

ĐÁNH GIÁ GIÁ TRỊ CỦA THANG ĐIỂM SORT TRONG TIÊN LƯỢNG BIẾN CHỨNG SAU PHẪU THUẬT LỚN

Trần Thị Vân¹, Nguyễn Hoàng Hải² và Vũ Hoàng Phương^{1,2,✉}

¹Bệnh viện Đại học Y Hà Nội

²Trường Đại học Y Hà Nội

Hàng năm, hơn 300 triệu cuộc phẫu thuật được tiến hành trên thế giới, vì thế cần phân tầng chính xác nguy cơ của người bệnh. Nghiên cứu nhằm đánh giá giá trị của thang điểm SORT trên người bệnh Việt Nam trong tiên lượng biến chứng sau phẫu thuật lớn. Chúng tôi tiến hành nghiên cứu tiền cứu mô tả trên 243 bệnh nhân phẫu thuật lớn tại Bệnh viện Đại học Y Hà Nội từ 1/4/ 2023 đến tháng 31/10/2023. Khả năng phân độ của SORT được đánh giá bằng test Hosmer - Lemeshow. Khả năng phân biệt của SORT được đánh giá bằng cách tính diện tích dưới đường cong AUROC. Tỷ lệ tử vong là 5,3%. Tỷ lệ tử vong dự báo theo SORT là 12,9%. Test Hosmer-Lemeshow cho kết quả $p < 0,001$. Diện tích dưới đường cong AUROC là 0,869. SORT có khả năng dự báo các biến chứng nhiễm trùng (AUC 0,7274), viêm phổi (AUC 0,7809), suy hô hấp (AUC 0,8216), ARDS (AUC 0,8012), chảy máu sau phẫu thuật (AUC 0,7266), tổn thương thận cấp (AUC 0,8753). Kết quả cho thấy mặc dù thang điểm SORT có giá trị tiên lượng tử vong cho bệnh nhân sau phẫu thuật lớn với khả năng phân biệt tốt nhưng SORT dự đoán quá mức tỷ lệ rủi ro.

Từ khóa: Phẫu thuật lớn, tử vong, SORT.

I. ĐẶT VẤN ĐỀ

Hàng năm, hơn 300 triệu cuộc phẫu thuật được tiến hành trên toàn thế giới, tỷ lệ biến chứng sau mổ khoảng 15% và tỷ lệ tử vong chu phẫu khoảng 1 – 3%;¹ vì thế cần phải phân tầng chính xác nguy cơ của người bệnh để tận dụng tối đa các nguồn lực nhằm cải thiện kết quả chu phẫu. Một số công cụ phân tầng rủi ro chỉ gồm các biến số trước phẫu thuật, trong phẫu thuật, sau phẫu thuật: Thang điểm của Hiệp hội bác sĩ gây mê Hoa Kỳ về Tình trạng thể chất (ASA-PS);² Thang đo rủi ro phẫu thuật (SRS);³ Điểm mức độ nghiêm trọng về sinh lý và phẫu thuật cho việc đo lường tỷ lệ tử vong và bệnh tật (POSSUM) và phiên bản Portsmouth tiếp theo (P-POSSUM).^{4,5} Mỗi thang điểm đều có những

hạn chế nhất định. Năm 2014, Protopapa KL và cộng sự đã giới thiệu một công cụ phân tầng rủi ro mới SORT (The Surgical Outcome Risk Tool), cho phép đánh giá tốt hơn những người bệnh có nguy cơ biến chứng và tử vong sau phẫu thuật không tim ở người lớn, làm rõ các nguy cơ và thảo luận tốt hơn trước khi phẫu thuật về các rủi ro.⁶ SORT là một mô hình phân tích sử dụng 6 thông số: mức độ nghiêm trọng của phẫu thuật, ASA-PS, mức độ khẩn cấp của phẫu thuật, chuyên khoa phẫu thuật có nguy cơ cao (tiêu hóa, lồng ngực, mạch máu), tuổi và ung thư, từ đó dự đoán xác suất tử vong sau 30 ngày phẫu thuật. Tại Việt Nam, việc áp dụng các công cụ phân tầng rủi ro sau phẫu thuật chưa được tiến hành thường xuyên và cũng chưa có nghiên cứu nào đánh giá giá trị của bảng điểm SORT. Vì vậy, chúng tôi tiến hành nghiên cứu này nhằm: Đánh giá giá trị của thang điểm SORT trong tiên lượng biến chứng sau phẫu thuật lớn.

Tác giả liên hệ: Vũ Hoàng Phương

Bệnh viện Đại học Y Hà Nội

Email: vuhuangphuong@hmu.edu.vn

Ngày nhận: 18/09/2024

Ngày được chấp nhận: 09/10/2024

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP

1. Đối tượng

Tiêu chuẩn chọn mẫu

Tất cả người bệnh có tuổi ≥ 16 , được phẫu thuật tại Bệnh viện Đại học Y Hà Nội. Bệnh nhân có ít nhất 1 tiêu chí thuộc phân loại phẫu thuật lớn theo nghiên cứu của David Martin⁷: có bệnh lý nghiêm trọng kèm theo trước phẫu thuật; trong phẫu thuật có: kẹp mạch máu hoặc thiếu máu cục bộ cơ quan, mất máu nhiều, sử dụng noradrenalin liều cao, thời gian phẫu thuật kéo dài, truyền máu trong mổ; sau mổ: có đáp ứng viêm hệ thống, cần điều trị tích cực sau mổ.

Tiêu chuẩn loại trừ

Không đánh giá được tình trạng sống sót sau 30 ngày phẫu thuật.

2. Phương pháp

Thiết kế nghiên cứu

Nghiên cứu tiến cứu mô tả.

Thời gian, địa điểm nghiên cứu

Từ tháng 1/4/2023 đến tháng 10/2023 tại khoa Gây mê hồi sức và chống đau, Khoa Ngoại và Khoa Hồi sức, Bệnh viện Đại học Y Hà Nội.

Cỡ mẫu và chọn mẫu

Áp dụng công thức tính cỡ mẫu cho 1 mẫu, xác định 1 tỷ lệ:

$$n = \frac{Z_{1-\alpha/2}^2 p (1 - p)}{d^2}$$

Trong đó:

n là cỡ mẫu, $Z_{1-\alpha/2}$ là giá trị phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê, lấy $\alpha = 5\%$ thì $Z_{1-\alpha/2} = 1,96$. P là tỷ lệ bệnh nhân mắc biến chứng sau mổ. Theo Zabolotskikh và cộng sự, tỷ lệ này khoảng 7,8%; d là mức sai số tuyệt đối chấp nhận. Chọn $d = P/2$. Từ các dữ kiện trên, ta tính được $n = 181$ bệnh nhân. Giả sử rằng có khoảng 15% số trường hợp bị loại khỏi nghiên cứu, thì cỡ mẫu tối thiểu cần cho nghiên cứu là 213 bệnh nhân.

Cỡ mẫu thu thập thực tế trong nghiên cứu này là 243 bệnh nhân.

Tiêu chí nghiên cứu

Đặc điểm chung (tuổi, giới, ASA, thời gian nằm viện, thời gian nằm, hồi sức, chuyên ngành phẫu thuật, mức độ nghiêm trọng của phẫu thuật, mức độ khẩn cấp của phẫu thuật, ung thư); thang điểm SORT (Cách tính điểm và tỷ lệ tử vong theo SORT: Truy cập trang web <http://sortsurgery.com/> và điền giá trị tương ứng của các biến, hệ thống sẽ tổng kết và tính ra phần trăm tiên lượng tử vong); tỷ lệ tử vong theo phân nhóm bảng điểm SORT; tỷ lệ biến chứng sau mổ, khả năng phân độ của thang điểm, khả năng phân biệt của thang điểm.

Quy trình tiến hành nghiên cứu

Thu thập các chỉ số, biến số của tất cả các bệnh nhân nghiên cứu. Tính tỷ lệ tử vong dự đoán sau 30 ngày phẫu thuật theo thang điểm SORT.

+ Truy cập trang web <http://sortsurgery.com/index.php?>

Surgical Outcome Risk Tool v2 (SORT)

Main Group

Sub Group

Procedure Description

Severity ?
 Minor Intermediate Major Xmajor/complex

ASA-PS ?
 1 2 3 4 5

Urgency ?
 Elective Expedited Urgent Immediate

Thoracics, gastrointestinal or vascular surgery
 Yes No

Cancer ? **Age**
 Yes No <65 65-79 >79

Clinical Risk Assessment
 Please select the clinical estimate of 30-day mortality; this should ideally be an assessment made by senior clinicians in the Multi-disciplinary perioperative care team.

Hình 1. Giao diện phần mềm thang điểm SORT

Theo dõi và ghi lại tình trạng tử vong sau 30 ngày kể từ ngày phẫu thuật và các biến chứng sau mổ cho đến khi ra viện.

Xử lý và phân tích số liệu

Các số liệu nghiên cứu được xử lý và phân tích bằng phần mềm SPSS 20.0. Giá trị $p < 0,05$ được coi là sự khác biệt có ý nghĩa thống kê. Tính diện tích dưới đường cong AUROC để đánh giá khả năng phân biệt. Sử dụng kiểm định Hosmer-Lemeshow X2 test để đánh giá khả năng phân độ.

3. Đạo đức nghiên cứu

Nghiên cứu được sự đồng ý của Bộ môn Gây mê hồi sức - Trường Đại học Y Hà Nội, ban lãnh đạo Khoa Gây mê hồi sức và chống

đau - Bệnh viện Đại học y Hà Nội, bệnh nhân và người nhà bệnh nhân tham gia nghiên cứu.

III. KẾT QUẢ

1. Đặc điểm chung

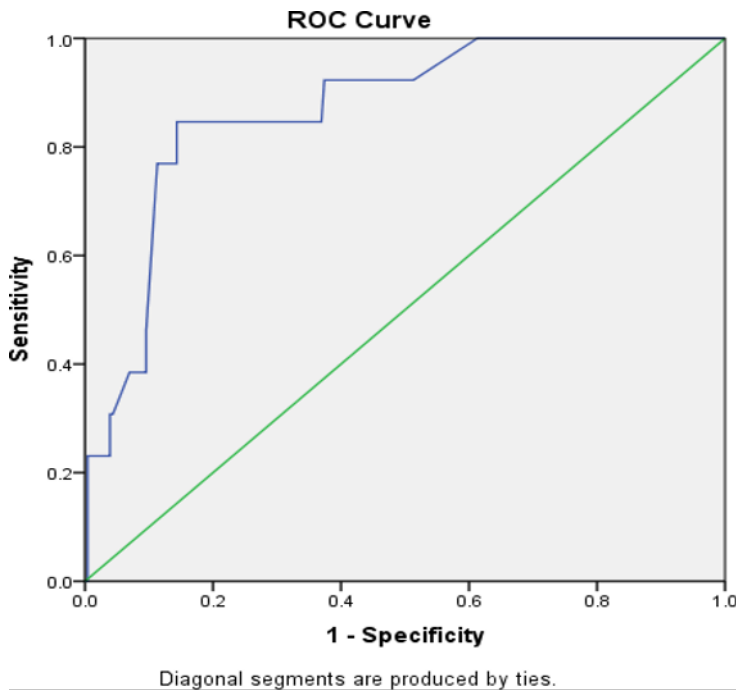
Các bệnh nhân trong nghiên cứu của chúng tôi có tuổi trung bình là $58,7 \pm 18,6$ tuổi (trong đó cao nhất là 99 tuổi, thấp nhất là 16 tuổi). Trong tổng số 243 bệnh nhân, 149 bệnh nhân là nam chiếm tỷ lệ 61% và 94 bệnh nhân là nữ chiếm tỷ lệ 39%. Phân loại ASA thường gặp nhất là mức độ III 47,7 %, thấp nhất là ASA I chiếm 5,8%.

2. Giá trị của thang điểm SORT trong tiên lượng biến chứng của phẫu thuật lớn:

Bảng 1. Điểm SORT ở nhóm sống và nhóm tử vong

Kết quả ĐT Thông số	Sống (n = 230)	Tử vong (n = 13)	p
SORT (%)	0,805	9,73	< 0,001
(trung vị, tứ phân vị)	[0,2 – 1,845]	[6,13 – 35,615]	Kiểm định Mann- Whitney
Thấp nhất – Cao nhất (%)	0,09 – 73,83	0,77 – 58,52	

Nhận xét: Điểm SORT ở nhóm tử vong cao hơn so với ở nhóm sống, khác biệt có ý nghĩa thống kê với $p < 0,001$.

**Biểu đồ 1. Diện tích dưới đường cong ROC của thang điểm SORT trong tiên lượng tử vong**

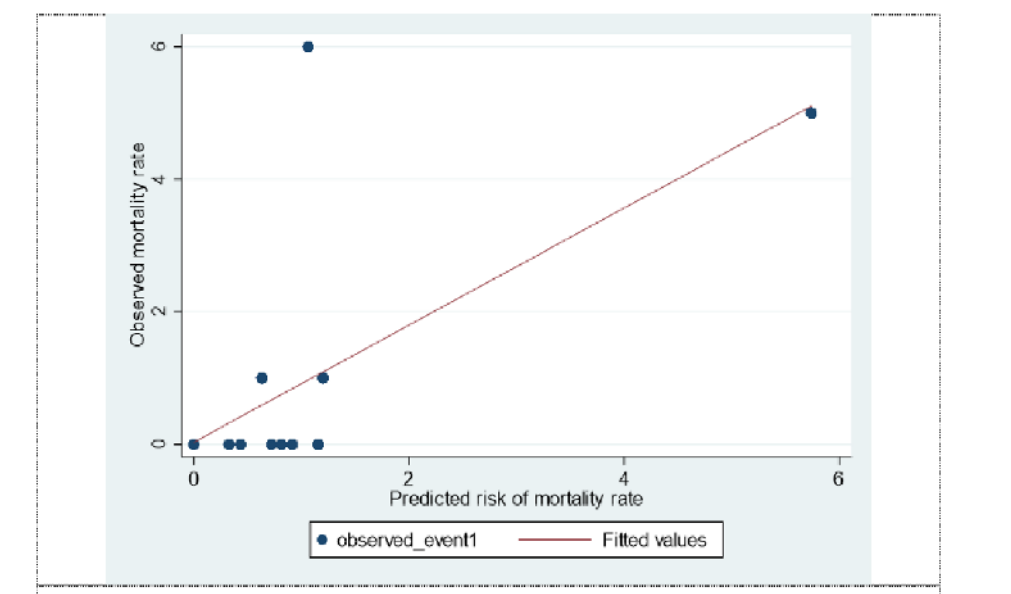
Nhận xét: Thống kê cho thấy thang điểm SORT khả năng tiên lượng tử vong với độ phân biệt tốt với AUROC là 0,869 (KTC 95%: 0,781 – 0,958) với $p < 0,001$.

Bảng 2. Tỷ lệ tử vong quan sát được so với dự đoán trong 30 ngày sau phẫu thuật ở các mức độ rủi ro khác nhau

Lượng tử	Số bệnh nhân	Xác suất tử vong ước tính trung bình theo SORT	Các trường hợp tử vong sau 30 ngày quan sát được, n (%)	Các trường hợp tử vong sau 30 ngày theo dự đoán	Hosmer - Lemeshow statistic
1	39	0,0296	0 (0%)	1,2	1,238095
2	11	0,0298	0 (0%)	0,3	0,308411

Lượng tử	Số bệnh nhân	Xác suất tử vong ước tính trung bình theo SORT	Các trường hợp tử vong sau 30 ngày quan sát được, n (%)	Các trường hợp tử vong sau 30 ngày theo dự đoán	Hosmer - Lemeshow statistic
3	24	0,0302	0 (0%)	0,7	0,72103
4	39	0,0309	1 (2,6%)	1,2	0,034392
5	14	0,0310	0 (0%)	0,4	0,411765
6	20	0,0319	1 (5%)	0,6	0,274914
7	28	0,0331	0 (0%)	0,9	0,929889
8	23	0,0371	0 (0%)	0,8	0,828829
9	21	0,0695	6 (28,6%)	1,1	23,03381
10	24	0,7885	5 (20,8%)	5,7	0,112741
Tổng	243		13 (5,3%)	12,9	

Nhận xét: Xác suất tử vong dự đoán trung bình dao động từ 2,96% đến 78,85%. Trong đa số các nhóm, SORT dự đoán quá mức tỷ lệ rủi ro.



Biểu đồ 2. Tỷ lệ tử vong trong 30 ngày được quan sát so với dự đoán ở các mức độ rủi ro khác nhau

Kiểm định Hosmer-Lemeshow cho kết quả $p < 0,001$ cho thấy tỷ lệ tử vong thực tế và tỷ lệ

tử vong ước tính trong mô hình chưa có sự phù hợp về mặt thống kê.

Bảng 3. Giá trị tiên lượng của thang điểm SORT trên các biến chứng

Biến chứng	AUROC	Standard error	95% CI
Nhiễm trùng	0,7274	0,0545	0,621 – 0,834
Viêm phổi	0,7809	0,0675	0,649 – 0,913
Nhiễm trùng vết mổ	0,5102	0,0392	0,433 – 0,587
Suy hô hấp	0,8216	0,0618	0,700 – 0,942
ARDS	0,8012	0,0703	0,663 – 0,939
Chảy máu sau phẫu thuật	0,7266	0,0681	0,593 – 0,860
Tổn thương thận cấp	0,8753	0,053	0,771 – 0,979

Nhận xét: Thang điểm SORT có khả năng phân biệt tốt các biến chứng suy hô hấp, ARDS, suy thận cấp; khả năng phân biệt trung bình với các biến chứng: nhiễm trùng, viêm phổi, chảy máu sau phẫu thuật; không có khả năng phân biệt với biến chứng nhiễm trùng vết mổ nông.

IV. BÀN LUẬN

Trong nghiên cứu của chúng tôi, điểm SORT trung bình là $4,77 \pm 10,72\%$, trung vị và tứ phân vị là $0,82\%$ [0,2 – 2,7], thấp nhất là $0,09\%$, cao nhất là $73,83\%$. Điểm SORT trong nghiên cứu của chúng tôi cao hơn các nghiên cứu khác trên thế giới. Wong và cộng sự nghiên cứu 22.631 bệnh nhân trên 18 tuổi trải qua phẫu thuật cần nằm viện ít nhất 1 đêm thu được kết quả điểm SORT trung vị, tứ phân vị là $0,4\%$ [0,2 – 1,6].⁸ Sự khác biệt này có thể được giải thích do đối tượng nghiên cứu của chúng tôi là bệnh nhân phẫu thuật lớn nên các thông số như ASA, mức độ nghiêm trọng của phẫu thuật đều cao hơn so với nghiên cứu của Wong và cộng sự.⁸ Điểm SORT ở nhóm tử vong là $9,73\%$ [6,13 – 35,615] cao hơn so với ở nhóm sống là $0,805\%$ [0,2 – 1,845], khác biệt có ý nghĩa thống kê với $p < 0,001$. Kết quả cũng phù hợp với kết quả nghiên cứu của Wong và cộng sự.⁸ Thang điểm SORT có giá trị tiên lượng tử vong cho bệnh nhân sau phẫu thuật lớn với khả năng phân biệt tốt với

diện tích dưới đường cong AUROC là $0,869$ (KTC 95%: $0,781 – 0,958$) với $p < 0,001$. Kết quả nghiên cứu của chúng tôi tương đồng với nhiều nghiên cứu trên thế giới. Nghiên cứu gốc của Protopapa phát triển thang điểm SORT, kết quả cho thấy thang điểm SORT có khả năng phân biệt rất tốt với AUROC là $0,91$ (KTC 95%: $0,88 – 0,94$; $p < 0,001$).⁶ Kiểm định Hosmer-Lemeshow cho kết quả $p < 0,001$ cho thấy tỷ lệ tử vong thực tế và tỷ lệ tử vong ước tính trong mô hình chưa có sự phù hợp về mặt thống kê. Kết quả của chúng tôi tương đồng với nghiên cứu của Wong và cộng sự tiến hành nghiên cứu trên 22.631 bệnh nhân trên 18 tuổi trải qua phẫu thuật cần nằm viện ít nhất 1 đêm, cho kết quả thang điểm SORT dự đoán quá mức rủi ro của bệnh nhân, kiểm định Hosmer-Lemeshow có $p < 0,001$.⁸ Năm 2021, Dimitrios E. Magouliotis và cộng sự đã tiến hành nghiên cứu đánh giá giá trị thang điểm SORT trong việc dự đoán tử vong sau phẫu thuật trên 526 bệnh nhân ung thư đại trực tràng tại Hy Lạp.⁹ Kết quả là thang điểm SORT có khả năng phân độ tốt. Trong tổng số 243 bệnh nhân nghiên cứu, có 115 bệnh nhân tương ứng $47,3\%$ xuất hiện biến chứng và có tổng 301 biến chứng. Kết quả này nhiều hơn so với kết quả nghiên cứu của Zabolotskikh và cộng sự tiến hành trên 1179 bệnh nhân phẫu thuật bụng theo chương trình, có 92 bệnh nhân

(7,8%) xuất hiện biến chứng và có tổng số 135 biến chứng.¹⁰ Nguyên nhân của sự khác biệt này là do nghiên cứu của Zabolotskikh và cộng sự tiến hành trên bệnh nhân mổ phôi và tiêu chuẩn loại trừ bao gồm¹⁰: mất máu khối lượng lớn cấp tính, co thắt phế quản, phản vệ và tăng thân nhiệt ác tính. Thống kê cho thấy thang điểm SORT có khả năng phân biệt tốt đối với biến chứng tổn thương thận cấp, với AUROC là 0,8753 (KTC 95%: 0,771 – 0,979). Kết quả này cao hơn kết quả nghiên cứu của Zabolotskikh và cộng sự có thể do mô hình bệnh tật và đối tượng nghiên cứu khác nhau. Biến chứng hô hấp xuất hiện ở 39 bệnh nhân, thống kê cho thấy thang điểm SORT có khả năng phân biệt tốt đối với biến chứng suy hô hấp, ARDS; khả năng phân biệt trung bình đối với biến chứng viêm phổi. Biến chứng nhiễm trùng xuất hiện trên 101 bệnh nhân, thống kê cho thấy thang điểm SORT có khả năng phân biệt trung bình đối với biến chứng nhiễm trùng, chảy máu sau phẫu thuật.

V. KẾT LUẬN

Nghiên cứu của chúng tôi trên 243 bệnh nhân phẫu thuật lớn cho thấy: Thang điểm SORT có giá trị tiên lượng tử vong cho bệnh nhân sau phẫu thuật lớn với khả năng phân biệt tốt với diện tích dưới đường cong là 0,869 (KTC 95%: 0,781 – 0,958), tuy nhiên thang điểm dự đoán quá mức tỷ lệ rủi ro với kiểm định Hosmer-Lemeshow cho kết quả $p < 0,001$. Thang điểm SORT có giá trị tiên lượng biến chứng tổn thương thận cấp, ARDS và suy hô hấp sau mổ cho bệnh nhân sau mổ phẫu thuật lớn với khả năng phân biệt tốt với diện tích dưới đường cong AUROC đều lớn hơn 0,8.

TÀI LIỆU THAM KHẢO

1. Weiser, T. G. et al. Estimate of the global volume of surgery in 2012: an assessment supporting improved health outcomes. *Lancet*

Lond. Engl. 2015. 385 Suppl 2, S11.

2. Cevik, B. & Yuce, Y. Risk Assessment in Surgical Patients: American Society of Anesthesiologist's (ASA) Classification vs Intraoperative Therapeutic and Diagnostic Interventions (I-ITS). *Am. J. Clin. Med. Res.* 2018. 6, 15-19.

3. Sutton, R., Bann, S., Brooks, et al. The Surgical Risk Scale as an improved tool for risk-adjusted analysis in comparative surgical audit. *Br. J. Surg.* 2002. 89, 763-768.

4. Copeland, G. P., Jones, D. & Walters, M. POSSUM: A scoring system for surgical audit. *Br. J. Surg.* 1991.78, 355–360.

5. Whiteley, M. S., Prytherch, D. R., Higgins, B., et al. An evaluation of the POSSUM surgical scoring system. *Br. J. Surg.* 1996. 83, 812–815.

6. Protopapa, K. L., Simpson, J. C., Smith, N. C. E. et al. Development and validation of the Surgical Outcome Risk Tool (SORT). *Br. J. Surg.* 2014. 101, 1774–1783.

7. Martin, D., Mantziari, S., Demartines, N. et al. Defining Major Surgery: A Delphi Consensus Among European Surgical Association (ESA) Members. *World J. Surg.* 2020. 44, 2211–2219.

8. Wong D. J. N. et al. Developing and validating subjective and objective risk-assessment measures for predicting mortality after major surgery: An international prospective cohort study. *PLOS Med.* 2020. 17, e1003253.

9. Magouliotis, D. E. et al. Validation of the Surgical Outcome Risk Tool (SORT) for Predicting Postoperative Mortality in Colorectal Cancer Patients Undergoing Surgery and Subgroup Analysis. *World J. Surg.* 2021. 45, 1940–1948.

10. I. B. Zabolotskikh, et al. Comparative evaluation of scales for predicting an unfavorable postoperative outcome: Preliminary results of the multicenter study 'The role of concomitant

diseases in the stratification of the risk of surgery STOPRISK, *Annals of Critical Care*, postoperative complications in abdominal 2022. no. 3, pp. 27–44.

Summary

ASSESSMENT OF SORT VALUE IN PREDICTING COMPLICATIONS AFTER A MAJOR SURGERY

As more than 300 million surgeries are performed worldwide yearly, it is necessary to accurately stratify patients's risk. The object of this study was to assess the performance of SORT in Vietnamese patients undergoing major surgery. We conducted a descriptive prospective study on 243 patients who underwent major surgery in Hanoi Medical University Hospital. Hosmer-Lemeshow test was used to assess calibration. Area under curve ROC was used to assess discrimination. Observed mortality was 5.3%. Mortality predicted by SORT was 12.9%. Calibration at the Hosmer-Lemeshow statistic was not good with $p < 0.01$. The AUC for SORT was 0.869. The SORT was predictive of infection (AUC 0.7274), pneumonia (AUC 0.7809), respiratory failure (AUC 0.8216), ARDS (AUC 0.8012), postoperative bleeding (AUC 0.7266), acute kidney injury (AUC 0.8753). This result shows that the SORT demonstrated good discrimination in predicting mortality, however mortality was over-predicted after major surgery.

Keywords: Major surgery, mortality, SORT.