Multidimensional Analysis of the Learning Curve for Cytoreductive Surgery and Hyperthermic Intraperitoneal Chemotherapy in Peritoneal Surface Malignancies

Shigeki Kusamura, MD, PhD, Dario Baratti, MD, and Marcello Deraco, MD

Objective: To evaluate the learning curve of cytoreductive surgery (CRS) and hyperthermic intraperitoneal chemotherapy (HIPEC) in treating peritoneal surface malignancies (PSM).

Summary and Background: CRS and HIPEC to treat PSM is a complex procedure with a significant morbidity. A long-lasting training program is required to acquire expertise in this type of operation.

Methods: We performed CRS using peritonectomy procedures. HIPEC through the closed abdomen technique employed cisplatin and mitomycin-C or cisplatin and doxorubicin. Risk-adjusted sequential probability ratio test was used to assess the learning curve on a series of 420 cases of PSM on the basis of rates of incomplete cytoreduction and G3-5 morbidity (NCI-CTCAE v3). We determined control limits setting the type I/II error rates and unacceptable odds ratios (ORs) for the outcomes being studied. We performed the risk adjustment using logistic regression model.

Results: Rates of incomplete cytoreduction, G3-5 morbidity, and postoperative mortality rates were 10.2%, 28.5%, and 2.1%, respectively. The risk-adjusted sequential probability ratio test curve crossed the lower control limit at the 137th and 149th case, respectively, for incomplete cytoreduction and G3-5 morbidity. At those points, the actual ORs are lower than the prespecified ORs for outcomes being studied. Therefore, we estimated that approximately 140 cases are necessary to ensure surgical proficiency in CRS and HIPEC.

Conclusions: CRS and HIPEC to treat PSM has a steep learning curve requiring 140 procedures to acquire expertise.

(*Ann Surg* 2012;255:348–356)

In the past, peritoneal surface malignancies (PSM) were considered a terminal disease, amenable only to palliation. However, recent reports describe curative treatment options for selected patients with PSM. Over the past 2 decades, a novel therapeutic approach to this clinical entity has emerged that combines cytoreductive surgery (CRS) with hyperthermic intraperitoneal chemotherapy (HIPEC). Theoretically, CRS is performed to treat macroscopic disease, and HIPEC is used to treat microscopic residual disease (RD). The combined treatment has been suggested as the standard of care for pseudomyxoma peritonei (PMP) and peritoneal mesothelioma (PM).[1–4] Promising results have been reported in a subset of peritoneal carcinomatosis from colorectal cancers, gastric cancers, and advanced epithelial ovarian cancers.[5–8]

The achievement of proficiency in the performance of surgical procedures requires a proper, long-lasting, and well-structured training program not only of the surgical staff but also of the multidisciplinary team that cares for the perioperative aspects of the patient. The combined procedure is technically demanding and carries a significant morbidity, even in referral centers.[9,10] Few studies addressing the learning curve of CRS and HIPEC have been conducted thus far.[11–14]

Sequential probability ratio test (SPRT) is a method originally conceived for quality control of military supplies during World War II. The SPRT represents one of the statistical process control tests that has been largely employed in medicine to monitor the safety of medical interventions.[15] It offers an advantage over other statistical process control methods by allowing formal hypothesis testing. This method incorporates selection of type I and II error rates and a threshold of an unacceptable odds ratio (OR) for an outcome. The SPRT is then able to determine whether the hypothesis has been accepted or rejected, or whether further information is required to determine the answer. Moreover, by providing a graphic summary of changes in performance with time, SPRT can alert a surgeon to suboptimal performance. SPRT is also well suited to monitoring surgical learning curves.[16]

The aim of this study was to evaluate the learning curve of a single surgeon undertaking CRS and HIPEC by analyzing, in a multidimensional perspective, the changes in surgical outcomes according to case sequence in a series of patients affected by PSM. Particular emphasis was given to the adjustment of potential confounders that may affect surgical outcomes, and the risk-adjusted (RA) SPRT model was used to assess the extent of surgical experience required to overcome the learning curve.

PATIENTS AND METHODS

All patients were treated under an institutionally approved protocol and provided written informed consent. The eligibility requirements for treatment were as follows: histologically confirmed diagnosis of PSM judged resectable on the basis of clinical and radiological data; age younger than 75 years; no distant metastasis; adequate renal, hematopoietic, and liver functions; and performance status (Eastern Cooperative Oncology Group = 0, 1, or 2).

We obtained data from the prospectively collected institutional database on PSM program of the National Cancer Institute of Milan. In total, 414 cases of PSM treated by 420 CRS and HIPEC procedures represented the study group. Six cases underwent the procedure twice for disease recurrence. The study period lasted from August 1995 to January 2011.

CRS and HIPEC

The technique of CRS has been described elsewhere.[17,18] Briefly, the surgical procedure was conducted with 1 or more of the following steps, depending on disease extension: (1) greater omentectomy, right parietal peritonectomy ± right colon resection; (2) pelvic peritonectomy ± sigmoid colon resection ± hystero-adnexectomy; (3) lesser omentectomy and dissection of the duodenal-hepatic ligament ± antrectomy ± cholecystectomy; (4) right upper quadrant peritonectomy ± Glisson capsule resection; (5) left upper quadrant

From the Department of Surgery, Peritoneal Surface Malignancy Program, National Cancer Institute, Milan, Italy.

Disclosure: The authors declare that there is no funding or financial support to the present work.

Reprints: Marcello Deraco, MD, Fondazione IRCCS Istituto Nazionale Tumori Milano, Via Venezian n.1, 20133 Milano, Italy. E-mail: marcello.deraco@istitutotumori.mi.it.

ISSN: 0003-4932/12/25502-0348
DOI: 10.1097/SLA.0b013e3182436c28

peritonectomy ± splenectomy; and (6) other intestinal resection and abdominal mass resection.

Peritoneal carcinomatosis was quantified according to the Peritoneal Cancer Index (PCI). The mean PCI was 19 (range: 0-39). We classified RD after surgery as follows: cc-0, no RD; cc-1, RD of 0 to 2.5 mm; cc-2, RD of 2.5 mm to 2.5 cm; and cc-3, RD of more than 2.5 cm.[19] We defined incomplete cytoreduction as RD of more than 2.5 mm (cc-2/3).

We performed the HIPEC with closed-abdomen technique using an extracorporeal device, which maintained the intra-abdominal temperature at 42 to 43°C [Performer LRT, RAND, Medolla (MO), Italy]. Chemotherapy regimens used were as follows: cisplatin (CDDP 25 mg · m^{-2} · L^{-1}) and mitomycin-C (MMC 3.3 mg · m^{-2} · L^{-1})[20,21] for colorectal cancer, gastric cancers, and PMP; and CDDP (43 mg/L of perfusate) and doxorubicin (Dx 15.25 mg/L of perfusate)[22] for ovarian cancer and PM. Patients age older than 70 years and those who had undergone previous chemotherapy received a 30% dose reduction of both drugs. The same surgical team guided by a same principal operator (D.M.) performed all operations.

Study Parameters

Adverse events [morbidity, anastomotic leak, systemic toxicity, and procedure-related mortality (PRM)] were graded according to the NCI CTCAE v3 criteria.[23] When more than 2 complications occurred during the same period in 1 patient, we considered the complication deemed as more serious for analysis. Anastomotic leak was defined according to the criteria established elsewhere.[24] The PRM was defined as death occurring during the in-hospital stay after CRS and HIPEC.[25] We defined G3-5 morbidity as the combination of (1) surgical or medical G3-5 morbidity; (2) PRM, and (3) G3-5 systemic toxicity. To assess surgical outcomes according to case sequence, they were divided into 7 subgroups of 50 for cases 1 to 350, whereas cases 351 to 420 were designated as another subgroup.

We assessed risk factors for incomplete cytoreduction among the following clinicopathological parameters: age, sex, ECOG performance status, Charlson Comorbidity Index (combined condition and age-related score),[26] previous systemic chemotherapy, previous surgical score (PSS), tumor histotype, PCI, bowel anastomosis, number of peritonectomy procedures, operating time, and subgroups of case sequence. Risk factors for G3-5 morbidity were searched among the following: age, sex, ECOG performance status, body mass index (BMI), preoperative serum albumin, Charlson Comorbidity Index, previous systemic chemotherapy, PSS, tumor histotype, PCI, bowel anastomosis, number of peritonectomy procedures, operating time, completeness of cytoreduction, HIPEC drug schedule, CDDP dose for HIPEC, and subgroups of case sequence.

Statistical Analysis

We compared each subgroup of case sequence with respect to the surgical outcomes. One-way analysis of variance and χ^2 test were used for analyzing changes of continuous and discrete variables, respectively, according to subgroups of case sequence.

The RA-SPRT plot was used to chart changes in the rates of G3-5 morbidity and rates of suboptimal cytoreduction (cc-2/3) across the case sequence. To elaborate the SPRT analysis, 4 parameters were defined: estimated probabilities of G3-5 morbidity and of cc-2/3 for each case, a prespecified OR for G3-5 morbidity and suboptimal cytoreduction, type I and II error rates. Probability of both type I and type II errors was set at 0.05. From these, 2 control limits [lower boundary of control limit (h_0) and upper boundary of control limit (h_1)] and the cumulative sum of log likelihood ratio with risk adjustment were calculated according to equations outlined in Table 1.[16]

The risk predictions for suboptimal cytoreduction and for G3-5 morbidity were elaborated by multivariate analysis, using the logistic-regression model. Clinical factors were selected as covariates when P values were less than 0.10 on univariate analysis. Using backward stepwise selection, independent factors were analyzed, for which the cutoff P value was set at 0.10 and the significance level for selection and staying was 0.05. The probability of event was calculated, which was adjusted with independent risk factors of incomplete cytoreduction or G3-5 morbidity from the logistic regression model. Discrimination was measured by the area under the receiver operating characteristic curve (area under the curve),[27] with values of 0.5 representing no discrimination and 1.0 representing perfect discrimination. Model fitness was assessed by the Hosmer-Lemeshow goodness-of-fit decile-based test,[28] with $P > 0.05$ indicating acceptable fit.

When creating the RA-SPRT curve, each case was plotted in sequence along the x axis. When a success occurred (no G3-5 morbidity or optimal cytoreduction), a log-likelihood ratio with risk

TABLE 1. Equations and Variables to Construct RA-SPRT Chart for Incomplete Cytoreduction and Morbidity G3-5 in Peritoneal Surface Malignancy

Variable	Meaning/Calculation[16]	cc-2/3	G3-5 Morbidity
OR_0	Odds ratio of acceptable outcome	1	1
OR_1	Odds ratio of unacceptable outcome	1.8	1.4
α	Probability type I error	0.050	0.050
β	Probability type II error	0.050	0.050
h_0	$-\ln[(1-\alpha)/\beta]/\ln(OR_1)$	−5.2	−8.9
h_1	$\ln[(1-\beta)/\alpha]/\ln(OR_1)$	5.2	8.9
$T^{cum}{}_i$	$T^{cum}{}_{i-1} + (O_i - s_i)$		
s_i	$\ln[(1-P_i) + (OR_1 \times P_i)]/\ln(OR_1)$	Variable according to P_i	Variable according to P_i
O_i	Observed outcome for ith case	0 or 1	0 or 1
P_i	Estimated probability of the outcome for the ith case as determined by the risk prediction model (multivariate analysis)	Variable according to the case	Variable according to the case

adjustment (s) was subtracted from the cumulative score. When a failure occurred, the constant $1 - s$ was added to the cumulative score. Thus, a positive slope in the RA-SPRT line indicated failure, whereas a negative slope indicated success.

We set the OR at 1.4 to detect unacceptable increase in the G3-5 morbidity rate, and it was estimated on the basis of the following literature data. The lower major morbidity rate was extracted from the review by Chua et al (28.8%), and the higher from a multicentric experience on PMP (40%).[1,29] The resulting ratio between these 2 values (40.0/28.8) was approximately 1.4.

We set the OR at 1.8 to detect unacceptable increase in the incomplete cytoreduction rate. Rates of incomplete cytoreduction were extracted from the multicentric study conducted by Glehen et al on PSM (15.3%) patients and from the study by Elias et al on PMP patients (27%).[1,30] The resulting ratio between these 2 values (27.0/15.3) was approximately 1.8.

After constructing the RA-SPRT curve if the line crossed the upper decision limit (h_1) from below, the actual OR for outcome is equal to or higher than the prespecified OR, with the probability of type I error of 0.05. If the line crossed the lower decision limit from above, the actual OR for the outcome being studied is less than the prespecified OR, with the probability of type II error of 0.05. When the line is between h_1 and h_0, no statistical inference could be made.

The minimum number of operated cases necessary to define the surgeon expert on CRS and HIPEC was estimated as the point in which the curve of cumulative sum of log likelihood ratio crossed the lower control limit (h_0) established by RA-SPRT model both for G3-5 morbidity and rates of incomplete cytoreduction.

Evolution of the multidisciplinary perioperative care was assessed analyzing the length of stay in the entire case sequence and in the subgroup presenting G3-5 morbidity, using the moving average (MA) method. Because trends in continuous variables could be obscured by individual outlying values, averaging of previous values filters this variation and accentuates any trend in data collected. Creating an average of the values that "moves" with the addition of new data results in a "smoothing" of the value action on the variables being analyzed. An MA order of 20 and 50 was used.[31]

The applied statistical software was SPSS 18.0 (SPSS, Inc., Chicago, IL), and the MA method and RA-SPRT model were calculated by Excel version 2003 (Microsoft Corporation, Redmond, WA). Statistical significance was set at $P < 0.05$.

RESULTS

Patient Characteristics in the Entire Series

The mean age was 52.4 years (range: 22–75 years). The male to female ratio was 155:265. The tumor histology distribution was as follows: PM, 31.2%; PMP, 36.9%; ovarian cancer, 12.6%; peritoneal sarcomatosis, 8.1%; peritoneal carcinomatosis from colorectal cancer, 5.0%; peritoneal carcinomatosis from gastric cancer, 2.9%; and other PSM, 3.6%. The mean Charlson Comorbidity Index was 3 (range: 2-7). Eighty-six percent of the cases presented ECOG performance status equal to 0, and 63% presented PSS less than 2. The mean BMI was 25 (range: 16-41), and the mean preoperative serum albumin level was 4.1 g/dL (range: 2.0-5.7). The mean number of peritonectomy procedures per patient was 7.2 (range: 0-15). The mean operating time was 563 minutes (range: 250-1320), and optimal cytoreduction was achieved in 89.7% of the cases. The mean length of stay was 22 days (range: 7-101). For further details of patients' clinicopathological characteristics and short-term surgical outcome, see Table 2.

Rates of G3-5 morbidity, G3-5 systemic toxicity, and PRM were 28.5%, 4.8%, and 2.1%, respectively. The most common G3-5 surgical complications were gastrointestinal anastomotic leak/perforation (10.4%), intra-abdominal abscess (3.6%), and postoperative bleeding (3.3%). The most common G3-5 systemic toxicities were renal failure (5.7%) and bone marrow suppression (4.5%), pulmonary toxicity (1.4%), and neutropenic infection (0.7%).

TABLE 2. Clinicopathological Characteristics and Short-Term Surgical Outcomes of 420 Peritoneal Surface Malignancy (PSM) Cases Undergoing CRS Plus HIPEC

Characteristic	Values
Total number of cases	414 (420)*
Mean age (range)	52.4 yr (22–75)
Male/female	155/265
ECOG performance status (0 vs 1/2)	353 vs 67
Mean Charlson Comorbidity Index† (range)	3.2 (2–9)
Mean body mass index (range)	25.0 kg/m^2 (16–41)
Mean preoperative serum albumin (range)	4.1 g/dL (2.0–5.7)
Previous surgical score (0/1 vs 2/3)	305 (72.6%) vs 115 (27.4%)
Previous systemic chemotherapy	185 (35.7%)
Tumor histotype	
Peritoneal mesothelioma	131 (31.2%)
Pseudomyxoma peritonei	155 (36.9%)
Ovarian cancer	53 (12.6%)
Peritoneal sarcomatosis	34 (8.1%)
Colorectal cancer	21 (5.0%)
Gastric cancer	12 (2.9%)
Other types of PSM	15 (3.6%)
Mean Peritoneal Cancer Index (range)	19 (0-39)
Surgical procedures	
Peritonectomy	
Diaphragmatic (left/right)	294/302
Glissonian capsule	133
Lesser omentum	282
Pelvic peritoneum	325
Greater omentum	309
Other surgical procedures	
Gastric resection	54
Colon resection	210
Small bowel resection	83
Cholecystectomy	150
Gastrointestinal anastomosis	402
Splenectomy	247
Hysterectomy and/or bilateral salpingo-oophorectomy	76
Para-aortic and/or pelvic lymphadenectomy	36
Ostomy	31
Diaphragmatic resection	15
Other resections	63
Mean no. of peritonectomy procedures per patient	7.2 (0–15)
Mean operating time (range)	563 min (250–1320)
Mean no. of red cell units transfused intraoperatively	2.7 (range: 0–37)
Incomplete cytoreduction (residual disease >2.5 mm)	43 (10.2%)
HIPEC drug schedule (CDDP + Dx vs CDDP + MMC)	181 (43.1%) vs 239 (56.9%)
Mean CDDP dose (range)	190 mg (100–300)
Mean length in the intensive care unit stay (range)	3 (0–10)
Mean length of in-hospital stay (range)	22 days (7–108)

*6 cases were operated twice due to disease relapse.
†Combined condition and age-related score.
CDDP indicates cisplatin; Dx, doxorubicin; MMC, mitomycin-C.

Incomplete Cytoreduction

We performed the risk assessment for incomplete cytoreduction and the univariate analysis identified the following variables: performance status (ECOG) equal to 0, previous systemic chemotherapy, PCI more than 20, mean number of peritonectomy procedures, operating time more than 600 minutes, and subgroups of case sequence.

Then we performed the multivariate analysis using the logistic regression model which identified ECOG performance status equal to 0 [OR = 0.37; 95% confidence interval (CI), 0.17-0.79; P = 0.010], PCI >20 (OR = 3.68; 95% CI, 1.72-7.87; P = 0.001), PM histology (OR = 2.34; 95% CI, 1.10-4.96; P = 0.027), previous systemic chemotherapy (OR = 2.54; 95% CI, 1.24-5.20; P = 0.011), and subgroups of case sequence (OR = 0.81; 95% CI, 0.66-0.99; P = 0.042) as independent risk factors for incomplete cytoreduction (Table 3).

Model evaluation resulted in a Hosmer-Lemeshow goodness-of-fit value of 12.53 (P = 0.129), indicating a good fit. The area under the receiver-operating characteristic curve in the logistic regression model was 0.83 (95% CI, 0.77-0.89), indicating that the model was adequate for discrimination (ie, the ability to assign the correct probability of outcome, eg, the ability to assign the higher probability of incomplete cytoreduction to patients who actually present cc-2/3).

Figure 1A shows the RA-SPRT chart for incomplete cytoreduction. The curve started presenting a positive slope (increase in the rates of cc-2/3) up to the case 31 (point A) and started decreasing after having crossed the upper boundary for unacceptable OR of incomplete cytoreduction (h_1: 5.2). Then it assumed a consistent negative

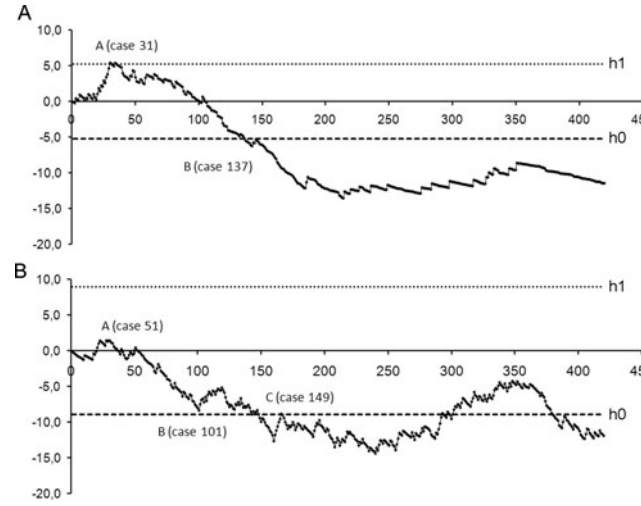

FIGURE 1. The *x* axis represents operation number, not calendar time. The *y* axis represents a cumulative score resultant from calculation involving risk-adjusted log likelihood ratios. h0: lower boundary of control limit; h1: upper boundary of control limit. The lower boundaries were surpassed by the RA-SPRT plots at the 137th (A) and 149th (B) cases, respectively, for incomplete cytoreduction and G3-5 morbidities. A, Incomplete cytoreduction rates (RA-SPRT). B, G3-5 postoperative morbidity rates (RA-SPRT).

TABLE 3. Risk Factors Correlated With Incomplete Cytoreduction and G3-5 Morbidity After CRS Plus HIPEC in 420 Patients Affected by PSM

	Dependent Variables							
	Incomplete Cytoreduction				G3-5 Morbidity			
	Univariate	Multivariate (logistic regression model)			Univariate	Multivariate (logistic regression model)		
Independent Variables	*P**	OR (adjusted)	95% CI	*P*	*P**	OR (adjusted)	95% CI	*P*
Age (>52 vs ≤52)	0.560	—	—	—	0.009	1.95	1.18–3.23	0.009
Gender (male vs female)	0.211	—	—	—	0.406	—	—	—
ECOG performance status (0 vs 1/2)	<0.0001	0.37	0.17–0.79	0.010	0.009	—	—	NS
Charlson Comorbidity Index (>3 vs ≤3)	0.079	—	—	—	0.151	—	—	—
Preoperative serum albumin <3.5 g/dL	—	—	—	—	0.008	2.22	1.14–4.32	0.019
BMI <20 kg/m^2	—	—	—	—	0.961	—	—	—
Tumor histotype	0.197†	2.34	1.10–4.96	0.027	0.080‡	—	—	—
Previous surgical score (0/1 vs 2/3)	0.074	—	—	—	0.604	—	—	—
Previous systemic chemotherapy (yes vs no)	0.002	2.54	1.24–5.20	0.011	0.291	—	—	—
PCI (>20 vs ≤20)	0.001	3.68	1.72–7.87	0.001	<0.0001	2.20	1.33–3.62	0.002
Bowel anastomosis (yes vs no)	0.167	—	—	—	<0.0001	—	—	NS
Mean no. peritonectomy procedures (>7 vs ≤7)	0.001	—	—	NS	0.021	—	—	NS
Completeness of cytoreduction (cc-0/1 vs 2/3)	—	—	—	—	0.124	—	—	—
HIPEC drug schedule (CDDP + Dx vs CDDP + MMC)	—	—	—	—	0.249	1.84	1.09–3.08	0.022
CDDP dose (>240 mg vs ≤240 mg)	—	—	—	—	<0.0001	4.09	2.05–8.13	<0.0001
Operating time (>600 min vs ≤600 min)	0.539	—	—	—	<0.0001	1.79	1.08–2.99	0.025
Subgroups of case sequence§	<0.0001	0.81	0.66–0.99	0.042	0.002	—	—	NS

* χ^2 test.
†PM vs other PSM.
‡Peritoneal mesothelioma and pseudomyxoma peritonei vs other peritoneal surface malignancies.
§Subgroups: 1–50, 51–100, 101–150, 151–200, 201–250, 251–300, 301–350, and 351–420.
CDDP indicates cisplatin; Dx, doxorubicin; MMC, mitomycin-C; OR, odds ratio.

slope and breach the lower boundary (h_0: –5.2) at the case 137 (point B). This point could be considered the breaking one at which the actual OR for incomplete cytoreduction is less than the prespecified OR of 1.8, with a probability of type II error of 0.05.

G3-5 Morbidity

We started identifying the clinicopathological variables with a statistically significant association with this adverse event by univariate analysis: age older than 52 years, performance status (ECOG) equal to 0, PCI more than 20, bowel anastomosis, mean number of peritonectomy procedures, dose of HIPEC CDDP more than 240 mg, operating time more than 600 minutes, and subgroups of case sequence.

Then we performed the multivariate analysis using the logistic regression model which identified age older than 52 years (OR = 1.95; 95% CI, 1.18-3.23; $P = 0.009$), preoperative serum albumin level less than 3.5 g/dL (OR = 2.22; 95% CI, 1.14-4.32; $P = 0.019$), PCI more than 20 (OR = 2.20; 95% CI, 1.33-3.62; $P = 0.002$), HIPEC drug schedule (CDDP and Dx) (OR = 1.84; 95% CI, 1.09-3.08; $P = 0.022$), dose of HIPEC CDDP more than 240 mg (OR = 4.09; 95% CI, 2.05-8.13, $P < 0.0001$), and operating time more than 600 minutes (OR = 1.79; 95% CI, 1.08-2.99; $P = 0.025$) as independent risk factors for G3-5 morbidity (Table 3).

Model evaluation resulted in a Hosmer-Lemeshow goodness-of-fit value of 6.72 ($P = 0.57$), indicating a good fit. The area under the receiver operating characteristic curve in the logistic regression model was 0.73 (95% CI, 0.68-0.78), indicating that the model was adequate for discrimination.

Figure 1B shows the RA-SPRT chart for G3-5 morbidity. After an initial phase of fluctuation, the curve assumed a negative slope from the case 51 (point A) down to the case 101 (point B). The curve fluctuated again, before crossing the inferior boundary ($h_0 = –8.9$) at point C, which corresponded to case 149. Point C could be considered the cutoff at which the actual OR for G3-5 morbidity is less than the prespecified OR of 1.4 with a probability of type II error of 0.05. At no time did the G3-5 morbidity rate breach the unacceptable threshold ($h_1 = 8.9$).

The clinicopathological variables, G3-5 morbidity, and rates of incomplete cytoreduction were analyzed regarding their distribution along the consecutive subgroups of case sequence (Table 4).

Evolution of the Multidisciplinary Perioperative Care

Analysis of the MA of length of stay revealed that this parameter gradually decreased along the case sequence from 19 days to 16 days (Fig. 2A). Two peaks are evident: the first at the 146th case and the second at the 368th case.

Taking into account that length of stay is a direct function of postoperative morbidity, we calculated the MA of this parameter in the subgroup of patients who presented a major complication after the procedure (Fig. 2B). We observed that the MA of length of stay decreased gradually from 37 days down to 17 days (64th complicated case) and then assumed a positive slope, reaching a steady state, around 30 days, from the 80th until the last case.

The Cutoff Point of Learning Curve

Taking into account the RA-SPRT for rates of G3-5 and incomplete cytoreduction, we estimated that the cutoff point for acquisition of proficiency in CRS and HIPEC is between 137 and 149 cases (~140 cases).

DISCUSSION

Our results suggest that approximately 140 cases are necessary for the surgeon to become an expert in CRS and HIPEC for the management of PSM. This cutoff is reproducible as was estimated, using a formal statistical tool (RA-SPRT). The learning curve was analyzed in a relatively large case mix of patients affected by PSM, which encompasses a large variety of tumor histologies. Moreover, all cases were operated by the same surgical team guided by the same principal operator (D.M.).

SPRT has been used for monitoring surgical outcomes and enables surgeons to follow their performance longitudinally against predetermined limits. Because it controls statistically for random variation if a surgeon has 5 major postoperative morbidities within the first 10 cases, it does not lead to an automatic "unacceptable" level as a simple event graph would. By monitoring patient outcomes in a continuous manner, surgeons will be warned if their results are breaching above their predetermined target values. SPRT has begun to be widely used in surgery to monitor postoperative binary outcomes, such as mortality rates.[15,32]

Few articles have addressed the issue of learning curve for CRS and HIPEC.[11,13,14] In a recently published review on the issue, Moradi and Esquivel[12] have pointed to a trend of better outcomes among almost all variables as the surgical team gains more experience with the operation. However, the data were extracted from very heterogeneous patient populations, through nonreproducible statistical tests, hampering any attempt of generalization or comparability of results among the different centers' experiences. The majority of the studies used simple graphs, with arbitrary splitting of the data into groups. Most of these studies performed only univariate analysis, usually without tests for trends. Multivariate techniques were used only in 2 studies, and neither of them used statistical process control tests.

The arbitrarily splitting method to find the cutoff point of the learning curve, if applied to our case sequence, would have generated contradictory results. We observed that the distribution of G3-5 morbidity rates along the subgroups of case sequence presented a progressive increase up to the 350th case, despite the expected progressive accumulation of experience (Table 4). A hypothetical analysis of learning curve, in this case, taking into account only the G3-5 morbidity rate, would falsely lead to the conclusion that the surgical team presented a progressive deterioration of performance until the 350th case and that the learning curve was not overcome until that point. This apparent conundrum could be cleared using the risk-adjustment technique.

In most surgical contexts, the risk of morbidity estimated preoperatively will vary considerably from patient to patient. An adjustment for prior risk is critical to ensure that morbidity rates that seem unusual and arise from differences in patient mix are not erroneously attributed to the surgeon. Some cases are so complex that the probability of adverse event is high even in the best surgical hands. Adjustment of SPRT based on prior risk can be done by adapting the magnitude of the score estimating the patient's surgical risk.

The RA-SPRT plot gives a visual representation on whether the cumulative events are above or below the predicted cumulative events, taking into account the expected risk associated with a particular case. For example, if we consider the RA-SPRT for each patient, the probability for outcome (incomplete cytoreduction or G3-5 morbidity) is determined by the logistic regression model, which, in turn, determines the entity by which the graph ascends or descends. For every successful case performed (complete cytoreduction or uneventful postoperative period), the graph descends and for every case that presented an incomplete cytoreduction or major morbidity, the graph ascends. The "penalty" in the case of unfavorable outcome in low-risk patients is substantially higher than that assigned in the case of unfavorable outcome emerging in high-risk case. There have been only some reports that have assessed learning curves for surgical intervention with risk-adjusted statistical process control test.[33,34]

TABLE 4. Distribution of Risk Factors Along the Case Sequence and Consecutive Subgroups

Subgroups	1–50 Cases	51–100 Cases	101–150 Cases	151–200 Cases	201–250 Cases	251–300 Cases	301–350 Cases	351–420 Cases	*P*
Mean age (yr)	49.4 ± 10.0	53.0 ± 12.9	53.2 ± 12.0	52.7 ± 12.4	54.5 ± 13.3	51.3 ± 13.3	55.7 ± 12.0	56.4 ± 13.8	0.079*
Gender (male vs female)	34%	32%	46%	24%	38%	30%	40%	47%	0.164†
ECOG performance status (0 vs 1/2)	60%	82%	82%	84%	92%	98%	82%	90%	<0.0001†
Mean Charlson Comorbidity Index	2.6 ± 0.7	3.0 ± 1.0	3.0 ± 1.0	3.1 ± 0.9	3.4 ± 1.1	3.0 ± 1.1	3.5 ± 1.3	3.6 ± 1.3	<0.0001*
Mean preoperative serum albumin (g/dL)	3.71 ± 0.5	3.89 ± 0.6	4.24 ± 0.4	3.94 ± 0.4	4.03 ± 0.6	4.51 ± 0.4	4.44 ± 0.6	4.26 ± 0.6	<0.0001*
Tumor histotype (PM and PMP vs other PSM)	28%	24%	62%	74%	86%	86%	88%	87%	0.097†
Previous surgical score (2/3 vs 0/1)	2%	10%	24%	22%	40%	46%	34%	37%	<0.0001†
Previous systemic chemotherapy (yes vs no)	60%	48%	52%	48%	26%	44%	28%	46%	0.008†
Mean PCI	14.4 ± 5.4	17.5 ± 8.3	22.5 ± 8.2	20.8 ± 9.8	19.0 ± 11.5	18.5 ± 11.6	20.8 ± 12.2	16.0 ± 9.2	<0.0001*
Bowel anastomosis	34%	56%	62%	56%	56%	60%	64%	56%	0.097†
Mean no. peritonectomy procedures	3.2 ± 2.6	3.4 ± 3.0	6.8 ± 3.8	7.1 ± 2.9	8.1 ± 2.8	8.9 ± 2.6	10.0 ± 3.0	9.8 ± 2.7	<0.0001*
Completeness of cytoreduction (cc-0/1 vs 2/3)	68%	86%	92%	96%	90%	94%	87%	99%	<0.0001†
HIPEC drug schedule (CDDP + Dx vs CDDP + MMC)	0%	50%	50%	56%	48%	34%	58%	47%	0.001†
Mean HIPEC CDDP dose (mg)	242 ± 37	199 ± 49	216 ± 48	199 ± 35	175 ± 31	182 ± 30	179 ± 41	147 ± 20	<0.0001*
Operating time (min)	412 ± 158	501 ± 173	612 ± 217	565 ± 123	645 ± 121	577 ± 96	613 ± 129	565 ± 106	<0.0001*
Mean number of red cell units transfused	0.8 ± 1.2	2.1 ± 3.0	3.9 ± 6.2	3.3 ± 4.0	2.9 ± 3.3	1.9 ± 2.0	3.4 ± 3.0	3.1 ± 2.5	<0.0001*
G3-5 morbidity rate	26%	12%	40%	34%	30%	32%	44%	14%	<0.0001†

*Analysis of variance.
†χ^2 test.
CDDP indicates cisplatin; Dx, doxorubicin; MMC, mitomycin-C.

The RA-SPRT for G3-5 morbidity demonstrated a negative slope from case 51 to 101 as an expression of an improvement in the surgical performance (Fig. 1B). Then, the curve suffered a positive inflection and fluctuated up to case 134. A definitive descending pattern reappears after that point and breaches the lower decision limit at the 149th case. The curve would have reached the lower boundary some cases before the 149th case if the positive inflection at the 101st case had not succeeded. A possible explanation to this delay in the definition of cutoff point could be found by a careful analysis of the data in Table 4. The proportion of PMP and PM in the first 100 cases is significantly lower than in the successive subgroup (24% from the 51st to 100th case, and 62% in the 101st to 150th case).

In fact, from the year 2000, the PSM program of our institute changed its policy and shifted the allocation of its resources from ovarian cancer, gastrointestinal cancer, and peritoneal sarcomatosis to the treatment of rare diseases, namely, PM and PMP. Although these entities are associated with higher G3-5 morbidity with a borderline significance ($P = 0.080$, Table 3), they could represent an indirect surrogate marker for major complication as they usually present a more widespread disease at diagnosis and require more aggressive surgeries. Approaching PM and PMP, the surgical team had to "reset" the learning process causing a positive inflection in the SPRT curve. Improvement in the performance reemerged from the case 116 and the cutoff point was overcome at the 149th case.

The second parameter employed to evaluate the learning curve was the rates of incomplete cytoreduction, which is a function of 2 major factors: preoperative selection of resectable cases and ability of the surgeon to cytoreduce. Rates of cc-2/3 could eventually increase when the surgeon indicates the procedure to treat a case that he is

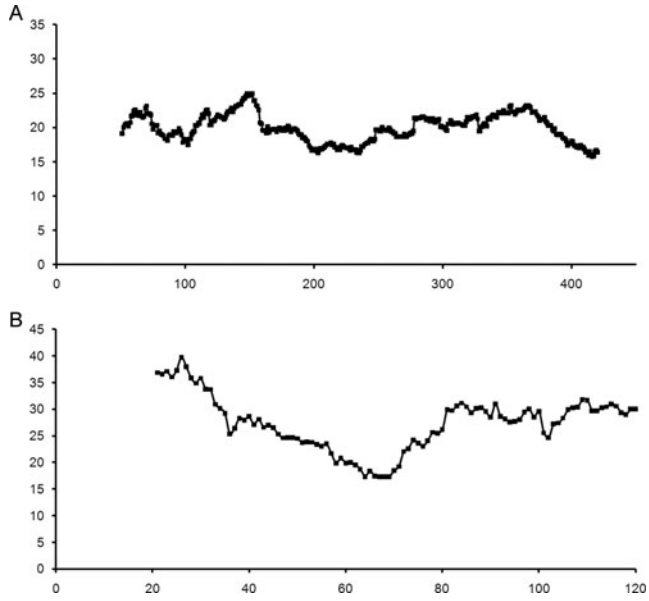

FIGURE 2. The *x* axis represents operation number. The *y* axis represents the moving average of length of stay (days). A, Moving average of length of stay (days) in PSM patients undergoing CRS + HIPEC. B, Moving average of length of stay (days) in 120 patients presenting G3-5 morbidity.

unable to completely cytoreduce. Both capacities, (correct selection of the case and technical proficiency to obtain optimal RD) are expected to improve along the case sequence. Thus, cases deemed unresectable in the beginning of the study could be reclassified as resectable by the same surgeon once he had completed the learning curve. Similarly, cases classified cytoreducible preoperatively in the beginning of the study period could be reclassified as not eligible if evaluated by the same surgeon after the acquisition of proficiency in the combined treatment.

Accordingly, the RA-SPRT revealed an initial increase in the rates of cc-2/3 up to case 31, when the unacceptable upper boundary was surpassed (Fig. 1A). At that moment, due to inability to correctly indicate the procedure or due to inability to cytoreduce completely, the rates of cases with gross RD increased so that the actual OR for incomplete cytoreduction is equal to or higher than the prespecified OR equal to 1.8. From the 32nd case, the rates of successful surgeries gradually increased and the learning curve, assuming a descending feature, crossed the lower boundary at 137th case. At that point, the actual OR for incomplete cytoreduction was less than the prespecified OR of 1.8 and the surgeon acquired the proficiency to adequately cytoreduce PSM.

Patient selection is of utmost importance to ensure maximal oncologic results. In our study, the level of operability was maintained relatively stable throughout the case sequence on the basis of patients' clinical conditions. In general, cases were characterized by good ECOG performance status (84% PS = 0), good nutritional status (mean BMI = 25 kg/m^2, mean preoperative serum albumin = 4.1 g/dL), and low Charlson Comorbidity Index (mean = 3.2). This trend was maintained along the case sequence (Table 4). Smeenk et al conducted a study on 390 cases of PMP and PSM from colorectal cancer.[13] They observed that the complexity of the cases decreased progressively along their series and attributed such tendency to an improvement in patient selection as part of learning process.[13]

In our case sequence, rather than a gradual decrease in the rates of high-risk cases, we observed a progressive expansion in the indication of the procedure with the inclusion of more complex cases, up to 350 cases, as outlined in the Table 4. This does not mean that the surgical team presented a decrease in their capacity to indicate the procedure along the case sequence. Although a patient with PSM from colorectal cancer with limited PCI is considered the best candidate for the combined treatment, the same does not hold true for PM and PMP. The extension of peritoneal carcinomatosis (PCI >20) in PM and PMP is not an exclusion criteria for the combined procedure.

After having crossed the inferior h_0 boundary, the curve continued to decrease until the 240th case. Then, it initiated to present a gradual positive slope, reaching a peak in the 350th case. This ascending behavior, represented a progressive increase in G3-5 morbidity rate (Table 4), which in any case did not breach the upper boundary h_1 of unacceptable G3-5 morbidity rate. This finding seems unexpected as once the surgical team has overcome the learning process, the SPRT curve is intuitively not expected to assume a positive slope any longer. These apparent contradictions could be attributed to the gradual shift toward more complex procedures and the concentration of high-risk patients in the later part of the series as the experience and confidence of the surgical team grew.

In fact, as outlined in Table 4, patients pertaining to the seventh period presented the highest G3-5 morbidity rate with respect to other subgroups of case sequence and were characterized by diseases with higher PCI that required more aggressive surgeries, with more visceral resections, and more anastomosis. Similar findings have been noted in previous studies that addressed other types of surgical interventions. Laparoscopic surgeons, after completing the learning phase, tend to be more liberal in their case selection and are more willing to approach high risk, challenging cases, leading to an apparent rise in the complication rates.[34,35]

Learning curve is a multidimensional concept that comprises several aspects such as patient selection for the procedure, mastering of surgical technique, and maturation of multidisciplinary team responsible for the perioperative care of the patient. The learning curve assessment is complex and should consider not only the short-term surgical outcomes (postoperative adverse event, length of stay) but also the oncologic results, such as disease recurrence and survival rates.[12,13]

To evaluate the effects on oncologic results, taking into account a case sequence composed by several histologies (characterized by a wide range of prognostic profiles), we had to elect the completeness of cytoreduction as an end point to construct the RA-SPRT. The RD is universally accepted as a surrogate marker of survival in all types of PSM. A more detailed assessment of oncologic outcome as a parameter of improvement of surgical performance would require a separate analysis of each tumor type, one by one, and this could be a topic for a future study.

Our study presents some methodological limitations. The composition of case sequence characterized by a high prevalence of PM and PMP (with respect to other forms of PSM) hampers applicability of our results in the analysis of other centers without experience in these malignancies. PM and PMP are very rare diseases with estimated incidences ranging from 0.5 to 3 cases per million per year.[36,37] Further separate analysis of learning curve for PM or PMP, or excluding them from the case mix, could be a reasonable strategy. A more homogeneous case mix in terms of risk profile could eventually provide more consistent results due to improvement of the risk models in predicting the outcomes.

Another limitation of the study is represented by imperfection of risk adjustment for G3-5 morbidity. The natural impossibility to contemplate all possible confounding factors conditions the risk

modeling. Several parameters, especially those related to perioperative care, were not included in our multivariate analysis.

In fact, the long study period encompasses different phases of perioperative care evolution concerning anesthesiological approach, nutritional support technology, and rehabilitation interventions. All these aspects may have exerted a positive effect on the prevention of G3-5 morbidity and the correlated parameters are unfortunately not retrievable due to the retrospective nature of the data. Nevertheless, the multidisciplinary team responsible for the perioperative care presented a continuous improvement translated by a gradual reduction of lengths of stay both in the group presenting G3-5 morbidity and in the entire series (Figs. 2A, B).

Another critical aspect to consider in our data before any attempt of generalization is that related to tutoring. The principal surgeon in our group has set up the PSM program in the early 1990s, after having accomplished fellowship periods in referral centers located in United States and France. The operations started and continued to be performed in our institute without the presence of a more experienced surgeon in peritonectomy procedures. The absence of a tutor at least in the early phase of the PSM program could have implied a delay in the achievement of proficiency in CRS plus HIPEC. In our study, 140 procedures (performed over 7 years) were necessary to ensure expertise in the combined treatment. A well-structured training program including a proficient surgeon tutoring could be crucial to enable the accomplishment of the learning process in a more acceptable time frame. To test such a hypothesis, a similar analysis in a more extensive multi-institutional case sequence would be advisable.

To conclude, using a reproducible statistical technique (SPRT) and taking into account the risk profile of the cases for G3-5 morbidity and rates of incomplete cytoreduction, the gaining of expertise in the surgical management of PSM requires the performance of at least 140 CRS and HIPEC procedures. The plurifactorial nature of the issue warrants further studies to clarify the role of single factor (patient selection vs mastering of the technique vs maturation of multidisciplinary correlated team) and its complex relationship. We hope that these results could be of assistance in setting up accreditation systems for PSM units in the future.

REFERENCES

1. Elias D, Gilly F, Quenet F, et al. Pseudomyxoma peritonei: a French multicentric study of 301 patients treated with cytoreductive surgery and intraperitoneal chemotherapy. *Eur J Surg Oncol.* 2010;36:456–462.
2. Baratti D, Kusamura S, Nonaka D, et al. Pseudomyxoma peritonei: biological features are the dominant prognostic determinants after complete cytoreduction and hyperthermic intraperitoneal chemotherapy. *Ann Surg.* 2009;249:243–249.
3. Yan TD, Deraco M, Baratti D, et al. Cytoreductive surgery and hyperthermic intraperitoneal chemotherapy for malignant peritoneal mesothelioma: multiinstitutional experience. *J Clin Oncol.* 2009;27:6237–6242.
4. Deraco M, Baratti D, Cabras AD, et al. Experience with peritoneal mesothelioma at the Milan National Cancer Institute. *World J Gastrointest Oncol.* 2010;2:76–84.
5. Cao C, Yan TD, Black D, et al. A systematic review and meta-analysis of cytoreductive surgery with perioperative intraperitoneal chemotherapy for peritoneal carcinomatosis of colorectal origin [review]. *Ann Surg Oncol.* 2009;16:2152–2165.
6. Esquivel J, Elias D, Baratti D, et al. Consensus statement on the loco regional treatment of colorectal cancer with peritoneal dissemination [review]. *J Surg Oncol.* 2008;98:263–267.
7. Yan TD, Black D, Sugarbaker PH, et al. A systematic review and meta-analysis of the randomized controlled trials on adjuvant intraperitoneal chemotherapy for resectable gastric cancer [review]. *Ann Surg Oncol.* 2007;14:2702–2713.
8. Raspagliesi F, Kusamura S, Campos Torres JC, et al. Cytoreduction combined with intraperitoneal hyperthermic perfusion chemotherapy in advanced/recurrent ovarian cancer patients: the experience of National Cancer Institute of Milan. *Eur J Surg Oncol.* 2006;32:671–675.
9. Kusamura S, Younan R, Baratti D, et al. Cytoreductive surgery followed by intraperitoneal hyperthermic perfusion: analysis of morbidity and mortality in 209 peritoneal surface malignancies treated with closed abdomen technique. *Cancer.* 2006;106:1144–1153.
10. Deraco M, Baratti D, Kusamura S. Morbidity and quality of life following cytoreduction and HIPEC [review]. *Cancer Treat Res.* 2007;134:403–418.
11. Yan TD, Links M, Fransi S, et al. Learning curve for cytoreductive surgery and perioperative intraperitoneal chemotherapy for peritoneal surface malignancy—a journey to becoming a Nationally Funded Peritonectomy Center. *Ann Surg Oncol.* 2007;14:2270–2280.
12. Moradi BN III, Esquivel J. Learning curve in cytoreductive surgery and hyperthermic intraperitoneal chemotherapy. *J Surg Oncol.* 2009;100:293–296.
13. Smeenk RM, Verwaal VJ, Zoetmulder FA. Learning curve of combined modality treatment in peritoneal surface disease. *Br J Surg.* 2007;94:1408–1414.
14. Mohamed F, Moran BJ. Morbidity and mortality with cytoreductive surgery and intraperitoneal chemotherapy: the importance of a learning curve [review]. *Cancer J.* 2009;5:196–199.
15. Matheny ME, Ohno-Machado L, Resnic FS. Risk-adjusted sequential probability ratio test control chart methods for monitoring operator and institutional mortality rates in interventional cardiology. *Am Heart J.* 2008;155:114–120.
16. Rogers CA, Reeves BC, Caputo M, et al. Control chart methods for monitoring cardiac surgical performance and their interpretation. *J Thorac Cardiovasc Surg.* 2004;128:811–819.
17. Sugarbaker PH. Peritonectomy procedures. *Ann Surg.* 1995;221:29–42.
18. Deraco M, Baratti D, Kusamura S, et al. Surgical technique of parietal and visceral peritonectomy for peritoneal surface malignancies. *J Surg Oncol.* 2009;100:321–328.
19. Jacquet P, Sugarbaker PH. Current methodologies for clinical assessment of patients with peritoneal carcinomatosis. *J Exp Clin Cancer Res.* 1996;15:49–58.
20. Fujimoto S, Takahashi M, Kobayashi K, et al. Combined treatment of pelvic exenterative surgery and intra-operative pelvic hyperthermochemotherapy for locally advanced rectosigmoid cancer: report of a case. *Surg Today.* 1993;23:1094–1098.
21. Baratti D, Kusamura S, Nonaka D, et al. Pseudomyxoma peritonei: clinical pathological and biological prognostic factors in patients treated with cytoreductive surgery and hyperthermic intraperitoneal chemotherapy (HIPEC). *Ann Surg Oncol.* 2008;15:526–534.
22. Rossi CR, Foletto M, Mocellin S, et al. Hyperthermic intraoperative intraperitoneal chemotherapy with cisplatin and doxorubicin in patients who undergo cytoreductive surgery for peritoneal carcinomatosis and sarcomatosis: phase I study. *Cancer.* 2002;94:492–499.
23. Younan R, Kusamura S, Baratti D, et al. Morbidity, toxicity, and mortality classification systems in the local regional treatment of peritoneal surface malignancy [review]. *J Surg Oncol.* 2008;98:253–257.
24. Younan R, Kusamura S, Baratti D, et al. Bowel complications in 203 cases of peritoneal surface malignancies treated with peritonectomy and closed-technique intraperitoneal hyperthermic perfusion. *Ann Surg Oncol.* 2005;12:910–918.
25. Kusamura S, Baratti D, Younan R, et al. Impact of cytoreductive surgery and hyperthermic intraperitoneal chemotherapy on systemic toxicity. *Ann Surg Oncol.* 2007;14:2550–2558.
26. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis.* 1987;40:373–383.
27. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143:29–36.
28. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol.* 1982;115:92–106.
29. Chua T, Yan TD, Saxena A, et al. Should the treatment of peritoneal carcinomatosis by cytoreductive surgery and hyperthermic intraperitoneal chemotherapy still be regarded as a highly morbid procedure? a systematic review of morbidity and mortality. *Ann Surg.* 2009;249:900–907.
30. Glehen O, Gilly FN, Boutitie F, et al. French Surgical AssociationToward curative treatment of peritoneal carcinomatosis from nonovarian origin by cytoreductive surgery combined with perioperative intraperitoneal chemotherapy: a multi-institutional study of 1,290 patients. *Cancer.* 2010;15:5608–5618.
31. Park IJ, Choi GS, Lim KH, et al. Multidimensional analysis of the learning curve for laparoscopic resection in rectal cancer. *J Gastrointest Surg.* 2009;13:275–281.
32. Steiner SH, Cook RJ, Farewell VT. Risk-adjusted monitoring of binary surgical outcomes. *Med Decis Making.* 2001;21:163–169.
33. Tekkis PP, Senagore AJ, Delaney CP, et al. Evaluation of the learning curve in laparoscopic colorectal surgery: comparison of right-sided and left-sided resections. *Ann Surg.* 2005;242:83–91.

34. Son GM, Kim JG, Lee JC, et al. Multidimensional analysis of the learning curve for laparoscopic rectal cancer surgery. *J Laparoendosc Adv Surg Tech A.* 2010;20:609–617.
35. Schlachta CM, Mamazza J, Seshadri PA, et al. Defining a learning curve for laparoscopic colorectal resections. *Dis Colon Rect.* 2001;44: 217–222.
36. Boffetta P. Epidemiology of peritoneal mesothelioma: a review. *Ann Oncol.* 2007;18:985–990.
37. Benedix F, Reimer A, Gastinger I, et al. Study Group Colon/Rectum Carcinoma Primary Tumor. Primary appendiceal carcinoma—epidemiology, surgery and survival: results of a German multi-center study. *Eur J Surg Oncol.* 2010;36:763–771.

