

原 著

消化器外科緊急手術における外科技術評価法 Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) および P-POSSUM の評価

熊本大学大学院医学薬学研究部消化器外科学¹⁾, 国立病院機構熊本医療センター外科²⁾

岩槻 政晃¹⁾²⁾ 芳賀 克夫²⁾ 片測 茂²⁾
池井 聰²⁾ 馬場 秀夫¹⁾

はじめに：イギリスで開発された外科技術評価法である Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (以下、POSSUM) および P-POSSUM の精度を、消化器外科緊急手術で検討する。方法：当科で行った消化器外科緊急手術症例 295 例を対象に、POSSUM および P-POSSUM の予測死亡率を算出した。Endpoint は原論文に従い、POSSUM では 30 日死亡、P-POSSUM では在院死亡とした。実際の死亡率と予測死亡率の比 (OE ratio) を各モデルの精度の指標とした。結果：全対象症例の術後合併症発生率は 20.7%、30 日死亡率は 6.5%、在院死亡率 7.1% であった。POSSUM は予測死亡率が 0~49% の低リスク群で実死亡率が 0.80% (2/249) であったにもかかわらず、予測死亡率は 13.3% (33/249) となり、過大に死亡率を予測した (OE ratio 0.061, 95% 信頼区間 0.015~0.25)。しかし、予測死亡率が 50~100% の高リスク患者では、実死亡率 37.7% (17/45)、予測死亡率 51.1% (23/45) と比較的正確に死亡率を予測した (OE ratio 0.74, 95% 信頼区間 0.46~1.12)。P-POSSUM では、予測死亡率が 0~59% の低リスク群では実死亡率が 2.2% (6/267) であるにもかかわらず、予測死亡率は 9.0% (24/267) であった (OE ratio 0.23, 95% 信頼区間 0.094~0.58)。予測死亡率が 60~100% の高リスク群では、実死亡率 55.6% (15/27)、予測死亡率 81.5% (22/27) であった (OE ratio 0.68, 95% 信頼区間 0.47~1.00)。考察：POSSUM および P-POSSUM では、低リスク患者の死亡率を過大に予測する傾向があり、我が国で使用する場合には注意を要する。

緒 言

病院間の外科技術水準の評価に術後死亡率を用いることは一見妥当性があるように思われるが、これではハイリスク患者を多く扱う病院や拡大手術を行う病院は不利となりうる。したがって、外科技術水準の評価には、患者のリスクと手術の質を補正する必要がある。イギリスで、Copeland ら¹⁾によって外科技術評価法 Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (以下、POSSUM) が開発された。その後 Whiteley ら²⁾は、POSSUM が術後

死亡率を過大に予測しうることから、その訂正版である P-POSSUM を発表した。これらイギリスで開発された POSSUM, P-POSSUM について、著者らは過去に予定消化器外科手術で検討しており³⁾、これらスコアは実際の死亡率よりも過大に死亡を予測しうることを報告した。今回は消化器外科緊急手術における、POSSUM, P-POSSUM の精度を検討したので報告する。

対象と方法

1. 対象：2001年1月から2003年12月までに当科で行った消化器外科緊急手術症例 294 例 (男性 172 例, 女性 122 例, 平均年齢 58.5 歳) を対象とした。緊急手術の定義は Copeland ら¹⁾に従い、受診後 24 時間以内に医学的必要性から緊急で

<2007年9月26日受理>別刷請求先：岩槻 政晃
〒860-8556 熊本市本荘 1-1-1 熊本大学大学院医学薬学研究部消化器外科学

Table 1 Indication for emergency operation in current subjects

Disease	N	Disease	N
Acute appendicitis	78	Incarcerated hernia	32
non-perforated	56	inguinal	10
perforated	22	femoral	8
Gastrointestinal perforation	71	umbilical	3
duodenal ulcer	33	incisional	4
gastric ulcer	9	obturator	7
perforation of large bowel	11	Blunt injury	10
perforation of small intestines	8	spleen	6
others	10	pancreas	3
Intestinal obstruction	68	liver	1
colorectal cancer	17	Gastrointestinal hemorrhage	9
strangulating obstruction	24	duodenal ulcer	3
obstruction due to adhesion	14	gastric ulcer	2
others	6	intestinal neoplasm	4
Acute intestinal ischemia	7	Others	26

行った手術とした。疾患および術式の内訳を、Table 1, 2 に示す。

2. 方法：全症例で POSSUM の Physiological score (PS) および Operative severity score (OSS) の項目 (Table 3) を prospective に入力した。以下の予測式に PS, OSS を代入し POSSUM および P-POSSUM の予測死亡率を算出した。実際の死亡率を予測死亡率で割った比 (OE ratio) を医療の質の指標とし、その 95% 信頼区間は Katz 法で算出した。POSSUM では exponential analysis で、P-POSSUM では linear analysis で求めた⁴⁾。end-point は原論文に従い、POSSUM は 30 日死亡、P-POSSUM は在院死亡とした。

POSSUM の予測 30 日死亡率 (R)¹⁾,

$$\ln[R/(1-R)] = -7.04 + (0.13 \times PS) + (0.16 \times OSS).$$

P-POSSUM の予測在院死亡率 (R)²⁾,

$$\ln[R/(1-R)] = -9.37 + (0.19 \times PS) + (0.15 \times OSS).$$

結 果

緊急手術症例 294 例中、術後合併症発生は 61 例 (20.7%) にみられた (Table 4)。その中で、30 日以内死亡は 19 例 (6.5%)、在院死亡は 21 例 (7.1%) にみられた。

また、POSSUM で予測 30 日以内死亡率の階層別に OE ratio を exponential analysis で調べる

と、予測死亡率が 0~49% の低リスク群では、実死亡数が 2 例に過ぎないにもかかわらず、予測死亡数が 33 例と死亡率を過大に予測することが判明した。一方、予測死亡率が 50~100% の高リスク群では、OE ratio が 0.74 で予測死亡数は実死亡数に近い値を示した (Table 5)。P-POSSUM で予測在院死亡率の階層別に OE ratio を linear analysis で求めると、POSSUM と同様に、予測死亡率が 0~19% の低リスク群では、実死亡例がないにもかかわらず、予測死亡数は 10 例と過大に死亡率を予測した (Table 6)。予測死亡率が 60~79%, 80~100% の高リスク群では、OE ratio は 0.7 前後で、予測死亡数は実死亡数に近い値を示した。一方で、急性虫垂炎の症例のみを対象にした場合、POSSUM, P-POSSUM とともに、実死亡数は 1 例のみであるにもかかわらず、予測死亡数がおのおの 6 例、4 例と死亡率を過大に予測した (Table 7, 8)。

さらに、消化管穿孔例のみを対象とした場合は、POSSUM では予測死亡率が 0~49% の低リスク群では、実死亡数がないにもかかわらず、予測死亡数が 3 例であった。予測死亡率が 50~100% の高リスク群では、OE ratio は 0.89 で、予測死亡数は実死亡数に近い値を示した (Table 9)。P-POSSUM でも同様に、予測死亡率が 0~19% の低リスク群では、実死亡例がないにもかかわらず、予測死亡数は 5 例と予測された。予測死亡率が

Table 2 Surgical procedure of the subjects

Procedure	N
Appendectomy	73
Omentopexy for perforated peptic ulcer	47
Hernia repair	32
Large bowel resection	25
Small bowel resection	23
Open drainage only	21
Adhesiotomy without resection	20
Ileostomy, colostomy without resection	19
Others	34
Total	294

Table 3 Physiological score (PS) and Operative severity score (OSS)

Physiological score (PS)	Operative severity score (OSS)
Age	Operative severity
Cardiac signs	Multiple procedures
Respiratory history	Total blood loss
Blood pressure (systolic)	Peritoneal soiling
Pulse	Presence of malignancy
Glasgow coma score	Mode of surgery
Haemoglobin	
White cell count	
Urea	
Sodium	
Potassium	
Electrocardiogram	

20~39%, 40~59%, 60~79%, 80~100%の高リスク群では、予測死亡数は実死亡数に比較的近い値を示した (Table 10)。全症例で OE ratio (95%信頼区間) を求めると、POSSUM は 0.34 (0.21~0.56)、P-POSSUM は 0.46 (0.28~0.75) であった。しかし、虫垂炎、ヘルニア嵌頓の小手術症例を除外した場合は、POSSUM は 0.64 (0.36~1.12)、P-POSSUM は 0.51 (0.31~0.85) であった。

考 察

イギリスでは外科技術水準を客観的に評価し、それを公開することを surgical audit と呼んでいる。その評価基準要因として術後合併症発生率や死亡率を用いるのは一見妥当のようであるが、誤った結論を導く可能性があり、危険である。イ

Table 4 Main complication in patients with post-operative morbidity

Main Complication*	N
Wound infection	20
Organ/space infection without evident leakage	3
Pneumonia	5
Urinary tract infection	2
Septicaemia	7
Superficial wound dehiscence	1
Anastomotic leakage	6
Deep vein thrombosis	1
Acute myocardial infarction	1
Heart failure	3
Others	12
Total	61

*indicates most severe complication of the complications in the complicated patients.

ギリスでは、1980年代に政府の下部機関である The National Confidential Enquiry into Patient Outcome and Death (NCEPOD) が、病院別の術後粗死亡率を公表したが、多くの外科医から猛烈な反発を招いた。術後死亡率を技術評価の指標に用いると高リスク患者を多く扱う施設や拡大手術を多く行う施設は不利となるからである。したがって、外科技術水準の評価には、患者のリスクと手術の質を補正する必要性が強調された。そのような中で、1991年に Copeland ら¹⁾は術後死亡率を補正する目的で、POSSUM scoring system を開発した。施設の技術評価の指標としては、実際の死亡率を予測死亡率で割った比 (OE ratio) を用いている。その後、Whiteley ら²⁾は POSSUM が鼠径ヘルニアなどの低リスク手術で術後死亡率は過大に予測することを指摘し、その改訂版である P-POSSUM を発表した。しかし、Wijesinghe ら⁴⁾は、Whiteley らが POSSUM を本来の exponential analysis で解析せず、linear analysis で解析したために、誤った結果を得ていると指摘した。自らの血管外科手術症例で、POSSUM と P-POSSUM を linear analysis で解析すると、OE ratio はおのおの 0.59, 0.89 であったが、POSSUM を exponential analysis, P-POSSUM を linear analysis で解析すると、OE ratio はおのおの 1.14, 0.89 で

Table 5 Accuracy of POSSUM scoring system in patients who underwent emergency operation

Predicted risk of mortality (%)	N	Observed No. of deaths	Expected No. of deaths	OE ratio* (95%CI)
0-49	249	2	33	0.06 (0.015-0.25)
10-49	120	2	12	0.17
20-49	53	2	11	0.18
30-49	26	2	8	0.25
40-49	11	1	4	0.25

50-100	45	17	23	0.74 (0.46-1.12)
60-100	33	16	20	0.80
70-100	23	13	16	0.81
80-100	18	10	14	0.71
90-100	9	6	8	0.75
Total	294	19	56	0.34 (0.21-0.56)

*Observed to expected 30-day mortality ratio calculated by an exponential analysis.

Table 6 Accuracy of P-POSSUM scoring system in patients who underwent emergency operation

Predicted risk of mortality (%)	N	Observed No. of deaths	Expected No. of deaths	OE ratio* (95%CI)
0-19	230	0	10	0.00
20-39	25	3	8	0.38
40-59	12	3	6	0.50
60-79	10	5	7	0.71
80-100	17	10	15	0.67
Total	294	21	46	0.46 (0.28-0.75)

*Observed to expected in-hospital mortality ratio calculated by a linear analysis.

Table 7 Accuracy of POSSUM scoring system in patients who underwent emergency operation for appendicitis

Predicted risk of mortality (%)	No. of patients	Observed No. of deaths	Expected No. of deaths	OE ratio* (95%CI)
0-29	72	0	3	0.00
10-29	9	0	1	0.00
20-29	1	0	1	0.00

30-49	3	0	1	0.00
50-79	2	0	1	0.00
60-79	1	0	1	0.00

80-100	1	1	1	1.00
90-100	0	0	0	0.00
Total	78	1	6	0.17 (0.021-1.35)

*Observed to expected 30-day mortality ratio calculated by an exponential analysis.

あったと報告している。すなわち、正しい解析法を用いれば、POSSUMは死亡率を過大予測することはないと述べている。POSSUMおよびP-

POSSUMの精度については、現在なお論争中である。本研究では、上記論争を踏まえ、POSSUMはexponential analysisで、P-POSSUMはlinear

Table 8 Accuracy of P-POSSUM scoring system in patients who underwent emergency operation for appendicitis

Predicted risk of mortality (%)	No. of patients	Observed No. of deaths	Expected No. of deaths	OE ratio* (95%CI)
0-19	73	0	1	0.00
20-39	2	0	1	0.00
40-59	2	0	1	0.00
60-79	0	0	0	—
80-100	1	1	1	1.00
Total	78	1	4	0.25 (0.29-2.19)

*Observed to expected in-hospital mortality ratio calculated by a linear analysis.

Table 9 Accuracy of POSSUM scoring system in patients who underwent emergency operation for gastrointestinal perforation

Predicted risk of mortality (%)	No. of patients	Observed No. of deaths	Expected No. of deaths	OE ratio* (95%CI)
0-49	53	0	3	0.00
10-49	32	0	3	0.00
20-49	12	0	2	0.00
30-49	5	0	2	0.00
50-100	18	8	9	0.89
60-100	13	7	8	0.88
70-100	11	7	8	0.88
80-100	9	5	7	0.71
90-100	5	4	5	0.80
Total	71	8	12	0.67 (0.29-1.53)

*Observed to expected 30-day mortality ratio calculated by an exponential analysis.

Table 10 Accuracy of P-POSSUM scoring system in patients who underwent emergency operation for gastrointestinal perforation

Predicted risk of mortality (%)	No. of patients	Observed No. of deaths	Expected No. of deaths	OE ratio* (95%CI)
0-19	50	0	5	0.00
20-39	6	1	2	0.50
40-59	2	1	1	1.00
60-79	6	2	4	0.50
80-100	7	4	7	0.58
Total	71	8	19	0.42 (0.20-0.90)

*Observed to expected in-hospital mortality ratio calculated by a linear analysis.

analysis で解析した. その結果, 両者とも低リスク群で死亡率を過大予測する傾向がみられたが, 高リスク群では比較的正確に死亡率を予測できることが判明した. 今回, 我々が医学中央雑誌(1983~2007年, キーワードは「POSSUM」, 「緊急手術」)で

検索しえた詳細が明らかな文献報告では, 本研究のように必ずしもリスク別に POSSUM の精度を検討していない. また, exponential analysis で OE ratio を検討した報告もない. その中で, 松田ら⁵⁾は 80 歳以上の高齢者腹部緊急手術症例を解析した

が、POSSUM の予測死亡率が50%以上の高リスク患者では比較的よく死亡率を反映しているが、全体の症例では POSSUM は過剰に死亡率を予測していた。つまり、低リスク群では POSSUM は過剰に死亡を予測することが示唆される。このことは、本研究結果を支持するものであろう。

本研究では、POSSUM は低リスク患者では死亡率を過大に予測する傾向があることが示唆された。本研究の虫垂切除における POSSUM の予測30日以内死亡率は7.7%であったが、実死亡率は1.3%であった。一方、穿孔性腹膜炎手術では、POSSUM の予測30日以内死亡率は16.9%であったが、実死亡率は11.3%であった。したがって、POSSUM, P-POSSUM で、施設間の外科技術評価を行うとき、虫垂切除などの小手術を含めると、OE ratio は小さくなり、その施設の技術を過大評価する可能性がある。本研究でも当院の OE ratio (95% 信頼区間) は、0.34 (0.21~0.56) と小さい値を示したが、虫垂切除、ヘルニア根治術を除いた場合は、0.64 (0.36~1.12) であった。また、対象患者にこのような小手術を受けた患者を除外したとしても、バイタルサインが安定したリスクの良い腹膜炎手術症例などの低リスク群においても、POSSUM は死亡率を過大に予測するので、技術評価法として用いるには問題がある。本研究の穿孔性腹膜炎手術症例の中で、POSSUM の予測死亡率が49%以下の低リスク群となったものは71例中53例にも上り、P-POSSUM で予測死亡率が19%以下となったものは71例中50例にも上ったので、無視できる population ではない。

このような低リスク群での過大予測の原因は、POSSUM, P-POSSUM が開発された背景が一因であると考えられる。POSSUM, P-POSSUM は1990年代にイギリスで開発されたものであり、当時に比べ、手術手技の向上、周術期管理の進歩により手術成績は向上していると思われる。また、社会背景、患者背景、国間での手術手技、術後管理のレベルなどの相違も考慮する必要があると思

われる。したがって、低リスク群患者においても精度の高い新たなモデルの開発が望まれる。

本研究では、POSSUM および P-POSSUM が高リスク患者では術後死亡率を比較的良く反映することが示唆されたが、これらは術後経過を予測する prediction rule として用いることができるのだろうか。POSSUM, P-POSSUM の予測死亡率を算出するには、30日以内に行った手術回数を入力しなければいけない。また、悪性腫瘍の場合は、リンパ節転移の有無の病理組織学的診断を待たなければいけない。したがって、手術直後に予測死亡率を算出することはできない。しかし、絞扼性イレウスや上腸間膜動脈閉塞症、十二指腸潰瘍穿孔などの良性疾患では、30日以内に行った手術回数を1回と仮定すれば、手術直後に予測死亡率を算出することができる。当院では上記の緊急手術を行ったときは、術当日または翌日に POSSUM スコアをコンピュータに入力し、予測死亡率を算出しているが、術後経過を予測するのに的中率も高く、役に立っている。今後、我々は対象となる疾患や術式をふまえたうえで、本邦においても施設間の技術評価のみならず、より簡便で精度が高い予測モデルを開発する予定である。

文 献

- 1) Copeland GP, Jones D, Walters M et al : POSSUM : a scoring system for surgical audit. *Br J Surg* **78** : 356-360, 1991
- 2) Whiteley MS, Prytherch DR, Higgins B et al : An evaluation of the POSSUM surgical scoring system. *Br J Surg* **83** : 812-815, 1996
- 3) Haga Y, Wada Y, Takeuchi H et al : Estimation of physiologic and surgical stress (E-PASS) for a surgical audit in elective surgery. *Surgery* **135** : 586-594, 2004
- 4) Wijesinghe LD, Mahmood T, Scott JA et al : Comparison of POSSUM and the Portsmouth predictor equation for predicting death following vascular surgery. *Br J Surg* **85** : 209-212, 1998
- 5) 松田充宏, 渡辺義二, 唐司則之ほか : 80歳以上の高齢者に対する腹部緊急手術例の POSSUM スコアによる評価. *臨外* **54** : 399-402, 1999

Evaluation of Surgical Audit Systems 'POSSUM and P-POSSUM' in Emergency Digestive Surgery

Masaaki Iwatsuki¹⁾²⁾, Yoshio Haga²⁾, Shigeru Katafuchi²⁾,
Satoshi Ikei²⁾ and Hideo Baba¹⁾

Department of Gastroenterological Surgery, Graduate School of Medical Sciences, Kumamoto University¹⁾
Department of Surgery, National Hospital Organization, Kumamoto Medical Center²⁾

Background : We evaluated the accuracy of surgical auditing systems, POSSUM and P-POSSUM, developed in the United Kingdom, in emergency digestive surgery. **Patients and Methods** : Using POSSUM and P-POSSUM, we calculated expected mortality in 294 patients who underwent emergency digestive surgery from January 2001 to December 2003. In the original literature, the endpoint for POSSUM was defined as 30-day mortality and that for P-POSSUM as the in-hospital mortality. The ratio of observed to expected mortality rates (OE ratio) was defined as a measure of accuracy. **Results** : Overall morbidity was 20.7%, 30-day mortality 6.5%, and in-hospital mortality 7.1%. POSSUM overpredicted mortality rates at a low-risk band where expected rates were 0 to 49%-observed-rate 2/249, 0.80%, vs. expected rate 33/249, 13.3% : OE ratio 0.061 with 95% confidence intervals (CI) of 0.015–0.25. Expected rates were comparable to observed rates at a high-risk band where expected rates were 50 to 100%-observed-rate 17/45, 37.7% vs. expected rate 23/45, 51.1% : the OE ratio 0.74 with 95% CI of 0.46–1.12. Similarly, P-POSSUM overpredicted mortality rates at a low-risk band where expected rates were 0 to 59%-observed rate 6/267, 2.2% vs. expected rate 24/267, 9.0% : OE ratio 0.23 with 95% CI of 0.094–0.58. In contrast, expected rates were closer to the observed rates at a high-risk band where the expected rates were 60 to 100%-observed rate 15/27, 55.6% vs. expected rate 22/27, 81.5% : OE ratio 0.68 with 95% CI of 0.47–1.00. **Conclusion** : Current single-center data suggests that both POSSUM and P-POSSUM may overpredict mortality rates in low-risk patients who undergo emergency digestive surgery.

Key words : POSSUM, P-POSSUM, emergency digestive surgery

[*Jpn J Gastroenterol Surg* 41 : 373–379, 2008]

Reprint requests : Masaaki Iwatsuki Department of Gastroenterological Surgery, Graduate School of Medical Sciences, Kumamoto University
1-1-1 Honjo, Kumamoto, 860-8556 JAPAN

Accepted : September 26, 2007