

一般口演

## 一般口演18

## 医療データ分析6（機械学習・テキストマイニング1）

2018年11月25日(日) 09:00 ~ 10:30 D会場 (4F 413+414)

## [4-D-1-2] 病院情報システムから消化管穿孔症例を抽出するアウトカム定義の開発

○伊豆倉 理江子<sup>1</sup>, 野原 康伸<sup>1</sup>, 山下 貴範<sup>1</sup>, 野尻 千夏<sup>5</sup>, 枝廣 圭太郎<sup>2</sup>, 谷口 大介<sup>2</sup>, 高田 敦史<sup>1</sup>, 朴 珍相<sup>1</sup>, 若田 好史<sup>1</sup>, 神田橋 忠<sup>1</sup>, 横井 英人<sup>3</sup>, 宇山 佳明<sup>4</sup>, 中島 直樹<sup>1</sup> (1.九州大学病院 メディカル・インフォメーションセンター, 2.九州大学病院 消化器科・総合外科, 3.香川大学医学部附属病院 医療情報部, 4.独立行政法人 医薬品医療機器総合機構, 5.株式会社 ケア・フォー)

[目的] 医薬品の副作用検知のために2018年度から MID-NETシステムが稼働しており、本研究では、医薬品リスク管理計画上で多く挙がっている有害事象の中から消化管穿孔を選定し MID-NETで用いる症例抽出スクリプト（以下、アウトカム定義）を開発し、その精度を検証した。[方法] 2011年4月から2016年12月において当院の病院情報システムに登録のある入院症例を対象とした。専門医の見解を基に抽出に用いるデータ項目として「資源病名等の全 DPC傷病名（以下 DPC病名）」「CT・X線検査実施」「手術実施・胃管挿入」「抗菌薬注射実施」を選定した。DPC病名を中心にこれらを組み合わせた数種類のアウトカム定義を作成し、当院の病院情報システムデータベースから該当症例を抽出した。DPC病名单独からの抽出症例のうち100例をランダムサンプリングして、判定基準に従い2名の専門医が独立して診療録を確認することにより真偽判定した。各アウトカム定義において妥当性評価として陽性的中度（以下 PPV）を算出した。[結果・考察] 「DPC病名单独」の PPVは77.0%（当院の病院情報システムデータベースからの抽出症例数：n=201）と高く、「CT実施情報」や「抗菌薬実施情報」を組み合わせることで PPVは上昇し、さらに「手術情報」を加えると、PPVはさらに改善したが抽出症例数は減った。これは、炎症性腸疾患など腸の脆弱性に伴う穿孔では必ずしも穿孔特有の手術を実施するとは限らず、術式の選択範囲を狭くすることでこれらの穿孔症例が抽出されない可能性が示唆された。本研究でのアウトカム定義は確実な穿孔症例を特定したいか、あるいは穿孔であろう症例を少しでも多く取りたいか等、目的に応じて使い分けていくことが重要である。一方で、これらのアウトカム定義を複数施設で検討した場合、PPVに大きな違いが認められた。DPCの運用などが施設間での違いに影響しているものと考えられ、今後のアウトカム定義の作成において十分に検討する必要がある。

## 病院情報システムから消化管穿孔症例を抽出するアウトカム定義の開発

伊豆倉 理江子<sup>\*1</sup>, 野原 康伸<sup>\*1</sup>, 山下 貴範<sup>\*1</sup>, 野尻 千夏<sup>\*6</sup>, 枝廣 圭太郎<sup>\*2</sup>, 谷口 大介<sup>\*2</sup>, 高田 敦史<sup>\*1</sup>, 朴 珍相<sup>\*1</sup>,  
若田 好史<sup>\*4</sup>, 神田橋 忠<sup>\*1</sup>, 横井 英人<sup>\*3</sup>, 宇山 佳明<sup>\*5</sup>, 中島 直樹<sup>\*1</sup>

1\*九州大学病院 メディカル・インフォメーションセンター, 2\*九州大学病院 消化器科・総合外科,  
3\* 香川大学医学部附属病院 医療情報部, 4\*徳島大学病院 病院情報センター,  
5\*独立行政法人 医薬品医療機器総合機構, 6\*株式会社 ケア・フォー

## The development of phenotyping algorithm for identifying gastrointestinal perforation disease from hospital information system

First Author<sup>\*1</sup>, Second Author<sup>\*2</sup>, Third Author<sup>\*3</sup>

Rieko Izukura<sup>\*1</sup>, Yasunobu Nohara<sup>\*1</sup>, Takanori Yamashita<sup>\*1</sup>, Chinatsu Nojiri<sup>\*5</sup>, Daisuke Taniguchi<sup>\*2</sup>, Keitaro Eda<sup>\*2</sup>,  
Atsushi Takada<sup>\*1</sup>, Jinsang Park<sup>\*1</sup>, Yoshifumi Wakata<sup>1</sup>, Tadashi Kandabashi<sup>\*1</sup>, Hideto Yokoi<sup>\*3</sup>,  
Yoshiaki Uyama<sup>\*4</sup>, Naoki Nakashima<sup>\*1</sup>

\*1 Medical Information center, Kyushu University Hospital, \*2 Department of Surgery and Science, Kyushu University Hospital, \*3 Department of Medical Informatics, Kagawa University Hospital, \*4 Hospital Information Center, Tokushima University Hospital, \*5 Pharmaceuticals and Medical Devices Agency, \*6 Care four, Inc.

Abstract in English comes here.

**[Aim]** We validated a phenotyping algorithm that can be deployed to identify gastrointestinal perforation disease in MID-NET. **[Method]** It was used diagnosis diagnostic codes of procedure combination-based payment system (DPC), a conduct of CT scan or X-ray, surgical procedures or equipment and medications. Hundred samples were randomly selected from the extracted data were sample and reviewed by two surgical experts. It was assessed the positive predictive values (PPV). **[Result & Discussion]** Using only DPC diagnostic codes, the PPV was 77.0% (n=201). As the conduct of radiographic examination added DPC diagnostic codes, the PPV was 83.1%. The addition of surgical procedures was the highest PPV in any combinations. However, in the case of the gastrointestinal perforation-focused procedures, the PPV was slightly lower than in the holistic gastrointestinal surgery, from 90.0% to 85.7%. For example, in the perforation with severe inflammatory bowel diseases with intestinal vulnerability sometimes performed stoma surgery, so it would not be much more extracted the potential perforation cases by limiting the surgical procedure data. It is important that the phenotyping algorithm of gastrointestinal perforation disease should be determined by sensitivity-focused or PPV-focused researches.

**Keywords:** validation, phenotyping algorithm, gastrointestinal perforation, DPC, PPV

### 1. 緒論

医療情報データベース事業 (MID-NET) では、厚生労働省と PMDA を中心に医薬品の副作用検知のためのデータベース構築を進めてきた。平成 30 年度からの本格稼働に伴い、副作用調査対象となる病態を発症した症例を、病院情報システム (HIS) のデータから一定のルールで正確に把握される精度により抽出すること、つまり「Phenotyping」の確立が急がれる。

本研究の目的は、医薬品リスク管理計画の中から、報告例の多い有害事象の一つである消化管穿孔を対象とし、HIS の構造化データを用いて疾患の発症例を抽出するための phenotyping アルゴリズム (以下、アウトカム定義) を開発し、MID-NET への適用可能性を検討することである。

### 2. 方法

#### 2.1 対象

2011 年 4 月 1 日から 2016 年 12 月 31 日の約 5 年間に於いて、九州大学病院 (以下、本院) に入院した全患者 68,515 名 (131,002 件) を対象とした。

#### 2.2 データの属性

アウトカム定義の作成に用いた構造化データ種目を表 1 に示す。MID-NET での利活用を考慮して、消化管穿孔に関連した DPC 傷病名、腹部 CT または X 線撮影の実施・加算、注射オーダ (抗菌薬)、手術実施及び特定器材データを用いた。

表 1 アウトカム定義の作成に用いた構造化データ種目

データ種目	変数	データソース
傷病名	128 (例: 胃/腸穿孔、腸瘻)	DPC
放射線検査 (腹部CT/X線撮影、CT/X線読影料)	82	DPC/レセプト
抗菌薬オーダ	1568	SS-MIX2
胃管挿入 (特定器材)	24	DPC/レセプト
手術情報実施	限定: 6 (例: 胃/十二指腸穿孔縫合術など) 広範囲: 354 (限定3変数を含む。 例: 腸閉塞症手術、胃腸管切除術など)	DPC/レセプト

#### 2.3 アウトカム定義の開発手順

- 1) 専門医の見解やガイドライン等をもとに構造化データを用いた数種類のアウトカム定義を作成する。
- 2) 抽出した症例からランダムサンプリングした症例に対して、

2名の疾患専門医が独立して判定基準に従い、診療録などを確認して真偽判定を行う。判定者間で評価が異なる場合は対象疾患に対する診療経験数の長い専門医の判定を採用する。専門医間の一致度の評価に Kappa 係数を算出する。

3) アウトカム定義の妥当性の指標として陽性的中度 (以下、PPV) を算出する。

### 2.3 倫理的配慮

本研究は九州大学医学研究院倫理審査委員会により承認された (承認番号 29-167 番)。

## 3. 結果

### 3.1 アウトカム定義の作成

DPC 病名のみでの 1 種類の広義のアウトカム定義 (定義 1) と消化管穿孔の確定診断に必須である腹部 CT の実施<sup>1-3)</sup> や抗生物質処方などのデータ種目を組み合わせた 9 種類の狭義のアウトカム定義 (定義 2~10)、計 10 種類を作成した (表 2)。また、index date を対象期間内の入院日に設定し、1 ケースを 1 入院と定義した。

### 3.2 専門医による真偽判定

定義 1 より 201 件 (175 名) が抽出され、うち 100 件 (95 名) をランダムサンプリングし、index date を含む入院期間中において消化管穿孔を発症しているかどうかを、2 名の消化器疾患専門医が独立して判定基準 (図 1) に従い 4 段階判定を行った。2 者間で判定が異なった症例に対しては、消化管穿孔に対する診療経験数の長い専門医の判定を採用するとともに、kappa 係数を算出して評価の信頼性を確認した (Kappa=0.59)。

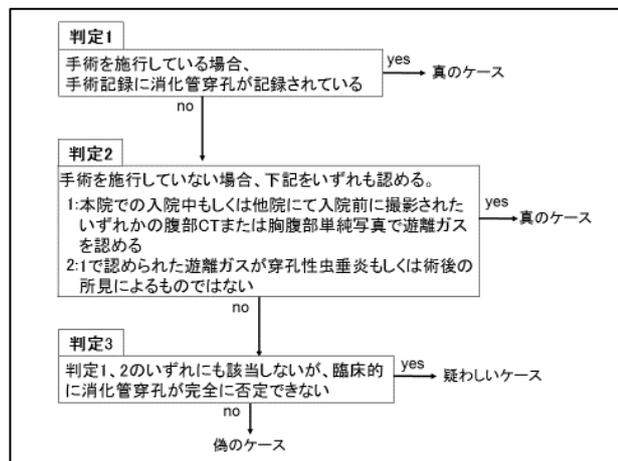


図 1 判定基準

### 3.3 抽出件数及び PPV

定義 1 のアウトカム定義において、傷病名のうち非外傷性腸穿孔 (ICD10: K631) が 206 件と最も多い該当数で、ランダムサンプリングに伴う真症例件数は 50 件中 39 件、PPV は 78.0%であった (表 3)。また、「入院の契機となった傷病名」の PPV が 87.8%と 6 種の DPC 病名分類の中で最も高く、「入院時併存症名」が 56.3%と最も低い精度であった (表 4)。

各アウトカム定義からの症例件数及び PPV を表 5 に示す。

DPC 傷病名のみを抽出条件とした場合 (定義 1)、ランダムサンプリング 100 件中真症例件数は 77 件、PPV は 77.0%であった。これに腹部 CT の実施や抗菌薬処方を加えた場合、100 件中 71 件と該当数は減ったが PPV は 83.1%まで上昇した (定義 2)。さらに、手術実施を加える (定義 3) と PPV は 90.0%に達した (該当数: 60 件、真症例件数: 54 件)。また、他院での CT 検査実施を想定して CT 読影料の有無を加えた場合 (定義 9) の PPV は 90.5% (該当数: 63 件、真症例件数: 57 件) であった。

なお、手術実施に加え胃管挿入 (特定器材) を条件に組み合わせても (定義 6)、胃管挿入を加えない定義 3 と該当症例件数、真症例数および PPV はほとんど変わらなかった (定義 3: PPV 90.0%, 定義 6: PPV 90.3%)。

手術実施において消化管穿孔に特化した術式に限定した場合 (定義 4)、該当数は 100 件中 7 件と広範囲に消化管手術の術式まで選定した場合 (定義 3) に比べ、該当数が 1/10 以下に減り、PPV も 85.7%とわずかに低下した。

## 4. 考察

本研究では、MID-NET に適用可能な構造化データを用いて消化管穿孔を抽出する 10 種類のアウトカム定義を開発し、各定義に対する精度を評価した。

DPC 病名分類の中で「入院時併存症名」に病名が登録されていた症例のほとんどが、他院での治療により症状が安定しているケースや既往症例で真の消化管穿孔症例ではなかった。穿孔を発症した場合、消化液や内容物の腹腔内漏出に伴う腹膜炎によりショック状態に陥るなど、重篤な合併症をきたすリスクが高く<sup>1)</sup>、何らかの緊急処置が優先的に実施されるケースが多い。本研究でも「主傷病名」や「入院の契機となった病名」の場合において PPV が高く、このような病態の観点からみても本結果は妥当であると考えられる。抽出条件として病態に応じて DPC 病名分類を予め選定することも一案である。

10 種類の定義の中で、DPC 病名のみを条件とした定義における PPV が最も低かったが、消化管穿孔の確定診断となる腹部 CT 検査実施<sup>2-3)</sup> や治療薬である抗菌薬処方といった疾患に特異的なデータ項目を抽出条件に加えることで、該当数は減るが PPV は改善した。

DPC 病名のみを条件とした定義からの抽出症例の中で偽症例と判定された 23 件のうち 5 件は、前医にて穿孔の確定診断及び初期治療が施され、緊急を脱した後に精査・加療のため本院に転院したケースであり、大学病院という本院の組織体制が症例抽出に関連していた可能性があった。また、ここでは言及しないが、同様のアウトカム定義を複数施設で検討した場合、PPV に大きな違いが認められ、DPC の運用などが施設間での違いに影響しているものと考えられた。このように、今後のアウトカム定義の作成において、施設の特性を踏まえた検討が必要である。

手術実施において腹腔鏡下穿孔縫合術など消化管穿孔に特化した術式に限定した場合、イレウス解除術など広範囲に消化管手術の術式まで選定した場合に比べ、PPV がわずかに低下した。炎症性腸疾患や腫瘍などでは腸管の脆弱性に伴い微小穿孔を生じることがあり<sup>4-8)</sup>、これが術中に発見され、穿孔修復処置も併用して実施されるケースがある。そのため、術式を限定することでこれらの真症例が抽出されない可能性がある。

また、保存的治療レベルの消化管穿孔症例まで広く特定

するため手術実施に胃管挿入を抽出条件に組み合わせても該当症例件数、真症例数および PPV はほとんど変わらなかった。これは、重症度の有無の判定には CT 検査が必須であり、胃管挿入を条件に加えなくとも CT 検査実施の条件で保存的治療の症例まで抽出されたものとする。

本研究で開発した 10 種類のアウトカム定義はいずれも PPV が 70%以上と決して精度が低いものではなかった。これらのアウトカム定義を利用するにあたり、確実な穿孔症例を特定したいか、あるいは穿孔であろう症例を少しでも多く取りたいか等、研究の目的に応じて使い分けていくことが重要である。

## 6. 結論

本研究で開発した 10 種類のアウトカム定義の精度はいずれも低いものではなかった。アウトカム定義の利用にあたり研究目的に応じて使い分けていくことが重要である。

## 7. 謝辞

本研究は、AMED 事業「アウトカム定義のバリデーション等に関する研究(番号:17mk0101088h0001) (代表:宇山佳明)」によって行った。関係各者に深謝する。

## 参考文献

- (1) 日本消化器病学会編. 消化性潰瘍診療ガイドライン 2015 改訂第 2 版. 南江堂, 2015.
- (2) Furukawa A, Sakoda M, Yamasaki M, et al (2005). Gastrointestinal tract perforation: CT diagnosis of presence, site, and cause. *Abdom Imaging* 2005 ; 30(5): 524-534.
- (3) Zissin R, Osadchy A, Gayer G. Abdominal CT findings in small bowel perforation. *Br J Radiol* 2009 ; 82(974):162-171
- (4) Merrell RC. The abdomen as source of sepsis in critically ill patients. *Crit Care Clin* 1995 ; 11(2): 255-272.
- (5) Roohullah A, Wong HL, Sjoquist KM. Gastrointestinal perforation in metastatic colorectal cancer patients with peritoneal metastases receiving bevacizumab. *World J Gastroenterol* 2015 ; 21(17): 5352-5338
- (6) Abdelrazeq AS, Scott N, Thorn C (2008). The impact of spontaneous tumour perforation on outcome following colon cancer surgery. *Colorectal Dis* 2008 ; 10(8): 775-780.
- (7) 舟山裕士. 炎症性腸疾患の外科治療. 2010 年(平成 22 年)度前期日本消化器外科学会集, 日本消化器外科学会 2010 : 23-30
- (8) Greenstein AJ, Aufses AH Jr (1985). Differences in pathogenesis, incidence and outcome of perforation in inflammatory bowel disease. *Surg Gynecol Obstet* 1985 ; 160(1): 63-69.

表 2 開発したアウトカム定義

定義	傷病名	注射オータ	放射線検査実施				特定器材	手術実施	
	DPC	SS-MIX2	DPC/レセプト				DPC	DPC/レセプト	
	穿孔または穿孔に関連した病名	抗菌薬	腹部CT	X線撮影	CT読影料	X線読影料	胃管挿入	広範囲	限定
1	○								
2	○	○	○						
3	○	○	○					○	
4	○	○	○						○
5	○	○	○				○	○	
6	○	○	○				○		○
7	○	○	○	○			○	○	
8	○	○	○	○			○		○
9	○	○	○		○		○	○	
10	○	○	○		○		○		○

結合条件: データ種目内は or 条件、データ種目間は and 条件

表 3 DPC 病名のみアウトカム定義(定義 1)からの抽出における病名別分類集計

病名及びICD10	全体 件数	ランダムサンプリング				
		判定対象 件数	真 件数	偽 件数	PPV 割合(%)	95%信頼区間 [下限値-上限値]
DPC傷病名(包括)	201	100	77	23	77	[67.5-84.8]
K251 胃潰瘍 急性、穿孔を伴うもの	3	2	2	0	100	[100.0-100.0]
K255 胃潰瘍 慢性又は詳細不明、出血を伴うもの穿孔を伴うもの	32	6	4	2	66.7	[22.3-95.7]
K261 十二指腸潰瘍 急性、穿孔を伴うもの	32	7	7	0	100	[100.0-100.0]
K265 十二指腸潰瘍 慢性又は詳細不明、出血を伴うもの穿孔を伴うもの	22	4	2	2	50	[6.8-93.2]
K285 胃空腸潰瘍 慢性又は詳細不明、出血を伴うもの穿孔を伴うもの	2	1	1	0	100	[100.0-100.0]
K572 穿孔及び膿瘍を伴う大腸の憩室性疾患	42	9	8	1	88.9	[51.8-99.7]
K631 腸穿孔(非外傷性)	206	50	39	11	78	[64.0-88.5]
K632 腸瘻(孔)	50	14	10	4	71.4	[41.9-91.6]
P780 周産期の腸穿孔	3	1	0	1	0	[0.00-0.00]
S363 胃損傷	3	1	1	0	100	[100.0-100.0]
S364 小腸損傷	6	1	1	0	100	[100.0-100.0]
S367 腹腔内臓器の多発性損傷	8	2	2	0	100	[100.0-100.0]
S369 詳細不明の腹腔内臓器の損傷	8	3	1	2	33.3	[0.8-90.6]

DPC 傷病名: 主傷病名、入院の契機となった傷病名、医療資源を 2 番目に投入した傷病名、入院時併存症名、入院後発症疾患名のいずれかを含むものとする

95%信頼区間: Clopper-Pearson 正確信頼区間により算出

PPV: 陽性的中度

表 4 DPC 病名のみアウトカム定義(定義 1)からの抽出における DPC 病名分類集計

DPC病名分類	全体 件数	ランダムサンプリング				
		判定対象 件数	真 件数	偽 件数	PPV 割合(%)	95%信頼区間 [下限値-上限値]
主傷病名	90	53	44	9	83	[70.2-91.9]
入院の契機となった傷病名	78	41	36	5	87.8	[73.8-95.9]
医療資源を最も投入した傷病名	92	53	44	9	83	[70.2-91.9]
医療資源を2番目に投入した傷病名	54	26	15	11	57.7	[36.9-76.7]
入院時併存症名	32	16	9	7	56.3	[29.9-80.3]
入院後発症疾患名	66	29	22	7	75.9	[56.5-89.7]

95%信頼区間: Clopper-Pearson 正確信頼区間により算出

PPV: 陽性的中度

表 5 各アウトカム定義の抽出件数及び PPV

定義	全体 件数	ランダムサンプリング				95%信頼区間 [下限値-上限値]
		判定対象 件数	真 件数	その他 件数	PPV 割合(%)	
1	201	100	77	23	77	[67.5-84.8]
2	141	71	59	12	83.1	[72.3-91.0]
3	117	60	54	6	90	[79.5-96.2]
4	12	7	6	1	85.7	[42.1-99.6]
5	126	62	56	6	90.3	[80.1-96.4]
6	101	48	43	5	89.6	[77.3-96.5]
7	168	82	68	14	83	[73.0-90.3]
8	135	63	54	9	85.7	[74.6-93.3]
9	127	63	57	6	90.5	[80.4-96.4]
10	102	49	44	5	89.8	[77.8-96.6]

95%信頼区間: Clopper-Pearson 正確信頼区間により算出

PPV: 陽性的中度

