規分布モデルによる結果

表1にEクラスタリングによる結果も含めた。正規性を示したのは、2型糖尿病、肺の悪性腫瘍、肘・膝の外傷、慢性腎炎、脳梗塞、二峰性を示したのは、肘・膝の外傷、慢性腎炎、

脳梗塞であった。

これらは、それぞれ肘・膝、慢性腎炎症候群、脳梗塞の部位の違いに起因することが予想される。本稿では、脳梗塞(DPCコード:010060x099030x)についての結果を以下に示す。図4に、脳梗塞におけるmixtoolsの適用結果を示した。

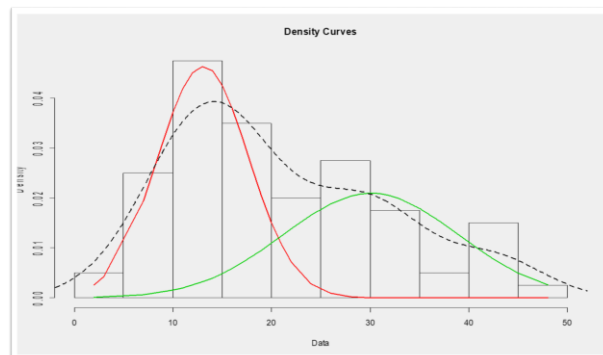


図4 脳梗塞に関する混合正規分布モデルのあてはめ

5.2 脳梗塞のパス作成

図3で示したパス生成プロセスを脳梗塞のデータに以下のように適用した。

- 図5と6に病院情報システムから抽出した脳梗塞(DPCコード:010060x099030x)の看護オーダー集計にDual-clusteringを適用した結果のデンドログラムを示した。図5から、看護オーダーは概ね3〜5つに分割された。図6から、病日に関しても3つのクラスタが得られた。ここから得られるパスは表2の通りである。
- 混合正規分布の推定(EM clustering)によって、元の症例を与えられたクラスタ数での確率モデルを生成する。この場合、分布は二峰性(クラスタ:2)が最もあてはめがよかった。Group1は入院日数の短い症例、Group2は長い症例からなる。
- 3.3で示した方法で分界点を求め、各症例を属すべきクラスタに配置する。
- 配置した各クラスタをsubgroupとし、subgroupに対して、Dual clusteringでパス生成を行った。このDual clusteringは再帰的にデータを分割することが可能であるが、表3,4に、それぞれ1回再帰的に適用した際の2つのクラスタで生成されたパスを示した(パスに表示されたものは、最も頻度の高いグループのみを示した。)

表2,3,4の比較から、より病態を反映したパスが生成できたと考えられる。なお、上記全プロセスについては、HP Z800 WorkStation (プロセッサ: Intel(R) Xeon(R) CPU X5650 @ 2.67GHz, メモリ: 24.0GB)にて、46.37秒かかった。

6. 考察

6.1 クラスタ数の調整

本手法では、EMクラスタリングと階層型クラスタリングという2つのクラスタリングを適用しているが、両手法とも分割数の決定が問題になる。前者では、確率密度の線型結合を考えて、各クラスタの事後確率を推定するため、クラスタ数を前も

って与える必要がある。後者については、得られたデンドログラムのどのレベルでの分割を利用するかという任意性がある。ともに、さまざまな解決方法[8]が提案されているが、決定的な方法は提案されていない。本提案手法では、この部分については、病日によるヒストグラム、デンドログラムの形状から大まかにクラスタ数を推定してから、パス作成を行っている。今後、これらのクラスタ数調整の自動化を実装していくことが必要であろう。

6.2 クラスタの意味づけ

クラスタリングは、(非)類似度の大きさを基準として、類似度の大きいものをグループとしてまとめる手続きである。したがって、まとめられたグループがどのような性質を持つかについては、類似度の計算で用いられた指標以外の情報を有さない。ドメインの意味づけをするには、それぞれのクラスタについての付加的情報を見る必要がある。

入院患者の場合、退院時要約を記述することが診療録管理上も重要であり、ほとんどの電子カルテシステムは退院時要約の電子化を実装している。退院時要約には、各症例の入院の経緯・診療行為がコンパクトに記述されているため、この中に、クラスタには含まれている付加情報が隠れていることが予想される。したがって、退院時要約のテキスト分析することによって、これらのクラスタの頻出キーワードのパターンを抽出することができると予想される。

著者らは、退院時要約に形態素解析および対応分析を適用し、実際に、上記脳梗塞の2グループの特徴付けを行った。結果から、入院日数の短いグループには、運動麻痺があるか少ない症例、入院日数の長いグループには、運動麻痺がある場合の症例からなることがわかった[10]。看護パス生成をより正確なものにするには、退院時要約のみならず、看護サマリをも含めて解析することが必要であり、クラスタの意味づけを系統的に与える手法としてのテキストマイニングを検討する予定である。

6.3 看護オーダー集計

本手法では、単純に実施看護オーダー数の集計を用いて、ユークリッド距離を計算している。看護オーダーによっては、1日2検、3検といった重複が許されているため、パス上においても、そのようなオーダー内容を呈示できた方がよい。これについては、各病日での入院数から、各病日でのオーダー数を求め、それによるクラスタリングを行う等の方法が考えられ、今後の検討を要する。

7. おわりに

本稿では、DPC コードの割り振られた症例について、混合正規分布モデルを使って帰属確率を推定、その推定確率を用いて、症例の分割を行い、分割されたデータで、Dual-clustering によって、看護パスの自動生成を行うための手法を開発し、実際に、病院情報システムから抽出した看護オーダーのデータを用いて検証した。本稿では、脳梗塞についての解析結果を示したが、今後、他疾患についても、本手法を適用し、評価を進めていく予定である。

表2 データ全体で生成された看護パス

	1日目	2日目	3-7日目	8-16日目	9-16日目	17-29日目	30-48日目
(Na)VS測定(セカ項目)	0	0	0				
(Na)体温・脈拍測定	0	0	0				
(Na)呼吸数	0	0	0				
ME機器作動状態の確認	0	0	0	0	0		0
SPO2			0	0	0		
しびれ			0				
その他症状			0	0	0		
体位の調整							0
体位変換			0	0	0	0	0
体温・脈拍測定			0	0	0	0	0
口腔清拭			0			0	
呼吸			0	0	0		
嘔気・嘔吐	0	0	0	0	0	0	0
室内環境調整			0				
座位保持介助	0	0	0	0	0	0	0
褥瘡			0	0	0	0	
情報提供							0
排泄ケア			0	0	0	0	0
排泄回数			0	0	0	0	0
更衣			0				
歯磨き介助			0	0	0	0	0
洗面			0	0	0	0	0
消化・腹部症状			0	0	0		
留置時管理	0	0	0	0	0		
疼痛			0	0	0	0	0
眼症状			0	0	0	0	
睡眠			0	0	0		
神経	0	0	0	0	0	0	0
移乗介助(車椅子)							0
義歯洗浄			0			0	
義歯装着							0
自己拭去防止ケ?			0				
血圧測定			0	0	0	0	0
起き上がり介助	0	0	0	0	0	0	0
転倒防止ケ?			0	0	0	0	0
転落防止ケ?			0	0	0	0	0
運動	0	0	0	0	0	0	0
陰部洗浄			0				
食事介助(ハシカ)			0				0
食事介助(コップカ)			0	0	0	0	
食事摂取量	0	0	0	0	0	0	0
飲水量							0
オムツ交換			0	0	0	0	0
パナ交換			0				0
ゴータブルトイレ排泄介助			0				

表 3 Group1 のデータ全体で生成された看護パス
(入院日数が短いグループ)

	1日目	2日目	3-9日目	10-14日目	15日目	16日目	17日目
(Ns)NS測定(外項目)	○	○	○				
(Ns)体温・脈拍測定	○	○					
(Ns)呼吸数	○	○					
ME機器作動状態の確認	○	○	○				
SPO2			○	○			
その他症状			○	○			
体温・脈拍測定			○	○	○	○	○
呼吸			○	○			
嘔気・嘔吐	○	○	○	○	○	○	○
座位保持介助	○	○	○				
褥瘡			○	○			
排尿ナース			○	○			
排便回数			○	○			
歯磨き介助			○	○			
洗面			○	○			
消化・腹部症状				○			
留置時管理	○	○	○				
疼痛			○	○			
眼症状			○				
睡眠			○	○			
神経	○	○	○	○	○	○	○
血圧測定	○	○	○	○	○	○	○
起き上がり介助	○	○	○				
転倒防止?	○	○	○	○	○	○	○
転落防止?			○	○			
運動	○	○	○	○	○	○	○
食事摂取量	○	○	○	○	○	○	○

表 4 Group2 のデータ全体で生成された看護パス
(入院日数が長いグループ)

	1日目	2日目	3-8日目	9-20日目	21-42日目	43-48日目
(Ns)NS測定(外項目)	○	○	○	○		
(Ns)体温・脈拍測定	○	○	○			
(Ns)呼吸数	○	○	○			
ME機器作動状態の確認	○	○	○	○		
SPO2			○	○		
その他症状			○	○		
体位の調整					○	
体位変換			○	○	○	
体温・脈拍測定			○	○	○	
口腔清拭			○	○		
呼吸			○	○		
嘔気・嘔吐			○	○	○	
室内環境調整			○	○		
座位保持介助	○	○	○	○		
褥瘡			○	○		
排尿ナース			○	○	○	
排便回数			○	○	○	
日常生活習慣の獲得を促す?					○	○
更衣			○	○		
歯磨き介助			○	○		
洗面			○	○	○	
消化・腹部症状			○	○		
留置時管理			○	○		
疼痛			○	○	○	
眼症状			○	○		
睡眠			○	○		
神経			○	○	○	
移乗介助(車椅子)					○	
経口吸引					○	
義歯洗浄			○	○		
義歯装着					○	○
自己抜去防止?			○			
血圧測定			○	○	○	
起き上がり介助	○	○	○	○		
転倒防止?			○	○	○	○
転落防止?			○	○	○	○
運動			○	○	○	
浴部洗浄			○	○		
食事介助(ハイス)				○	○	
食事介助(ロウズ)			○	○		
食事摂取量	○	○	○	○	○	○
オムツ交換			○	○	○	
パット交換			○	○	○	

謝辞

この研究は日本学術振興会、科学研究費基盤(B)15H2750 の補助を受けた。

参考文献

- 1) 岩田春子、津本周作、“病院情報システムに蓄積された看護行為に関する時系列マイニング—看護パス作成支援に向けて—”、第31回 医療情報学会連合大会(第12回日本医療情報学会学術大会)、2011.
- 2) 岩田春子、津本周作、“病院情報システムに蓄積された看護行為に関する時系列マイニング—看護パス作成支援に向けて(その2)—”、第16回日本医療情報学会春季学術大会、2012.
- 3) 岩田春子、津本周作、“病院情報システムに蓄積された看護行為に関する時系列マイニング—看護パス作成支援に向けて(その3)—”、第32回 医療情報学会連合大会(第13回日本医療情報学会学術大会)、2012 .
- 4) 岩田春子、津本周作、“病院情報システムに蓄積された看護行為に関する時系列マイニング—看護パス作成支援に向けて(その4)—”、第35回医療情報学連合大会(第16回日本医療情報学会学術大会)、2015.
- 5) Iwata H, Hirano S and Tsumoto T, Maintenance and Discovery of Domain Knowledge for Nursing Care using Data in Hospital Information System, Fundamenta Informaticae 137, pp.237-252, 2015.
- 6) Tsumoto Y, Iwata H, Hirano S, Tsumoto T, Construction of Clinical Pathway Using Dual Clustering, Neuroscience and Biomedical Engineering, 3(1): 49 - 56, 2015.

- 7) McLachlan GJ and Peel D, Finite Mixture Models, Wiley, New York, 2000.
- 8) Everitt, BS, Landoau, S. and Leese, M., Cluster Analysis Fourth Edition, 2001.
- 9) Benaglia T , Chauveau D, Hunter DR, and Young D, mixtools: An r package for analyzing finite mixture models, Journal of

Statistical Software, 32(6):pp.1-29, 2009.

- 10) 津本周作, 平野章二, 岩田春子, 木村知広, “クラスタ分析に基づく同一 DPC コード内の症例の類型化”, 第2回人工知能学会医用人工知能研究会抄録集 (SIG-AIMED-002-12), 2016.

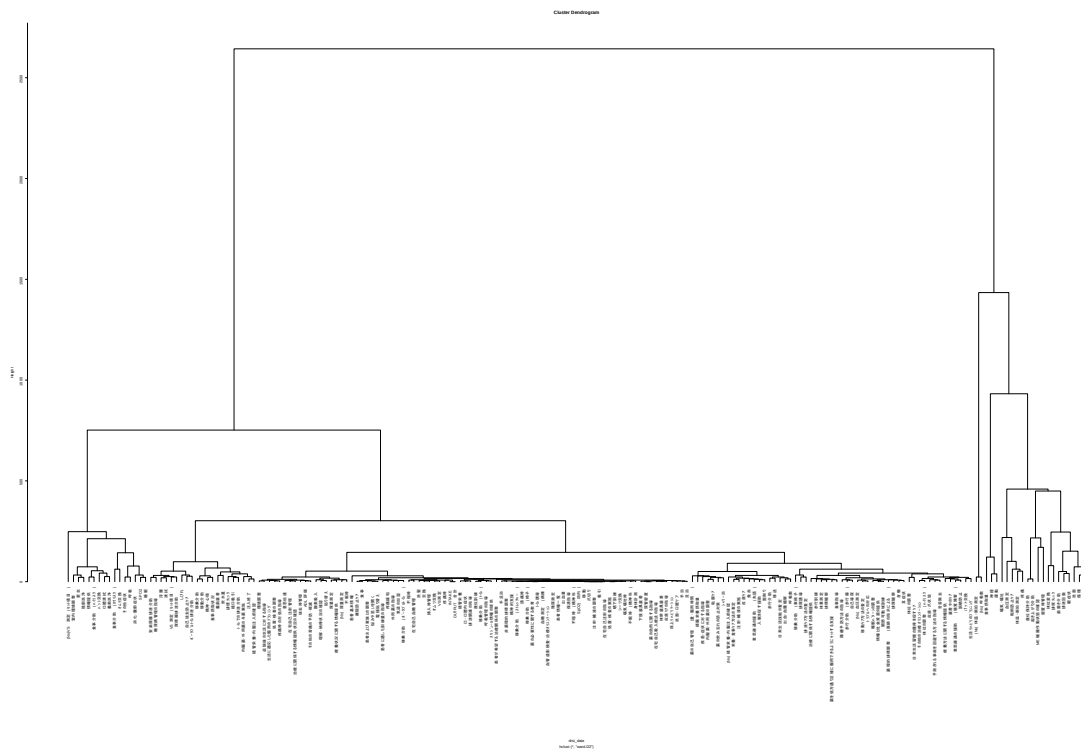


図5 データ全体を用いた場合の看護オーダーのデンドログラム

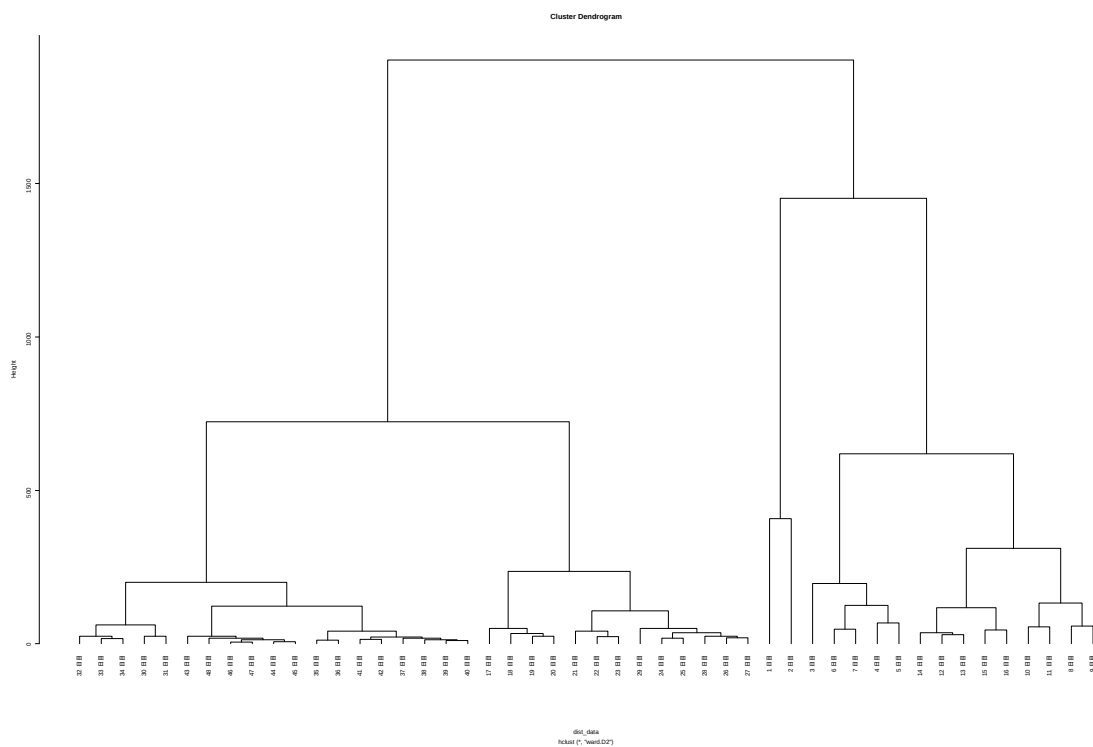


図5 データ全体を用いた場合の看護オーダーのデンドログラム

