



# 日本麻酔科学会第55回学術集会 共催セミナー22

**日時** 平成20年6月13日(金) 12:00~13:00

**会場** パシフィコ横浜 第10会場(414-415号室)

**座長**

**河本 昌志 先生**

(広島大学大学院医歯薬学総合研究科 病態制御医科学講座 麻酔蘇生学 教授)

**演題**

**医療用データマイニングツールOrangeによる  
データマイニングの実践:**

阪神大震災におけるデータに基づいた  
クラッシュ症候群リスクの予測モデル構築の例

**演者**

**青木 則明 先生**

(NPO法人 ヘルスサービスR&Dセンター)

**共催：社団法人 日本麻酔科学会  
フクダ電子株式会社**



# 医療用データマイニングツールOrangeによる データマイニングの実践:

## 阪神大震災におけるデータに基づいた クラッシュ症候群リスクの予測モデル構築の例

1) NPO法人 ヘルスサービスR&Dセンター 2) School of Health Information Sciences, University of Texas – Houston  
3) 東京大学大学院 医学系研究科 公共健康医学専攻 青木則明<sup>1,3</sup>、酒井未知<sup>1</sup>、高橋由光<sup>2</sup>、大田祥子<sup>1,2</sup>

本セミナーでは、医学におけるデータマイニング特長と注意点を説明した上で、  
医学用データマイニングツールOrangeを利用した  
阪神大震災のデータに基づくデータマイニング研究を紹介する。

### 1 データマイニング的アプローチへのニーズ

データマイニングは、「現存するデータ」を分析することで、データに含まれる情報の抽出やパターン認識を行い、新たな仮説の創出や、意思決定への活用を目的とするデータ分析の手法である。このようなアプローチは「探索的」と呼ばれ、1990年代後半までは、恣意的な結果を導くと考えられることも多かったが、近年、データ

量の爆発的な増加、コンピュータの処理能力の向上、「あるがままの現状」の分析へのニーズの増大、などの理由から、データマイニングは、医学研究に限らず、医療機関におけるマネジメントへの活用という視点でも注目を集めている。

### 2 医療におけるデータマイニングでの特徴<sup>1</sup>

医療におけるデータマイニングにおいては生物医学つまり病態生理学という知識基盤を十分に考慮する必要がある。データマイニングの結果、新たに発見されたルールが確立されている病態生理学と全く相反する結果をもたらした場合、それが全く新しい発見である可能性はゼロではないが、その解釈には慎重な姿勢が必要である。また、マイニングプロセスの自動化は特に必須の事項ではない。分析の過程で、ドメインエキスパートの意見を反映させたり、病態

生理学的に意味のある因子をあえて変数選択のプロセスで残したり、新たな変数を生成するなどといった作業も検討に値するプロセスである。

特に臨床データでは欠損値や外れ値の取り扱いも注意が必要である。臨床データマイニングの場合、欠損値に対して、クリニカルコンテキストから推測するという方法をとる意味がある場合も少なくはない。また、医学上、外れ値に意味がある可能性もある。

### 3 データマイニングツール『Orange』を用いた解析の紹介

Orangeは医療データ分析を目的として開発されたデータマイニングツールである。<sup>2</sup> 現時点でも開発・改訂が加わっているが、商業用のデータマイニングソフトに近いインターフェイスと分析機能を持っている。Orangeを用いたデータマイニングの一例として、阪神大震災のデータに対する研究を紹介する。<sup>3</sup>

阪神大震災でcrush injuryと診断された372名を対象に、重度の合併症である挫滅症候群 (crush syndrome) を起こす可能性のある患者を判別するための予後予測モデルをデータマイニング的アプローチに基づいて構築した。21種類の変数の中から、フィールドで入手可能な測定項目をロジスティック回帰モデルでモデル化した。フィールドで入手可能なデータに基づいたトリアージモデ

ルを表1に示す。この予測モデルのArea under the ROC curveは0.73であった。

表1. フィールド用 トリアージテーブル (J Traumaの論文<sup>3</sup>の表を改変)  
(挫滅症候群を起こす予測可能性 [95%信頼区間])

尿の色	脈拍数	救出までの時間	
		< 3hour	≥ 3hour
正常	< 120/min	4% [<9]	18% [10-25]
	≥ 120/min	1% [<4]	23% [11-34]
異常	< 120/min	16% [4-28]	45% [38-51]
	≥ 120/min	19% [5-32]	100% [>99]

### 4 おわりに

今後、データマイニング的アプローチの果たす役割は、さらに重要になってくると考えられる。治験に代表されるような、厳密にデザインされた前向き多施設臨床研究が、グローバルなエビデンスを創出していくために重要な役割を果たす一方、各地域や各施設、あるいは特定の集団に特有な意思決定を考えていくための仮説ある

いは根拠となる情報を提供するという意味で、後ろ向きのデータマイニングの重要性も増していくのではないかと考える。また、臨床データに加え、医療マネジメント上、有益な情報を抽出するという目的でもデータマイニングが果たす役割は大きいと考えられる。

1. 青木則明, 平憲司, 福井次矢. Applied Data Mining –臨床判断におけるデータマイニングの役割. BME 2002;16:37-43.

2. Demsar J, Zupan B, Kattan M, Aoki N, Beck JR. Orange and Decisions-at-Hand: Bridging Medical Data Mining and Decision Support. In: Proceedings in Intelligence Data Analysis in Medicine and Pharmacology (IDAMAP); 2001; London, UK; 2001. p. 87-92.

3. Aoki N, Demsar J, Zupan B, et al. Predictive model for estimating risk of crush syndrome: a data mining approach. J Trauma 2007;62:940-5.