

Predictive validity of American College of Surgeons: National Surgical Quality Improvement Project risk calculator in patients with ovarian cancer undergoing interval debulking surgery

Beryl Manning-Geist,^{1,2} Ann M Cathcart,³ Mackenzie W Sullivan,^{1,2} Andrea Pelletier,¹ Stephanie Cham,⁴ Michael G Muto,^{1,5} Marcela Del Carmen,² Whitfield B Growdon,² Rachel Clark Sisodia,² Ross Berkowitz,^{1,5} Michael Worley Jr^{1,5}

For numbered affiliations see end of article.

Correspondence to

Dr Beryl Manning-Geist, Department of Obstetrics, Gynecology and Reproductive Biology, Harvard Medical School, Boston, USA; ManningB@mskcc.org

Received 12 May 2021

Accepted 16 August 2021

Published Online First

13 September 2021

HIGHLIGHTS

- The ACS-NSQIP calculator was calibrated for post-operative complications among all patients undergoing interval debulking surgery.
- The ACS-NSQIP calculator did not accurately predict patient-specific risk of post-operative complications.
- The ACS-NSQIP calculator has limited utility in informing pre-operative counseling or surgical planning for interval debulking surgery.

ABSTRACT

Introduction In gynecologic patients, few studies describe the accuracy of the American College of Surgeons-National Surgical Quality Improvement Project (ACS-NSQIP) pre-operative risk calculator for women undergoing surgery for ovarian cancer.

Objective To determine whether the ACS-NSQIP risk calculator accurately predicts post-operative complications and length of stay in patients undergoing interval debulking surgery for advanced stage epithelial ovarian cancer.

Methods For this multi-institutional retrospective cohort study, pre-operative risk factors, post-operative complication rates, and Current Procedural Terminology codes were abstracted from records of patients with ovarian cancer managed with open interval debulking surgery from January 2010 to July 2015. A power calculation was done to estimate the minimum number of complications needed to evaluate the accuracy of the ACS-NSQIP risk calculator. Predicted risk compared with observed risk was calculated using logistic regression. The predictive accuracy of the ACS-NSQIP risk calculator in estimating post-operative complications or length of stay was assessed using c-statistics and Briar scores. Complications with a c-statistic of >0.70 and Briar score of <0.01 were considered to have high discriminative ability.

Results A total of 261 patients underwent interval debulking surgery, encompassing 21 unique Current Procedural Terminology codes. Readmission (n=25), surgical site infection (n=35), urinary tract infection (n=12), and serious post-operative complications (n=57) met the minimum event threshold (n>10). All predicted complication rates fell within the IQR of the observed incidence rates. However, the ACS-NSQIP calculator demonstrated neither discriminative ability nor accuracy

for any post-operative complications based on c-statistics and Briar scores. The calculator accurately predicted length of stay within 1 day for only 32% of patients and could not accurately predict which patients were likely to have a prolonged length of stay (c-statistic=0.65).

Conclusion Among patients undergoing interval debulking surgery, the ACS-NSQIP did not accurately discriminate which patients were at increased risk of complications or extended length of stay. The risk calculator should be considered to have limited utility in informing pre-operative counseling or surgical planning.

INTRODUCTION

Approximately 20 000 women will undergo upfront or interval debulking surgery for ovarian cancer in the United States this year.¹ For the surgeons caring for these women, estimating surgical risk is a vital component of informed consent, surgical planning, and post-operative care. Prediction tools such as the web-based American College of Surgeons-National Surgical Quality Improvement Project (ACS-NSQIP) risk calculator can help to generate patient and procedure-specific risks of post-operative complications, such as surgical site infection, length of stay, and readmission.

The ACS-NSQIP risk calculator has undergone limited evaluation in gynecologic oncology. Although 57.3% of patients included in the ACS-NSQIP are female and up to one-third of women will undergo hysterectomy by age 60, only 5.3% of cases included in the calculator are gynecologic patients. Two studies have investigated the ACS-NSQIP's accuracy in gynecologic surgery for benign and malignant



© IGCS and ESGO 2021. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Manning-Geist B, Cathcart AM, Sullivan MW, et al. *Int J Gynecol Cancer* 2021;**31**:1356–1362.

conditions, but neither study specifically evaluated women with ovarian cancer, a subset of patients that undergoes complex gynecologic surgeries.^{2,3} Furthermore, neither study offered data on the accuracy of the ACS-NSQIP risk calculator in patients with ovarian cancer triaged to neoadjuvant chemotherapy followed by interval debulking surgery, a group that is vulnerable to pre-operative factors omitted from the calculator, including anemia, frailty, and advanced disease at presentation.^{4,5}

Evidence suggests that rates of neoadjuvant chemotherapy followed by interval debulking surgery have increased in recent years, underscoring the urgency to validate current risk-based stratification tools that inform patient consent, surgical management, and post-operative care planning.⁵ This study evaluates whether the ACS-NSQIP risk calculator accurately predicts post-operative complications and length of stay in patients undergoing neoadjuvant chemotherapy followed by interval debulking surgery for advanced stage epithelial ovarian cancer.

METHODS

After institutional review board approval, a retrospective chart review was conducted of patients at Brigham and Women's Hospital and Massachusetts General Hospital undergoing neoadjuvant chemotherapy and interval debulking surgery for advanced-stage (International Federation of Gynecology and Obstetrics (FIGO) stage III C–IV) epithelial ovarian/fallopian tube/primary peritoneal carcinomas between January 1, 2010 and July 31, 2015, as previously published.⁶ Exclusion criteria were: low-grade serous histology (n=12), primary debulking surgery (n=240), or minimally invasive surgery (n=9). Decision to undergo neoadjuvant chemotherapy followed by interval debulking surgery, rather than primary debulking surgery, was based on surgeon discretion. Neoadjuvant chemotherapy regimens were platinum- and taxane-based and administered according to standardized protocols. The intention of neoadjuvant treatment was three to four cycles of chemotherapy prior to interval debulking surgery, at which time patients underwent a CT scan to determine resectability. If deemed unresectable, patients received additional cycles of chemotherapy.

To gather variables for calculator input, a retrospective review of the electronic medical record was conducted and integrated with a pre-existing surgical outcomes database. All risk factors included in the ACS-NSQIP definitions were used: age group (<65, 65–74, 75–84, ≥85); sex (male, female); functional status (independent, partially dependent, totally dependent) as extrapolated from Eastern Cooperative Oncology Group (ECOG) scores, with 0–2 classified as independent, 3 classified as partially dependent, and 4 classified as totally dependent; emergency case (yes, no); Anesthesia Society of America class (healthy patient, mild systemic disease, severe systemic disease, severe systemic disease/constant threat to life, moribund/not expected to survive surgery); steroid use for chronic condition (yes, no); ascites within 30 days prior to surgery (yes, no); systemic sepsis within 48 hours prior to surgery (none, systemic inflammatory response syndrome, sepsis, septic shock); ventilator dependent (yes, no); disseminated cancer (yes, no); diabetes (no, oral, insulin); hypertension requiring medication (yes, no); congestive heart failure in 30 days prior to surgery (yes, no); dyspnea (no, with moderate exertion, at rest); current smoker within 1 year (yes,

no); history of severe chronic obstructive pulmonary disease (yes, no); dialysis (yes, no); acute renal failure (yes, no); body mass index. Wound class and previous cardiac event are no longer included in the risk calculator, which differs from the calculator examined in prior studies.^{2,3}

Next, surgical procedures were reviewed, and Current Procedural Terminology codes assigned. In multiprocedure surgeries, such as radical debulking (58 953) and partial colectomy with primary anastomosis (44 140), the calculator was run for each code and the highest per cent predicted risk was selected. This methodology was consistent with prior studies.² Only planned surgical procedures were evaluated. Patient charts were reviewed to gather information on eight post-operative complications predicted by the ACS-NSQIP risk calculator (death, reoperation, readmission, cardiac event, renal failure, pneumonia, surgical site infection, urinary tract infection, venous thromboembolic event/pulmonary embolus) and length of stay. A composite of serious post-operative complications, defined by the ACS-NSQIP, was compiled, including cardiac arrest, myocardial infarction, pneumonia, progressive renal insufficiency, acute renal failure, venous thromboembolic event/pulmonary embolus, return to the operating room, deep incisional surgical site infection, organ space, surgical site infection, systemic sepsis, unplanned intubation, urinary tract infection, and wound disruption.⁷ To reach the minimum statistical threshold for evaluation of accuracy, a power calculation using a simple sample size calculation for univariate logistic regression was done. This determined that with a sample of 261, outcomes with <10 events (3.8%) would not have enough power (80%) to reach significance (p<0.05) in logistic regression, per Concato et al.⁸

Predicted risk was compared with actual risk of the outcome using logistic regression. The accuracy of the ACS-NSQIP risk calculator was assessed using the c-statistic and Brier score. The ability of a risk calculator to stratify patients as high or low risk, which is often relevant to surgical planning, relies on the model's discriminative ability, represented by a c-statistic. The usefulness of a risk calculator in counseling patients about risk in general is measured by the model's calibration. Brier scores combine both model discrimination and calibration to measure model accuracy. For the c-statistic score, a threshold of >0.7 was used to indicate a model with high discriminative ability. For the Brier score, a theoretical threshold of 0.01, or 90% accuracy, was used as the generally accepted threshold for a high predictive model.⁷ Stata version 15.0 was used to perform all statistical analyses (Cary, North Carolina, USA).

RESULTS

A total of 261 patients met inclusion criteria. **Table 1** presents demographics and clinical characteristics. Most patients were <65 (50.6%) or 65–74 (39.1%) years of age, median body mass index was 25.2 kg/m², and the majority of patients were white (87%). Overall, 92.7% had serous histology and 55.9% had stage III C disease. Most patients (97%) underwent three cycles of neoadjuvant chemotherapy (range 2–8 cycles) and 100.0% received a platinum/taxane doublet. A minority of patients had diabetes on medication/insulin (6.1%), hypertension requiring medication (38.3%), or active steroid use (2.7%). Most (64.8%) patients were American Society of

Original research

Table 1 Pre-operative demographics and clinical characteristics

Characteristics	Patients (n=261)	%
Age group (years)		
<65	132	50.6
65–74	102	39.1
75–84	22	8.4
>85	5	1.9
Functional status		
Independent	261	100.0
Partially dependent	0	0
Totally dependent	0	0
ASA class		
Healthy patient (ASA 1)	0	
Mild systemic disease (ASA 2)	169	64.8
Severe systemic disease (ASA 3)	92	35.2
Severe systemic disease/constant threat to life (ASA 4)	0	
Moribund (ASA 5)	0	
Steroid use for chronic condition		
No	254	97.3
Yes	7	2.7
High (≥ 4)	0	
Ascites		
No	170	65.1
Yes	91	34.9
Diabetes		
No	245	93.9
Yes-on oral agent	10	3.8
Yes-on insulin	6	2.3
Hypertension requiring medication		
No	161	61.7
Yes	100	38.3
Congestive heart failure within 30 days prior to surgery		
No	261	100.0
Yes	0	0.0
Dyspnea		
No	257	98.5
With moderate exertion	4	1.5
At rest	0	0.0
Smoker		
No	190	72.8
Yes	71	27.2
History of severe COPD		
No	259	99.2

Continued

Table 1 Continued

Characteristics	Patients (n=261)	%
Yes	2	0.8
Stage		
IIIC	146	55.9
IV	115	44.1
BMI (median, range); n=260	25.2 (22.5–28.4)	
Histology		
Serous	242	92.7
Mucinous	1	0.4
Endometrioid	2	0.8
Carcinosarcoma	5	1.9
Clear cell	4	1.5
Mixed	1	0.4
Operative time, min (median, range); n=65	152 (116–242)	
EBL, mL (median, range); n=259	300 (150–500)	
Surgical complexity		
Low	154	59.0
Moderate	90	34.5
High	17	6.5
Residual disease; n=260		
Complete surgical resection	166	63.8
Optimal (≤ 1 cm)	79	30.4
Suboptimal (> 1 cm)	15	5.8
Length of stay, days (median, range); n=261	5 (3–6)	
Post-operative complications		
Surgical site infection (deep incisional or organ space)	35	13.4
Systemic sepsis	0	0.0
Post-operative VTE/PE	9	3.4
Pneumonia	5	1.9
Unplanned intubation	0	0.0
Myocardial infarction/cardiac event	0	0.0
UTI	12	4.6
Acute renal failure/insufficiency	0	0.0
Readmission	25	9.6
Reoperation	4	1.5
Any serious complication	57	21.8

ASA, American Society of Anesthesiologists; BMI, body mass index; COPD, chronic obstructive pulmonary disease; EBL, estimated blood loss; UTI, urinary tract infection; VTE/PE, venous thromboembolic event/pulmonary embolus.

Table 2 Interval debulking surgeries

Operation	CPT* codes	Patients (n)
Total abdominal hysterectomy,±bilateral salpingo-oophorectomy,±nodal evaluation	58 951–58 954, 58 956	201
Bilateral salpingo-oophorectomy,±omentectomy,±nodal evaluation	58 943, 58 950	2
Colectomy with ostomy formation	44 143, 44 144, 44 150, 44 160, 44 320	11
Colectomy with anastomosis	44 140	19
Small bowel resection with anastomosis	44 120	3
Low anterior resection	44 145	4
Laparotomy with biopsy	49 000	2
Partial hepatectomy	47 120	3
Splenectomy	38 100, 38 102	13
Partial pancreatectomy	48 140, 48 146	2

*CPT, current procedural terminology.

Anesthesiologists (ASA) class 2. Overall, 59.0% of patients underwent a low complexity surgical operation, and 63.8% achieved complete surgical resection. Additional surgical characteristics are reported in [Table 1](#). In total, 201 women underwent abdominal hysterectomy, bilateral salpingo-oophorectomy, and radical debulking (Current Procedural Terminology codes 58 951–58 954, 58 956). Of 37 patients undergoing bowel resection, 11 had ostomy formation. Additional procedures and corresponding codes are detailed in [Table 2](#).

Post-operative complications included death (n=0), reoperation (n=4), readmission (n=25), cardiac event (n=0), renal failure (n=0), pneumonia (n=5), surgical site infection (n=35), UTI (n=12), VTE (n=9), as well as composite serious post-operative complications (n=57). Post-operative complications with >10 occurrences included readmission, surgical site infection, UTI, and serious post-operative complications. The accuracy of the risk calculator was investigated for all complications ([Table 3](#)). The median predicted risk of reoperation was 3.5%, whereas the incidence risk was 1.5%. Similarly, the median predicted risk for readmission was 10.6%, whereas the incidence risk was lower at 9.6%. Neither overestimation was significant (p=0.245 and p=0.283, respectively). The risk calculator was found to underestimate all other risks—however, all observed risks fell within the IQR of predicted risks ([Table 3](#)). The risk calculator was found to have high discriminative ability, as defined by a c-statistic >0.70 only for reoperation, however it is noted that this complication did not meet the prespecified minimum number of events. The risk calculator was not accurate for any outcome, as defined by a Brier score of <0.01. The risk calculator had neither discriminative ability (c-statistic=0.555) nor accuracy (Brier score=0.171) for estimating a composite of serous post-operative outcomes ([Figures 1 and 2](#)).

Post-operatively, the ACS-NSQIP risk calculator predicted a median length of stay of 6.0 days (IQR 5.0–7.0) compared with the actual length of stay of 5.0 days (IQR 3.0–6.0). The risk calculator overestimated the length of stay by >1.0 days in 51% (n=132) of patients and underestimated the length of stay by >1.0 days in 17% (n=45) of patients. To examine the accuracy of predicted length of stay, we then dichotomized patients into those staying >75% of the predicted length of stay (over 6 days) versus those staying ≤75th centile of the

median length of stay. Patients with a longer predicted length of stay based on the risk calculator had a statistically significant increase in staying >6 days (OR=1.39, 95% CI 1.17 to 1.66); however, this did not meet the c-statistic threshold for significance (c-statistic=0.65).

DISCUSSION

Summary of Main Results

The ACS-NSQIP risk calculator was appropriately calibrated for post-operative complications and length of stay but did not accurately discriminate which patients were at increased risk. In other words, ACS-NSQIP could generate an estimated magnitude of risk for patients undergoing interval debulking surgery, in general, but could not accurately predict which specific patients were at higher risk of a post-operative complication. The risk calculator should thus be considered to have limited utility in pre-operative counseling or surgical planning for interval debulking surgery. Unfortunately, there is a lack of other calculators that could readily be used for the purpose of risk stratification for patients with ovarian cancer. A nomogram generated from ACS-NSQIP data is available for risk of grade 4–5 complications, but similar tools do not exist for less serious, but still significant, complications, such as readmission or extended length of stay.⁹

Results in the Context of Published Literature

In 2013, the ACS-NSQIP introduced a pre-operative planning tool across an array of surgical specialties, including gynecologic oncology. This tool represented a milestone in systems-based practice and quality improvement, as it used robust datasets and statistical rigor to predict frequently encountered, costly post-operative complications. The ACS-NSQIP calculator was heralded as advancing risk stratification for three reasons. First, the calculator was based on a vast input of pre-operative risk factors and post-operative outcomes, including 1 414 006 surgeries. Second, large-scale, internal review using c-statistics and Brier scores validated the calculator's accuracy.^{7 10} Finally, multicenter external studies further validated the accuracy of the tool for an array of procedures, including colectomy and small bowel resection with reanastomosis.^{11 12}

Table 3 Comparison of calculated risk and outcomes for the NSQIP risk calculator in patients with ovarian cancer (n = 261)

Outcome	Did not have an event		Had an event		OR (95% CI)*	c-statistic*	Brier score†
	n (%)	Median risk (min-max)	n (%)	Median risk (Mmin-max)			
Serious complication	204 (78.2)	14.2 (11.5–18.0)	57 (21.8)	15.5 (11.3–19.6)	1.04 (0.98 to 1.10)	0.555	0.171
Pneumonia	256 (98.1)	1.2 (0.8–2.2)	5 (1.9)	1.9 (1.9–2.6)	1.21 (0.64 to 2.29)	0.616	0.019
SSI	226 (86.6)	6.6 (5.7–8.2)	35 (13.4)	7.4 (6.1–11.6)	1.19 (1.06 to 1.34)‡	0.628	0.117
UTI	249 (95.4)	3.2 (2.6–3.5)	12 (4.6)	3.4 (2.2–4.1)	1.03 (0.53 to 1.99)	0.529	0.044
VTE/PE	252 (96.6)	3.3 (2.7–4.0)	9 (3.4)	2.9 (2.9–3.1)	0.75 (0.36 to 1.56)	0.573	0.034
Reoperation	257 (98.5)	2.4 (2.0–3.1)	4 (1.5)	3.5 (2.5–4.9)	1.35 (0.81 to 2.26)	0.713	0.015
Readmission	236 (90.4)	10.9 (8.3–13.4)	25 (9.6)	10.6 (10.1–14.1)	1.07 (0.95 to 1.21)	0.548	0.087
Prolonged LOS§	200 (76.6)	5.5 (4.5–6.5)	61 (23.4)	6.5 (5.0–8.0)	1.39 (1.17 to 1.66)‡	0.645	0.206

*c-statistic: range 0.5–1.0, where 0.5 is no better than chance and 1.0 indicates that the model perfectly predicts the outcome. Models are considered reasonable when >0.7, and strong when >0.8.

†The Brier score describes the mean squared differences between the predicted risk and actual outcome. The Brier score is 0 when the model perfectly predicts the outcome.

‡P value < 0.05 for all outcomes.

§>75th centile or >6 days.

LOS, length of stay; SSI, surgical site infection; UTI, urinary tract infection; VTE/PE, venous thromboembolic event/pulmonary embolus.

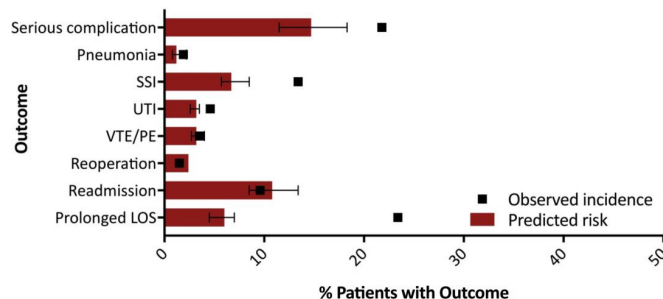


Figure 1 American College of Surgeons-National Surgical Quality Improvement Project (ACS-NSQIP) predicted risk and observed incidence of post-surgical outcomes for 261 patients. Solid bars represent median risk and error bars represent IQR. LOS, length of stay; SSI, surgical site infection; UTI, urinary tract infection; VTE/PE, venous thrombotic effect/pulmonary embolus.

Only two studies have analyzed the ACS-NSQIP risk calculator in patients with gynecologic oncology. One report found that the calculator accurately predicted death, cardiac complications, and renal failure, but did not accurately predict other outcomes.² Notably, these events occurred infrequently in that study, with 9 occurrences of death, 12 of cardiac complications, and 18 of renal failure.² Further, only 23.8% of patients in that study underwent debulking surgery for any malignancy. A second study evaluated 628 patients undergoing laparotomy and laparoscopy and found that while the calculator accurately predicted death, pneumonia, urinary tract infection, and venous thromboembolic event/pulmonary embolus among all patients with gynecologic cancer, it did not accurately predict complications among a subset of 94 patients with ovarian cancer.³ Again, these occurrences were infrequent, with 3 deaths, 7 cases of pneumonia, 11 cases of urinary tract infection, and 7 cases of venous thromboembolic event/pulmonary embolus.³ Both studies demonstrated some accuracy of the ACS-NSQIP calculator for gynecologic patients, but neither presented significant data on ovarian cancer outcomes, or interval debulking surgery.

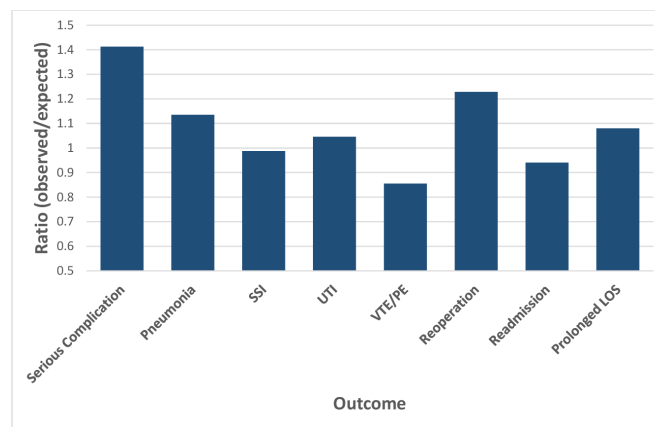


Figure 2 American College of Surgeons-National Surgical Quality Improvement Project (ACS-NSQIP) ratio of observed to predicted risk incidence of post-surgical outcomes for 261 patients. LOS, length of stay; SSI, surgical site infection; UTI, urinary tract infection; VTE/PE, venous thrombotic effect/pulmonary embolus.

In addition to the gap in ACS-NSQIP validation for ovarian cancer, to our knowledge, only one other study has specifically assessed the accuracy of the calculator in patients undergoing surgery after neoadjuvant treatment. This study in laryngeal cancer found that the ACS-NSQIP calculator performed well for patients receiving primary surgery, but lost calibration and discrimination abilities among patients undergoing surgery after neoadjuvant chemoradiation.¹³

In this study, the ACS-NSQIP calculator was found to be adequately calibrated—that is the observed risk fell within the range of predicted risks. However, the ACS-NSQIP did not accurately discriminate risk for any outcomes with >10 events meeting adequacy for analysis, including serious post-operative complications, surgical site infections, urinary tract infection, readmission, or prolonged length of stay. Thus, the calculator did not reliably predict which patients would be at higher risk of these complications based on their pre-operative risk factors.

There are several reasons why the ACS-NSQIP risk calculator might have demonstrated poor discriminative ability, including limitations due to homogeneous populations, lack of gynecologic cancer-specific data used to create the ACS-NSQIP risk calculator, and the fact that gynecologic cancer surgeries often span colorectal, hepatobiliary, and gynecologic boundaries. With regards to c-statistics, it should be noted that the c-statistic is affected by a model, and also by the homogeneity of cases in a sample. Discriminative models can be expected to perform more poorly as cases become more similar, and it is expected that c-statistics for a more homogeneous population will be smaller than those calculated for a heterogeneous population.^{14 15} It is difficult to disentangle the effects of case homogeneity from true inadequacy of the ACS-NSQIP calculator model specification for the patient subset included in the present study. It is also possible that the ACS-NSQIP calculator is poorly specified for those patients undergoing interval debulking surgery. The ACS-NSQIP calculator was developed from the NSQIP database, which does not distinguish between primary and interval debulking surgery. Further, the calculator used NSQIP data from January 2009 to June 2012, a time span when interval debulking surgery was employed in less than 25% of patients.⁵ Additionally, only 5.3% of operations used in the creation of the ACS-NSQIP risk calculator were gynecologic (74737 of the 1414006 surgeries included in the model), and the proportion of surgeries completed by gynecologic oncologists is not specified. Finally, the ACS-NSQIP risk calculator allows users to input only one Current Procedural Terminology code, potentially limiting a field where surgeries often include multiple operations. For example, interval debulking surgery may include a total abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy, and splenectomy, but because splenectomy yields the highest predicted risk, only this code is considered when calculating post-operative risks. Procedural input limitations thus disallow for the possibility of additive, or even synergistic, risk.

Strengths and Weaknesses

This study offers the only data available on the ACS-NSQIP risk calculator exclusive to ovarian cancer surgery. Further, it is one of the few studies on interval debulking surgery and represents a unique opportunity to offer insights applicable across surgical specialties. A strength of our study is that our entire cohort is made up of patients undergoing surgery for ovarian cancer: previous investigations

using ACS-NSQIP in gynecologic oncology have included other disease sites and surgery for benign disease.^{2 3} Furthermore, our data spans more than 10 surgeons across two institutions over half a decade. Although this study is limited by event frequency and limitations inherent to retrospective research in general, this study is unique from others in that it included a power calculation. While other studies reported the accuracy of the ACS-NSQIP in predicting outcomes with as few as three events, this study sought to abide by a minimum statistical threshold for evaluation of accuracy. Recent literature has suggested that external validation studies estimating performance measures using c-statistics and calibration require a minimum of 100 events for unbiased and precise estimation; the outcome with the greatest number of events in the present study fell short of this recommendation with 57 events.¹⁶ As such, this study did not reach the recommended minimum (100 events) for unbiased and precise estimation of performance using c-statistics and calibration.

Implications for Practice and Future Research

For patients triaged to neoadjuvant chemotherapy, the extent of baseline disease and/or medical co-morbidities are important considerations in pre-operative counseling and planning. Surgical risk calculators that consider both patient-specific and procedure-specific risk may aid in pre-operative counseling and planning. However, it is important that available risk stratification tools undergo independent validation. Results presented here suggest that the ACS-NSQIP surgical risk calculator should not be used to individualize counseling for expected risk of post-operative complications, readmission, or prolonged length of stay associated with interval debulking surgery. Instead, our findings underpin the need to develop accurate surgical risk calculators that can be used in this unique and complex patient population.

CONCLUSIONS

In summary, the ACS-NSQIP risk calculator did not accurately discriminate which patients were at increased risk of complications or extended length of stay following interval debulking surgery. Consequently, the risk calculator should be considered to have limited utility in informing pre-operative counseling or surgical planning, including decisions on extent of cytoreduction or the triaging of patients to minimally invasive cytoreductive options. The findings in this study underscore the need for a better validated surgical risk calculator to guide surgical decision-making among patients receiving neoadjuvant chemotherapy followed by interval debulking surgery.

Author affiliations

¹Department of Obstetrics and Gynecology, Brigham and Women's Hospital, Boston, Massachusetts, USA

²Division of Gynecologic Oncology, Vincent Obstetrics and Gynecology, Massachusetts General Hospital, Boston, Massachusetts, USA

³Harvard Medical School, Boston, Massachusetts, USA

⁴Division of Gynecologic Oncology, Brigham and Women's Hospital, Boston, Massachusetts, USA

⁵Dana-Farber Cancer Institute, Boston, Massachusetts, USA

Contributors Conceptualization: BM-G; data curation: BM-G, AMC; formal analysis: AP, AMC, BM-G; methodology: BM-G, AP, MW; roles/writing original draft: BM-G, AMC, MWS; writing – review and editing: all authors.

Original research

Funding This research was supported in part by award Number T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

Competing interests MW Jr.: CONMED Corporation (consulting and honoraria).

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data are available from the corresponding author upon reasonable request.

REFERENCES

- 1 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019;69:7–34.
- 2 Rivard C, Nahum R, Slagle E, et al. Evaluation of the performance of the ACS NSQIP surgical risk calculator in gynecologic oncology patients undergoing laparotomy. *Gynecol Oncol* 2016;141:281–6.
- 3 Szender JB, Frederick PJ, Eng KH, et al. Evaluation of the National Surgical Quality Improvement Program universal surgical risk calculator for a gynecologic oncology service. *Int J Gynecol Cancer* 2015;25:512–20.
- 4 Manning-Geist BL, Alimena S, del Carmen MG, et al. Infection, thrombosis, and oncologic outcome after interval debulking surgery: does perioperative blood transfusion matter? *Gynecol Oncol* 2019;153:63–7.
- 5 Horner W, Peng K, Pleasant V, et al. Trends in surgical complexity and treatment modalities utilized in the management of ovarian cancer in an era of neoadjuvant chemotherapy. *Gynecol Oncol* 2019;154:283–9.
- 6 Manning-Geist BL, Hicks-Courant K, Gockley AA, et al. A novel classification of residual disease after interval debulking surgery for advanced-stage ovarian cancer to better distinguish oncologic outcome. *Am J Obstet Gynecol* 2019;221:326.e1–326.e7.
- 7 Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. *J Am Coll Surg* 2013;217:833–42.
- 8 Concato J, Peduzzi P, Holford TR, et al. Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. *J Clin Epidemiol* 1995;48:1495–501.
- 9 Cham S, Chen L, St Clair CM, et al. Development and validation of a risk-calculator for adverse perioperative outcomes for women with ovarian cancer. *Am J Obstet Gynecol* 2019;220:571.e1–571.e8.
- 10 Liu Y, Cohen ME, Hall BL, et al. Evaluation and enhancement of calibration in the American College of Surgeons NSQIP surgical risk calculator. *J Am Coll Surg* 2016;223:231–9.
- 11 Hu W-H, Chen H-H, Lee K-C, et al. Assessment of the addition of hypoalbuminemia to ACS-NSQIP surgical risk calculator in colorectal cancer. *Medicine* 2016;95:e2999.
- 12 Mogal HD, Fino N, Clark C, et al. Comparison of observed to predicted outcomes using the ACS NSQIP risk calculator in patients undergoing pancreaticoduodenectomy. *J Surg Oncol* 2016;114:157–62.
- 13 Cao A, Khayat S, Cash E, et al. ACS NSQIP risk calculator reliability in head and neck oncology: the effect of prior chemoradiation on NSQIP risk estimates following laryngectomy. *Am J Otolaryngol* 2018;39:192–6.
- 14 Merkow RP, Hall BL, Cohen ME, et al. Relevance of the C-statistic when evaluating risk-adjustment models in surgery. *J Am Coll Surg* 2012;214:822–30.
- 15 Cohen ME, Liu Y, Ko CY, et al. An examination of American College of Surgeons NSQIP surgical risk calculator accuracy. *J Am Coll Surg* 2017;224:787–95.
- 16 Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016;35:214–26.